Pattern Recognition 48 (2015) 772–784

Contents lists available at ScienceDirect

Pattern Recognition

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

Non-rigid visible and infrared face registration via regularized Gaussian fields criterion

Jiayi Ma [a], Ji Zhao [b], Yong Ma [a,*], Jinwen Tian [b]

[a] Electronic Information School, Wuhan University, Wuhan 430072, China
[b] School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China

ARTICLE INFO

Article history:
Received 26 June 2014
Received in revised form
15 August 2014
Accepted 8 September 2014
Available online 21 September 2014

Keywords:
Registration
Image fusion
Infrared
Non-rigid
Face recognition
Gaussian fields

ABSTRACT

Registration of multi-sensor data (particularly visible color sensors and infrared sensors) is a prerequisite for multimodal image analysis such as image fusion. Typically, the relationships between image pairs are modeled by rigid or affine transformations. However, this cannot produce accurate alignments when the scenes are not planar, for example, face images. In this paper, we propose a regularized Gaussian fields criterion for non-rigid registration of visible and infrared face images. The key idea is to represent an image by its edge map and align the edge maps by a robust criterion with a non-rigid model. We model the transformation between images in a reproducing kernel Hilbert space and a sparse approximation is applied to the transformation to avoid high computational complexity. Moreover, a coarse-to-fine strategy by applying deterministic annealing is used to overcome local convergence problems. The qualitative and quantitative comparisons on two publicly available databases demonstrate that our method significantly outperforms the state-of-the-art method with an affine model. As a result, our method will be beneficial for fusion-based face recognition.

1. Introduction

Multi-sensor data often provides complementary information about the region surveyed, and image fusion which aims to create new images from such data offering more complex and detailed scene representation has then emerged as a promising research strategy for the purposes of human visual perception, object detection, as well as target recognition [1–3]. Particularly, for the face recognition problem, the use of multi-sensor data such as visible and thermal infrared (IR) images has been shown to be able to achieve higher recognition rate [4]. For example, thermal IR images are not affected by illumination variation or face disguise, while visible information is better for establishing a discriminative face model. However, successful image fusion requires an essential and challenging step that the image pairs to be fused have to be correctly co-registered on a pixel-by-pixel basis. In this paper, we focus on registration of visible and thermal IR face images for fusion-based face recognition.

Registration of multi-sensor images typically can be implemented by either hardware or software. Due to elevated cost and low availability, a special-purpose imaging sensor assembly that generates co-registered image pairs may not be practical in many face recognition scenarios. In this context, software-based registration may be more appropriate for large-scale deployment, where off-the-shelf low cost visible and infrared cameras can be utilized and no additional hardware is needed. Although the pixel-to-pixel alignment accuracy could be reduced compared to the hardware-based registration, the salient structures within the images can still be matched sufficiently well in a software-based registration process. In this paper, we are interested in software-based registration techniques.

The first step of software-based registration is to choose the information type that should be extracted and used to represent the images. Most registration methods assume global statistical dependence of the images to be aligned [5,2], and hence appearance features such as graylevels/colors, textures (e.g. Gabor filters [6]) and gradient histograms (e.g. SIFT [7] and HOG [8]) are preferred for the purpose of matching. However, for visible and thermal IR face images which are manifestations of two different phenomena, these appearance features will not likely match; instead, certain features representing salient structures could be adopted, such as points of high curvature, line intersections, strong edges, structural contours and silhouettes within the images. Here we choose the edge map, which is a significant common feature that might be preserved in face images [9,4].

To register edge maps of visible and IR face images, Kong et al. [4] proposed to use a straightforward criterion, named the Gaussian

* Corresponding author.
E-mail addresses: jiayima@whu.edu.cn (J. Ma), zhaoji84@gmail.com (J. Zhao), mayong@whu.edu.cn (Y. Ma), jwtian@hust.edu.cn (J. Tian).

http://dx.doi.org/10.1016/j.patcog.2014.09.005
0031-3203/

fields criterion [10,11], which consists of a Gaussian mixture depending on distances as well as point attributes such as image intensities and local shape descriptors. The criterion then maximizes the overlap between the salient structures such as edges that are present in both images. This method possesses several characteristics, for example, the criterion is differentiable and hence standard gradient-based optimization techniques could be employed. Besides, the method does not need any explicit set of point correspondences. However, the method is in general designed for rigid or affine registration, where the underlying relation between the data can be modeled by a parametric model. But a human face is a non-rigid shape and the deformation of the point set sampled from a face is restricted by physical constraints of bones and muscles. Therefore, the parametric model such as affine transformation used in [4] cannot produce accurate alignment. To address this problem, in this paper we generalize the Gaussian fields criterion, and propose a regularized Gaussian fields criterion for the purpose of non-rigid registration.

More precisely, we model the non-rigid transformation in a functional space, called the reproducing kernel Hilbert space (RKHS) [12], in which the function has an explicit kernel representation. To ensure well-posedness of the Gaussian fields criterion, we introduce a regularization term to enforce smoothness of the transformation. This regularized Gaussian fields criterion keeps the property of differentiable, and, it is more promising in handling data that involves non-rigid motion. Feature descriptors such as shape context [13] are used as attributes to help recover the point correspondence. Moreover, a sparse approximation based on a similar idea as the subset of regressors method [14] is also introduced to improve the computational efficiency.

Our contribution in this paper includes the following three aspects. Firstly, we analyze the robustness of the Gaussian fields criterion both in theory and experiment, which provides support for why it can be successfully used to the point set registration problem. Secondly, we generalize the Gaussian fields criterion from rigid to the non-rigid case, and propose a regularized Gaussian fields criterion for non-rigid point set registration, which enables us to deal with more real-world matching problems. Thirdly, we customize and apply the proposed regularized Gaussian fields criterion to visible and thermal IR face image registration. Compared to previous methods, it can increase the registration accuracy, and hence is able to improve the reliability of fusion-based face recognition systems.

The rest of the paper is organized as follows. Section 2 describes background material and related work. Section 3 describes the Gaussian fields criterion for rigid point set registration, and discuss the robustness of the criterion as well. In Section 4, we present our regularized Gaussian fields criterion, and apply it to visible and thermal IR face image registration. Section 5 illustrates our method on face landmark set registration and then tests it for visible and thermal IR face image registration on a public available dataset with comparisons to other approaches, followed by some concluding remarks in Section 6.

2. Related work

There are in general two types of approaches for image registration: area-based methods and feature-based methods, as discussed in recent survey papers [2,15]. These approaches are typically developed for visible stereo, remote sensing, or medical image pairs. The application to visible/infrared camera configuration is not straightforward due to the different spectral sensitivity of the sensors. For instance, visible sensors capture reflected light, while IR sensors capture principally thermal radiations emitted by objects; and hence appearance features such as textures in a visible image often get lost in the corresponding IR image since texture seldom influences the heat emitted by an object. Next we briefly review the above mentioned two types of approaches, especially in the context of multimodal registration such as visible and thermal IR images.

2.1. Area-based methods

Area-based methods deal directly with the image intensity values without attempting to detect salient structures. These methods can be broadly classified into three types: correlation-like methods, Fourier methods, and mutual information (MI) methods [2].

Correlation-like methods such as cross-correlation and its modifications are a classical representative of the area-based methods [16]. The main idea of these methods is to compute the similarities of window pairs in two images, and consider the one with the largest similarity as a correspondence. To perform multimodal image registration, correlation ratio-based methods have been developed under the assumption that intensity dependence can be represented by some function [17]. The correlation-like methods suffer from some drawbacks such as the flatness of the similarity measure in textureless regions and high computational complexity. However, due to their easy hardware implementation which is beneficial for real-time applications, these methods are still often in use.

Fourier methods are the second type of approaches. These methods exploit the Fourier representation of images in the frequency domain [18]. Compared to correlation-like methods, they have some advantages in computational efficiency and are also robust to frequency-dependent noise. A representative application exploiting the Fourier transform for multimodal registration is described in [19]. It computes the correlation in the frequency domain, which is able to handle multimodal images such as visible and IR images when applied to the edge representations rather than the original graylevel images.

Finally, area-based methods also include mutual information methods. The mutual information provides an attractive metric for maximizing the dependence between two images, and it is particularly suitable for multimodal registration [20]. However, for visible/infrared image pairs with quite different textures, the mutual information is good usually only on a small portion of the images. Therefore, mutual information is typically used only on a selected region of an image [21], on a detected foreground [22], or on region with similar edge density [23]. Furthermore, the mutual information methods not only work directly with image intensities, but also with extracted features such as points of the area borders [24].

2.2. Feature-based methods

The second approach for image registration is based on the extraction of salient structures – features – in the images, which can be represented as compact geometrical entities at different levels, such as points, line segments, curves, and surfaces [5,2]. The registration problem then reduces to determining the correct correspondence and to find the underlying spatial transformation between two sets of extracted features. In general, features at higher levels are more difficult to extract reliably; the point feature is the simplest form of feature, and it is also more general since lines and surfaces can be discretized as a set of points. In this sense, point matching serves as the basis for the feature-based methods.

To address the point registration problem, a variety of techniques have been developed in the fields of computer vision, remote sensing, medical imaging, etc. A popular strategy is to iteratively solve the two unknown variables, i.e., the correspondence and the transformation. This kind of algorithms includes iterated closest point (ICP) [25], robust point matching using thin-plate spline (TPS-RPM) [26], coherence point drift (CPD) [27], vector field consensus