Pattern Recognition Letters 70 (2016) 87–92

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

# Accuracy improved image registration based on pre-estimation and compensation☆

Xiangguo Li*

*College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, PR China*

ARTICLE INFO

*Article history:*
Received 4 February 2015
Available online 28 November 2015

*Keywords:*
Image registration
Accuracy
Fourier–Mellin
Lucas–Kanade
Compensation

ABSTRACT

In this paper, we propose to improve the registration accuracy by pre-estimation and compensation. The idea is motivated by the observation of some registration algorithms that, for a given algorithm, the accuracies of the translation-only model are much higher than those of other complex models. Therefore, it seems that, if pre-estimation and compensation are performed and the residual model is close to translation only, the following estimation could achieve improved accuracy. To verify the idea, we implement two algorithms in the rotation-translation (RT) model. We use the Fourier–Mellin transform to isolate and convert the rotation into translation, then apply the classical Lucas–Kanade algorithm to obtain the high accuracy rotation estimation. In the following, the one takes account into the incomplete rotation compensation, and use the Keren algorithm for the residual model; the other assumes the rotation compensation is complete, and uses the second Lucas–Kanade algorithm. Finally, we perform simulations using typical test images, and the results confirm the accuracy improvement.



## 1. Introduction

Image registration is a fundamental preprocessing component in many image processing applications, such as super-resolution, image mosaicing, image fusion, and computer vision [26]. Among several evaluation metrics, accuracy may be the most important one because it can largely determine the performance of the following processing components. For instance, in satellite image mosaicing [20], one pixel in the image data may correspond to a ground block ranging from several meters to tens of meters on the Earth, e.g., the spatial resolution of a Landsat-8 image is about 30 m; therefore, pixel-level registration of such images will provide $\pm$ 15 m resolution, while 0.1 pixel registration can provide $\pm$ 1.5 m resolution. Besides, for the success of super-resolution reconstruction, accurate registration is even more critical [16], and otherwise super-resolution using inaccurately registered images is no better than interpolation of a single image.

The idea of this paper is mainly motivated by the following observation. In the experimental results of some registration algorithms, it can be observed that, if we consider a given algorithm and compare the estimation accuracies between the results under different transformation models, the accuracies under translation-only model are much higher than those under other complex models, e.g., the rotation-translation (RT) model (the detailed can refer to Section 2). This hints that, for the RT model, if the rotation could be firstly accurately estimated and well compensated, the translation would be estimated with improved accuracy, and even if the rotation is not completely compensated, since the residual rotation is quite small, it is still expected to achieve accuracy improved estimation under the residual RT model. We assume that this observation and hypothesis is generally true, and this is the main start point of this paper.

To verify the idea, we implement it in the RT model. In this case, the critical task is the high accuracy estimation of rotation. To do this, we choose to isolate the rotation out of the RT model and then obtain its high accuracy estimation. We use the Fourier–Mellin transform, which can be regarded as performing the Fourier transform followed by converting the magnitude spectrum into the log-polar coordinate system. According to the properties of the Fourier transform, only the rotation factor is left in the magnitude, and the following log-polar conversion can further convert the rotation factor into translation. Then, we need an “accurate enough” algorithm to perform the high accuracy estimation, and we choose the classical Lucas–Kanade algorithm, which is based on linear Taylor approximation, thus can achieve sub-pixel estimation without need of upsampling interpolation. After the rotation compensation, we perform the residual estimation. We implement two algorithms, the one assumes the residual model as translation-only, and the other treats the residual model as RT but the residual rotation is quite small; as the motivation, the translation accuracies in both the two are expected to be improved than the traditional approach.

☆ Paper has been recommended for acceptance by Dr. D. Coeurjolly.
* Tel.: +86 153 0372 0201.
*E-mail address:* xiangguoli@gmail.com

http://dx.doi.org/10.1016/j.patrec.2015.11.017
0167-8655/

Being a feasible isolation technique, the Fourier–Mellin transform has been successfully used in the extension of phase correlation to the rotation-scaling-translation model [6,10,18]. This two-stage strategy provides a framework for image registration. For example, Ho et al. [11] proposed to apply Argyriou's sub-pixel extension of phase correlation [2] in this framework, and Tzimiropoulos et al. [21] proposed to implement the normalized gradient correlation in this framework. We also implement in this framework, thus the implementations are formally similar to the methods above. However, we aim to achieve the rotation estimation as accurate as possible in the Fourier–Mellin domain, while others employ the Fourier–Mellin transform and then perform two common successive estimations. This is a subtle difference between ours and others.

As a classical gradient-based algorithm, the Lucas–Kanade algorithm can obtain high accuracy sub-pixel estimation directly, by contrast, most algorithms achieve sub-pixel accuracy by interpolation [20,24,25]. Since the seminal work of Lucas and Kanade [13], it has been extended into more general models [4], in which a notable extension was proposed by Keren et al. [12]. Moreover, it has been active till now, and new variants were presented in recent years [1,7,14,17,19]. Although the Lucas–Kanade algorithm has been proposed for more than 30 years, it can still keep the state-of-the-art sub-pixel accuracy. In this paper, we use the classical Lucas–kanade algorithm for "accurate enough" estimation.

In addition to the accuracy improvement, owing to the two-stage strategy, the proposed algorithms can broaden the rotation range. On the other hand, since the Lucas–Kanade algorithm is used to estimate the rotation in the Fourier–Mellin domain, the image magnitude feature may affect the Lucas–Kanade performance. To evaluate the algorithms, test images with typical magnitude features are used in the experiments, and the results verify the motivation.

The remainder of the paper is organized as follows. In Section 2, the motivation is introduced to clearly express the idea. Section 3 describes the classical Lucas–Kanade algorithm, the Keren algorithm, and the Fourier–Mellin transform. Section 4 presents the implementations in detail, and Section 5 gives results and discussion. Finally, Section 6 summarizes this paper.

## 2. Motivation

To clearly express the idea, we introduce the motivation here as a separate section. We will cite experimental data from Ref. [23], in which Vandewalle et al. presented a high accuracy registration algorithm for super-resolution in a two-stage fashion. They also used the magnitude spectra, but, instead of the log-polar transform, they performed radial integral to convert the magnitudes into one-dimensional functions, and determined the rotation by one-dimensional correlation; after rotation compensation, they performed least squares on the phase difference plane to obtain the translation estimation.

To test their algorithm, they compared with others under the RT model and the translation-only model. Since their proposed algorithm, Marcel et al. [15], and Keren et al. [12] are common under the two models, we prefer to compare the performances of the same algorithm under the two models. Therefore, we extract and rearrange the data as shown in Table 1, in which $\mu$ and $\sigma$ are the average absolute error and the standard deviation of the error, respectively. To keep data intact, we cite both the translation ($\hat{\mathbf{x}}_0$) and rotation ($\hat{\theta}_0$).

Note we do not compare between the algorithms but only concern the same algorithm under the two models. It is clear that, for all the three algorithms, the translation accuracies under the translation-only model ('T' in the Table) outperform their own counterparts under the RT model, and especially for Vandewalle's algorithm, the difference is dramatic. Furthermore, we notice that the three algorithms work in different principles. The algorithm by Vandewalle et al. is a hybrid one as described before, the algorithm by Marcel et al. is an extension of phase correlation, and the algorithm by Keren et al. is an extension of classical Lucas–Kanade algorithm. Therefore, since the three different algorithms exhibit common trends, we assume this phenomenon is common for general image registration, and propose the hypothesis in the introduction.

**Table 1**
Comparison of the estimation accuracies under the RT model and the translation model.

| | | Vandewalle et al. | | Marcel et al. | | Keren et al. | |
|---|---|---|---|---|---|---|---|
| | | T | RT | T | RT | T | RT |
| $\hat{\mathbf{x}}_0$ | $\mu$ | 3.2E-15 | 0.029 | 0.3126 | 1.999 | 4.1E-3 | 0.019 |
| | $\sigma$ | 3.9E-15 | 0.038 | 0.3803 | 11.522 | 6.0E-3 | 0.027 |
| $\hat{\theta}_0$ | $\mu$ | – | 0.126 | – | 19.003 | – | 0.053 |
| | $\sigma$ | – | 0.191 | – | 79.086 | – | 0.071 |

## 3. Related work

### 3.1. Classical Lucas–Kanade algorithm

Consider two images $f(\mathbf{x})$ and $g(\mathbf{x})$, where $\mathbf{x} = [x, y]^{\mathrm{T}}$, and assume that the two images are related by the translation-only model:

$$g(\mathbf{x}) = f(\mathbf{x} - \mathbf{x}_0), \tag{1}$$

where $\mathbf{x}_0 = [x_0, y_0]^{\mathrm{T}}$, and $x_0$ and $y_0$ are the horizontal translation and vertical translation, respectively.

To estimate $\hat{\mathbf{x}}_0$ of $\mathbf{x}_0$, the classical Lucas–Kanade algorithm aims to minimize the sum of squared error (SSE) function [13]:

$$\begin{aligned}\hat{\mathbf{x}}_0 &= \arg\min_{\mathbf{x}_0} \{E(\mathbf{x}_0)\} \\ &= \arg\min_{\mathbf{x}_0} \left\{ \sum_{\mathbf{x}} [f(\mathbf{x} - \mathbf{x}_0) - g(\mathbf{x})]^2 \right\}.\end{aligned} \tag{2}$$

First, we take the linear approximation of $f(\mathbf{x} - \mathbf{x}_0)$ by Taylor's series:

$$f(\mathbf{x} - \mathbf{x}_0) \approx f(\mathbf{x}) - \nabla f(\mathbf{x})^{\mathrm{T}} \mathbf{x}_0 \tag{3}$$

with $\nabla f(\mathbf{x}) = [\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}]^{\mathrm{T}}$.

Then, we compute the partial derivatives of $E(\mathbf{x}_0)$ with respect to $\mathbf{x}_0$ and set them to zero:

$$\begin{aligned}\mathbf{0} &= \frac{\partial}{\partial \mathbf{x}_0} \left\{ \sum_{\mathbf{x}} \left[ f(\mathbf{x}) - \nabla f(\mathbf{x})^{\mathrm{T}} \mathbf{x}_0 - g(\mathbf{x}) \right]^2 \right\} \\ &= \sum_{\mathbf{x}} 2 \nabla f(\mathbf{x}) \left[ f(\mathbf{x}) - \nabla f(\mathbf{x})^{\mathrm{T}} \mathbf{x}_0 - g(\mathbf{x}) \right],\end{aligned} \tag{4}$$

from which we can obtain the estimation:

$$\hat{\mathbf{x}}_0 = \left[ \sum_{\mathbf{x}} \nabla f(\mathbf{x}) \nabla f(\mathbf{x})^{\mathrm{T}} \right]^{-1} \left[ \sum_{\mathbf{x}} \nabla f(\mathbf{x}) [f(\mathbf{x}) - g(\mathbf{x})] \right]. \tag{5}$$

### 3.2. Keren algorithm

Keren et al. [12] extended the classical Lucas–Kanade algorithm into the RT model from a special manner, and they performed two successive Taylor approximations. We define the following RT model:

$$g(x, y) = f(x \cos\theta_0 - y \sin\theta_0 - x_0, x \sin\theta_0 + y \cos\theta_0 - y_0). \tag{6}$$

The first Taylor approximation is performed on $\sin\theta_0$ and $\cos\theta_0$ with the first two terms, and the second Taylor approximation is