Pattern Recognition 48 (2015) 3466–3476

Contents lists available at ScienceDirect

Pattern Recognition

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

Efficient and accurate set-based registration of time-separated aerial images

Ognjen Arandjelović*, Duc-Son Pham, Svetha Venkatesh

Centre for Pattern Recognition and Data Analytics (PRaDA), School of Information Technology, Deakin University, Geelong 3216 VIC Australia

ARTICLE INFO

Article history:
Received 23 May 2014
Received in revised form
25 March 2015
Accepted 8 April 2015
Available online 17 April 2015

Keywords:
Remote
Sensing
Alignment
Map
Graph
Constraints

ABSTRACT

This paper addresses the task of time-separated aerial image registration. The ability to solve this problem accurately and reliably is important for a variety of subsequent image understanding applications. The principal challenge lies in the extent and nature of transient appearance variation that a land area can undergo, such as that caused by the change under illumination conditions, seasonal variations, or the occlusion by non-persistent objects (people, cars). Our work introduces several major novelties (i) unlike previous work on aerial image registration, we approach the problem using a set-based paradigm; (ii) we show how image space local, pair-wise constraints can be used to enforce a globally good registration using a constraints graph structure; (iii) we show how a simple holistic representation derived from raw aerial images can be used as a basic building block of the constraints graph in a manner which achieves both high registration accuracy and speed; (iv) lastly, we introduce a new and, to the best of our knowledge, the only data corpus suitable for the evaluation of set-based aerial image registration algorithms. Using this data set, we demonstrate (i) that the proposed method outperforms the state-of-the-art for pair-wise registration already, achieving greater accuracy and reliability, while at the same time reducing the computational cost of the task and (ii) that the increase in the number of available images in a set consistently reduces the average registration error, with a major difference already for a single additional image.

1. Introduction

The goal of the present work is to achieve accurate registration of time separated aerial images of primarily urban/suburban scenes. Accurate registration is an important pre-processing step required by a wide range of practical applications including semantic labelling of images [1] and the detection of meaningful (high-level), structural changes [2–4]. Our aim is to achieve registration accuracy of approximately 0.5 m given a set of images with the interscan period in the range from a single day up to several months. This target accuracy is motivated by our aim to use registered images to detect and classify changes of interest to bodies such as local councils or the police. These changes include the erection of unauthorized structures such as sheds or extensions, overgrown bushes or trees, well drilling, solar panel installation, etc.

The key challenge of the registration addressed in this paper emerges as a consequence of the potentially large transient appearance variability, such as that which may be caused by a combination of different illumination conditions, seasonal variations, and occlusions by objects with a non-permanent presence (e.g. people, cars, and lawnmowers). A sample of images taken from the novel evaluation data set introduced herein is shown in Fig. 1. Since the human visual system – additionally assisted by our ability to understand and interpret visual data at a high semantic level – is highly adapted to the processing of scenes such as these, we highlight some of the aforementioned challenges in Fig. 2 by showing the change in Canny edges between two typical images. It can be readily seen that the edges which can be considered as reliable cues for registration and which correspond to more permanent structures such as houses, are vastly outnumbered by spurious edges which correspond to practically unpredictable changes in appearance, for example of plants. Moreover, it can be observed that even those edges which are salient in this context are not preserved well due to illumination changes. Fig. 3 additionally illustrates the difficulties posed by objects that move between acquisitions, as well as strong cast shadows.

Registration, as a general problem of geometric normalization, is pervasive in computer vision. Unsurprisingly, the corpus of relevant previous work is rich and varied, often with a high degree of domain specificity [5]. In aerial imaging applications, most registration approaches described in the literature typically focus on man-made structures, a priori choosing to exploit the presence of line features [6,7], rectangular buildings [8,9], or roads [10,11] in images. All of these methods register images in a pair-wise manner [12,13], either

* Corresponding author. Tel.: +61 3522 73079.
E-mail address: ognjen.arandjelovic@gmail.com (O. Arandjelović).
URL: http://mi.eng.cam.ac.uk/~oa214 (O. Arandjelović).

http://dx.doi.org/10.1016/j.patcog.2015.04.011
0031-3203/

Fig. 1. Input images to our algorithm (for easier visualization, only 400×400 pixel sub-regions of 1500×1500 pixel images used as an actual input are shown). These correspond to approximately the same land area imaged from an aeroplane on different days and at different times of the day, and registered using a state-of-the-art commercial registration system which uses both GPS and image data. Notice that despite the registration substantial registration errors remain (compare images (h) and (i), for example – the misalignment error between them is 78 pixels, Sections 3.1 and 3.3 for details). Also, observe the extent and the nature of appearance changes across images: cars move in and out, shadows change in their shape, direction, and intensity, grass and leaves change color across seasons.

aligning an aerial image to an aerial image, or an aerial image to a map. No previous work on aerial image registration operates directly on image sets as an input, nor is readily extended to this problem setup.

To place the present work in broader context and better appreciate our contributions, it is worth noting that set-based registration methods (sometimes also referred to as groupwise methods) have been described in other application domains of computer vision, most notably for medical image registration [14]. However, an examination of these approaches shows that they too cannot be readily applied to the problem we consider in this paper. Indeed, medical images are often taken under calibrated conditions, consistent across acquisitions, while aerial images exhibit extreme variations due to uncontrollable illumination and seasonal effects, among others. For example, the set-based method using the Havrda–Carvat cumulative residual entropy proposed by Chen et al. [15] requires the shapes of objects of interest to be known in advance. The approaches described by Lord et al. [16] and Li et al. [17] suffer from a similar limitation in this context, given that they require accurate contour information. Their target domain being void of such challenges, Wachinger and Navab [18] do not consider the issues of illumination change or the potential presence of transient structures and occlusions. Similar observations hold for a number of other approaches such as those of Thevenaz et al. [19], Foroosh et al. [20], and Bartoli [21]. Orchard and Mann [22], and Špiclin et al. [23] describe group-wise methods for the registration of images of the *same* object acquired using different modalities. In contrast, in the registration setting herein the actual content of the scene changes – trees grow or lose foliage, mobile objects leave the scene etc. Another noteworthy recent development in the field draws from the advances in sparsity learning, often with the specific focus on the alignment of images of faces [24–26]. Because of their computational cost, these methods are limited in their application to images