J. Vis. Commun. Image R. 25 (2014) 1558–1565

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Gradient-based subspace phase correlation for fast and effective image alignment

Jinchang Ren [a,*], Theodore Vlachos [b], Yi Zhang [c], Jiangbin Zheng [d], Jianmin Jiang [e,*]

[a] Centre for Excellence in Signal and Image Processing, University of Strathclyde, Glasgow, United Kingdom
[b] Department of Audiovisual Arts, Ionian University, Corfu, Greece
[c] School of Computer Software, Tianjin University, China
[d] School of Computer Software and Microelectronics, Northwestern Polytechnical University, Xi'an, China
[e] School of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China

ARTICLE INFO

Article history:
Received 20 September 2013
Accepted 3 July 2014
Available online 14 July 2014

Keywords:
Image registration
Sub-pixel alignment
Phase correlation
Interference terms
Subspace projection
Fourier transform
Motion estimation
Fast algorithm

ABSTRACT

Phase correlation is a well-established frequency domain method to estimate rigid 2-D translational motion between pairs of images. However, it suffers from interference terms such as noise and non-overlapped regions. In this paper, a novel variant of the phase correlation approach is proposed, in which 2-D translation is estimated by projection-based subspace phase correlation (SPC). Conventional wisdom has suggested that such an approach can only amount to a compromise solution between accuracy and efficiency. In this work, however, we prove that the original SPC and the further introduced gradient-based SPC can provide robust solution to zero-mean and non-zero-mean noise, and the latter is also used to model the interference term of non-overlapped regions. Comprehensive results from synthetic data and MRI images have fully validated our methodology. Due to its substantially lower computational complexity, the proposed method offers additional advantages in terms of efficiency and can lend itself to very fast implementations for a wide range of applications where speed is at a premium.

1. Introduction

Registration of images plays a crucial role in the analysis of multi-dimensional visual data in the digital domain, where at least two images captured under different circumstances, such as from different sensors or at different times, need to be aligned for consistent measurement and processing. This can benefit a wide range of applications, including remote sensing [1,2], motion detection [3], image mosaicking [4,36,37], object recognition [5], medical imaging [6] and super-resolution for data visualization [7] as well as surveillance and video compression [7,8]. As a consequence the literature is enormous and any attempt to provide an account of it would be out of the context of this paper. Apart from a comprehensive survey in [8], some more recent papers can be found in [9–12], though they mainly focus on particular topics such as deformable medical image registration [9], remote sensing image registration [10], evolutionary image registration methods for 3D modeling [11] and 2D/3D registration methods for image-guided interventions [12].

Among many approaches proposed, phase correlation is a well-known technique for image registration [8,14], and has been successfully used in many applications such as object recognition [13] and motion estimation [15,16]. Further applications include verification in video shot detection [17] and motion extraction for video summarization and retrieval [18,19]. The baseline method utilizes the Fourier shift theorem, according to which shifts in the spatial domain correspond to linear phase changes in the frequency domain. Phase correlation is then further extended to estimate changes of rotation and scale using the Fourier–Mellin transform and the so-called pseudo-polar Fourier transform [20,21]. However, estimation of shifts between images with high accuracy remains a fundamental problem, in which potential exists for further research and improvement in terms of sub-pixel registration [22,23], video frame alignment [24] and blur-invariant registration [25].

Although pixel-level registration is adequate for some applications, higher accuracy sub-pixel registration is generally beneficial to most applications [26,27]. The need for sub-pixel registration arises from the simple fact that actual displacements between images are oblivious to the discrete grid employed at the image

* Corresponding authors.
E-mail addresses: Jinchang.Ren@strath.ac.uk (J. Ren), t.vlachos@ionio.gr (T. Vlachos), yizhang@tju.edu.cn (Y. Zhang), zhengjb@nwpu.edu.cn (J. Zheng), jianmin.jiang@szu.edu.cn (J. Jiang).

http://dx.doi.org/10.1016/j.jvcir.2014.07.001
1047-3203/

acquisition stage. Additionally, in other applications such as magnetic resonance imaging (MRI), data are usually sampled of non-integer offsets in the spatial Fourier domain before reconstruction and sub-pixel registration by phase correlation is a natural approach in such a context. Detailed comparison of several sub-pixel schemes can be found in [28].

Typically, 2-D Fourier transform is utilized by existing phase correlation approaches in estimating shifts between images. Its complexity under fast implementation, however, still remains an issue for many applications, where massive amount of data are involved. In addition, 2-D approaches perform less robustly, especially in estimating sub-pixels shifts in the presence of noisy data. To this end, a more accurate and robust solution of sub-pixel accuracy is desirable, which forms the motivation of the work described in this paper.

The main contributions are highlighted as follows. Firstly, we derive a fast solution using projection-based subspace phase correlation to estimate 2-D shifts in images, which is shown more robust to zero-mean noise than existing conventional approaches using 2-D phase correlation. Secondly, gradient based subspace phase correlation is proposed to deal with non-overlapped regions between the images under registration. These regions are taken as non-zero-mean noise in projected signals and it is proved that they have less influence using the proposed scheme than otherwise. In addition, we also demonstrate that the proposed method will yield higher peak than its 2-D counterpart.

The remaining part of this paper is organized as follows. Section 2 contains introductory concepts related to the phase correlation approach and problem formulation. In Section 3, subspace phase correlation and its gradient based variant are presented and their robustness is also demonstrated. Experimental results are given in Section 4 using synthetic data and MRI image data. Comparisons with existing techniques are also provided to confirm the superiority against proposed techniques. Finally, brief conclusions are drawn in Section 5.

2. Problem formulation

The baseline method of phase correlation is based on the Fourier shift theorem, which states that a shift in spatial domain will lead to linear phase differences in Fourier domain. Let $r(x,y)$ and $g(x,y)$ be two images satisfying $r(x,y) = g((x-x_0)\oplus M, (y-y_0)\oplus N)$ in which the images are $M \times N$ in size and $\oplus$ refers to the modulo operator. Accordingly, the Fourier transforms $R(u,v)$ and $G(u,v)$ of the images should satisfy

$$R(u,v) = G(u,v)e^{-j2\pi(ux_0/M+vy_0/N)} \tag{1}$$

Then, the phase difference can be obtained using the normalized cross-power spectrum as given below

$$P(u,v) = \frac{R(u,v)G^*(u,v)}{G(u,v)G^*(u,v)} = e^{-j2\pi(ux_0/M+vy_0/N)} \tag{2}$$

where $*$ is the complex conjugate, $j=\sqrt{-1}$, and $P(u,v)$ is referred to as the cross power spectrum of the two images.

If we apply the inverse Fourier transform $\Im^{-1}$ to $P(u,v)$, a phase correlation surface (PCS) $p(x,y)$ can be obtained as follows, which is essentially a Dirac function centered at (x_0,y_0).

$$p(x,y) = \Im^{-1}(P(u,v)) = \delta(x-x_0, y-y_0) \tag{3}$$

If the two images under consideration are not perfect replicas of each other hence the surface is noisy due to interference terms such as noise and non-overlapped regions. The latter will cause inconsistency due to the fact that the real shift is not a strict mod operator. However, crucially it still contains a dominant peak whose location corresponds to the shift parameters and can be recovered below, though the peak value can be less than unity as expected.

$$(x_0,y_0) = \arg\max_{x,y}|p(x,y)| \tag{4}$$

Due to sub-pixel shifts, the peak value can also be substantially degraded since the peak energy can be distributed to several adjacent neighboring peaks [22]. Peak height is an indication of confidence to the estimate obtained especially in the presence of the interference terms mentioned above. To enhance the peak identification accuracy in these cases pre-processing in the shape of windowing or filtering is often considered. In our paper, however, such pre-processing measures have not been considered in order to obtain cleaner and more straight-forward comparisons with competing methods i.e. comparisons which are not conditional upon using a specific pre- or post-processing regime. Nevertheless, results using spatial windowing from conventional approaches are also presented for evaluation purposes as discussed in Section 4.1.

Let $n(x,y)$ denote the effect of the interference terms, then the original two images satisfy

$$r(x,y) = g((x-x_0)\oplus M, (y-y_0)\oplus N) + n(x,y) \tag{5}$$

Let $C_{rg}(x_d, y_d) = E[r(x,y)g(x+x_d, y+y_d)]$ be the correlation function between two functions r and g, where E refers to the mathematical expectation, the corresponding cross-power spectrum becomes [30]

$$\begin{aligned} P_n(u,v) &= \frac{\Im[C_{rg}(x_d,y_d)]}{\Im[C_{gg}(x_d,y_d)]} \\ &= \frac{\Im[C_{gg}(x_d+x_0,y_d+y_0)] + \Im[C_{ng}(x_d,y_d)]}{\Im[C_{gg}(x_d,y_d)]} \\ &= \frac{[G(u,v)e^{-j2\pi(ux_0/M+vy_0/N)} + N(u,v)]G^*(u,v)}{G(u,v)G^*(u,v)} \\ &= e^{-j2\pi(ux_0/M+vy_0/N)} + N(u,v)/G(u,v) \end{aligned} \tag{6}$$

As can be seen, $P_n(u,v)$ is no longer a simple phase difference and it will approach $P(u,v)$ only if we have $N(u,v)/G(u,v) \to 0$, i.e. under a high signal to noise ratio (SNR). However, this requirement cannot be always satisfied, especially when there are non-overlapped regions between images. As a result, we have proposed the projection-based subspace phase correlation to address this difficulty. Although the concepts of subspace and projection are not new in phase correlation, the essence of our proposed algorithm still is original, considering the fact that most existing work either needs 2-D phase correlation to enable subspace identification of displacement [15,29] or shows lack of robustness [30]. In [15], based on the fact that the noise-free model for the phase correlation matrix is a rank one matrix, a subspace extension is proposed to identify subpixel shifts from the correlation matrix. A least-square fit to the phase components is employed with its slope determined as the subpixel shifts. However, the phase correlation matrix needs still be generated through 2D phase correlation. For the phase correlation function obtained from 2D Fourier transform, a masking operator is proposed in [29] to generate a projected one while rejecting components that are unrelated to the estimated shifts. Then, Hoge's approach [15] is applied to the so-called projected matrix to determine subpixel shifts for improved accuracy. In [30], projection based subspace phase correlation is used for pixel-level image registration. A windowing function is applied to the raw data before registration to avoid the failure of the approach. Without image gradient, this approach can only deal with small displacements. On the contrary, our proposed algorithm uses only 1-D phase correlation to estimate 2-D offsets in a robust way as explained in the next section. As