Advanced Engineering Informatics 38 (2018) 203–215

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Review article

Automated thermal 3D reconstruction based on a robot equipped with uncalibrated infrared stereovision cameras

T. Sentenac[a,b,*], F. Bugarin[b], B. Ducarouge[a], M. Devy[a]

[a] LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France
[b] Institut Clément Ader (ICA), Université de Toulouse, CNRS, Mines Albi, UPS, INSA, ISAE-SUPAERO, Campus Jarlard, F-81013 Albi CT Cedex 09, France

ARTICLE INFO

Keywords:
3D reconstruction
Hand-eye calibration
Rectification
Vision
Robotic
Infrared cameras
Uncalibrated cameras

ABSTRACT

In many industrial sectors, Non Destructive Testing (NDT) methods are used for the thermomechanical analysis of parts in assemblies of engines or reactors or for the control of metal forming processes. This article suggests an automated multi-view approach for the thermal 3D reconstruction required in order to compute 3D surface temperature models. This approach is based only on infrared cameras mounted on a Cartesian robot.

The low resolution of these cameras associated to a lack of texture to infrared images require to use a global approach based first on an uncalibrated rectification and then on the simultaneous execution, in a single step, of the dense 3D reconstruction and of an extended self-calibration.

The uncalibrated rectification is based on an optimization process under constraints which calculates the homographies without prior calculation of the Fundamental Matrix and which minimizes the projective deformations between the initial images and the rectified ones.

The extended self-calibration estimates both the parameters of virtual cameras that could provide the rectified images directly, and the parameters of the robot. It is based on two criteria evaluated according to the noise level of the infrared images. This global approach is validated through the reconstruction of a hot object against a reference reconstruction acquired by a 3D scanner.

1. Introduction

This article addresses the problem of a fully automated 3D thermal reconstruction [1–3] from sensors embedded on a robotic system. Such a method can be suitable for performing diagnostics on mechanical assemblies, such as nuclear reactors [4], or for improving energy efficiency in building construction [5] or for monitoring forming processes [6]. The first sub-problem that arises is to define the system architecture using heterogeneous sensors. The second sub-problem is selecting the dense 3D reconstruction methods with overlaid thermal data. The third sub-problem is to make the inspection task totally automatic with a self-extended calibration (i.e. a calibration without a specific target, covering all the geometric parameters of the robot-sensor system, including the intrinsic and extrinsic sensor parameters and the robot parameters).

The most conventional architecture is based on a 3D laser scanner and infrared cameras mounted on a robot [7–9]. To overcome the significant cost of the 3D laser scanner, several articles [5,10,11] have suggested an architecture using only cameras. Inexpensive and readily available digital visible cameras (CCD camera, color camera, Kinect, etc.) give images processed by a 3D modeler, while infrared cameras provide the thermal data mapped on the 3D model. One successful system is the HeatWave system [3,12], i.e. a hand-operated device consisting of rigidly attached infrared and color cameras. These multi-sensory architectures face the difficulty of fusing 3D data provided by a 3D modeler and temperature data acquired by infrared cameras into a common coordinate frame. A joint geometric calibration of heterogeneous sensors [13] must be performed, which requires finding a pattern that is completely visible by both the 3D sensors and the IR cameras. This could be a tricky task, because these heterogeneous sensors have different spectral sensitivities, spatial resolutions and fields of view. The ideal architecture is then only based on infrared cameras for a direct 3D thermal reconstruction. Assuming that thermal methods already described in [14–16] are not in the scope of this paper, the challenge is then to provide a dense image-based 3D reconstruction [17,18] with infrared cameras.

Several image-based 3D reconstruction algorithms have been proposed using visible cameras. The first step, image registration, requires the detection and matching of features between images. Many feature detectors (e.g., Harris, SIFT, SURF, FAST, ORB, …) automatically and

* Corresponding author at: CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France.
E-mail addresses: sentenac@laas.fr (T. Sentenac), florian.bugarin@univ-tlse3.fr (F. Bugarin), benoit.ducarouge@laas.fr (B. Ducarouge), michel.devy@laas.fr (M. Devy).

https://doi.org/10.1016/j.aei.2018.06.008
Received 28 November 2017; Received in revised form 13 June 2018; Accepted 21 June 2018
1474-0346/

correctly extract and match interest points on infrared images. Then two classes of methods can simultaneously build a sparse 3*D* model and the camera trajectory using only these matched interest points. The Robotics community has developed several Vision-based Simultaneous Localization And Mapping (VSLAM) techniques [19,20], taking advantage of other proprioceptive data acquired from the robot (odometry, IMU…), but assuming generally that the intrinsic camera parameters are known. The Vision community has proposed Structure from Motion (SfM) [21] approaches (Bundler, OpenMVG…) in order to recover from an image sequence both the 3*D* environment structure and the camera Motion; extrinsic and intrinsic camera parameters can be estimated simultaneously when computing the 3*D* point positions. The recovered parameters should be consistent with the reprojection error (i.e., the sum of distances between the projections of each set of 3*D* corresponding feature points and its corresponding image features). This minimization problem can be formulated as a non-linear least squares problem and solved from a Bundle Adjustment (BA) algorithm [22]. Exploiting VSLAM or SfM methods, an accurate and dense 3*D* model could be incrementally and gradually built and refined with, typically, a sequence of one thousand images, either from an Iterative Closest Point (ICP) algorithm (stereovision) or by a Multi-View Stereo (MVS) [23] technique (monocular vision).

The paper proposes an automated thermal 3*D* reconstruction based on an architecture composed of a Cartesian robot equipped only with uncalibrated infrared cameras. The architecture requires a coupled method that deals simultaneously with a multi-view 3*D* thermal reconstruction and a self-extended geometric calibration. An infrared stereo vision rig provides a compensation to the lower spatial resolution of infrared images. It also improves the number of reconstructed points and thus the density of the 3*D* model. Moreover, it gives an initial guess for the 3*D* position of every point. For the first step of the method, a reasonable amount of stable and tractable matched points is obtained through a specific method for infrared images based on the phase congruency model [24,25] which is combined with more classical feature detectors. With few and low-textured infrared images, the result would be limited to a sparse 3*D* reconstruction. Next, the simultaneous reconstruction and self-calibration with uncalibrated infrared cameras is solved by the minimization of a cost function which integrates all the geometrical calibration parameters for both the cameras and the robot. Estimations variables are: four intrinsic parameters for each camera, six for the relative position and orientation between the two cameras and six for the rotational and the translational components of the Euclidean transformation. This latter transformation is named hand-eye, between hand (robot gripper) and eye (camera). The total number of parameters is twenty if it is assumed that geometrical distortions due to the lens are corrected beforehand. This assumption is a good trade-off between the accuracy and the computation time. Indeed, it decreases the accuracy, but it avoids additional degrees of freedom and high non-linearities in the geometric model which are consuming in computation times. For the targeted application in a robotic context, a real-time processing of the calibration, compatible with the speed of the robot, is preferred even if the accuracy is not optimal. Finally, these parameters have to be estimated on-line from features extracted and matched from images acquired on the unknown object from two or more positions of the robot. These positions are known accurately in the robot reference frame from its forward kinematics.

A projective rectification method is first introduced for improving the matching problem between pixels on the left and right images, limiting the search space from the whole image to only a line. Another objective is to reduce the number of parameters of the cost function. The projective rectification method, applied to uncalibrated stereovision infrared cameras, requires a specific algorithm. Two homographies, applied to rectify the left and right images, have to be estimated in a single step to cope with the problem of noisy and low-textured infrared images. A new cost function is proposed which is minimized under geometric constraints by a non-linear optimization process. These constraints are defined to keep the structure and the skewness of the images, which are essential to preserving their geometries.

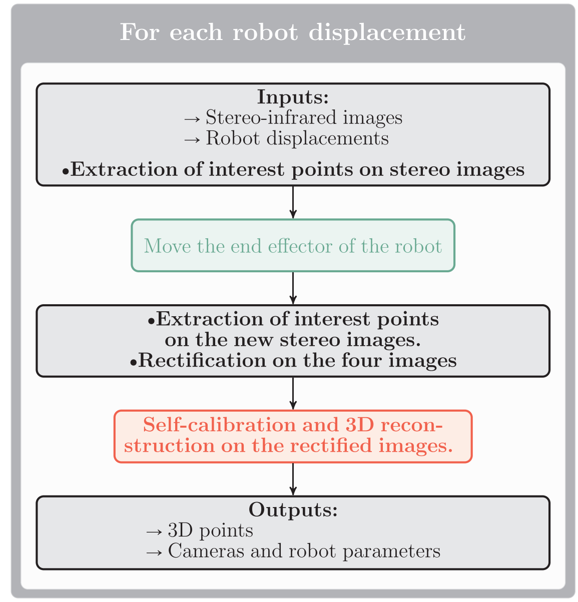

Fig. 1. Flowchart of the suggested method.

The set of points matched from rectified images is then used to perform simultaneously the extended self-calibration and the multi-view infrared 3*D* reconstruction. It is based on the minimization of two different functions depending on the observed noise level in infrared images. The first objective function is based on the minimization of the reprojection errors. When the noise level increases, a second objective function is expressed in the projective space which takes advantage of the epipolar constraint between two images. This second objective function is expressed using the intrinsic camera parameters and the essential matrix between two robot positions which depends itself on the hand-eye parameters and the robot motion. Fig. 1 summarizes the suggested coupled method which merges an extended self-calibration and a multi-view infrared 3*D* reconstruction applied on rectified images.

The paper is organized as follows. Section 2 briefly describes existing literature on rectification and places our suggested method with uncalibrated infrared images in this context. The method is then described and the results are compared to those in the literature. Section 3 outlines the formulation of the multi-view infrared 3*D* reconstruction simultaneous to the extended self-calibration. The method is evaluated on synthetic data. Finally, Section 4 summarizes the results of the fully automated 3*D* reconstruction performed from multiple views acquired by an uncalibrated infrared stereo rig mounted on a Cartesian robot. The whole approach is evaluated comparing the 3*D* model of a reconstructed object with a reference *CAD* model.

2. Suggested rectification method of uncalibrated cameras

The suggested rectification method, applied to uncalibrated stereovision cameras, takes the advantage of calculating the homographies, projective transformations applied to rectify images, without a previous calibration and in a single step to cope with the problem of noisy and low-textured infrared images. These homographies are then calculated with only one non-linear optimization process under geometric constraints.

The section begins with a short description of the background for calculating the homographies and works related to the rectification problem. The suggested projective rectification method is then