Elsevier Editorial System(tm) for Computer Vision and Image Understanding

Manuscript Draft

Manuscript Number: CVIU-16-93R1

Title: Frequency domain subpixel registration using HOG Phase Correlation

Article Type: Research paper

Keywords: Phase Correlation; registration in frequency domain; subpixel; Fourier; Histogram of Oriented Gradients.

Corresponding Author: Dr. Vasileios Argyriou, PhD

Corresponding Author's Institution: Kingston University

First Author: Vasileios Argyriou, PhD

Order of Authors: Vasileios Argyriou, PhD; Georgios Tzimiropoulos

Abstract: We present a novel frequency-domain image registration technique, which employs histograms of oriented gradients providing subpixel estimates. Our method involves image filtering using dense Histogram of Oriented Gradients (HOG), which provides an advanced representation of the images coping with real-world registration problems such as non-overlapping regions and small deformations. The proposed representation retains the orientation information and the corresponding weights in a multi-dimensional representation. Furthermore, due to the overlapping local contrast normalization characteristic of HOG, the proposed Histogram of Oriented Gradients - Phase Correlation (HOG-PC) method improves significantly the estimated motion parameters in small size blocks. Experiments using sequences with and without ground truth including both global and local/multiple motions demonstrate that the proposed method outperforms the state-of-the-art in frequency-domain motion estimation, in the shape of phase correlation, in terms of subpixel accuracy and motion compensation prediction for a range of test material, block sizes and motion scenarios.

1. Cover Letter

Dear Review Coordinator,

We would like to submit the attached manuscript, "Frequency domain subpixel registration using HOG Phase Correlation," for consideration for possible publication in Elsevier Computer Vision and Image Understanding journal.

This paper has not been published or accepted for publication from another journal or conference. The major contributions of this paper include a novel high-performance version of the phase correlation algorithm based on histogram of oriented gradients (HOG-PC).

Sincerely,

Vasileios Argyriou
Kingston University London
Faculty of Science, Engineering and Computing.
Penrhyn Road
Kingston upon Thames
Surrey KT1 2EE, UK
Email: Vasileios.Argyriou@kingston.ac.uk
Phone: +44 (0) 208 417 2591

Frequency domain subpixel registration using HOG Phase Correlation

June 14, 2016

1 General Comments

We would like first to thank the reviewers for their valuable comments and suggestions.

2 Reviewer #1

In this paper, combination of histograms of oriented gradients (HOG) and phase correlation was used for image registration, which is a reasonable extension of phase correlation approach for more robustness. Although slightly improved results are reported, the paper needs address the following issues to improve its presentation.

1) The motivation: as the improvement in comparison to other approaches is less significant, the author need further justify the motivation of the combination;

The main motivation of our work is to propose a dense HOG-based PC method for sub-pixel translation estimation that is also invariant to small deformations, and performs well when the assumption for translation invariance breaks. To the best of our knowledge this is one of the most important problems in block-based motion estimation, as the problem of noise has been addressed by many authors in the past. The robustness and accuracy of the proposed scheme is particularly evident when small blocks (8×8) are used. Please notice that in this case (see tables 1, 4, 7 and Figs 5-10), the proposed HOG phase correlation significantly outperforms other methods.

More details were added in line 45, *The main point of this work is to propose...*

2) Need compare the complexity with other relevant algorithms, especially those from Ren, GC, NGC and Xiaohua.

Comments on the complexity of the relevant algorithms were added and a new table with times was included (line 306, table 14).

Overall the complexity of the proposed HOG-PC is higher compared to the other approaches due to the computational power required for the pre-processing stage and the estimation of the dense HOG transform. In the current architecture we did not considered any parallel implementation but if a GPU HOG [1] was used it could be no actual difference among them. [1] Victor Prisacariu and Ian Reid, "FastHOG - a real-time GPU implementation of HOG", Department of Engineering Science, Oxford University", 2310/09", 2009

3) Show results from Ren in Figs 11, 13 and Fig 14. You can test on cases with more noise for better comparisons

All the above figures were updated including more noise (pages 21, 26 and 27). In the case of real videos the proposed method outperforms the other methods. Regarding the MRI images since the image/block size is significantly higher most of the approaches are very close in terms of accuracy with Ren's method to outperform especially in very large levels of noise.

4) Show results under different noise rather than Gaussian

Experiments were performed using different levels of motion blur noise. From the obtained results it can be observed that the proposed HOG-PC method outperforms the others especially in very small block sizes.

The updates are in lines 250 and 301, *Furthermore, experiments with motion blur present were performed...* and tables 4-9.

Furthermore, experiments were performed with 8 different levels of motion blur. In each case... and table 13.

5) Any way to extend the proposed approach to deal with other motions such as rotation and zooming?

Thank you for the suggestion, this is something that we have also thought about, however it is not clear how to apply the principles of scale and rotation estimation in the frequency domain for multi-channel representations like the dense HOG (for up to 2-channel representation one can use a complex number representation). This is an interesting extension and we hope to pursue it in the future.

6) If the estimated motion is block by block, how do you distinguish from background motion and foreground motion? Also will you do block by block motion compensation?

We did not consider to distinguish foreground background motion in this work and more details were added explaining that the estimated motion is block based and that the motion compensation was block by block too.

7) Check the format and English usage.

Typos were fixed and some parts were rephrased.

3 Reviewer #2

This paper presented a new motion estimation method by combining the power of HOG representation and phase correlation. Generally this paper is well written and easy to follow. I have one concern regarding the computational time, which is an important factor in video processing. However, I didn't see the results on computation cost.

As it was mentioned above comments on the complexity of the relevant algorithms were added and a new table with times was included (line 306, table 14).

Overall the complexity of the proposed HOG-PC is higher compared to the other approaches due to the computational power required for the pre-processing stage and the estimation of the dense HOG transform. In the current architecture we did not considered any parallel implementation but if a GPU HOG [1] was used it could be no actual difference among them. [1] Victor Prisacariu and Ian Reid, "FastHOG - a real-time GPU implementation of HOG", Department of Engineering Science, Oxford University", 2310/09", 2009

Please give more detail information about the medical images used in the last experiment. For example, what is the image resolution, are they 3D images or the video clip? In medical imaging, it is very difficult to get the ground truth. Please explain how do you obtain the ground truth in sub-voxel accuracy.

More details on the MRI images were added in line 218.

The images show real MRI data from a grapefruit that was acquired using a production quality Fast Spin Echo (FSE) sequence on a GE (Faireld, CT, USA) Signa Lx 1.5 Tesla MRI scanner. The 256×256 *pixel images cover a 16* cm^2 *FOV corresponding to a 0.0625mm square per pixel. Five images were acquired with the fruit at dierent positions in the FOV, by manually moving the scanner table.*

Also, Gaussian noise is not common in real clinical applications. On the contrary, the noise is highly related with machine (scanner). Actually motion blur is more common than noise. It would be great if the authors can show the motion estimation accuracy in that scenario.

As it was mentioned above experiments were performed using noise generated by blur filter. From the obtained results it can be observed that the proposed HOG-PC method outperforms the others especially in very small block sizes.

The updates are in lines 250 and 301, *Furthermore, experiments with motion blur present were performed...* and tables 4-9.

Furthermore, experiments were performed with 8 different levels of motion blur. In each case... and table 13.

Highlights

A dense HOG Phase Correlation method

Invariant to small deformations

Performs well when the assumption for translation invariance breaks

The robustness and accuracy is particularly evident when small blocks

Frequency domain subpixel registration using HOG Phase Correlation

Vasileios Argyriou and Georgios Tzimiropoulos

Kingston University and University of Nottingham

Abstract

We present a novel frequency-domain image registration technique, which employs histograms of oriented gradients providing subpixel estimates. Our method involves image filtering using dense Histogram of Oriented Gradients (HOG), which provides an advanced representation of the images coping with real-world registration problems such as non-overlapping regions and small deformations. The proposed representation retains the orientation information and the corresponding weights in a multi-dimensional representation. Furthermore, due to the overlapping local contrast normalization characteristic of HOG, the proposed Histogram of Oriented Gradients - Phase Correlation (HOG-PC) method improves significantly the estimated motion parameters in small size blocks. Experiments using sequences with and without ground truth including both global and local/multiple motions demonstrate that the proposed method outperforms the state-of-the-art in frequency-domain motion estimation, in the shape of phase correlation, in terms of subpixel accuracy and motion compensation prediction for a range of test material, block sizes and motion scenarios.

Keywords: Phase Correlation, registration in frequency domain, subpixel, Fourier, Histogram of Oriented Gradients.

1. Introduction

A critical component of various high-level computer vision and video processing systems is motion estimation and registration. To perform image registration, we usually assume that the input images are related by a parametric

geometrical transformation. Then, in order to obtain the unknown motion parameters, an optimisation approach is applied on a matching criterion. Pure translation is assumed in this work, which is fundamental in a number of applications such as standards conversion, noise reduction, image super-resolution, medical image registration, restoration, and compression. In such systems, motion compensated prediction is widely used for filtering and redundancy reduction purposes. International standards for video communications such as MPEGx and H.26x employ motion compensation prediction, which is based on regular block-based partitions of incoming frames.

Recently there has been a lot of interest in motion estimation techniques operating in the frequency domain. Perhaps the best-known method in this class is phase correlation [1, 2], which has become one of the motion estimation methods of choice for a wide range of professional studio and broadcasting applications [3]. Phase Correlation (PC) and other frequency domain approaches (that are based on the shift property of the Fourier Transform (FT)) offer speed through the use of FFT routines and enjoy a high degree of accuracy featuring several significant properties: immunity to uniform variations of illumination, insensitivity to changes in spectral energy and excellent peak localization accuracy. Furthermore, it provides sub-pixel accuracy that has a significant impact on motion compensated error performance and image registration for super-resolution and other applications, as theoretical and experimental analyses have suggested [4]. Sub-pixel accuracy mainly can be achieved through the use of bilinear interpolation, which is also applicable to frequency domain motion estimation methods.

One of the main issues of frequency domain registration methods is that in order to obtain reliable motion estimates large blocks of image data are required. Although this requirement is not an issue when there is a single motion, it causes problems when multiple motions are present and affects the accuracy and the overall motion compensated error (especially at the motion borders). On the other hand, reducing the block size increases the sensitivity to noise and reduces the amount of useful image information. Therefore to circumvent the problem,

selecting useful and reliable features is essential. In computer vision and image processing, histogram of oriented gradients (HOG) [5] is a feature descriptor that is invariant to geometric and photometric transformations used mainly for object recognition. Histogram of oriented gradients describe local shapes within an image by the distribution of intensity gradients. The image is divided into cells, and for the pixels within each cell, a histogram of gradient directions is calculated. The local histograms can be normalized by calculating a measure of the intensity across a larger block over a set of neighbouring cells providing invariance to changes in illumination and shadowing.

The main point of this work is to propose a dense HOG-based PC method that is invariant to small deformations, and performs well when the assumption for translation invariance breaks. To the best of our knowledge this is one of the most important problems in block-based motion estimation, as the problem of noise has been addressed by many authors in the past. Additionally, the limitations of frequency based methods when small blocks are used is key part of the motivation of the combination, since HOG transform provides an extra advantage in very small block sizes. In more details, in this paper we introduce a novel high-performance version of the phase correlation algorithm based on histogram of oriented gradients (HOG-PC). The key advances introduced by this paper are the use of a dense histogram of oriented gradients to represent the images. Note that the proposed dense representation is quite different from the traditional representation of a block (or patch) based on HOG. The latter achieves invariance to small translational displacements and hence does not appear to be suitable for motion estimation. In contrast, we propose to use a very dense representation by calculating a descriptor per pixel. This allows us to interpret the obtain representation as a multi-channel block representation. Then, motion estimation is performed by correlating the multi-channel representations from two blocks. Our main contribution lies in showing that this representation not only can recover translational motion very accurately but is also better able to cope with real-world registration problems such as non-overlapping regions, small deformations but also white noise. Furthermore,

due to the overlapping local contrast normalization characteristic of HOG, the proposed HOG-PC method improves significantly the estimated motion parameters in smaller size blocks. Finally, subpixel accuracy is obtained through the use of simple interpolation schemes [6, 7]. Experiments with ground truth data, noisy MR images, and real video sequences have shown that our scheme performs significantly better than recently proposed subpixel extensions to the phase correlation method.

This paper is organised as follows. In Section 2, we review the state-of-the-art in sub-pixel motion estimation using phase correlation. In Section 3, we discuss the principles of the proposed HOG-PC and the key features of this method are analysed. In Section 4 we present experimental results while in Section 5 we draw conclusions arising from this paper.

2. Related work

In this section, a brief review of current state-of-the-art Fourier-based methods for image registration is presented [8]. In many practical encoder implementations, sub-pixel motion estimation is achieved by straightforward extensions to the baseline integer-pixel block-matching algorithm mainly through the use of bilinear interpolation. Interpolation in the data domain is also applicable to frequency domain motion estimation methods such as phase correlation. Moreover such an approach cannot provide estimates of true floating-point accuracy, only approximations to the nearest negative power of two. To circumvent the above difficulties associated with interpolation, alternative approaches have been developed.

Recently, several subpixel extensions have been proposed [9, 10, 11, 12, 13, 14]. In [15], Hoge proposes to perform the unwrapping after applying a rank-1 approximation to the phase difference matrix. In more detail, Hoge presents a so-called Subspace Identification Extension method, which is based on the observation that a 'noise-free' phase correlation matrix (i.e. a matrix computed from shifted replicas of the same image) is a rank one, separable-variable matrix.

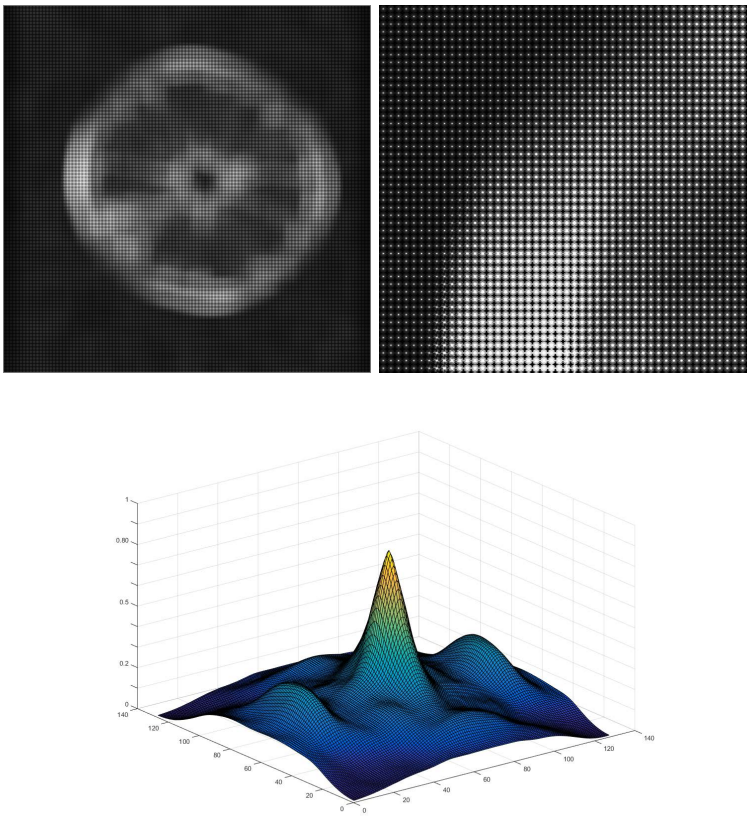

Figure 1: An example of the dense HOG features represented with orientation histograms (top) and a obtained correlation surface (bottom).

For a "noisy phase correlation matrix (i.e. a matrix computed from consecutive frames of a moving sequence), the sub-pixel motion estimation problem can be recast as finding the rank one approximation to that matrix. This can be achieved by using Singular Value Decomposition (SVD) followed by the identification of the left and right singular vectors. These vectors allow the construction of a set of normal equations, which can be solved to yield the required estimate. The work in [16] is a noise-robust extension to [15], where noise is assumed to be AWGN. The authors in [17] derive the exact parametric model of the phase difference matrix and solve an optimization problem for fitting the analytic model to the noisy data.

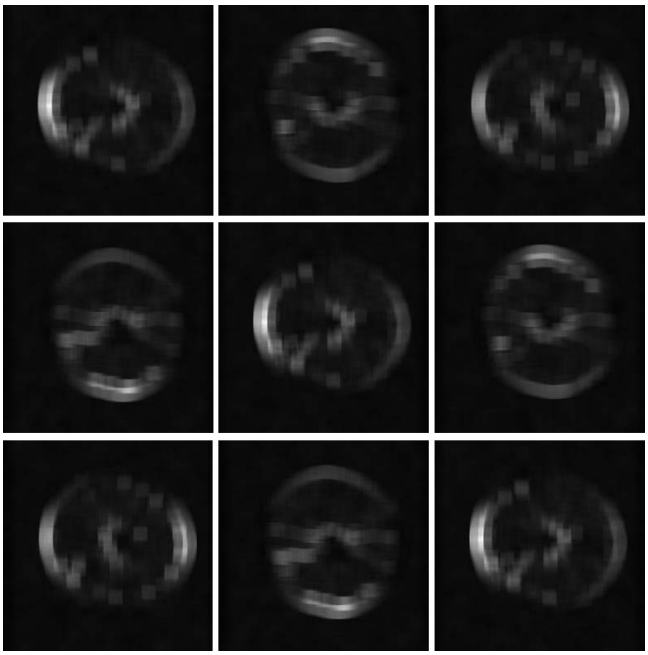

Figure 2: The first $\theta = 9$ channels of the dense HOG that were used in the proposed HOG-PC.

To estimate the subpixel shifts, Stone et al. [18] fit the phase values to a 2-D linear function using linear regression, after masking out frequency components corrupted by aliasing. The method inevitably requires 2-D phase unwrapping which is a difficult ill-posed problem, while the parameters controlling masking are arbitrarily chosen and require fine tuning. Thus, after obtaining an integer-precision alignment of the input images their method takes steps towards alias cancellation by eliminating certain spectral components of each of the two input images. Elimination is based on two criteria: (i) radial distance of a spectral component from the component located at the origin and (ii) magnitude of a spectral component in relation to a threshold. The latter is dynamically determined as follows. Spectral components are sorted by magnitude and are progressively eliminated starting with the lowest. The authors claim that there exists a range in which the accuracy of the computed motion estimate becomes stable and independent of the degree of progressive elimination. This stability range is indirectly used to determine the required threshold. A plane fitting

operation on the frequencies that have survived the above two criteria yields the required motion estimates. An extension to the method for the additional estimation of planar rotation has been proposed in [19].

Foroosh et al. [20] showed that the phase correlation function is the Dirichlet kernel and provided analytic results for the estimation of the subpixel shifts using the *sinc* approximation. According to [20], images mutually shifted by a sub-pixel amount can be assumed as having been obtained by an integer pixel displacement on a higher resolution grid followed by subsampling. This assumption allows the analytic computation of the normalised cross-power spectrum as a polyphase decomposition of a filtered unit impulse. The authors demonstrate that the signal power of the resulting phase correlation surface is not concentrated in a single peak but is distributed to several coherent peaks adjacent to each other. The authors further show that this amounts to a Dirichlet kernel, which can be closely approximated by a *sinc* function. This approximation allows for the development of a closed-form solution for the sub-pixel shift estimate.

Finally, a fast method for subpixel estimation based on FFTs has been proposed in [21]. Notice that the above methods either assume aliasing-free images [20, 22, 21, 17], or cope with aliasing by frequency masking [18, 16, 15, 19], which requires fine tuning.

3. HOG-PC for Subpixel Registration

Let $I_i(\mathbf{x})$, $\mathbf{x} = [x, y]^T \in \mathcal{R}^2$, $i = 1, 2$ be two image functions, related by an unknown translation $\mathbf{t} = [t_x, t_y]^T \in \mathcal{R}^2$

$$I_2(\mathbf{x}) = I_1(\mathbf{x} - \mathbf{t}) \tag{1}$$

To estimate the translational displacement, we use phase based correlation schemes. Each image $I_i(\mathbf{x})$ can be considered as a continuous periodic image function with period $T_x = T_y = 1$, [23]. The Fourier series coefficients of I are given by

$$F_I(\mathbf{k}) = \int_\Omega I(\mathbf{x}) e^{-j\omega_0 \mathbf{k}^T \mathbf{x}} d\mathbf{x} \tag{2}$$

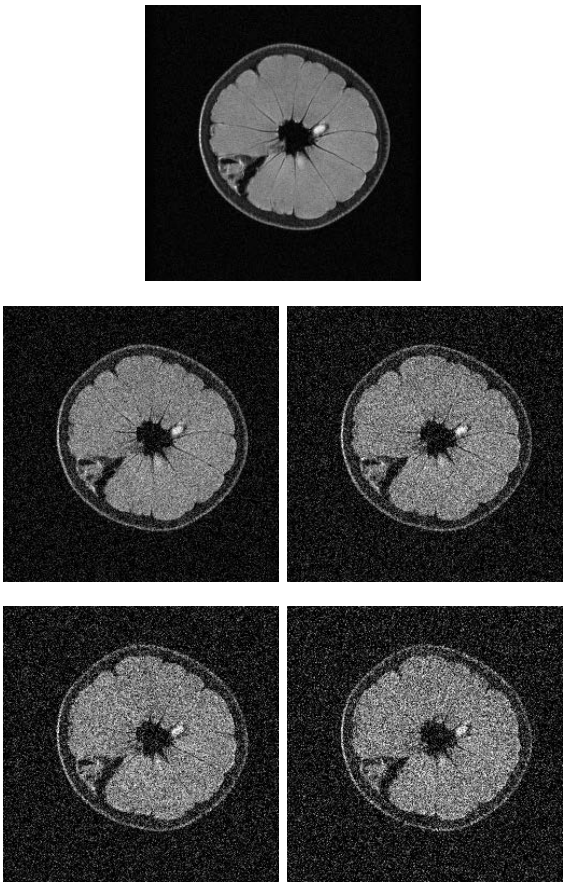

Figure 3: An example of the MRI data without and with noise of different levels (0.01,0.02,0.03,0.04).

where $\Omega = \{\mathbf{x} : -1/2 \leq \mathbf{x} \leq 1/2\},\ \mathbf{k} = [k,l]^T \in Z^2$ and $\omega_0 = 2\pi$. If we sample I at a rate N with a 2-D Dirac comb function $D(\mathbf{x}) = \sum_{\mathbf{s}} \delta(\mathbf{x} - \mathbf{s}/N)$, we obtain a set of $N \times N$ discrete image values $I_1(\mathbf{m}) = I(\mathbf{m}/N),\ \mathbf{m} = [m,n]^T \in \mathcal{R}^2$ and $-N/2 \leq \mathbf{m} < N/2$, [23]. Using D, we can write the DFT of I_1 as

Figure 4: A frame of each video sequence that was used in our evaluation process.

$$\begin{aligned}\hat{I}_1(\mathbf{k}) &= \sum_{\mathbf{m}} I_1(\mathbf{m}) e^{-j(2\pi/N)\mathbf{k}^T\mathbf{m}} \\ &= \int_{\Omega} D(\mathbf{x}) I(\mathbf{x}) e^{-j(2\pi/N)\mathbf{k}^T\mathbf{x}} d\mathbf{x} \\ &= F_I(\mathbf{k}) \star \sum_{\mathbf{s}} e^{-j(2\pi/N)\mathbf{k}^T\mathbf{s}/N} \\ &= N^2 \sum_{\mathbf{s}} F_I(\mathbf{k} - \mathbf{s}N) \end{aligned} \tag{3}$$

where $-N/2 \leq \mathbf{k} < N/2$ and $\star$ denotes convolution.

Moving to the shifted version of the image [23], given by the equation (1) with $\mathbf{t} = [t_x, t_y]^T, \{\mathbf{t} : -1 < N\mathbf{t} < 1\}$. Sampling with D in a similar fashion we get I_2 and its DFT is given based on the Fourier shift property by

$$\hat{I}_2(\mathbf{k}) = N^2 \sum_{\mathbf{s}} F_I(\mathbf{k} - \mathbf{s}N) e^{-j(2\pi/N)(\mathbf{k}-\mathbf{s}N)^T(N\mathbf{t})} \tag{4}$$

Table 1: Average PSNR (dB) values for all the video sequences and block size 8×8.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	42.4208	43.0771	**43.2201**	39.8234	40.8600	40.9810	38.8748	41.7708
Flower	23.6991	25.2763	**26.0815**	17.5719	19.8677	20.0458	15.9774	20.5315
Football	18.7215	18.7584	**18.8040**	17.9245	18.2426	18.2003	17.7605	18.3034
Foreman	27.5536	28.3746	**28.6088**	24.6746	26.1103	26.0809	24.2605	35.9339
Highway	31.5640	32.1402	**32.3362**	30.4101	31.2490	30.9609	29.0454	32.8063
MobCal	21.6909	23.5442	**23.9349**	16.6177	19.3245	19.4276	14.9527	21.8954

Table 2: Average PSNR (dB) values for all the video sequences and block size 16×16.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	43.1677	**43.2094**	43.1455	43.1980	41.4381	41.8490	41.2989	41.0237
Flower	28.3028	28.5663	**28.7076**	23.8995	25.9038	25.3230	15.7029	24.4162
Football	19.6813	19.7728	**20.0338**	18.1636	18.6105	18.4471	17.6603	18.5527
Foreman	29.5387	29.7872	**30.2192**	25.9039	27.3208	28.0282	24.2521	36.3688
Highway	32.5355	32.8818	**33.6166**	31.3321	31.7945	30.7668	28.5041	32.8148
MobCal	24.5285	24.8592	**24.9101**	21.7679	23.0760	23.6421	14.6892	22.7079

Assuming no aliasing and combining equations (3) and (4) we have

$$\hat{I}_2(\mathbf{k}) = \hat{I}_1(\mathbf{k}) e^{-j(2\pi/N)\mathbf{k}^T(N\mathbf{t})} \tag{5}$$

Note that the well-known shift property of the DFT refers to integer shifts and does not assume aliasing-free signals. Hereafter, we assume that our sampling device eliminates aliasing. Traditionally to estimate the translational displacement, we use phase correlation (PC), which is perhaps the most widely used correlation-based method in image registration. It looks for the maximum of the phase difference function which is defined as the inverse FT of the normalized cross-power spectrum [1]

$$\mathrm{PC}(\mathbf{u}) \triangleq F^{-1}\left\{\frac{\hat{I}_2(\mathbf{k})\hat{I}_1^*(\mathbf{k})}{|\hat{I}_2(\mathbf{k})||\hat{I}_1^*(\mathbf{k})|}\right\} = F^{-1}\{e^{j\mathbf{k}^T\mathbf{t}}\} = \delta(\mathbf{u}-\mathbf{t}) \tag{6}$$

where $*$ denotes complex conjugate and F^{-1} the inverse Fourier transform.

Table 3: Average PSNR (dB) values for all the video sequences and block size 32×32.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	42.1412	**42.1554**	41.9947	41.0056	39.8609	41.4001	38.2170	40.5871
Flower	28.3463	**28.3894**	28.3483	27.0887	27.8201	27.1752	15.6074	25.6091
Football	20.4393	20.5912	**20.7832**	18.7080	19.5177	18.9756	17.5200	18.6555
Foreman	30.7678	31.1017	**31.4673**	27.6488	29.5744	29.4516	24.2495	36.2586
Highway	32.9365	33.1902	**33.8099**	32.1031	32.3265	29.5379	28.5122	33.8376
MobCal	24.1923	**24.2241**	24.1245	23.7131	23.5444	23.5672	14.5033	23.4697

Table 4: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 8×8 with 0.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	46.9358	47.9122	**48.1894**	43.8985	44.0877	43.5653	43.4136	46.3062
Flower	28.4947	30.4891	**31.5629**	22.6101	24.8113	22.4272	21.7711	26.0737
Football	18.2997	18.2785	**18.3126**	17.3936	17.6461	17.3321	17.3271	17.8680
Foreman	31.4301	32.5690	**33.3006**	28.3002	29.1089	28.1744	27.9264	30.5642
Highway	36.2005	36.8582	**37.1551**	34.6884	35.0371	34.3573	33.5344	36.0504
MobCal	25.3502	27.3564	**27.6531**	20.4433	22.0091	20.0361	19.5922	23.6230

3.1. Proposed methodology for HOG-PC

In this section, we introduce the proposed phase correlation algorithm based on histogram of oriented gradients (HOG-PC). Note that the proposed dense representation is quite different from the traditional representation of a block (or patch) based on HOG. The latter achieves invariance to small translational displacements and hence does not appear to be suitable for motion estimation. In contrast, we propose to use a very dense representation by calculating a descriptor per pixel. This allows us to interpret the obtain representation as a multi-channel block representation. Then, motion estimation is performed by correlating the multi-channel representations from two blocks. Our main contribution lies in showing that this representation not only can recover translational motion very accurately but is also better able to cope with real-world registration problems such as non-overlapping regions small deformations but also white

Table 5: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 16×16 with 0.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	47.7864	47.7916	**47.8869**	44.5982	43.9858	44.7131	43.2949	46.1797
Flower	36.2843	36.8024	**37.0960**	24.6397	28.5770	32.9379	21.2003	30.4665
Football	19.2769	19.3445	**19.6467**	17.4480	17.6951	17.9852	17.2224	17.8443
Foreman	33.8094	33.8268	**34.7385**	28.4528	28.8829	31.1467	27.7902	30.1301
Highway	37.0921	37.7399	**38.3794**	35.4120	35.6794	34.8676	33.3544	36.5686
MobCal	29.0098	29.5434	**29.7630**	22.9852	22.6976	26.2229	19.3175	24.6576

Table 6: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 32×32 with 0.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	**47.2598**	47.0844	46.9395	42.8979	43.2420	45.2531	42.8459	45.1704
Flower	**37.7471**	37.7145	37.5289	34.4633	31.8037	35.6403	21.0538	33.7585
Football	20.0454	20.2501	**20.4676**	17.9476	17.4846	18.2341	17.0028	17.5864
Foreman	34.4031	34.5012	**35.1448**	29.4785	28.9129	32.9204	27.6481	29.9641
Highway	38.3238	38.8231	**39.3635**	36.4025	36.3173	34.2663	33.3791	36.9849
MobCal	29.3901	**29.4741**	29.3422	27.5164	23.2643	27.5576	18.9994	25.1790

noise. Furthermore, due to the overlapping local contrast normalization characteristic of HOG, the proposed HOG-PC method improves significantly the estimated motion parameters in smaller size blocks. Finally, subpixel accuracy is obtained through the use of simple interpolation schemes [6, 7].

We first describe the traditional HOG descriptor. HOG uses the normalized combination of gradient vectors from a given number of pixels to build up a histogram of binned angles that relate to the feature. The process begins by breaking the image up into set features spaces f comprised of a number of cells c, which in turn is made up of pixels. For each pixel within a cell the filter mask $[-1, 0, 1]$ is applied to its neighbouring pixels giving us the gradient vector $\vec{g}$.

The magnitude $\|\vec{g}\|$ of the gradient vector is obtained and its orientation

Table 7: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 8×8 with 1.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	50.2757	51.0925	**51.3390**	47.4929	47.0069	47.3463	47.2890	49.0431
Flower	31.4674	32.8165	**33.3934**	27.2786	27.6245	27.2046	26.9964	29.2145
Football	**19.4725**	19.3277	19.3736	18.5606	18.6270	18.5354	18.5142	18.8796
Foreman	32.3824	33.2128	**33.7498**	29.9563	29.9954	29.8935	29.6478	31.4852
Highway	38.8627	39.1613	**39.4571**	37.5952	37.3573	37.5161	36.8784	38.2206
MobCal	28.1125	29.4835	**29.6976**	24.4869	24.6221	24.3584	24.1381	26.2775

Table 8: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 16×16 with 1.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	51.4247	51.3765	**51.6485**	47.3513	46.8294	50.0535	47.1346	48.8104
Flower	38.2331	39.0902	**39.7536**	27.5034	27.9538	35.9841	27.0223	29.7296
Football	20.2406	20.3819	**20.6140**	18.4652	18.5598	19.0224	18.4073	18.7527
Foreman	34.5928	34.2964	**35.3176**	29.8665	29.7574	33.3645	29.6204	30.9343
Highway	40.1471	40.3989	**41.0301**	37.7915	37.6942	38.7673	36.9909	38.5279
MobCal	31.6078	32.2799	**32.4236**	24.8317	24.6213	29.8854	24.1093	26.2672

expressed using angle θ.

$$\theta = \tan^{-1}(g_y, g_x) \tag{7}$$

Additionally a weight w is defined for each pixel, which is used to scale its contribution to its cell's histogram. This is given by the mean value of the pixels within a given 2D kernel indicating the density over this area. By applying this weight, the proposed approach provides accurate estimates also in the presence of noise.

Once these values are established the pixels within each cell are binned into a histogram H according to their θ angle. The value added to a bin is given as the weighted magnitude of the vector $w\|\vec{g}\|$. Finally all cell histograms within a multi-dimensional feature H_j are normalised using the L_2 norm.

Table 9: Average PSNR (dB) values for all the video sequences (50 first frames) and block size 32×32 with 1.75 variance motion blur.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Akiyo	**51.2539**	50.7153	50.6264	40.8783	46.4685	49.6346	46.6632	47.7458
Flower	**43.1483**	42.6464	42.5386	30.6144	28.6057	41.1037	27.2853	30.8018
Football	21.1975	21.3593	**21.7731**	18.5759	18.3757	18.6508	18.2522	18.5016
Foreman	35.4305	35.2909	**36.2796**	29.6047	29.6118	33.8368	29.4914	30.6552
Highway	41.0594	41.7453	**42.0343**	37.8387	38.3018	39.9322	37.2725	39.0443
MobCal	**33.2915**	33.2234	33.2526	26.9309	24.4869	30.5041	24.1541	26.2379

$$H_j \rightarrow \frac{H_j}{\sqrt{\|\vec{g}\|_2^2 + e^2}} \tag{8}$$

The obtained features are then vectorised as a $\theta-$dimensional descriptor

$$\vec{d} = \{H_1, ..., H_\theta\} \tag{9}$$

Having defined HOG for a single cell, we now turn to the proposed dense HOG representation. For I_i, $i = 1, 2$, we extract d from (9) at each pixel location $I_i(\mathbf{m})$:

$$H_i(\mathbf{m}) = \{H_{i,1}(\mathbf{m}), H_{i,2}(\mathbf{m}), ..., H_{i,\theta}(\mathbf{m})\} \tag{10}$$

The resulting histograms can be re-arranged as a multi-channel feature representation (see figures 1 and 2).

To estimate the subpixel shift $\mathbf{t}$ from (1) using HOG-PC, we simply compute the correlation between the two multi-channel representation:

$$HOGPC(\mathbf{m}) = \sum_{j=1}^{\theta} H_{1,j}(\mathbf{m}) \star H_{2,j}(-\mathbf{m}) \tag{11}$$

and find $\mathbf{t} = \arg\max_{\mathbf{m}} HOGPC(\mathbf{m})$. We can estimate sub-pixel accuracy registration $\mathbf{t}_0 = (x_0, y_0)$ by fitting a $1D$ kernel to the vicinity of the maximum on the correlation surface. A parametric kernel is used, which can adapt its shape to fit the correlation functions as well as to provide accurate estimates of the

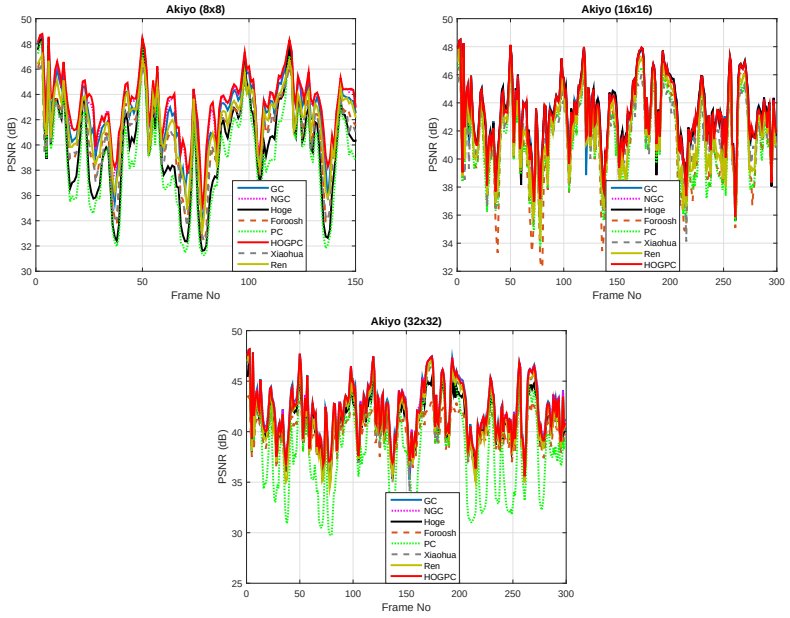

Figure 5: The PSNR values for the Akiyo sequence versus the frame number for all the block sizes.

subpixel shifts. Based on the work in [23] a reasonable choice for our kernel is given by

$$K_{1D}(x;\{x_0,\mathbf{p}\}) = p_1\{1-(p_2(x-x_0))^2\}\frac{1}{\sqrt{2\pi}p_3}e^{\frac{-(x-x_0)^2}{2p_3^2}} \tag{12}$$

which is a simple modification of the mexican hat wavelet [24]. To estimate y_0, we set up a similar problem with the kernel defined as

$$K_{1D}(y;\{y_0,\mathbf{q}\}) = q_1\{1-(q_2(y-y_0))^2\}\frac{1}{\sqrt{2\pi}q_3}e^{\frac{-(y-y_0)^2}{2q_3^2}} \tag{13}$$

Our algorithm estimates the kernel parameters $\{x_0, \mathbf{p} = [p_1, p_2, p_3]^T\}$ and $\{y_0, \mathbf{q} = [q_1, q_2, q_3]^T\}$ in a least-squares sense.

4. Results

To evaluate and illustrate the efficiency of the proposed scheme a comparative study was performed with state of the art frequency domain based tech-

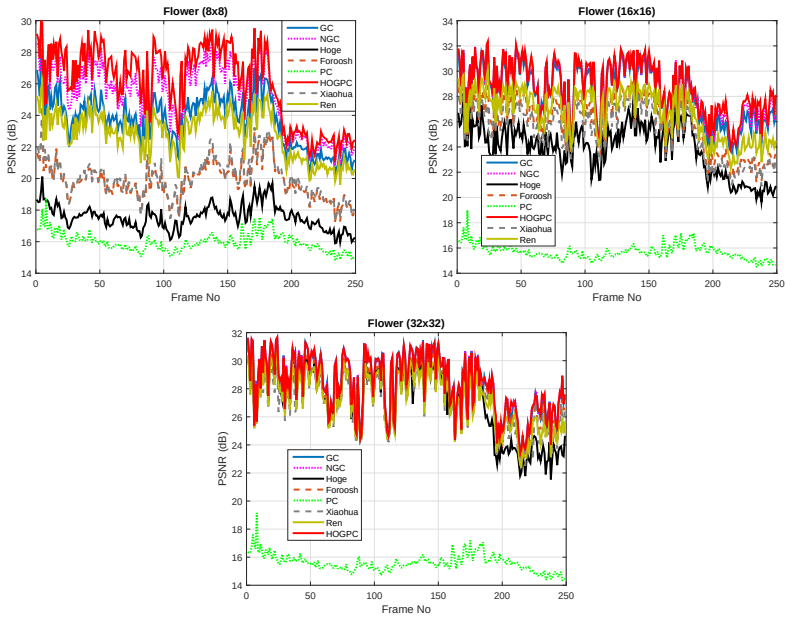

Figure 6: The PSNR values for the Flower sequence versus the frame number for all the block sizes.

niques. Both data with ground truth and video sequences have been used for evaluating the performance. A set of MRI images are employed which have undergone sub-pixel displacement and it is available by the authors in [15] (see figure 3). The images show real MRI data from a grapefruit that was acquired using a production quality Fast Spin Echo (FSE) sequence on a GE (Faireld, CT, USA) Signa Lx 1.5 Tesla MRI scanner. The 256×256 pixel images cover a 16 cm^2 FOV corresponding to a 0.0625mm square per pixel. Five images were acquired with the fruit at dierent positions in the FOV, by manually moving the scanner table. Regarding the real videos the well-known sequences of 'Akiyo', 'Flower', 'Football', 'Foreman', 'Highway' and 'MobCal' were used including $150 - 300$ frames each (see figure 4).

4.1. Video sequences without ground truth

Regarding the real video sequences without ground truth, in order to evaluate the accuracy of the proposed method the visual quality (fidelity) of the

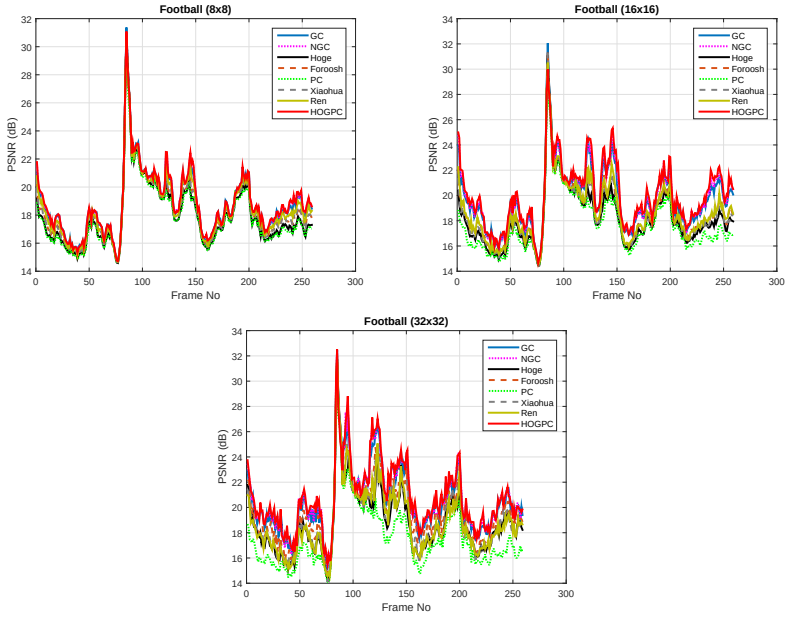

Figure 7: The PSNR values for the Football sequence versus the frame number for all the block sizes.

motion compensated sequence is considered. It is defined as the closeness between the motion compensated frames and the original ones, and the peak signal to noise ratio (PSNR) is used in this work defined by

$$PSNR = 10 \log \left(\frac{255^2}{MSE_I} \right) \tag{14}$$

where MSE_I is the mean square error of the original and motion compensated frames.

The performance of the proposed $HOGPC$ scheme is compared with more than five popular PC based methods [15, 20, 22, 6, 17, 7, 23, 9]. Foroosh's method [20] estimates the subpixel shifts by fitting a *sinc* function to the available correlation samples. Hoge's and Xiaohua's [15, 14] methods are based on frequency masking, phase unwrapping and linear regression, while Ren's [22] approach applies a linear weighting of the height of the main peak on the one hand and the difference between its two neighboring side-peaks on the other.

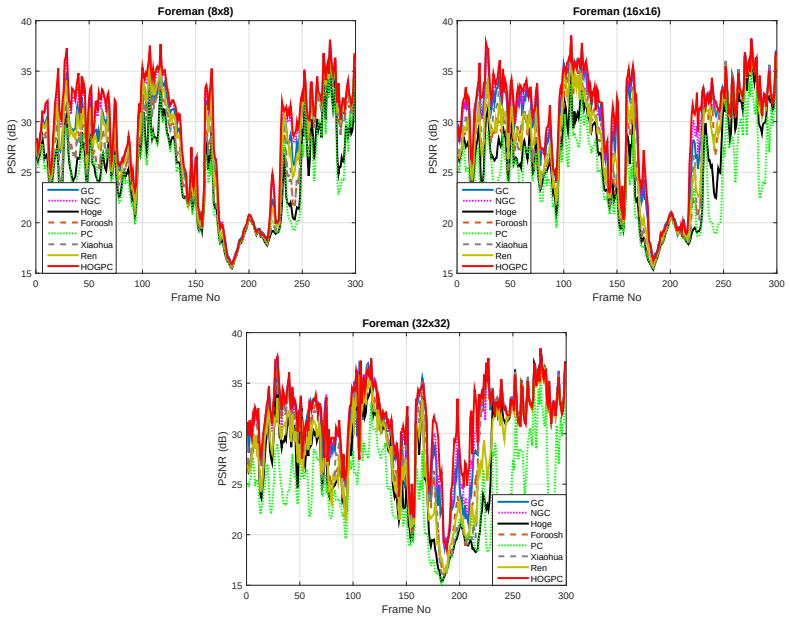

Figure 8: The PSNR values for the Foreman sequence versus the frame number for all the block sizes.

In the second part of our evaluation process, experiments were performed using read video sequences and applying block based motion estimation. The selected block sizes were 32×32, 16×16 and 8×8 pixels and the motion compensated prediction error was estimated for each block size over all the sequences. The average PSNR values are shown in Tables 1,2 and 3 and it can be observed that the proposed approach results the highest values indicating better visual quality. In figures 5, 6, 7, 8, 9, and 10 the PSNR values over time for the video sequences are shown with the proposed scheme to be the most accurate and consistent in comparison with the other state-of-the-art methods. Furthermore, experiments with motion blur present were performed indicating the accuracy of the proposed method especially in the case of small block sizes (e.g. 8×8). The average PSNR values are shown in Tables 4,5 and 6 for motion blur variance equal to 0.75 and in Tables 7,8 and 9 for motion blur variance equal to 1.75.

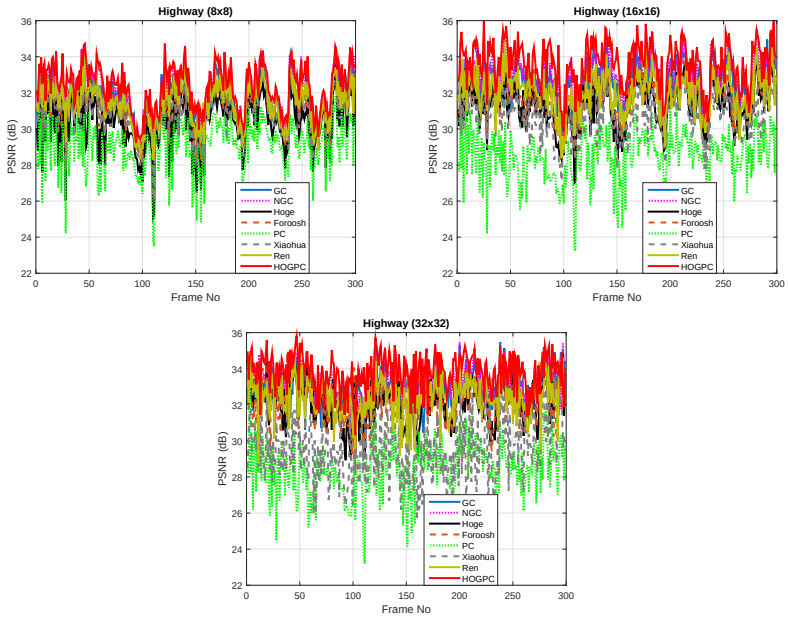

Figure 9: The PSNR values for the Highway sequence versus the frame number for all the block sizes.

Finally, in figure 11 we can see the gain of the *HOGPC* method as a ratio over the other approaches moving from larger to smaller block sizes. As it was expected the ratio increases due to the characteristics of our scheme and HOG. So, since HOG is utilising neighboring information (i.e. surrounding cells) even for small blocks *HOGPC* scheme contains more information allowing more accurate estimates especially if larger motions are present. Furthermore, observing the results in Tables 1,2 and 3 focusing on the proposed method and especially for the Akiyo sequence that is characterised of small motion vectors in average, it shows that HOGPC provides the best results for the case of 8×8 pixels. Also, it outperforms other methods used over larger blocks such as 16×16 pixels, indicating the accuracy of the proposed HOGPC method that exploits the overlapping local contrast normalization characteristic of dense HOG.

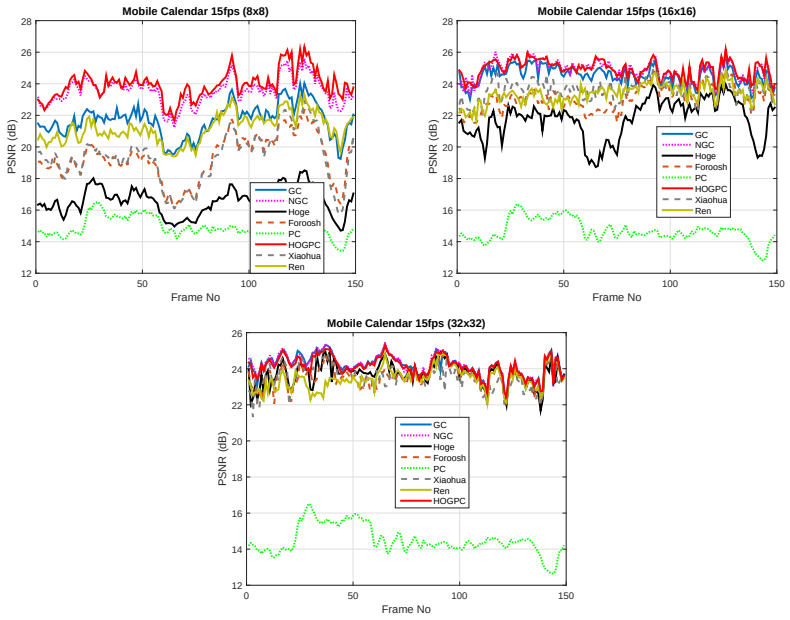

Figure 10: The PSNR values for the MobCal sequence versus the frame number for all the block sizes.

4.2. Real data with ground truth

In the case that ground truth is available, the mean square error (MSE) between the estimated subpixel motion vectors and the ground truth is used as a performance measure. Considering two vectors $\mathbf{u}$ and $\mathbf{v}$ representing the original (ground truth) and the estimated one, respectively, then

$$MSE_{MV} = \frac{1}{n} \sum_{i=x,y} (u_i - v_i)^2 \tag{15}$$

Consequently, a good quality estimate is expected to minimize MSE, which provides the accuracy of the estimates.

In more details, a set of five 256×256 pixel real MR images [15] was used and a sample of them is shown in figure 3. The 5 images yield a total of 10 possible pairwise registrations and the ground truth of the subpixel translations is provided.

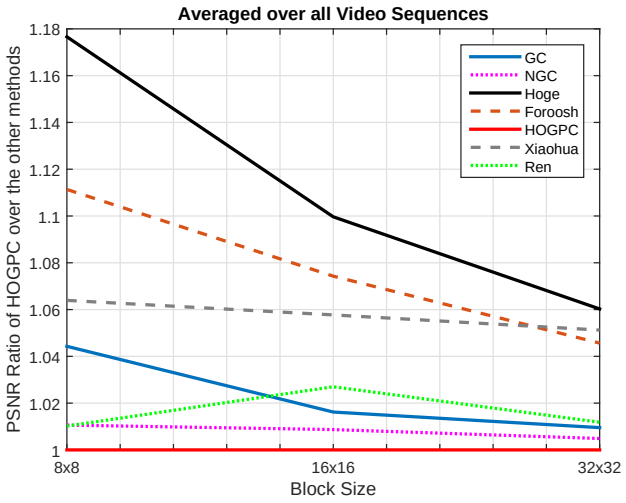

Figure 11: The PSNR ratio of the proposed HOGPC scheme over all the other methods for the different block sizes.

The estimated shifts and the corresponding measurements of their average MSE are shown in Tables 4-12. Observing the results, the proposed method provides the most accurate overall estimates with the lowest mean square error. Furthermore, since ground truth measurements can be significantly biased [15]; the performance of each method was assessed by computing the peak signal-to-noise ratio (PSNR) of the motion compensated prediction error. Figure 12 shows the obtained results for each method and all the image pairs. The proposed scheme achieves marginally the best registration accuracy in comparison with *NGC* [23], while the difference with the other methods is higher.

Additionally, the five MR images were used to evaluate the performance of each method in the presence of additive white Gaussian noise. In this case we assume that the correct shift is given by the corresponding noise-free estimate for each method and image pair. In figure 13 the mean value of the registration error for noise variance in the range $[0.005, 0.045]$ is illustrated for each method. Observing the results it can be seen that the proposed method is one of the most stable at high noise variances and provides the lowest overall MSE error. In the

Table 10: Average MSE with the corresponding PSNR values, and the estimated motion vectors for the 10 image pairs of the MRI data (Part 1).

Image pairs	[1,2]	[1,3]	[1,4]	[1,5]
GT	(-2.40,4.00)	(-4.80,8.00)	(-7.20,4.32)	(-7.20,12.00)
Hoge	(-2.03,4.01)	(-4.13,8.01)	(-6.81,4.17)	(-6.82,12.02)
Foroosh	(-2.22,4.23)	(-4.36,8.24)	(-6.59,4.41)	(-6.59,12.26)
Balci	(-2.11,4.10)	(-3.90,8.05)	(-6.22,4.34)	(-6.39,12.15)
Gaussian	(-2.07,4.02)	(-4.33,8.01)	(-6.57,4.37)	(-6.57,12.06)
Quadratic	(-2.03,4.01)	(-4.18,8.00)	(-6.73,4.25)	(-6.74,12.03)
Sinc	(-2.00,4.00)	(-4.12,8.00)	(-6.72,4.12)	(-6.73,12.00)
ESinc	(-2.00,4.00)	(-4.25,8.00)	(-6.54,4.31)	(-6.54,12.04)
Ren	(-2.09,4.02)	(-4.34,8.01)	(-6.58,4.38)	(-6.59,12.08)
GC	(-2.04,4.02)	(-4.24,8.00)	(-6.67,4.30)	(-6.68,12.03)
NGC	(-2.04,4.02)	(-4.24,8.00)	(-6.67,4.30)	(-6.68,12.02)
Xiaohua	(-2.04,3.95)	(-4.23,7.97)	(-6.66,4.36)	(-6.68,12.06)
HOGPC	(-2.06,4.04)	(-4.25,8.03)	(-6.67,4.33)	(-6.67,12.04)

case of the other methods, the error rapidly increases for noise beyond a certain level, since they do not always provide the correct pixel accuracy. The proposed *HOGPC* scheme exploiting the accuracy of HOG over noisy data allows precise estimates even for noise variance over the above range. Also, the PSNR was used to further compare the proposed scheme with the other state-of-the-art methods in the case of noise and the obtained results are shown in figure 14 demonstrating further the accuracy of *HOGPC* in terms of motion compensated prediction error. Furthermore, experiments were performed with 8 different levels of motion blur. In each case the variance was increased moving from 0.25 up to 2 and for each level five repetitions were performed. The overall results are in Table 13 showing that most of the methods to have similar performance with the one in [22] and the proposed HOG-PC to result the best performance.

Overall the complexity of the proposed HOG-PC is higher compared to most

Table 11: Average MSE with the corresponding PSNR values, and the estimated motion vectors for the 10 image pairs of the MRI data (Part 2).

Image pairs	[2,3]	[2,4]	[2,5]	[3,4]
GT	(-2.40,4.00)	(-4.80,0.32)	(-4.80,8.00)	(-2.40,-3.68)
Hoge	(-2.10,3.99)	(-4.28,0.15)	(-4.78,8.00)	(-2.17,-3.84)
Foroosh	(-2.32,3.75)	(-4.55,0.39)	(-4.55,8.24)	(-2.40,-3.61)
Balci	(-2.18,3.86)	(-4.16,0.30)	(-4.13,7.92)	(-2.34,-3.62)
Gaussian	(-2.26,3.97)	(-4.55,0.35)	(-4.56,8.01)	(-2.43,-3.66)
Quadratic	(-2.13,3.98)	(-4.65,0.22)	(-4.65,8.00)	(-2.25,-3.78)
Sinc	(-2.09,4.00)	(-4.72,0.11)	(-4.71,8.00)	(-2.27,-3.89)
ESinc	(-2.19,4.00)	(-4.59,0.28)	(-4.60,8.00)	(-2.46,-3.72)
Ren	(-2.27,3.96)	(-4.54,0.36)	(-4.54,8.01)	(-2.40,-3.65)
GC	(-2.17,3.98)	(-4.59,0.27)	(-4.60,8.00)	(-2.31,-3.71)
NGC	(-2.17,3.98)	(-4.59,0.27)	(-4.60,8.00)	(-2.31,-3.71)
Xiaohua	(-2.18,3.96)	(-4.59,0.34)	(-4.58,8.04)	(-2.39,-3.64)
HOGPC	(-2.19,3.99)	(-4.60,0.28)	(-4.60,8.00)	(-2.35,-3.68)

of the other approaches due to the computational power required for the preprocessing stage and the estimation of the dense HOG transform. In this work all the methods were implemented in Matlab and the average required time per method is shown in table 14. In the current architecture we did not considered any parallel implementations, but if a GPU-HOG transform [25] was used it could be no significant difference among them.

5. Conclusion

In this paper, a phase correlation technique based on histograms of oriented gradients that operates in the frequency domain for subpixel image registration was presented. The proposed method takes full account of all the advantages of HOG filter providing especially higher accuracy in small block sizes. One of the most attractive features of the proposed scheme is that it retains the ori-

Table 12: Average MSE with the corresponding PSNR values, and the estimated motion vectors for the 10 image pairs of the MRI data (Part 3).

Image pairs	[3,5]	[4,5]	Average MSE (x,y)	PSNR dB
GT	(-2.40,4.00)	(0.00,7.68)	(0.0000, 0.0000) ⇒ 0.0000	0.0000
Hoge	(-2.18,4.51)	(0.01,7.85)	(0.3667, 0.1914) ⇒ 0.5581	30.2380
Foroosh	(-2.41,3.76)	(-0.18,7.61)	(0.3368, 0.1945) ⇒ 0.5313	30.3865
Balci	(-2.49,4.07)	(-0.03,7.66)	(0.5857, 0.0841) ⇒ 0.6697	30.0364
Gaussian	(-2.44,4.00)	(-0.01,7.64)	(0.3558, 0.0324) ⇒ 0.3882	30.7528
Quadratic	(-2.27,4.00)	(-0.01,7.78)	(0.3334, 0.0602) ⇒ 0.3936	30.6963
Sinc	(-2.27,4.00)	(0.00,7.87)	(0.3490, 0.1281) ⇒ 0.4771	30.5317
ESinc	(-2.47,4.00)	(0.00,7.54)	(0.3834, 0.0494) ⇒ 0.4329	30.7081
Ren	(-2.41,4.00)	(-0.02,7.64)	(0.3488, 0.0403) ⇒ 0.3892	30.7583
GC	(-2.32,4.02)	(-0.01,7.73)	(0.3367, 0.0297) ⇒ 0.3664	30.7835
NGC	(-2.32,4.02)	(-0.01,7.73)	(0.3366, 0.0299) ⇒ 0.3664	30.7835
Xiaohua	(-2.39,4.04)	(-0.02,7.70)	(0.3399, 0.0411) ⇒ 0.3810	30.7700
HOGPC	(-2.35,4.04)	(0.01,7.72)	(0.3301, 0.0301) ⇒ **0.3601**	**30.7901**

entation information and the corresponding weights of HOG filter and exploits its robustness to noise. HOG phase correlation yields very accurate subpixel motion estimates for a variety of test material and motion scenarios and outperforms techniques, which are the current registration methods of choice in the frequency domain.

[1] C. Kuglin, D. Hines, The phase correlation image alignment method, in: Proc. IEEE Conf. Cyber. and Soc., 1975, pp. 163–165.

[2] J. Pearson, D. Hines, S. Goldsman, C. Kuglin, Video rate image correlation processor, Proc. SPIE Application of Digital Image Processing 119.

[3] G. Thomas, Television motion measurement for datv and other applications, BBC Res. Dept. Rep., No. 1987/11.

[4] B. Girod, Motion-compensating prediction with fractionalpel accuracy, IEEE Trans. Comm. 41 (4) (1993) 604.

Table 13: Average MSE of the estimated motion vectors with the corresponding PSNR values, for the 10 image pairs of the MRI data using 8 different motion blur levels and 5 repetitions for each one.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	Ren
PSNR	39.4322	39.4368	**39.4384**	36.7849	38.5084	39.4287	39.3615
MSE	0.0220	0.0117	0.0114	0.3692	0.0183	0.0102	**0.0079**

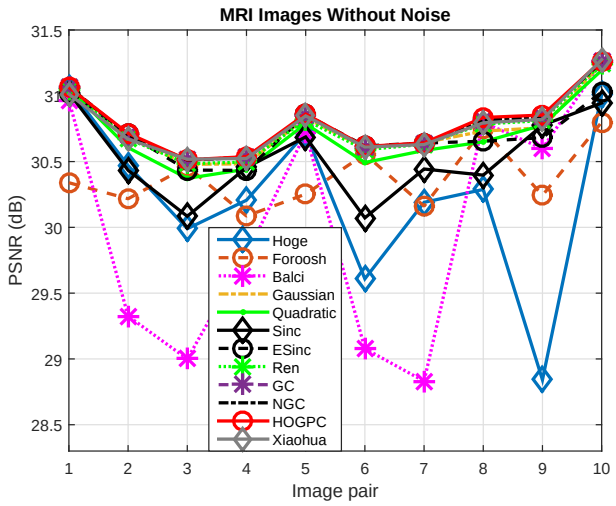

Figure 12: The PSNR values in dB over all image pairs.

[5] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01, CVPR '05, 2005, pp. 886–893.

[6] I. Abdou, Practical approach to the registration of multiple frames of video images, in Proc. SPIE Conf. Vis. Commun. Image Process. 3653 (1999) 371–382.

[7] V. Argyriou, T. Vlachos, A study of sub-pixel motion estimation using phase correlation, in Proc. Brit. Mach. Vis. Assoc. (2006) 387–396.

[8] S. Kruger, A. Calway, A multiresolution frequency domain method for es-

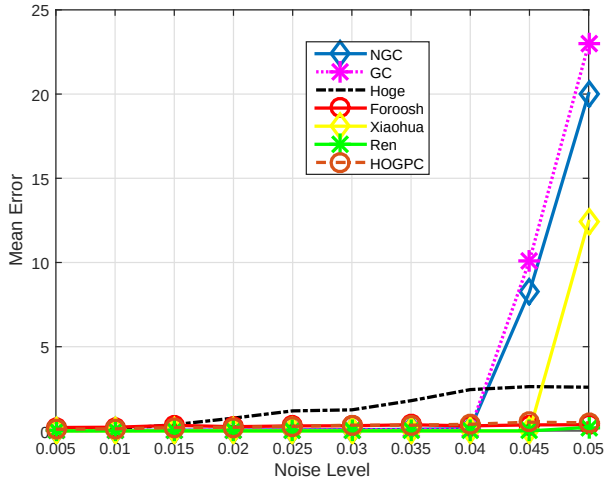

Figure 13: The Mean Error over all image pairs with different levels of noise for the top five methods.

Table 14: Average complexity for each method.

Method	GC	NGC	HOGPC	Hoge	Foroosh	Xiaohua	PC	Ren
Time (sec)	0.3824	0.4340	0.6988	0.1826	0.0701	1.5374	0.0372	0.0389

timating affine motion parameters, In Proc. IEEE International Conf. on Image Processing (1996) 113116.

[9] X. Tong, Z. Ye, Y. Xu, S. Liu, L. Li, H. Xie, T. Li, A novel subpixel phase correlation method using singular value decomposition and unified random sample consensus, Geoscience and Remote Sensing, IEEE Transactions on 53 (8) (2015) 4143–4156. doi:10.1109/TGRS.2015.2391999.

[10] L. Zhongke, Y. Xiaohui, W. Lenan, Image registration based on hough transform and phase correlation, Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on 2 (2003) 956–959.

[11] V. Maik, E. Chae, L. Eunsung, P. Chanyong, J. Gwanghyun, P. Sunhee,

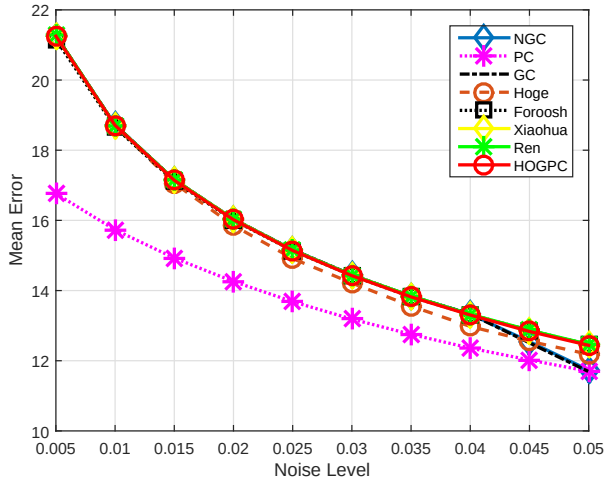

Figure 14: The PSNR values in dB over all image pairs with different levels of noise.

H. JinHee, J. Paik, Robust sub-pixel image registration based on combination of local phase correlation and feature analysis, Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on (2014) 1–2.

[12] M. Uss, B. Vozel, V. Dushepa, V. Komjak, K. Chehdi, A precise lower bound on image subpixel registration accuracy, Geoscience and Remote Sensing, IEEE Transactions on 52 (6) (2014) 3333–3345.

[13] P. Cheng, C.-H. Menq, Real-time continuous image registration enabling ultraprecise 2-d motion tracking, Image Processing, IEEE Transactions on 22 (5) (2013) 2081–2090.

[14] X. Tong, Y. Xu, Z. Ye, S. Liu, L. Li, H. Xie, F. Wang, S. Gao, U. Stilla, An improved phase correlation method based on 2-d plane fitting and the maximum kernel density estimator, Geoscience and Remote Sensing Letters, IEEE 12 (9) (2015) 1953–1957.

[15] W. Hoge, Subspace identification extension to the phase correlation method, IEEE Trans. Med. Imag. 22 (2) (2003) 277280.

[16] Y. Keller, A. Averbuch, A projection-based extension to phase correlation image alignment, Signal Process. 87 (2007) 124–133.

[17] M. Balci, H. Foroosh, Subpixel estimation of shifts directly in the fourier domain, IEEE Trans. Image Process. 15 (7) (2006) 1965–1972.

[18] H. Stone, M. Orchard, E. Chang, S. Martucci, A fast direct fourier-based algorithm for subpixel registration of images, IEEE Trans. Geosci. Remote Sens. 39 (10) (2001) 2235–2243.

[19] P. Vandewalle, S. Susstrunk, M. Vetterli, A frequency domain approach to registration of aliased images with application to superresolution, EURASIP J. Appl. Signal Process. (2006) 1–14.

[20] H. Foroosh, J. Zerubia, M. Berthod, Extension of phase correlation to subpixel registration, IEEE Trans. Image Process. 11 (2) (2002) 188–200.

[21] J. Ren, T. Vlachos, J. Jiang, Subspace extension to phase correlation approach for fast image registration, in Proc. IEEE ICIP (2007) 481–484.

[22] J. Ren, J. Jiang, T. Vlachos, High-accuracy sub-pixel motion estimation from noisy images in fourier domain, Image Processing, IEEE Transactions on 19 (5) (2010) 1379–1384.

[23] G. Tzimiropoulos, V. Argyriou, T. Stathaki, Subpixel registration with gradient correlation, Image Processing, IEEE Transactions on 20 (6) (2011) 1761–1767.

[24] S. Mallat, A wavelet tour of signal processing, 2nd ed. New York: Academic.

[25] V. Prisacariu, I. Reid, Fasthog - a real-time gpu implementation of hog, Department of Engineering Science, Oxford University 09 (2310).