

Received July 29, 2016, accepted September 16, 2016, date of publication November 23, 2016,
date of current version December 8, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2629987

Heterogeneous Multi-View Information Fusion: Review of 3-D Reconstruction Methods and a New Registration with Uncertainty Modeling

**HADI ALIAKBARPOUR¹, V. B. SURYA PRASATH¹,
KANNAPPAN PALANIAPPAN¹, (Senior Member, IEEE),
GUNA SEETHARAMAN², (Fellow, IEEE), AND JORGE DIAS^{3,4}, (Senior Member, IEEE)**

¹Computational Imaging and VisAnalysis Laboratory, Department of Computer Science, University of Missouri–Columbia, Columbia, MO 65211, USA

²Advanced Computing Concepts, U.S. Naval Research Laboratory, Washington, DC 20375, USA

³Faculty of Science and Technology, Institute of Systems and Robotics, University of Coimbra, 3000-315 Coimbra, Portugal

⁴Robotics Institute, Khalifa University of Science, Technology and Research, Abu Dhabi 127788, United Arab Emirates

Corresponding author: V. B. S. Prasath (prasaths@missouri.edu)

This work was supported in part by the U.S. Air Force Research Laboratory under Grant AFRL FA8750-14-2-0072 and in part by the Portuguese Foundation for Science and Technology.

ABSTRACT We consider a multisensor network fusion framework for 3-D data registration using inertial planes, the underlying geometric relations, and transformation model uncertainties. We present a comprehensive review of 3-D reconstruction methods and registration techniques in terms of the underlying geometric relations and associated uncertainties in the registered images. The 3-D data registration and the scene reconstruction task using a set of multiview images are an essential goal of structure-from-motion algorithms that still remains challenging for many applications, such as surveillance, human motion and behavior modeling, virtual-reality, smart-rooms, health-care, teleconferencing, games, human–robot interaction, medical imaging, and scene understanding. We propose a framework to incorporate measurement uncertainties in the registered imagery, which is a critical issue to ensure the robustness of these applications but is often not addressed. In our test bed environment, a network of sensors is used where each physical node consists of a coupled camera and associated inertial sensor (IS)/inertial measurement unit. Each camera-IS node can be considered as a hybrid sensor or fusion-based virtual camera. The 3-D scene information is registered onto a set of virtual planes defined by the IS. The virtual registrations are based on using the homography calculated from 3-D orientation data provided by the IS. The uncertainty associated with each 3-D point projected onto the virtual planes is modeled using statistical geometry methods. Experimental results demonstrate the feasibility and effectiveness of the proposed approach for multiview reconstruction with sensor fusion.

INDEX TERMS Structure-from-motion, image registration, 3D reconstruction, heterogeneous information fusion, homography, coupled sensors, inertial measurement unit (IMU), sensor network, geometric uncertainty, virtual reality.

I. INTRODUCTION

Performing three dimensional (3D) object and scene reconstruction using a set of 2D planar images is one of the key problems in computer vision. A network of cameras, whose usage and availability have been increasing in the last decade, can provide images from different views of the scene. Recently, inertial sensor (IS)/inertial measurement unit (IMU) is becoming cheaper and widely available, for example most smart-phones are equipped with both camera and IS. A full scale 3D reconstruction is useful for many engineering areas and recently it picked up a lot of interest due

to the proliferation of affordable sensors. These areas include visual surveillance, human motion modeling, behavior monitoring, virtual-reality, games, teleconferencing, human–robot interaction, medical imaging, and general object recognition. Recently, research in 3D reconstruction and information fusion is of interest due to the proliferation various imaging devices with increasing availability of computing power. The use of 3D information in the field of cultural heritage is investigated by Vergauwen and Van Gool [1] where a web-based 3D reconstruction service is proposed. More recently, Agarwal *et al.* [2] presented a city scale 3D reconstructions

based on mined photographs from web users, see also [3]. Feldmann *et al.* [4] studied on-line full body motion tracking based on volumetric 3D reconstruction. Brice *et al.* [5] investigated the use of multi-view geometry to human model and pose reconstruction. Luo *et al.* [6] introduced a method to estimate human pose for multiple persons based on volume reconstruction. Capturing of complex human movements from multiple views is studied by Kehl [7]. 3D reconstruction of natural underwater scenes using an stereo-vision is studied by Brandou *et al.* [8]. A human body posture estimation method based on back projection of human silhouette images is proposed by Takahashi *et al.* [9]. Uriol [10] used a camera network to reconstruct human and synthesis an avatar. A system for real-time 3D human visual hull reconstruction and skeleton voxels extraction is proposed by Yang *et al.* [11]. One of the main problem in human body 3D reconstruction is the data registration from different planes which is different from multisensor image fusion methods. We next provide a explorative review of data registration frameworks proposed in recent literature based in terms of 3D reconstruction, the involved geometric relations, and handling underlying uncertainties.

A. REVIEW OF MULTI-IMAGES BASED 3D RECONSTRUCTION

Scene reconstruction using a set of images is a classic problem in computer vision and still remains an active field of research. In what follows we review the important literature in 3D reconstruction, modeling, registration and broadly classify them under geometric relations or uncertainty modeling. With respect to specific problems in 3D, we refer to the following specialized review papers for more details: Remondino and El-Hakim [12] provide a review of 3D modeling using terrestrial imagery. Kordelas *et al.* [13] provided a survey for existing 3D model reconstruction algorithms where the approaches are categorized by the data acquisition devices (laser range finder and camera). Different 3D reconstruction methods based on visual hull approach are compared and evaluated by Fredriksson [14]. A survey on motion-parallax-based 3-D reconstruction techniques is provided by Lu *et al.* [15]. Efficient methods for the 3D reconstruction of static and dynamic scenes from stereo images, stereo image sequences, and images captured from multiple viewpoints are explored by [16]–[18]. Different multi-view stereo reconstruction algorithms are compared and evaluated on a common ground truth by Seitz *et al.* [19]. A Comparison between different computer vision methods for real-time 3D reconstruction for the use in mobile robots has been done by Dornauer *et al.* in [20]. Multi-image 3D reconstruction in cultural heritage was reviewed by Koutsoudis *et al.* [21]. Fathi *et al.* [22] reviewed image-based 3D reconstruction of civil infrastructure. Gimenez *et al.* [23] provide a review of 3D reconstruction in buildings from 2D scanned plans. We next review some of the salient works on 3D reconstruction with geometric relations, and uncertainty handling occurring in such geometric relations.

1) HOMOGRAPHY BASED METHODS FOR 3D RECONSTRUCTION

Khan *et al.* [24] study a homographic framework for the fusion of multi-view silhouettes using three vanishing points of the reference plane in the scene. Later Khan [25] proposed further algorithms to track, reconstruct and object classification by using a homographic occupancy constrain. A similar approach undertaken by Michoud *et al.* [26] who introduced a marker-less 3D human motion capturing approach using multiple views. Zhang *et al.* [27] introduced an algorithm for 3D projective reconstruction based on infinite homography. Their contribution is an improvement of classical 4 points based method [28], [29] using a linear algorithm based on 3 points on a reference plane which is visible in all views. Homography-based mapping is used to implement a 3D reconstruction algorithm by Zhang and Hanson in [30]. Wada *et al.* [31] studied a 3D reconstruction method using homography transformation. They presented a parallel volume intersection method based on plane-to-plane homography for real-time 3D volume reconstruction using active cameras where the focus is on the acceleration of back-projection from silhouette images to 3D space. Lai and Yilmaz [32] utilized uncalibrated cameras for performing projective reconstruction of buildings based on Shape From Silhouette (SFS) approach with building's structures are used to compute vanishing points. Homography transformations are used to achieve a metric reconstruction up to a scale factor. Lee and Yilmaz [33] applied a 3D reconstruction method using photo consistency in images taken from uncalibrated multiple cameras. Zhang and Li [34] proposed a dynamic calibration and 3D reconstruction method with homography transformations. Tang *et al.* [35] investigated metric 3D reconstruction for large structures from uncalibrated images and homography transformations.

2) NON-HOMOGRAPHY BASED METHODS FOR 3D RECONSTRUCTION

Some researchers advocated the use of non-homography methods for 3D reconstruction. Sorman *et al.* [36] presented a volumetric graph-cuts based multi-view reconstruction. Guerchouche *et al.* [37] proposed a multi-resolution volumetric 3D object reconstruction. Azevedo *et al.* [38] proposed an object reconstruction using uncalibrated images from single off-the-shelf cameras. The reconstructions were achieved using active computer vision method for the 3D reconstruction of objects from image sequences. Structure From Motion (SFM) was used to recover the 3D shape of an object based on the relative motion and photo-consistency (voxel coloring) was used to perform the volumetric reconstructions [39]. 3D reconstruction of urban areas using a fast and robust Structure-from-Motion algorithm was proposed in our recent works [40], [41]. Jethwa [42] proposed a method to perform efficient voxel-based reconstruction of urban environments using a large set of images. Color and silhouette information from multiple views are fused by Khan and Shah [43] for reconstructing articulated objects

in monocular video. Sudipta [44] studied how silhouettes extracted from images and video can help both multi-view camera calibration and 3D surface reconstruction from multiple images. Remondino [45] studied static human body shape from image sequences using least squares matching and involves calibration and orientation of the images based on photogrammetric techniques. Maitre *et al.* [46] investigated a method to perform multi-view reconstruction of a scene by using camera arrays. Ruwwe *et al.* in [47] proposed an approach for image registration based on reconstructed 3D octrees by voxel carving. Díez *et al.* [48] provide a qualitative review on coarse 3D registration methods and Oliveira and Tavares [49] review automated image registration methodologies that have been used in the medical field.

3) UNCERTAINTY MODELING

In the context of uncertainty modeling in a homography transformation there are prior art. To the best of our knowledge, these methods are for cases where the homography transformation is estimated by using point correspondences and not when the homography is analytically calculated from its principal equations. In one of the earliest works, Criminisi *et al.* [50] utilized a plane measuring device and discussed about the uncertainty of homography mappings. The uncertainty analysis is done with respect to number point correspondences and the uncertainties in localization of those points. Meidow *et al.* [51] provided general geometric reasoning with uncertain 2D point and lines. Negahdaripour *et al.* [52] studied the accuracy of planar homography in applications such as video frame-to-frame homography. Ochoa and Belongie [53] presented an approach to determine a search region to be used in guided matching under projective mappings. The problem of finding optimal point correspondences between images related by homography transformation was addressed by Chum *et al.* [54]. Their work explored how to determine a pair of points that are exactly consistent with the homography, given an homography transformation and a pair of matching points, and also minimize the overall geometric error. Baker *et al.* [55] studied the parametrization of homography to maximize the plane estimation performance. They compared their method with the classical estimation process using a parametrization that combines 4 fixed points in one of the images with 4 variable points in the other image. A direct method to estimate planar projective transformation was introduced by Chi *et al.* [56]. They considered a method to register 2D points set which reduces the search space for the homographies from eight-dimensional space to a three-dimensional case.

The use of inertial sensors (IS) for various computer vision applications is recently gaining traction as the sensors become much cheaper and more readily available commercially. Moreover, the availability of smart-phones with MEMS (Micro-Electro-Mechanical-Systems) chipsets, provides valuable source of IS and cameras. Dias *et al.* [57] and Lobo *et al.* [58] investigated visual and inertial cooperation for 3D vision problems. Lobo and Dias [59] proposed

an efficient method for relative pose calibration between visual and IS. Lobo *et al.* [60] used IS with a stereo camera for the purpose of world feature detection and mapping. Mirisola *et al.* [61] used a rotation-compensated imagery along with IS data for trajectory recovery of an airship. Bleser *et al.* [62], [63] utilized fusion of image and inertial data for tracking in the mobile augmented reality scenarios. Ababsa [64] used global positioning system (GPS) position alone with orientation from inertial sensor for real time 3D reconstruction of urban scenes. Zendjebil *et al.* [65] used three different sensors (GPS, IMU, camera) with synchronized calibration for 3D localization of an outdoor mobile robots. Hogue *et al.* [66] used combined IS and stereo vision for underwater environment reconstruction. Hsieh *et al.* [67] considered concurrent multiple camera calibration using IS with no overlapping FOVs (field of views) for robot localization. Clark *et al.* [68] proposed to augment inertial data with monocular video to perform 3D environment reconstruction.

Fusion of image data with location and orientation sensor data streams for the purpose of camera trajectory recovering and scene reconstruction was investigated by Gat *et al.* [69]. They fused inertial information together with geographical data and images from a video stream recorded by a mobile camera in order to reconstruct the camera trajectory for the purpose of consumer video applications. Beşdok [70] proposed a method to calibrate a pair of cameras using IS attached to a calibration pattern with RBF neural networks used for training the system. Randeniya *et al.* [71] studied a method to estimate the intrinsic parameters for the camera and the extrinsic parameters among the camera and IS. Their approach is mentioned to be effective and precise for intelligent transportation Systems applications with large field of view and capable of functioning in maneuvers. In Mirzaei and Roumeliotis [72] a Kalman filter-based algorithm to calibrate coupled IS and camera was proposed. Okatani and Deguchi [73] studied robust estimation of camera translation between two images using a 3D orientation sensor. In a similar way, Labrie and Hebert [74] recovered 3D camera motion recovery using IS. Brodie *et al.* [75] studied static accuracy and calibration of IS to reduce 3D orientation errors. Kalantari *et al.* [76] solve the relative orientation problem between two cameras using only 3 points and vertical direction information.

B. CONTRIBUTIONS

In this work, we investigate the fusion of inertial sensor (IS) and camera as a coupled synergistic system. We see from the above literature review on 3D reconstruction methods, an important assumption is to have a planar ground to mark some points on the ground. This is required for estimating the homography matrix among the image plane and ground, see [24], [30], [31], [33], [38]. However this assumption may not hold in due to two main reasons: (a) In outdoor scenarios a flat ground plane may not exist, and (b) in textured grounds

(partial or full) identifying geometrically points with either automatic or manual marking is difficult. Some authors have assumed to have a set of vertical parallel lines in the scene in order to estimate a vertical vanishing point, see for e.g. [32], [77]. This assumption may also not be satisfied when there is a lack of enough vertical lines in the scene, especially for natural scenes.

In this paper, we consider network of smart sensors (coupled camera and inertial sensors) fusion for 3D data registration and human movement analysis. Note that one of the outputs of IS is 3D earth cardinal orientation (North-East-Down), this orientation information can be exploited in a variety of ways, see [78]–[82]. In these previous works, we focused on using the 3D orientation information provided by IS in three particular cases: (1) vertical and earth-aligned sensors based virtual camera which acts as a hybrid/smart sensor, (2) set of Euclidean virtual planes computed from the scene, including the virtual ground plane, for active vision tasks, and (3) calibration of extrinsic parameters for a network of cameras. Here, we develop the concepts of a general framework further and additionally provide two principal contributions:

- 1) We derive the underlying geometric relations among the different inertial planes within the framework. The achieved equations can promote and expedite a further investigation and utilization of the framework.
- 2) We model and analyze the involved uncertainties within the framework, which can be exploited in further applications.

Moreover, we provide detailed derivations of the multisensor fusion registration framework. The geometric relations, or in other words homography transformations, among different projective and Euclidean inertial-based planes, within the proposed framework, are investigated and for each particular case a parametric homography function is obtained. Such homography transformations are used for 3D registration of the scene onto inertial-based Euclidean planes and a volumetric reconstruction algorithm is proposed. A set of experiments in 3D human and object reconstructions are carried out to present the usability of the proposed method for different areas such as scene understanding and human movement analysis. Overall method is implemented using CUDA enabled GP-GPU, thereby allowed us to obtain objects/person 3D reconstructions in realtime.

The rest of the paper is organized as follows. Section II provides details of the multi-plane 3D data registration framework. Section III derives the geometric relations among different Euclidean, projective planes, and 3D data registration using the obtained equations. The visibility problem in the proposed framework are also introduced. Section IV details the uncertainty modeling of inertial based homography. Section V presents the experimental results on the 3D reconstruction of human and objects in the scene. We also discuss the camera coverage problem in detail. Finally, Section VI conclusions are given.

II. MULTI-PLANE 3D DATA REGISTRATION FRAMEWORK

In this section, we illustrate the main components of the proposed 3D data registration framework with multisensor fusion. Figure 1 shows the overall approach we consider for 3D data registration framework using fusion of network of camera and IS sensors. We use these two types of sensors, a camera for image grabbing, and inertial sensor (IS), for obtaining 3D orientation. Each camera is rigidly coupled to an IS to obtain a hybrid sensor using infinite homography. This leads to fused downward-looking virtual cameras whose axes are aligned to the earth cardinal direction North-East-Down. The 3D orientation from IS is further used for defining a set of virtual and parallel Euclidean inertial planes. Using well-defined homography transformations, all the image planes of virtual cameras are projected onto this set of inertial planes, and the 3D points of the scene (person or object) are registered using parallax geometry.

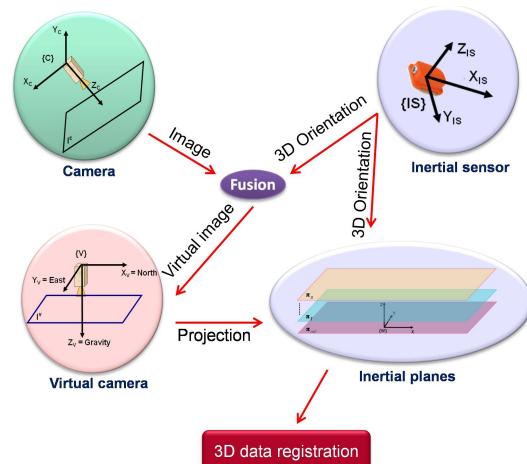


FIGURE 1. Overall fusion framework for 3D data registration: Infinite homography transform is used to fuse the 3D orientation from IS and image data from camera, to obtain a downward looking virtual camera. The axes of this virtual camera is aligned to the earth cardinal direction, North-East-Down. Further, 3D orientation information from IS is used to define a set of 2D Euclidean parallel inertial planes in the scene. The final 3D data registration is computed by projecting the obtained set of virtual images onto this set of inertial planes.

We assume the classical pinhole camera model, in which a 3D point $\mathbf{X} = [X \ Y \ Z]^T$ from the scene is projected onto an image plane of a camera as a 2D point, $\mathbf{x} = [x \ y \ 1]^T$. The relationship can be described by the following relation,

$$\mathbf{x} = K(R\mathbf{X} + \mathbf{t}), \quad (1)$$

where K is the *camera calibration matrix* which also known as the *intrinsic parameter matrix*, R and \mathbf{t} are respectively rotation matrix and translation vector between the world and camera coordinate systems, see the monograph [28] for basic details on multi-view geometry and notations. Here the camera matrix K is given by,

$$K = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

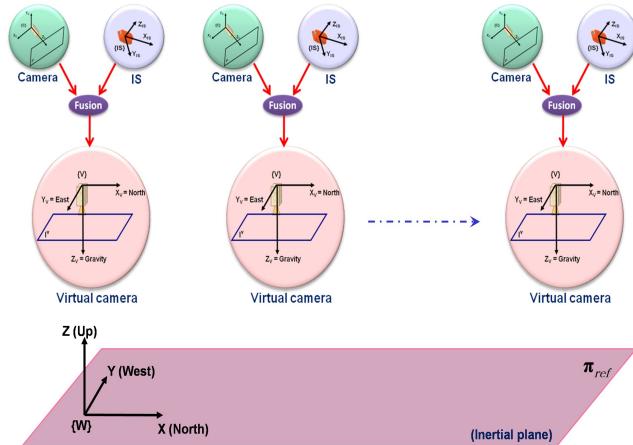


FIGURE 2. We utilize a distributed network of cameras and inertial sensors and fuse them to obtain a set of virtual cameras in which the 3D scene data is obtained. The Euclidean plane, π_{ref} , is defined as a virtual reference inertial plane which is a virtual ground for the 3D world.

in which (f_x, f_y) represent the focal length of the camera in the directions of (x, y) , and u_0 and v_0 are the elements of the principal point vector P . We use a homography transformation [28] to map points from one plane to another while preserving collinearity. Let us assume that a 3D plane, 2D plane in 3D space, is observed by two cameras and $\mathbf{x}_1, \mathbf{x}_2$ are two image points of a 3D point \mathbf{X} on the 3D plane. The points \mathbf{x}_1 and \mathbf{x}_2 are called a pair of corresponding points and the relation between them is given as $\mathbf{x}_2 = H\mathbf{x}_1$ where H is a 3×3 planar homography matrix induced by the 3D plane. This can be written as (up to scale, c.f. [83]),

$$H = K' \left(R + \frac{1}{d} \mathbf{t} \mathbf{n}^T \right) K^{-1}, \quad (3)$$

where R and \mathbf{t} are respectively rotation matrix and translation vector between the two cameras centers, \mathbf{n} is normal of the 3D plane, d is the orthogonal distance between the 3D plane and the camera center, K and K' are intrinsic parameters of the two cameras, here the first camera coordinate system is assumed as the world reference.

A. MULTISENSOR FUSION

Figure 2 illustrates a sensor network with multiple camera nodes, where π_{ref} is the inertial plane (written as π in some equations for simplicity), defined by the 3D orientation of IS, which is common for all cameras. Here $\{W\}$ denotes the world reference frame. In this setup, as mentioned before, each camera is rigidly coupled with an IS. Referring to Figure 4, the aim is to register a 3D point \mathbf{X} , observed by camera C , onto the reference plane π_{ref} as $\pi\mathbf{x}$ (2D), by using homography and fusing with inertial data. A virtual image plane is considered for each camera. Such a virtual image plane is defined (using inertial data) as a horizontal image plane at a distance f below the camera sensor, f being the focal length [61]. In other words, it can be thought that beside each real camera C in the setup, a virtual camera V exists whose center, $\{V\}$, coincides to the center of the real

camera $\{C\}$ (see Figure 4 below). So that the transformation matrix between $\{C\}$ and $\{V\}$ has a rotation part with the translation part is zero vector. The involved reference frames in this framework are indicated in Figure 3.

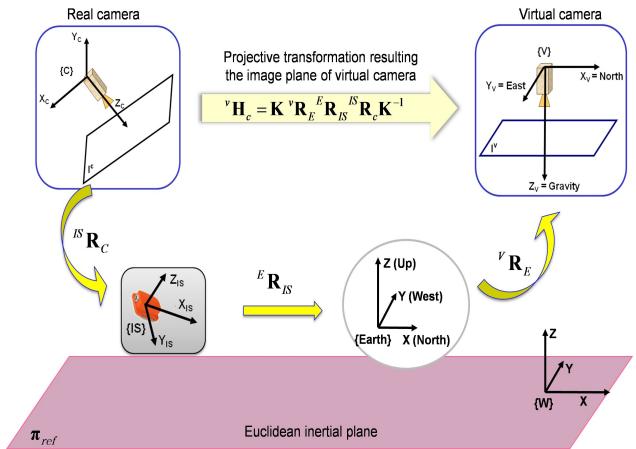


FIGURE 3. Schematic of a virtual camera: A virtual camera is created from a IS-camera pair by using infinite homography. Different coordinate systems are involved in this definition. **{Earth}**: Earth cardinal coordinate system, **{IS}**: Inertial reference frame expressed in **{Earth}**, **{W}**: world reference frame of the framework, **{C}**: camera reference frame, **{V}**: reference frame of the virtual camera corresponding to **{C}**. Based on the definition, the virtual camera has a horizontal image plane (projective), parallel to the euclidean image plane π_{ref} .

The motivation is to fuse the 3D orientation provided by IS to register image data on the reference plane π_{ref} defined in $\{W\}$, the world reference frame. The reference Euclidean plane π_{ref} is defined such a way that it spans the X and Y axis of $\{W\}$ and it has a normal parallel to the Z , see Figure 3 for the schematics. In this proposed method, the idea is not to use any real Euclidean plane inside the scene for estimating homography. Hence we assume there is no a real 3D plane available in the scene, so our $\{W\}$ becomes a virtual reference frame and consequently π_{ref} is a horizontal virtual plane on the fly. Although $\{W\}$ is a virtual reference frame, it needs to be specified and fixed in the 3D space. Therefore, we next define $\{W\}$ and as a result π_{ref} . Without loss of generality we place \mathbf{O}_W , the center of $\{W\}$, in the 3D space such a way that \mathbf{O}_W has a height d with respect to the first virtual camera, V_0 . Again with no loss of generality we specify its orientation the same as $\{E\}$ (earth fixed reference). As a result we can describe the reference frame of a virtual camera $\{V\}$ with respect to $\{W\}$ via the following homogeneous transformation matrix,

$${}^W T_V = \begin{bmatrix} {}^W R_V & \mathbf{t} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}, \quad (4)$$

where ${}^W R_V$ is a rotation matrix defined as (see Figure 3):

$${}^W R_V = \begin{bmatrix} \hat{\mathbf{i}} & -\hat{\mathbf{j}} & -\hat{\mathbf{k}} \end{bmatrix}, \quad (5)$$

with $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ being the unit vectors of the X , Y and Z axis, respectively, and \mathbf{t} is a translation vector of the

V 's center w.r.t $\{W\}$. Note that using the preceding definition and conventions, for the first virtual camera we have $\mathbf{t} = [0 \ 0 \ d]^T$.

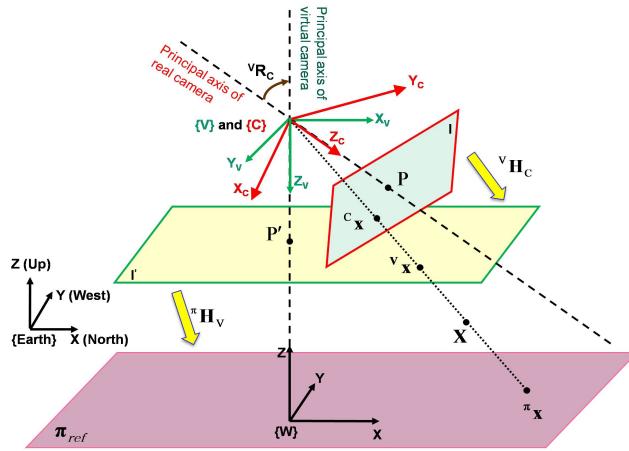


FIGURE 4. The concept of infinite homography is used to fuse inertial-visual information and define an earth cardinal aligned virtual camera. π_{ref} is well defined by using 3D orientation from inertial sensor. The red and green coordinate systems correspond to the real and virtual camera coordinates respectively. The principal points are shown by P and P' for two image planes. X is a 3D point from the scene and c_x is its projection on the real image plane I . I' depicts the projective image plane of the virtual camera. By applying v_H_c , the Homography transformation from I to I' , c_x gets registered on I' as v_x . π_H_V is homography transformation between the projective plane I' and euclidean inertial plane π_{ref} . By applying π_H_V , v_x can be registered on π_{ref} as π_x .

B. REGISTRATION USING HOMOGRAPHY

Based on the above observations, a 3D point X from the scene can be registered on π_{ref} as π_x using the following equation (see Figure 4):

$$\pi_x = \pi_H_V v_H_c K (RX + \mathbf{t})X,$$

where $K(RX + \mathbf{t})$ is the camera projective model, v_H_c is the homography matrix from camera image plane to virtual camera image plane and π_H_V is the homography from virtual camera image plane to the inertial plane π_{ref} . The 3×3 homography matrix v_H_c which transforms a point c_x from the real camera image plane I to its corresponding point v_x on the virtual camera image plane I' as $v_x = v_H_c c_x$ can be obtained from Eqn. (3) by considering \mathbf{t} equal to zero:

$$v_H_c = K v_R_c K^{-1}, \quad (6)$$

with v_R_c being the rotation between C and V and can be obtained by three consecutive rotations as:

$$v_R_c = {}^V R_E {}^E R_{IS} {}^{IS} R_c, \quad (7)$$

where ${}^{IS} R_C$ is rotation matrix from camera to IS, ${}^E R_{IS}$ is rotation matrix from IS to earth fixed reference and ${}^V R_E$ is the rotation matrix from earth to downward-looking virtual camera [84]. ${}^{IS} R_C$ can be obtained through a IS-camera calibration procedure. ${}^E R_{IS}$, is given by the IS sensor w.r.t $\{E\}$. ${}^V R_E = [\hat{\mathbf{i}} - \hat{\mathbf{j}} - \hat{\mathbf{k}}]$, since the $\{E\}$ has the Z upward but

the virtual camera is defined to be downward-looking with a downward Z.

We formalize the homography matrix π_H_V that transforms points from a virtual camera image plane I' to the common world 3D plane π_{ref} (see Figure 4). In general, a homography among a world plane and image plane can be expressed as,

$$\pi_H_V^{-1} = K[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}], \quad (8)$$

with \mathbf{r}_1 and \mathbf{r}_2 being the first and second columns of the 3×3 rotation matrix and \mathbf{t} is the translation vector between the π_{ref} and camera center [28]. In the case of our framework, π_{ref} is our world plane, \mathbf{r}_1 and \mathbf{r}_2 are the first and second columns of the rotation matrix ${}^W R_V$. By substituting \mathbf{r}_1 with $\hat{\mathbf{i}}$ and \mathbf{r}_2 with $-\hat{\mathbf{j}}$ (first and second columns of ${}^W R_V$ defined in Eqn. (5)), and $\mathbf{t} = [t_1 \ t_2 \ t_3]^T$ as the translation vector and eventually replacing K from its equivalence defined in Eqn. (2), Eqn. (8) becomes,

$$\pi_H_V^{-1} = \begin{bmatrix} f_x & 0 & f_x t_1 + u_0 t_3 \\ 0 & -f_y & f_y t_2 + v_0 t_3 \\ 0 & 0 & t_3 \end{bmatrix}. \quad (9)$$

As one can see, this homography equation depends on the translation vector \mathbf{t} among the cameras. Such a transformation can be either measured by a GPS (Global positioning system) e.g. in outdoor scenarios or can be estimated using two-points-based method presented in our previous work [79].

III. PARAMETRIC HOMOGRAPHY AND 3D RECONSTRUCTION

The inertial-based data registration and fusion framework was introduced in the previous section. This current section is dedicated to exploring geometric relationships among different virtual planes in the framework. As result, a set of equations, parametric homography functions, shall be obtained which then generate the appropriate homography matrix to perform 3D data registration.

A. EXPLORING PARAMETRIC HOMOGRAPHIES AMONG VIRTUAL CAMERA AND INERTIAL-PLANES

In the preceding section the homography matrix from the image plane of a virtual camera V to the world 3D plane π_{ref} was obtained as π_H_V (see Eqn. (9)). For the sake of multi-layer reconstruction, it is desired to obtain the homography matrix from a virtual image to another world 3D plane parallel to π_{ref} once we already have π_H_V . Lets consider π' as a 3D plane which is parallel to π_{ref} and has a height Δh w.r.t it (see Figure 5). $\pi' H_V$ denotes the homography transformation which maps the points of the image plane of V onto π' . By substituting t_3 in the Eqn. (9) with $t_3 + \Delta h$, $\pi' H_V$ can be expressed as a function of π_H_V and Δh as follows,

$$\pi' H_V^{-1}(\Delta h) = \pi_H_V^{-1} + \Delta h P \hat{\mathbf{k}}^T, \quad (10)$$

where $P = [u_0 \ v_0 \ 1]^T$ is the principal point of the camera V and $\hat{\mathbf{k}}$ is the unit vector of the Z axis.

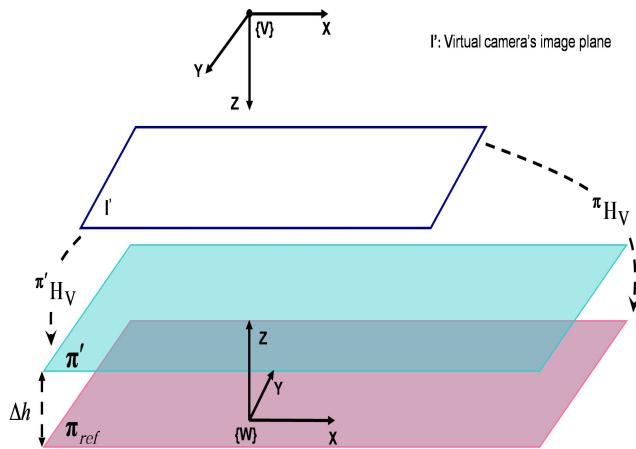


FIGURE 5. Extending homography for planes parallel to π_{ref} . $\pi' H_V$ is the available homography matrix among virtual image plane I' and the first inertial-based virtual plane π_{ref} . π' is another inertial-based virtual plane, parallel to π_{ref} . Δh is the distance among π and π' . The idea is to obtain $\pi' H_V$, the homography between the image plane and π' , having the $\pi' H_V$ and Δh .

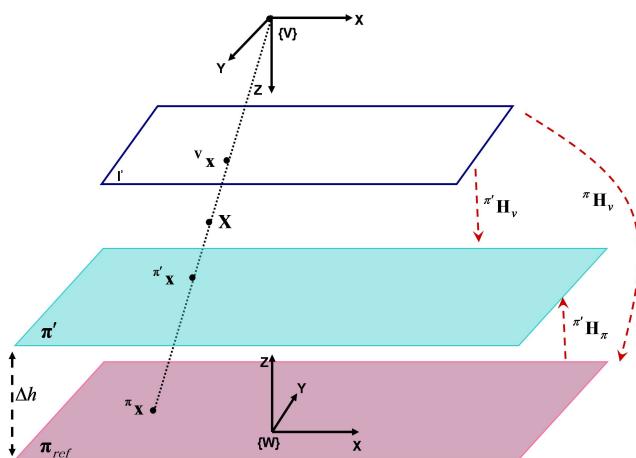


FIGURE 6. Homographic relation among an inertial-plane π' and the reference inertial-plane π_{ref} : The homography transformation $\pi' H_{\pi_{\text{ref}}}$, induced by image plane of virtual camera V , which maps points from π_{ref} onto π' can be expressed as a function of Δh , Δh being the euclidean distance between two inertial-planes.

B. HOMOGRAPHY RELATION AMONG EUCLIDEAN INERTIAL-PLANES

Suppose π' is an inertial-plane with an Euclidean distance Δh to the reference inertial plane π_{ref} . $\pi' H_{\pi}$ denotes the homography transformation among the two inertial-planes, induced by the image plane of a virtual camera, and is desired to be obtained (see Figure 6). Such a homography transformation can be expressed by the following equation,

$$\pi' H_{\pi} = \pi' H_V \pi H_V^{-1}, \quad (11)$$

where $\pi' H_V$ is the homography transformation among the image plane of a virtual camera V and the inertial-plane π' , and πH_V is the homography transformation between the image plane of V and the reference inertial-plane π_{ref} .

By substituting $\pi' H_V$ with Eqn. (10), Eqn. (11) becomes,

$$\pi' H_{\pi} = (\pi H_V^{-1} + \Delta h P \hat{\mathbf{k}}^T)^{-1} \pi H_V^{-1}. \quad (12)$$

The term $(\pi H_V^{-1} + \Delta h P \hat{\mathbf{k}}^T)^{-1}$ in above equation can be written in an equivalent form using the *Sherman-Morrison-Woodbury*¹ formula [85], [86] as follows,

$$(\pi H_V^{-1} + \Delta h P \hat{\mathbf{k}}^T)^{-1} \equiv \pi H_V - \frac{\pi H_V P \hat{\mathbf{k}}^T \pi H_V}{\alpha + \hat{\mathbf{k}}^T \pi H_V P}, \quad (13)$$

where $\alpha = \frac{1}{\Delta h}$. Eventually, Eqn. (12) after simplifications can be expressed as a function of the distance between two inertial-planes,

$$\pi' H_{\pi}(\alpha) = \mathbf{I}_{3 \times 3} - f(\alpha) \Gamma, \quad (14)$$

where $f(\alpha)$ is a scalar function of the vertical distance between two inertial plane,

$$f(\alpha) = \frac{1}{\alpha t_3 + 1}, \quad (15)$$

and Γ is a 3×3 matrix equal to,

$$\Gamma = [-t_1 \quad t_2 \quad 1]^T \hat{\mathbf{k}}^T. \quad (16)$$

Note that Γ is constant for all inertial planes induced by the camera V (assuming no movement for the cameras) and also independent of the camera intrinsic parameters. Eqn. (14) is interesting in the sense that once a basic homography $\pi' H_V$ to project an image to the reference inertial plane π_{ref} is obtained, a direct projection, which is independent to the intrinsic parameters, can be performed from π_{ref} to any arbitrary inertial plane namely π' with just knowing the Euclidean distance (Δh) among them for the purpose of 3D data registration.

While Eqn. (14) expresses the projective relation among the reference plane π_{ref} and other inertial planes, we are interested to obtain an equation which could express the projective relation among any two consecutive inertial planes, namely $\pi_k H_{\pi_{k-1}}$. Figure 7 depicts a set of inertial planes where the Euclidean distance among any two consecutive planes is equal to Δh_0 . Suppose $\pi_k H_{\pi}$ expresses the homography projection, induced by a virtual image, from the reference plane π_{ref} to k -th inertial plane. Such a transformation is written as,

$$\pi_k H_{\pi} = (\pi_k H_{\pi_{k-1}})(\pi_{k-1} H_{\pi}),$$

and then,

$$\pi_k H_{\pi_{k-1}} = (\pi_k H_{\pi})(\pi_{k-1} H_{\pi}^{-1}). \quad (17)$$

In Eqn. (17), by substituting the terms $\pi_k H_{\pi}$ with its equivalence from Eqn. (14) and $\pi_{k-1} H_{\pi}$ with Eqn.(12) we have (considering $\alpha_0 = 1/\Delta h_0$),

$$\pi_k H_{\pi_{k-1}} = \left(\mathbf{I}_{3 \times 3} - \frac{k}{\alpha_0 t_3 + k} \Gamma \right) \quad (18)$$

¹Considering A as a square matrix and U and V as two column vectors, the *Sherman-Morrison-Woodbury* formula gives $(A + UV^T)^{-1} = A^{-1} - \frac{A^{-1}UV^TA^{-1}}{1+V^TA^{-1}U}$. Other alternative formulae exists as well, see [85].

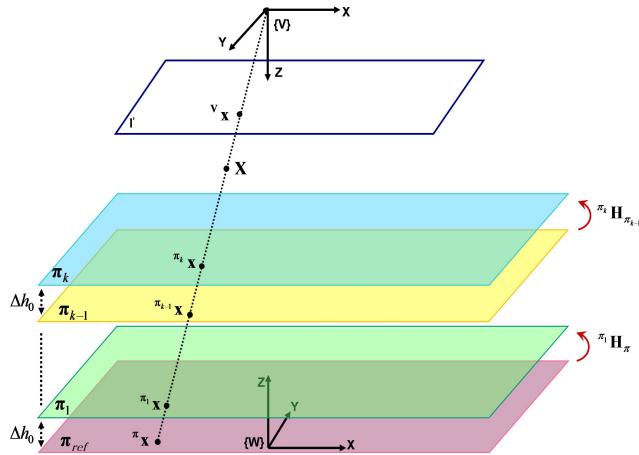


FIGURE 7. Parametric homography between two consecutive inertial-planes induced by a virtual image. The image shows a set of Euclidean inertial-planes where the distance between two consecutive planes is equal to Δh_0 . In this case, the homography transformation among any two consecutive inertial-planes can be expressed as a function of Δh_0 and the index of the plane, see Eqn. (20).

$$\begin{aligned} & \left[(\pi^* H_v^{-1} + \frac{k-1}{\alpha_0} P \hat{\mathbf{k}}^T)^{-1} \pi^* H_v^{-1} \right]^{-1} \\ &= \left(\mathbf{I}_{3 \times 3} - \frac{k}{\alpha_0 t_3 + k} \Gamma \right) \left(\mathbf{I}_{3 \times 3} + \frac{k-1}{\alpha_0 t_3} \Gamma \right), \quad (19) \end{aligned}$$

which after simplification becomes,

$$\pi^* H_{\pi_{k-1}}(k, \alpha_0) = \mathbf{I}_{3 \times 3} - g(\alpha_0, k) \Gamma, \quad (20)$$

where $g(\alpha_0, k)$ is a scalar function whose inputs are the index of inertial plane π_k and the vertical resolution factor α_0 as follows,

$$g(\alpha_0, k) = \frac{1}{\alpha_0 t_3 + k}. \quad (21)$$

The Γ is a constant 3×3 matrix which is independent of the camera intrinsic parameters. Therefore the obtained homography $\pi^* H_{\pi_{k-1}}$ in Eqn. (20) expresses a homography matrix which transforms the 2D points from inertial plane π_{k-1} to its consecutive inertial plane π_k , independent of the camera intrinsic parameters (as expected), and is a function of k and α_0 .

C. HOMOGRAPHIC RELATION AMONG IMAGE PLANES OF VIRTUAL CAMERAS

In the previous section, the homography transformation between image plane of a virtual camera and an Euclidean virtual plane (π) was obtained. Here we explain the homography transformation between the images of two virtual cameras in a parametric form. Figure 8 depicts two virtual cameras V_i and V_j with their reference frames. Without loss of generality, we consider V_i as the world reference frame here. The idea is to obtain ${}^j H_i^\pi$, the homography matrix among V_i and V_j , induced by an inertial plane such as π . Based on Eqn. (3), ${}^j H_i^\pi$ can be expressed as:

$${}^j H_i^\pi = K_j \left(R + \frac{1}{d} \Delta \mathbf{t} \mathbf{n}^T \right) K_i^{-1}, \quad (22)$$

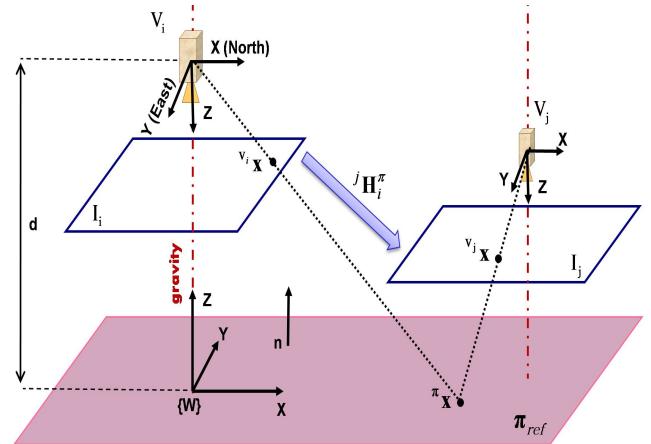


FIGURE 8. Homography between image planes of two virtual cameras.

where K_j and K_i are the camera calibrations matrices, respectively for V_j and V_i . Since there is no rotation among the virtual cameras then R becomes equal to the identity matrix ($I_{3 \times 3}$). $\Delta \mathbf{t}$ is a 3-elements vector describing the translation from V_i to V_j . $\mathbf{n} = [0 \ 0 \ -1]^T$ is the normal of plane π and d is the distance between π and $\{V_i\}$ along the Z axis of $\{V_i\}$. Therefor, after substitutions and simplifications, Eqn. (22) can be expressed as:

$${}^j H_i^\pi = K_j [\hat{\mathbf{i}} \ \hat{\mathbf{j}} \ (\hat{\mathbf{k}} - \frac{\Delta \mathbf{t}}{d})] K_i^{-1}, \quad (23)$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are the unit vectors for X, Y and Z axes, respectively. Assuming no changes in camera parameters, Eqn. (23) generates the homography matrix related to the parameter d , the Euclidean distance between the inertial-plane and $\{V_i\}$.

Note that Eqn. (23) expresses the homography relation among the image planes of two cameras, induced by the reference inertial plane π_{ref} . It is interesting to obtain the homography between two image planes induced by another inertial plane (π') using the basic relation from Eqn. (23). Such a homography matrix can be notated as ${}^{v_2} H_{v_1}^{\pi'}$ and is depicted in Figure 9. One can write this homography as,

$${}^{v_2} H_{v_1}^{\pi'} = ({}^{v_2} H_{\pi'}) ({}^{\pi'} H_{v_1}) = ({}^{\pi'} H_{v_2}^{-1}) ({}^{\pi'} H_{v_1}).$$

The terms ${}^{\pi'} H_{v_2}^{-1}$ and ${}^{\pi'} H_{v_1}$ can be replaced by their equivalences using from Eqns. (10) and (13), respectively:

$${}^{v_2} H_{v_1}^{\pi'} = ({}^{\pi'} H_{v_2}^{-1} + \frac{1}{\alpha} P_2 \hat{\mathbf{k}}^T) \left({}^{\pi'} H_{v_1} - \frac{{}^{\pi'} H_{v_1} P_1 \hat{\mathbf{k}}^T {}^{\pi'} H_{v_1}}{\alpha + \hat{\mathbf{k}}^T {}^{\pi'} H_{v_1} P_1} \right), \quad (24)$$

where P_1 and P_2 are respectively the principal vectors of the virtual cameras V_1 and V_2 , α is equal to the inverse of Δh , the distance among the two inertial planes π_{ref} and π' . After simplification and replacing ${}^{\pi'} H_{v_2}^{-1}$ with ${}^{v_2} H_{v_1}^{\pi'}$ we have,

$${}^{v_2} H_{v_1}^{\pi'} = {}^{v_2} H_{v_1}^{\pi'} (I_{3 \times 3} - f(\alpha) P_1 \hat{\mathbf{k}}^T) + f(\alpha) P_2 \hat{\mathbf{k}}^T, \quad (25)$$

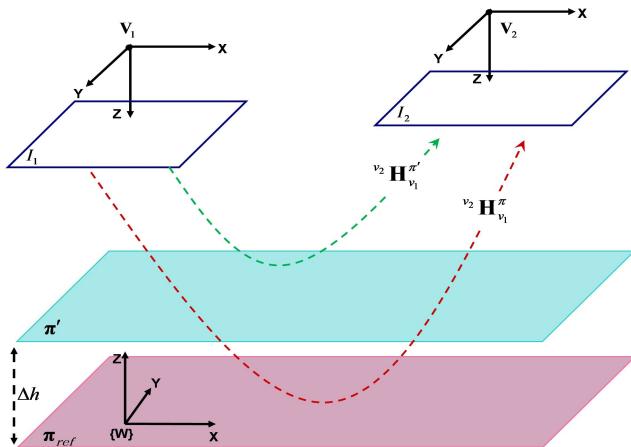


FIGURE 9. Homography between the image planes of two virtual cameras, induced by an inertial plane π' parallel to the reference inertial-plane π_{ref} .

where $f(\alpha)$ was previously defined in Eqn (15). As one can see, the Eqn. (25) expresses the homography between two virtual cameras induced by an inertial plane π' parallel to π_{ref} , by using a linear equation of the homography among the same virtual cameras but induced through the reference inertial plane π_{ref} .

D. VOLUMETRIC RECONSTRUCTION

The geometric models for projecting 3D data onto a set of virtual horizontal planes based on the concept of homography was previously introduced. The homography transformation can be interpreted as shadow on each inertial-based virtual plane created by a light source located at the camera position. Considering several cameras (remembering light sources) which observe the object, then different shadows will appear on the inertial planes. Conceptually, the intersection between each one of these planes and the observed object can be thus obtained by using the intersections of all shadows.

There is a geometrical explanation to support this interpretation. Figure 10 demonstrates a person being observed by two virtual cameras V_1 and V_2 . π' is an inertial plane which passes across the person. X and Y are two 3D points from the person surface. X lies on the plane π' and Y is off the plane. The 3D points X and Y are imaged as x_1, y_1, x_2 and y_2 on the image planes of V_1 and V_2 , respectively (using the proposed homography methods). Suppose $\pi' x_1, \pi' y_1, \pi' x_2$ and $\pi' y_2$ are respectively the projections of the imaged points x_1, y_1, x_2 and y_2 onto π' . As seen in Figure 10, for an on the plane point such as X , all three points $X, \pi' x_1, \pi' x_2$ are coincident and meet on π' . In contrary, for the point Y which is off the plane, the three points $Y, \pi' y_1, \pi' y_2$ are distinct. y'_2 denotes the image of $\pi' y_1$ on the image plane of V_2 . The vector between y'_2 and y_2 is called *parallax*. Indeed the line through y'_2 and y_2 is the image of the ray passing through the center of V_1 and Y (which is also an *epipolar line*). For all points off the inertial plane π' , the norm of their parallax is bigger than zero and for those points which are on π' , there is no parallax or in other words their parallax's norm is zero.

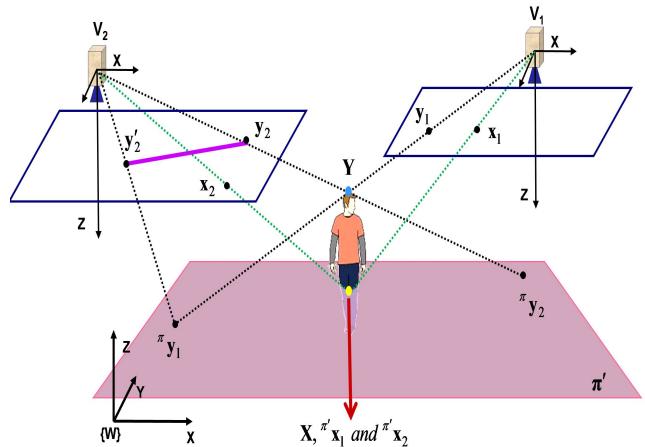


FIGURE 10. Geometric interpretation of the intersection among a person and an inertial plane π' : X and Y are two exemplary 3D points belonging to the person's body, being observed by two virtual cameras V_1 and V_2 . X lies on π' and Y is off the plane π' . The 3D points get projected on π' using the proposed homographic method. The homographic projections of the point which are on π' such as X are coincident ($\pi' x_1, \pi' x_2$) whereas for 3D points off the plane (such as Y) their projections on π' are distinct ($\pi' y_1$ and $\pi' y_2$). In other words, for the points off π' there is a parallax (like the vector through y_2 and y'_2).

Algorithm 1 summarizes the procedure of the proposed approach for 3D data registration by using the obtained parametric homography functions. In this algorithm, `ThreeDimRegistration()` is the main function. In ‘main’, some variables are initialized and also the image planes of the virtual cameras are obtained. Then the obtained virtual images get projected onto the reference Euclidean plane $\pi_{ref}^{(v_i)}$. Please note that for the sake of practical implementation the process of finding the intersection of the projected points onto the Euclidean planes (intersection of ‘shadows’ described previously), for each virtual camera v_i we have considered a temporary Euclidean plane $\pi_{ref}^{(v_i)}$. After initialization, for each camera, `RegisterRecursive()` function is called. This function recursively projects and registers the image data onto the consecutive inertial planes in the scene. In this function, `Warp()` is a function which performs the operation of usual homography warping. The function `GenerateNextHomography(k)` accepts k as the index for k -th Euclidean plane and generates a homography matrix for mapping points to the next Euclidean plane. The last operation in ‘main’ is to perform the cell-wise intersection among the temporary Euclidean planes corresponding to the same level.² This operation is basically a production of the corresponding binary cells as is illustrated in Figure 11.

IV. UNCERTAINTY MODELING OF INERTIAL-BASED HOMOGRAPHY

To be aware of the uncertainties 3D data registration is important for further applications, especially when the data are registered by fusion from different sources. The introduced

²We refer to the Supplementary PDF for algorithmic explanations, including an efficient GP-GPU implementation and computational complexity of our approach.

ALGORITHM 1 3D Data Registration Using Inertial-Planes in a Recursive Form. k Is the Index of Inertial Plane and α_0 Is the Inverse of the Euclidean Distance Between Two Consecutive Inertial Planes

Inputs: Images I_{c_i} , Rotation R_{c_i} , Intrinsic K_i , translation t_{c_i} for $i = 1, \dots, N_c$

Function ThreeDimRegistration()

begin

/* Initialization */

for $i \leftarrow 1$ to N_c do

$v_i H_{c_i} \leftarrow K_i v_i R_{c_i} K_i^{-1}$

$I_{v_i} \leftarrow v_i H_{c_i} I_{c_i}$

$\pi_{ref}^{(v_i)} H_{v_i} \leftarrow \text{inv}(K_{c_i} [\hat{\mathbf{i}} - \mathbf{J} t_{c_i}])$ Eqn. (9)

$\pi_{ref}^{(v_i)} \leftarrow \pi_{ref}^{(v_i)} H_{v_i} I_{v_i}$

$\Gamma_{v_i} \leftarrow [-t_1 \ t_2 \ 1]^T \hat{\mathbf{k}}^T$ Eqn. (16)

/* Performing recursive registration for each camera */

for $i \leftarrow 1$ to N_c do

$\pi_{N_\pi-1}^{(v_i)} \leftarrow \text{RegisterRecursive } (N_\pi - 1,$

$\pi_{N_\pi-1}^{(v_i)})$

/* Performing cell-wise intersection, see Figure 11 */

for $i \leftarrow 1$ to N_π do

$\pi_i \leftarrow \prod_{j=1}^{N_c} \pi_j^{(v_i)}$

Output: 2D registered planes π_i for $i = 1, \dots, N_\pi$

Function RegisterRecursive ($k, \pi^{(v_i)}$)

begin

if $k == 0$ then

return $\pi_{ref}^{(v_i)}$

else

$\pi_k H_{\pi_{k-1}} \leftarrow$

GenerateNextHomography ($k - 1$)

$\pi_k^{(v_i)} \leftarrow \text{Warp}(H, \text{RegisterRecursive}$

$(k - 1, \pi_{k-1}^{(v_i)}))$

return $\pi_k^{(v_i)}$

Function GenerateNextHomography (k)

begin

$\pi_k H_{\pi_{k-1}}(k, \alpha_0) = \mathbf{I}_{3 \times 3} - g(k, \alpha_0) \boldsymbol{\Gamma}$ Eqn. (20)

return H

inline Function $g(k, \alpha_0)$ return $\frac{1}{\alpha_0 t_3 + k}$ Eqn. (21)

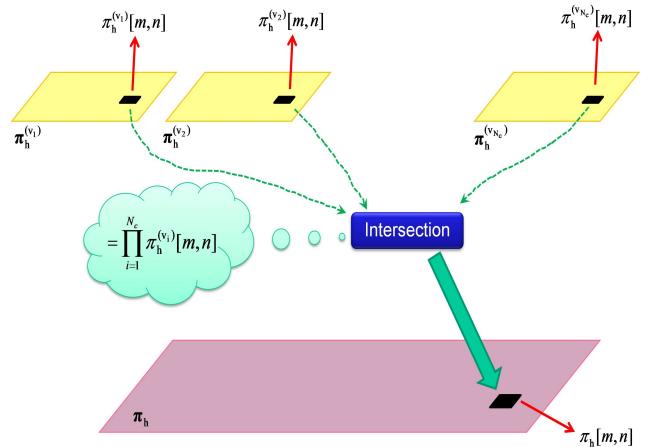


FIGURE 11. Cell-wise intersection of the projections of the virtual images onto an inertial-plane π_h : As seen in Algorithm 1, the images of all virtual cameras firstly get projected onto a temporary inertial plane. $\pi_h^{(v_i)}$ indicates the temporary inertial-plane corresponding to the virtual camera V_i . Then the corresponding cells of all temporary inertial-planes are fused using an AND operator in order to provide the final registration on the inertial-plane π_h , m and n indicate the indices of a cell. Note that the images are considered as binary.

analyze uncertainty of registered data using statistical geometry. First we model the uncertainty for the image plane of a virtual camera, and then model uncertainties of the points registered on an Euclidean inertial plane.

A. UNCERTAINTY OF IMAGE PLANE OF VIRTUAL CAMERAS

The infinite homography presented by Eqn. (6) depends to the 3D orientation measured by IS. Such an orientation can be presented by a random vector,

$$\mathbf{s} = [\theta_r \quad \theta_p \quad \theta_y]^T, \quad (26)$$

where θ_r , θ_p and θ_y denote the three elements of the Euler angles (respectively roll, pitch and yaw). We assume that \mathbf{s} has a mean equal to zero and a covariance of,

$$\boldsymbol{\Sigma}_s = \text{diag}\{\delta_r^2, \delta_p^2, \delta_y^2\},$$

where δ_r , δ_p and δ_y are respectively the standard deviations for θ_r , θ_p and θ_y . In the homography formula of Eqn. (6), ${}^v H_c$, can be expressed as a linear function of the orientation vector $f : \mathbf{s} \mapsto {}^v H_c$ where it maps the input three angles into the 9-elements homography matrix (R^3 into R^9). For simplicity, we express the homography matrix H as,

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix},$$

and assume \mathbf{h} as a vector form of H . Then we consider ${}^v \mathbf{h}_s$ as a random vector and are interested to model its uncertainty. Using a first-order Taylor approximation, as presented in [87], the uncertainty of H can be obtained as,

$$\boldsymbol{\Sigma}_h = \mathbf{J}_{h,s} \boldsymbol{\Sigma}_s \mathbf{J}_{h,s}^T,$$

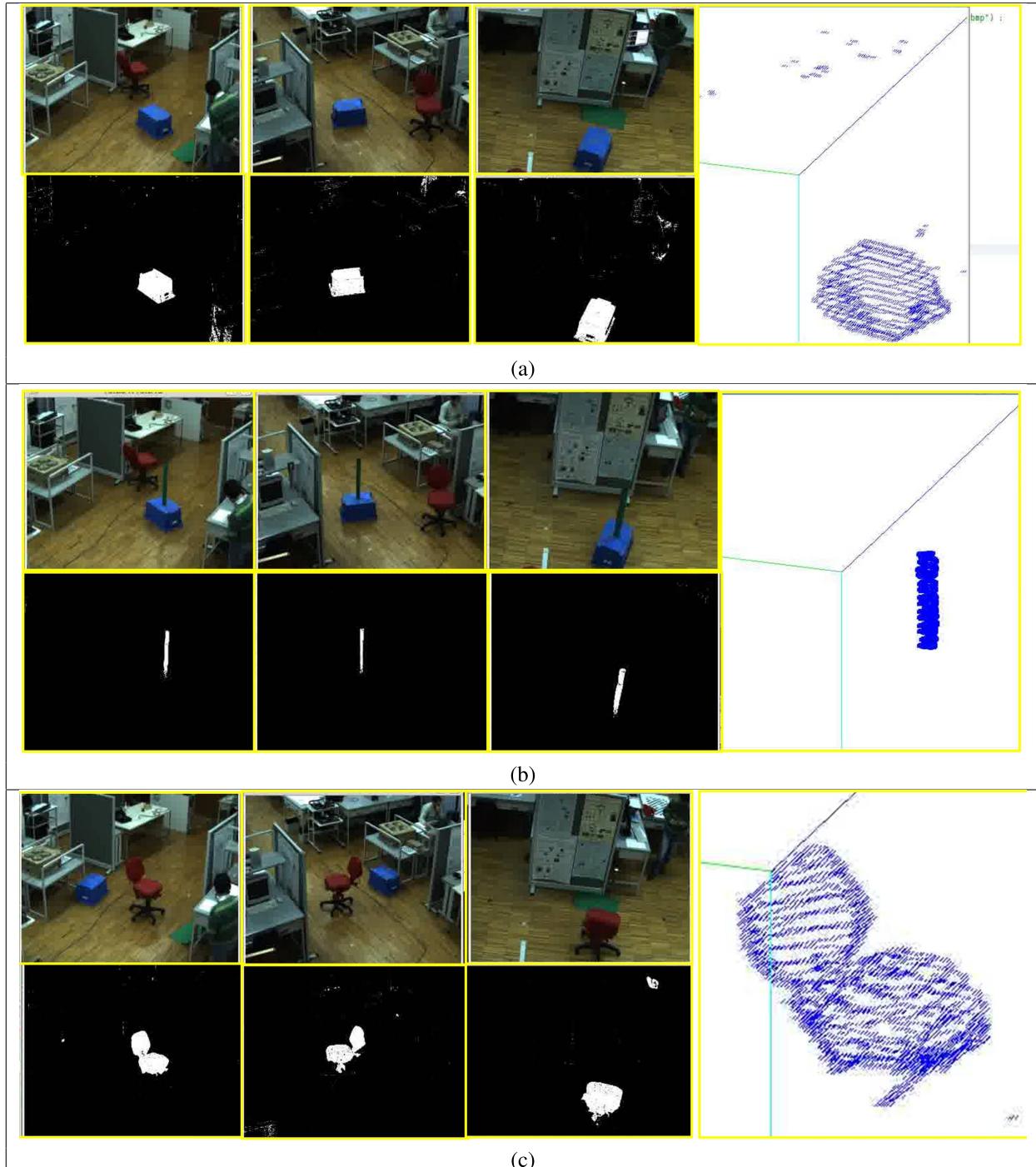


FIGURE 12. Results of the proposed multi-layer 3D data registration. Three experiments, each one for a different object, are carried out. (a) represents the result for a semi-rectangular blue box. (b) depicts the result for a small cylindrical green object on the top of a box. (c) demonstrates the result for the red covered parts of a chair. In all these experiments a color-based background subtraction is performed. The image at the right demonstrate the visualized result.

where \mathbf{J} is a Jacobian matrix:

$$\mathbf{J}_{\mathbf{h}, \mathbf{s}} = \begin{bmatrix} \partial h_1 / \partial \theta_r & \partial h_1 / \partial \theta_p & \partial h_1 / \partial \theta_y \\ \partial h_2 / \partial \theta_r & \partial h_2 / \partial \theta_p & \partial h_2 / \partial \theta_y \\ \vdots & \vdots & \vdots \\ \partial h_9 / \partial \theta_r & \partial h_9 / \partial \theta_p & \partial h_9 / \partial \theta_y \end{bmatrix}.$$

The homography transformation ${}^v H_c$ maps points from real camera to virtual camera's image plane. Having the uncertainty of the homography matrix ${}^v H_c$, we consequently can characterize the uncertainty for the mapped points. The points on virtual camera's image plane, ${}^v \mathbf{x}$, are obtained by the

$$\Sigma_{\pi} \mathbf{h}_v = \frac{1}{t_3^4} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \delta_{t_1}^2 t_3^2 + \delta_{t_3}^2 t_1^2 & 0 & 0 & -\delta_{t_3}^2 t_1 t_2 & 0 & 0 & -\delta_{t_3}^2 t_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\delta_{t_3}^2 t_1 t_2 & 0 & 0 & -\delta_{t_2}^2 t_3^2 + \delta_{t_3}^2 t_2^2 & 0 & 0 & \delta_{t_3}^2 t_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\delta_{t_3}^2 t_1 & 0 & 0 & \delta_{t_3}^2 t_2 & 0 & 0 & \delta_{t_3}^2 \end{bmatrix} \quad (31)$$

$$\mathbf{J}_{\pi} \mathbf{h}_{\pi}, \mathbf{t} = f(\alpha) \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -\alpha t_1 f(\alpha) & 0 & 0 & \alpha t_2 f(\alpha) & 0 & 0 & \alpha f(\alpha) \end{bmatrix}^T \quad (34)$$

$$\Sigma'_{\pi} \mathbf{x} = f^4(\alpha) \begin{bmatrix} \delta_1^2 f^{-2}(\alpha) + (\alpha \delta_3)^2 (t_1 + x))^2 & -(\alpha \delta_3)^2 (t_1 + x)(t_2 - y) \\ -(\alpha \delta_3)^2 (t_1 + x)(t_2 - y) & \delta_2^2 f^{-2}(\alpha) + (\alpha \delta_3)^2 (t_2 - y))^2 \end{bmatrix} \quad (35)$$

$$\Sigma_{\pi_k} \mathbf{h}_{\pi_{k-1}} = \frac{\alpha_0^2 \delta_3^2}{g^4(\alpha_0, k)} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & t_1^2 + \frac{\delta_1^2}{\alpha_0^2 \delta_{t_3}^2 g^2(k, \alpha_0)} & 0 & 0 & -t_1 * t_2 & 0 & 0 & -t_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t_1 * t_2 & 0 & 0 & t_2^2 + \frac{\delta_2^2}{\alpha_0^2 \delta_{t_3}^2 g^2(k, \alpha_0)} & 0 & 0 & t_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t_1 & 0 & 0 & t_2 & 0 & 0 & 1 \end{bmatrix} \quad (37)$$

$$\Sigma'_{\pi} \mathbf{x} = g^4(\alpha_0, k) \begin{bmatrix} g^{-2}(k, \alpha) \delta_1^2 + \alpha_0^2 (t_1 + x)^2 \delta_3^2 & -\alpha_0^2 \delta_3^2 (t_1 + x) (t_2 - y) \\ -\alpha_0^2 \delta_3^2 (t_1 + x) (t_2 - y) & g^{-2}(k, \alpha) \delta_2^2 + \alpha^2 (t_2 - y)^2 \delta_3^2 \end{bmatrix} \quad (38)$$

following mapping,

$${}^v \mathbf{x} = {}^v H_c \mathbf{x}, \quad (27)$$

with \mathbf{x} being a point from real camera's image plane. Assuming no uncertainty in real camera's image, the uncertainty of the points on the virtual image plane can be expressed as following (according to [88]):

$$\Sigma_v \mathbf{x} = (\mathbf{I} \otimes {}^v \mathbf{x}^T) \Sigma_h (\mathbf{I} \otimes {}^v \mathbf{x}), \quad (28)$$

where \mathbf{I} is a 3×3 identity matrix and \otimes denotes Kronecker product.

B. UNCERTAINTY OF EUCLIDEAN INERTIAL-PLANES

Earlier, we used Eqn. (9), ${}^{\pi} \mathbf{x} = {}^{\pi} H_v {}^v \mathbf{x}$, in order to project points from image plane of virtual camera onto Euclidean inertial planes. The uncertainty for such projected points is influenced by first the uncertainty of the points from virtual image, and then by the uncertainty of the homography transformation ${}^{\pi} H_v$. We first assume the points on

virtual image and consider uncertainty for the homography matrix ${}^{\pi} H_v$ alone. Afterwards, we take into account the uncertainties of the virtual image's points and propagate them.

1) UNCERTAINTY OF HOMOGRAPHY FROM VIRTUAL IMAGE TO EUCLIDEAN PLANE

We next consider the uncertainty of ${}^{\pi} H_v$, Eqn. (9). Such a homography is considered as a linear function of the translation vector $\mathbf{t} = [t_1 \ t_2 \ t_3]^T$. We assume the uncertainty of \mathbf{t} with following covariance matrix,

$$\Sigma_{\mathbf{t}} = \text{diag}\{\delta_{t_1}^2, \delta_{t_2}^2, \delta_{t_3}^2\},$$

where δ_{t_1} , δ_{t_2} and δ_{t_3} denote the standard deviations for the three elements of the translation vector. Again we use the first-order of Taylor approximation [87] and express the uncertainty of ${}^{\pi} \mathbf{h}_v$ (the vector form of ${}^{\pi} H_v$) as,

$$\Sigma_{\pi} \mathbf{h}_v = \mathbf{J}_{\pi} \mathbf{h}_v, \mathbf{t} \Sigma_{\mathbf{t}} \mathbf{J}_{\pi}^T \mathbf{h}_v, \mathbf{t}, \quad (29)$$



FIGURE 13. Results of the proposed multi-layer 3D data registration for an object (small cylindrical green) in the hand of a person. The image at the right demonstrate the visualized result.

where $\mathbf{J}_{\pi} \mathbf{h}_v, \mathbf{t}$ is a Jacobian matrix:

$$\mathbf{J}_{\pi} \mathbf{h}_v, \mathbf{t} = \begin{bmatrix} \partial h_1 / \partial t_1 & \partial h_1 / \partial t_2 & \partial h_1 / \partial t_3 \\ \partial h_2 / \partial t_1 & \partial h_2 / \partial t_2 & \partial h_2 / \partial t_3 \\ \vdots & \vdots & \vdots \\ \partial h_9 / \partial t_1 & \partial h_9 / \partial t_2 & \partial h_9 / \partial t_3 \end{bmatrix}. \quad (30)$$

After simplification, Eqn. (30) can be written as (31), as shown at the top of the previous page.

Having the uncertainty of the homography transformation πH_v , the uncertainty for points to be mapped via this homography is given by,

$$\Sigma'_{\pi' \mathbf{x}} = (\mathbf{I} \otimes {}^v \mathbf{x}^T) \Sigma_{\pi} \mathbf{h}_v (\mathbf{I} \otimes {}^v \mathbf{x}),$$

provided that the points ${}^v \mathbf{x}$ are certain.

2) PROPAGATION OF UNCERTAINTIES ON EUCLIDEAN INERTIAL PLANES

We next take into account the uncertainties for points from virtual camera image plane and their propagations on the Euclidean inertial plane π . In this case, the two covariance matrices, $\Sigma_{\pi} \mathbf{h}_v$ and $\Sigma_{\pi' \mathbf{x}}$ get augmented [88] and the uncertainty for a registered point on the Euclidean inertial plane, ${}^{\pi} \mathbf{x}$, becomes,

$$\Sigma_{\pi \mathbf{x}} = {}^{\pi} H_v \Sigma_{\pi' \mathbf{x}} {}^{\pi} H_v^T + \Sigma'_{\pi' \mathbf{x}}. \quad (32)$$

3) PROPAGATION OF UNCERTAINTIES BETWEEN TWO INERTIAL PLANES

Next we model the uncertainties between two parallel Euclidean inertial planes. Having the geometric relations between two Euclidean inertial planes, the next step is to obtain the uncertainty for the homography transformation which maps the 2D points among them. This uncertainty can be expressed as,

$$\Sigma'_{\pi' \mathbf{h}_{\pi}} = \mathbf{J}_{\pi' \mathbf{h}_{\pi}, \mathbf{t}} \Sigma_t \mathbf{J}_{\pi' \mathbf{h}_{\pi}, \mathbf{t}}^T. \quad (33)$$

where $\mathbf{J}_{\pi' \mathbf{h}_{\pi}, \mathbf{t}}$ is the following Jacobian matrix (34), as shown at the top of the previous page.

Now, the uncertainty for a 2D point ${}^{\pi} \mathbf{x}$ lying on the inertial plane π which is through the homography ${}^{\pi'} H_{\pi}(\alpha)$ mapped onto another inertial plane as ${}^{\pi} x$ is written as,

$$\Sigma'_{\pi' \mathbf{x}} = (\mathbf{I} \otimes {}^{\pi'} \mathbf{x}^T) \Sigma_{\pi' \mathbf{h}_{\pi}} (\mathbf{I} \otimes {}^{\pi'} \mathbf{x}).$$

This equation after simplification becomes (35), as shown at the top of the previous page.

The uncertainty of the homography of Eqn. (20) is given as,

$$\Sigma_{\pi_k H_{\pi_{k-1}}} = \mathbf{J}_{\pi_k H_{\pi_{k-1}}, \mathbf{t}} \Sigma_t \mathbf{J}_{\pi_k H_{\pi_{k-1}}, \mathbf{t}}^T, \quad (36)$$

where $\mathbf{J}_{\pi_k H_{\pi_{k-1}}, \mathbf{t}}$ is a Jacobian matrix. Eqn. (36) after simplifications is given in (37), as shown at the top of the previous page.

The uncertainty for a point ${}^{\pi_k} \mathbf{x}$ which is mapped on the inertial sensor π_k via its previous plane π_{k-1} is,

$$\Sigma'_{\pi_k \mathbf{x}} = (\mathbf{I} \otimes {}^{\pi_k} \mathbf{x}) \Sigma_{\pi_k \mathbf{h}_{\pi_{k-1}}} (\mathbf{I} \otimes {}^{\pi_k} \mathbf{x}).$$

By substituting the term $\Sigma_{\pi_k \mathbf{h}_{\pi_{k-1}}}$ in above equation and after simplification we obtain Eqn. (38), as shown at the top of the previous page.

V. EXPERIMENTS

Different experiments are introduced here to demonstrate the proposed approaches in this paper. Firstly a set of experiments, related to 3D reconstruction of human and objects in the scene, is presented. Overall method is implemented using CUDA (with 256 cores) enabled GP-GPU, which allowed us to obtain objects/person 3D reconstructions in realtime. Without CUDA implementation the processing time for our method was well over 10 hours in a standard MATLAB environment. Then some simulated results from the proposed camera coverage problem in the proposed geometric fusion framework are studied.

A. MULTI-LAYER 3D DATA REGISTRATION

A set of experiments were carried out to demonstrate the proposed 3D registration approach. Note that a planar ground is visible in the experiments we show here, but it is not required

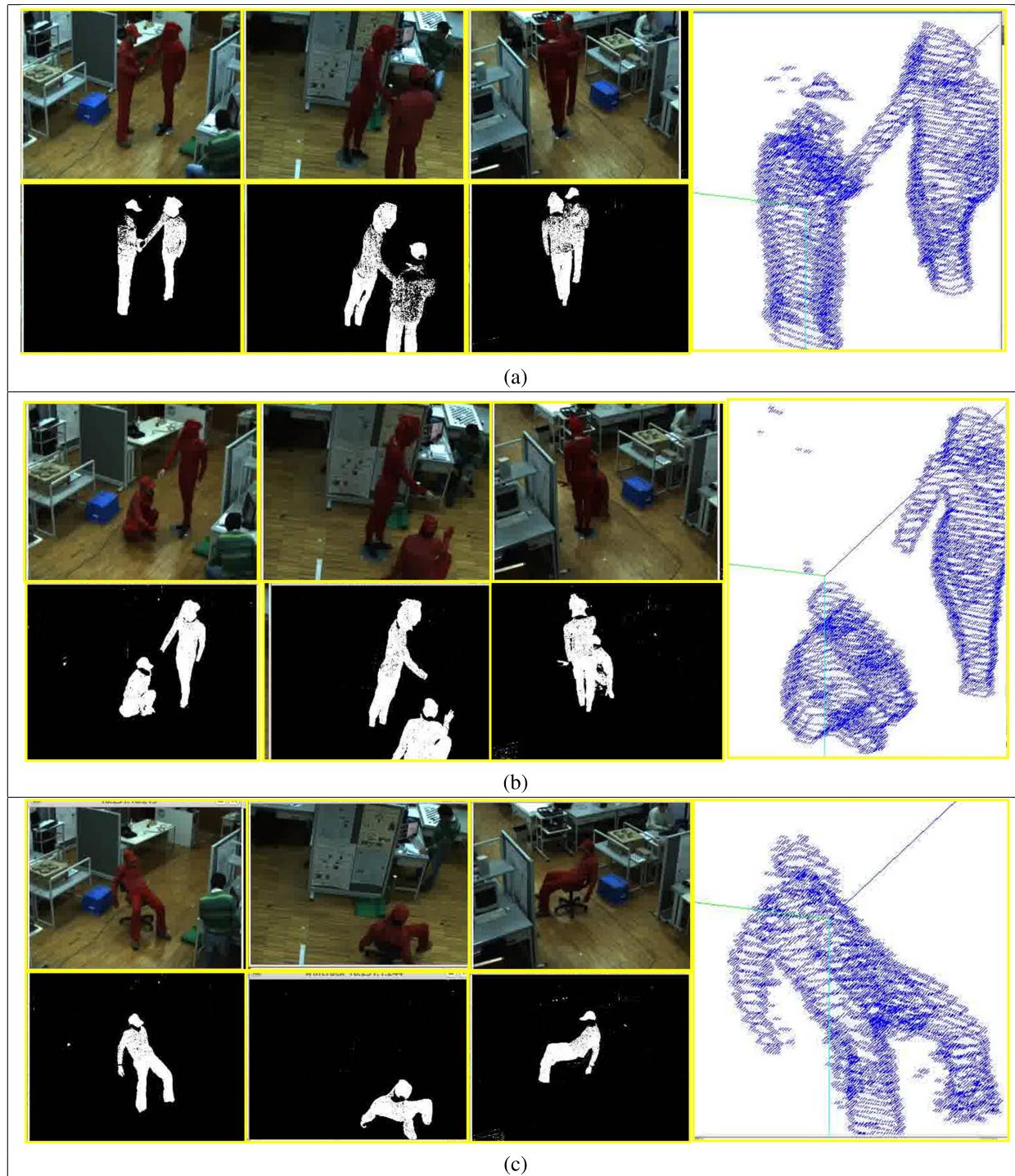


FIGURE 14. Proposed multi-layer 3D data registration works well for various dynamic scenarios. The first experiment, (a), stands for a scene where a person is hand-shaking with a manikin. In the second one, (b), the person is seated in front of the manikin. (c) shows a case where the person is seated on a chair. A set of euclidean inertial planes are used (with interval of 35 mm) to register the data in 3D. The image at the right demonstrate the visualized result.

by our fusion framework and instead we use the inertial plane. The scene is covered by a network of cameras and inertial sensors. The cameras are AVT Prosilica GC650C,³ synchro-

nized by hardware. Each camera is rigidly coupled with an IS from Xsens MTx.⁴ As shown earlier in our geometric derivations, the primary use of the inertial sensors is to have

³Allied Vision Tech Prosilica GC650C GigE Camera 659 x 493, <http://1stvision.com/cameras/AVT/Prosilica-GC650-GC650C.html>

⁴Xsens Motion Technologies, <https://www.xsens.com/functions/inertial-sensor-modules/>

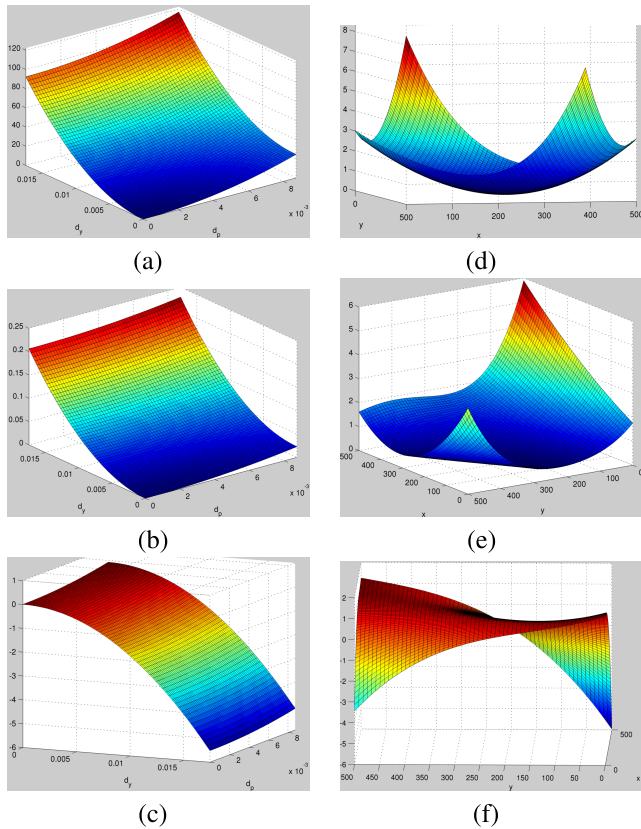


FIGURE 15. Plots for the elements of the covariance matrix of a virtual camera's image plane. (a), (b) and (c): Depict the covariance matrix's elements of for an exemplary pixel $[800 \ 200 \ 1]^T$. They correspond to a case where the homography matrix is obtained from the attached IS with the angles $roll = 0$, $pitch = \pi/4$ and $yaw = 0$. The standard deviation for the roll angle (δ_r) is assumed zero. The elements of the covariance matrix for the mentioned pixel ($\Sigma_{v_x,xx}$, $\Sigma_{v_x,yy}$ and $\Sigma_{v_x,xy}$) are plotted with respect to the variations to the value of standard deviations of the other two angles (δ_p and δ_y). (e),(f) and (g): The covariance matrix's elements for the different pixels of the virtual image plane. They correspond to a case where the homography matrix is obtained from the IS with the angles $roll = \pi/4$, $pitch = 0$ and $yaw = \pi/8$. The dimension of the image plane is assumed as 500×500 pixels. The covariance matrix of IS observation is considered as $\Sigma_s = diag\{0.25, 0.25, 1.0\}$ in degrees.

3D orientation with respect to earth, and obtain virtual camera that define virtual horizontal planes. Intrinsic parameters of the cameras are estimated using a camera calibration toolbox.⁵ In order to obtain extrinsic parameters among IS and camera (${}^I S R_c$) we used a toolbox from Lobo and Dias [59].

In all the experiments, first a background subtraction (based on color-histogram) is performed and then the binary images are fed to the data registration algorithm. A set of inertial-based euclidean planes are used to register 3D data. Figure 12 shows a set of experiments performed on some objects in the scene. In Figure 12-(a) and Figure 12-(b), a semi rectangular blue box and a cylindrical green object are reconstructed, respectively. A chair which is partially covered in red is registered (the red part) in Figure 12-(c). Figure 13 demonstrates a case where the cylindrical object is

⁵J.-Y. Bouguet's camera calibration toolbox for MATLAB®: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

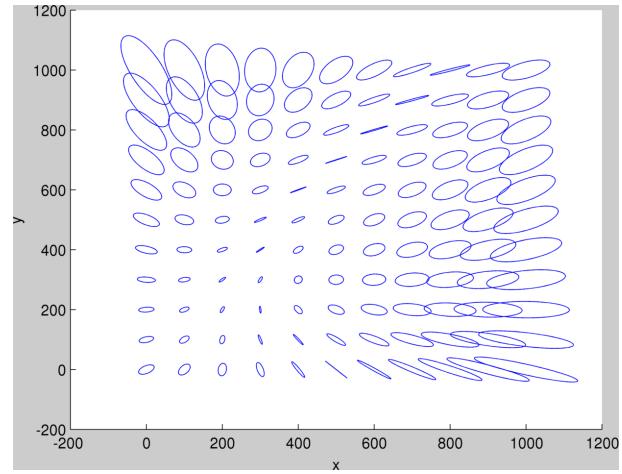


FIGURE 16. The covariance matrices, $\Sigma_{v_x,x}$, for different pixels of the virtual camera's image plane (related to Figure 15-(d),(e) and (f)) are demonstrated by ellipses, where they are scaled 1000 times for clarity.

in the hand of a person. Reconstructions of human actions are carried out and are shown in Figure 14 for different scenarios. Figure 14-(a) and Figure 14-(b) show the result for a scene including a person and a manikin. The person seated on a chair is reconstructed and shown in Figure 14-(c). Note that in these scenarios our proposed framework was able to capture all the dynamic objects faithfully.

B. UNCERTAINTY ANALYSIS

We next carryout experiments to demonstrate the uncertainty values in different scenarios. In these experiments we simulate a set of IS-camera couples where the cameras have the following calibration matrix,

$$K = \begin{bmatrix} 150 & 0 & 250 \\ 0 & 150 & 250 \\ 0 & 0 & 1 \end{bmatrix}.$$

1) ANALYZING UNCERTAINTY IN VIRTUAL CAMERA'S IMAGE PLANE

Experiments are carried out for an exemplary IS-camera couple in order to analyze the uncertainty of the imaged points on the virtual image plane, when the homography transformation is obtained using the inertial sensor. Figure 15-(a),(b) and (c) indicate the variation for the elements of the covariance matrix ($\Sigma_{v_x,xx}$, $\Sigma_{v_x,yy}$ and $\Sigma_{v_x,xy}$) for an exemplary pixel $[800 \ 200 \ 1]^T$ of the image plane of virtual camera. They are for a case where the homography matrix is obtained from the attached IS with the angles $roll = 0$, $pitch = \pi/4$ and $yaw = 0$. The standard deviation for the roll angle, is assumed zero ($\delta_r = 0$) and the elements of the covariance matrix for the mentioned pixel is plotted with respect to the variations to the values of standard deviations of the other two angles (δ_p and δ_y). As expected, the uncertainties of the point get increased with increasing uncertainties of the IS observation.

In another experiment, Figure 15-(d),(e) and (f) represent the uncertainty for all the points of the virtual image plane,

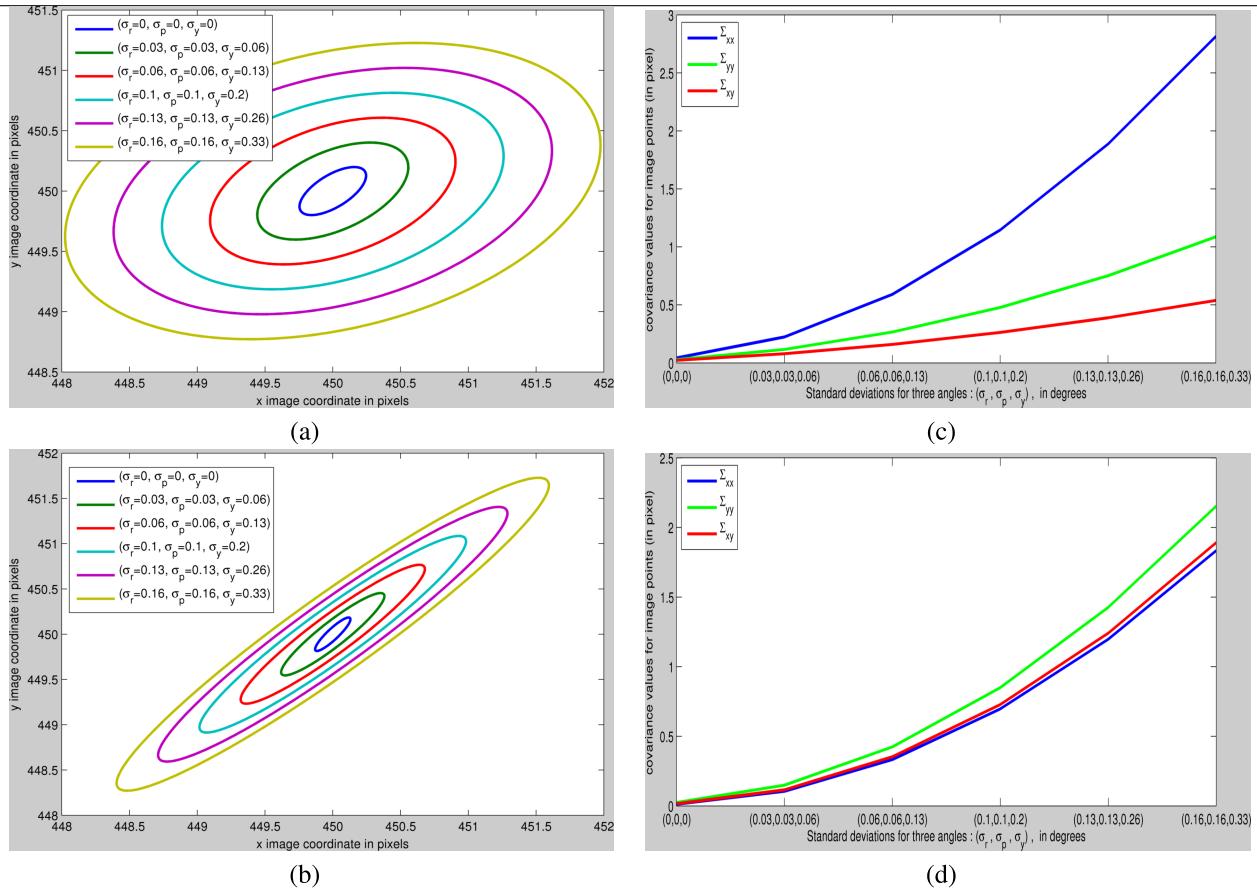


FIGURE 17. Uncertainties for an exemplary pixel, $x = [450 \ 450 \ 1]^T$, of a virtual image plane. The first and second rows correspond to two different values of IS observation: $s = [0 \ \pi/4 \ \pi/8]^T$ and $s = [\pi/2 \ -\pi/2 \ 0]^T$, respectively. For each of these two cases, the pixel uncertainties related to different noise level of IS are shown.

where its dimension is 500×500 pixels. In this case, the IS's observation vector (see Eqn. (26)) is given as,

$$s = [\pi/4 \ 0 \ \pi/8]^T.$$

Based on [76] we assume the covariance matrix of IS observation as $\Sigma_s = \text{diag}\{0.25, 0.25, 1.0\}$ in degrees. One can see that the uncertainties for the pixels close to the center of the image (principal point) is minimum and they increase in the other image coordinates with respect to the configuration. The same image uncertainties are shown by ellipses in Figure 16 where the values are scaled 1000 times for clarity.

The changes of uncertainties on the image plane of two other IS-camera couples have been analyzed with respective to increasing the uncertainties in their coupled inertial sensor observations. Figure 17 shows the progress of the uncertainties on an exemplary point, $x = [450 \ 450 \ 1]^T$, on the virtual camera's image plane. The IS's observation given as,

$$s = [0 \ \pi/4 \ \pi/8]^T.$$

Six incremental covariance matrices for IS observation are considered, starting from $\Sigma_s = \text{diag}(0, 0, 0)$ and finishing by $\Sigma_s = \text{diag}(0.16^2, 0.16^2, 0.33^2)$. The uncertainty matrices for the particular point x of the virtual plane are shown by some

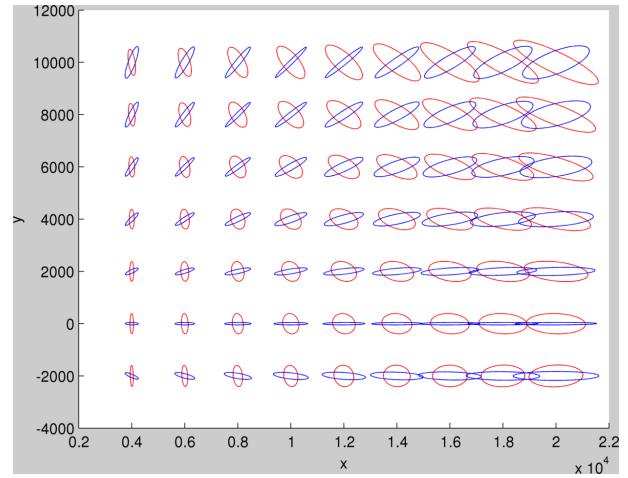


FIGURE 18. The covariance matrices, Σ_x , for different registered points on the Euclidean inertial plane, demonstrated by ellipses. The blue and red ellipses stand for points registered by the first and second camera, respectively. For the sake of clarity the covariance values are scaled 500 and 600 times, respectively for the first and second cameras.

ellipses in Fig. 17-(a),(b). Also the values for each element of the same covariance matrices are plotted in Fig. 17-(c),(d). Similar experiment is carried out for the same IS-camera

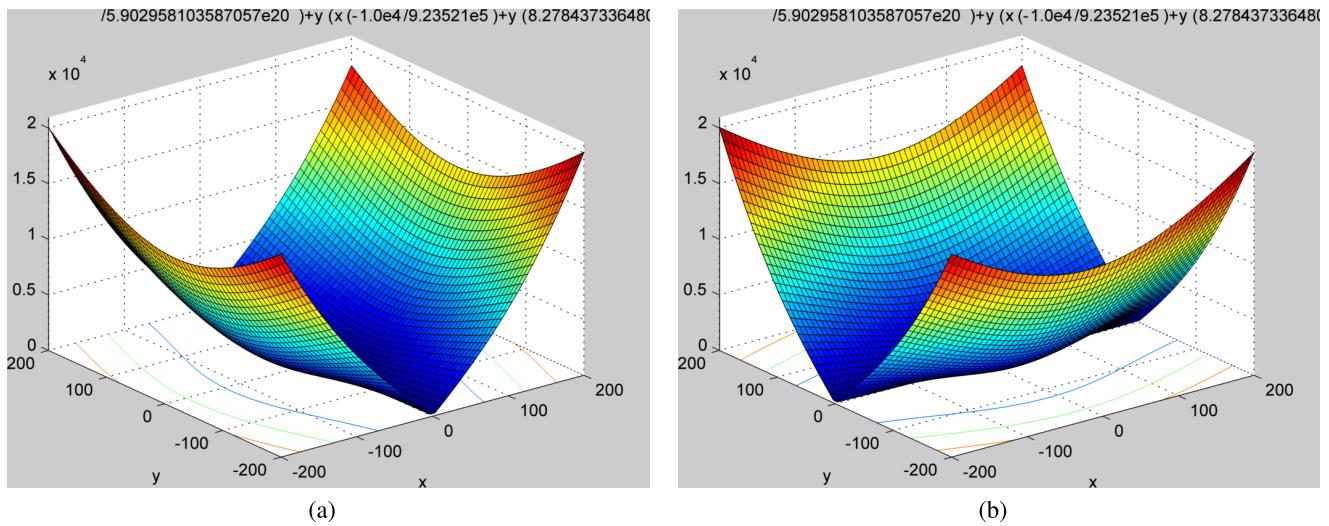


FIGURE 19. Uncertainties in transforming points between two inertial planes with an Euclidean distance h . The camera's translation vector is assumed $[1000 \ 1000 \ 3000]^T$. The horizontal axes express x and y position on the inertial plane and the vertical axis expresses h . The left and right figures depict the two elements of the covariance matrix $\Sigma_{\pi'x}^T . XX$ and $\Sigma_{\pi'x}^T . YY$, respectively.

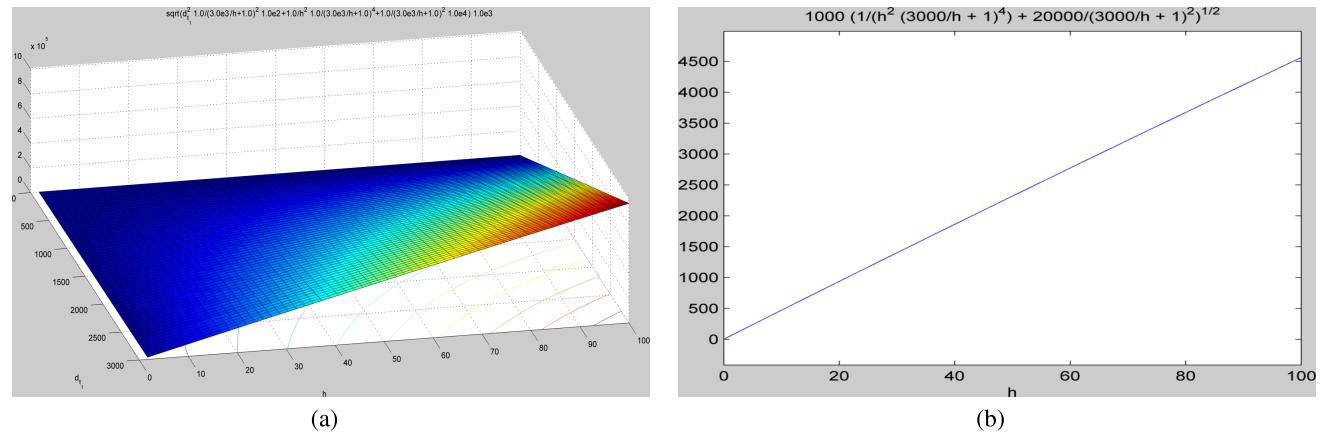


FIGURE 20. (a) Variation of one of the elements of the covariance matrix, $\Sigma_{\pi'x}^T . XX$, is plotted with respect to variation in the Euclidean distance to the reference plane h (inverse of α) and the uncertainty in the X element of the estimated t . (b) The progress of uncertainty with respect to increasing the Euclidean among the planes. The horizontal axis indicates the Euclidean distance h and the vertical one shows one of the elements of the covariance matrix ($\Sigma_{\pi'x}^T . XX$).

couple, where the IS observation is given as,

$$\mathbf{s} = [\pi/2 \ -\pi/2 \ 0]^T. \quad (39)$$

As can be seen, the uncertainties of the mapped point increase with increasing the uncertainties in IS observation.

2) ANALYZING UNCERTAINTY IN EUCLIDEAN INERTIAL PLANE

To analyze the uncertainties of the points registered on an Euclidean inertial plane two couples of IS-camera are used. The translations for first camera and second camera respectively are,

$$\mathbf{t}_1 = [100 \ -100 \ 500]^T,$$

and

$$\mathbf{t}_2 = [-1500 \ 7000 \ 6500]^T,$$

given in mm . The uncertainty covariances for these two translation vectors are assumed as $\Sigma_{\mathbf{t}_1} = diag\{200^2, 300^2, 500^2\}$ and $\Sigma_{\mathbf{t}_2} = diag\{100^2, 500^2, 300^2\}$ again in mm . The observation vectors for the first and second inertial sensors are,

$$\mathbf{s}_1 = [\pi/4 \ 0 \ \pi/8]^T, \quad (40)$$

and

$$\mathbf{s}_2 = [\pi/4 \ \pi/2 \ 0]^T, \quad (41)$$

where the covariance matrices for both IS_1 and IS_2 are assumed as $\Sigma_{\mathbf{s}_1} = \Sigma_{\mathbf{s}_2} = diag\{0.25, 0.25, 1.0\}$ in degrees [76].

For this experiment, the uncertainty values for the registered points on the Euclidean plane are calculated by using

Eqn. (32). The obtained uncertainties are presented by covariance ellipses in Figure 18. The blue ellipses are for points which are mapped through to the first IS-camera couple and the red ones are for the point mapped through to the second couple. The covariance ellipses for the first and second cameras are respectively scaled 500 and 600 times for clarity. The Euclidean plane is also scaled to 10^4 times and is shown in mm.

The uncertainties in transforming points among the reference inertial plane, π_{ref} , and another inertial plane, π' , with an interval distance h was already expressed by Eqn. (35). Figure 19 shows the visualization of an exemplary case. The uncertainty covariance for the translation vector is assumed as $\Sigma_t = \text{diag}\{10^2, 10^2, 10^2\}$ in mm. The translation for the camera is,

$$\mathbf{t} = [1000 \quad 1000 \quad 3000]^T,$$

given in mm. The two elements of the covariance matrix (the uncertainties), $\Sigma'_{\pi'X}XX$ and $\Sigma'_{\pi'X}YY$, are respectively plotted in Figure 19-(a) and Figure 19-(b). Figure 20-(a) shows another experiment in which the variation of one of the elements of the covariance matrix, $\Sigma'_{\pi'X}XX$, is plotted with respect to variation in the Euclidean distance to the reference plane h (inverse of α) and the uncertainty in the X element of the estimated \mathbf{t} . As can be seen, the value of $\Sigma'_{\pi'X}XX$ sharply increases when both the h and the uncertainty in $t.X$ increase. The progress of uncertainty with respect to just increasing the Euclidean among the planes is plotted in Figure 20-(b).

VI. DISCUSSION AND CONCLUSIONS

This work presented geometric relations among the virtual planes in a multisensor fusion registration framework, where the planes are obtained from inertial sensor. The framework includes a network of coupled IS-cameras. For each IS-camera couple a downward-looking virtual camera was defined with each couple can be thought as a smart or hybrid sensor. A set of Euclidean inertial plane in the scene was defined to register 3D data without any planar ground assumption. The geometric relations among Euclidean inertial planes of the scene and projective planes of images were explored in detail, based on homography transformation. A geometric 3D data registration algorithm was proposed using the obtained transformations. In the context of data registration it is important to have the uncertainty of each registered geometric entity. In the introduced framework we used homography transformations to map and register the data. For each IS-camera couple, the 3D orientation acquired from the IS and the translation vector, obtained by an estimation method is directly used to compute the homography transformations. Several experiments have been carried out to illustrate our proposed approach. They demonstrate the effectiveness and feasibility of the proposed 3D reconstruction where persons and objects were fully reconstructed.

Following are the salient points of the geometric framework presented here:

- We note that for the background subtraction method used here as a preliminary step, we utilized color histogram based segmentation from OpenCV.⁶ Any reasonably performing background subtraction method from color video data is sufficient.
- In the experimental setup shown here although there is a planar ground visible, it is not used in the estimation process. Instead the IS is used to define virtual ground, this makes our approach applicable to general scenarios.
- Thanks to the fusion of IS with camera data, the rotations among all virtual cameras are relaxed. Therefore in the aspect of having extrinsic parameters of the camera network, what remains is to have only the translation vector among cameras. For an outdoor scenario the translation part can be obtained using a GPS coupled to each sensor.
- The use of GPS can improve the accuracy of IS in its orientation angle until 0.01° [76]. In case unavailability of coupled GPS or in indoor scenario, we can estimate the translation vectors among the virtual cameras using [79].
- The IS-camera couples used in the presented scenarios are all static, however the proposed algorithm has potential to be used in cases where some of the IS-camera couples have movement capability e.g. a mobile agent equipped with IS, camera and GPS.
- Due to imperfection of sensors observation and estimation algorithms, the obtained homographies contain some uncertainties. As a consequence these uncertainties (of homographies) will get propagated to the points which are mapped through. To be aware of the degree of uncertainty for each data which is registered by an IS-camera couple is very important specially where a network of sensors is used. We modeled the uncertainties of the mapped points in the framework by using statistical geometry analysis.
- Knowing the uncertainties for a homography transformation is of importance for data registration. This issue has been already investigated by some researchers [50], [52], [54], [55], [89]. But to the best of our knowledge, all of the studies are for cases when the homographies are estimated from point correspondences, however the homography transformations in the presented framework are directly obtained using the IS data.

Following are the drawbacks of our proposed 3D reconstruction algorithm:

- The proposed reconstruction method requires no feature extraction since it uses only the silhouette of the object or person. This can be useful in cases where there is difficulty to extract features and match them across cameras. However having occlusion in one of the

⁶OpenCV available at: <http://opencv.willowgarage.com/>

cameras can result in a failure. This we believe can be avoided if an appropriate probabilistic fusion method is utilized.

- Our method requires that the cameras must have non-trivial intersection among their field of views. We do not cover the study on the placement of the cameras. To have a better coverage of objects in the scene a genetic algorithm based optimal solution can be used [90].
- In the presented work, the fusion of the observations obtained from different nodes (IS-camera couple) was performed by using an intersection (see Figure 11). However having the uncertainties for the points, which are projected onto the inertial Euclidean planes through different nodes, grants the possibility of using a more sophisticated method, such as probabilistic fusion, with the ability of taking into the account the covariance matrix (uncertainty) of each point.

These last two drawbacks can be eliminated if we utilize sophisticated methods to find the intersection among the views e.g. using a probabilistic method instead of the proposed deterministic one. The quality of reconstruction using the proposed method mainly depends on the three parameters: (1) number of cameras, (2) camera configurations (e.g. positions) and (3) accuracy of the applied background subtraction method. The number of cameras needed depends on the application and available budget. In this work, we did not go through the details of background subtraction methods since it does not affect our geometric analysis framework. However the second parameter, the camera configuration, specifically their positions in the scene was investigated and a geometric method to find an optimal configuration was done. In this work we characterized the uncertainty associated with the analytical homography transformations by taking into account the noise of IS. As a result we provide uncertainty models for each node within the sensor network; however we did not explore a proper fusion scheme which could fully take advantage of such obtained uncertainty models. In order to fuse the individual uncertainties associated with multiple sensor measurements in the final Euclidean registration plane we intend to explore probabilistic solutions and covariance consistency methods. Despite camera is an essential sensors in data registration and scene perception with ability to sense colors, it has some weaknesses, for example, it is highly dependent on light conditions, shadows and homogeneous textures. A precise active sensor like laser range finder (LRF) is able to provide 3D information of a scene without dependency on texture, but they do not yield color information. Integration of range data within the proposed inertial-based data registration framework, along with probabilistic fusion [91] is thus a promising avenue to explore. Nevertheless, the primary step before being able to register range and image measurements is to have these two sensors (LRF and camera) precisely calibrated. The reader is referred to our LRF-camera calibration work in [92].

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

ACKNOWLEDGMENTS

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of AFRL, NRL or the U.S. Government.

REFERENCES

- [1] M. Vergauwen and L. Van Gool, "Web-based 3D reconstruction service," *Mach. Vis. Appl.*, vol. 17, pp. 411–426, Oct. 2006.
- [2] S. Agarwal *et al.*, "Building rome in a day," *Commun. ACM*, vol. 54, no. 10, pp. 105–112, Oct. 2011.
- [3] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: Exploring photo collections in 3D," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 835–846, 2006.
- [4] T. Feldmann, I. Mihailidis, S. Schulz, D. Paulus, and A. Wörner, "Online full body human motion tracking based on dense volumetric 3D reconstructions from multi camera setups," in *Advances in Artificial Intelligence*, vol. 6359, R. Dillmann, J. Beyerer, U. D. Hanebeck, and T. Schultz, Eds. Springer, 2010, pp. 74–81.
- [5] M. Brice, G. Erwan, and B. Saïda, "Human model and pose reconstruction from," in *Proc. Int. Conf. Mach. Intell.*, Nov. 2005, pp. 788–795.
- [6] X. Luo, B. Berendsen, R. T. Tan, and R. C. Veltkamp, "Human pose estimation for multiple persons based on volume reconstruction," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Washington, DC, USA, Oct. 2010, pp. 3591–3594.
- [7] R. Kehl, "Markerless motion capture of complex human movements from multiple views," Ph.D. dissertation, Dept. Inf. Technol. Elect. Eng., Swiss Federal Inst. Technol. Zurich, Zürich, Switzerland, 2005.
- [8] V. Brandou *et al.*, "3D reconstruction of natural underwater scenes using the stereovision system IRIS," in *Proc. OCEANS—Europe*, 2007, pp. 674–679.
- [9] K. Takahashi, Y. Nagasawa, and M. Hashimoto, "Remarks on 3D human body's feature extraction from voxel reconstruction of human body posture," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2007, pp. 121–126.
- [10] M. C. V. Uriol, "Video-based avatar reconstruction and motion capture," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. California, Irvine, CA, USA, 2005.
- [11] T. Yang, Y. Zhang, M. Li, D. Shao, and X. Zhang, "A multi-camera network system for markerless 3D human body voxel reconstruction," in *Proc. Int. Conf. Image Graph.*, Sep. 2009, pp. 706–711.
- [12] F. Remondino and S. El-Hakim, "Image-based 3D modelling: A review," *Photogram. Rec.*, vol. 21, no. 115, pp. 269–291, 2006.
- [13] G. Kordelas, J. D. P.-M. Agapito, J. M. V. Hernandez, and P. Daras, "State-of-the-art algorithms for complete 3D model reconstruction," Engage Summer School, Zermatt, Switzerland, Sep. 2010.
- [14] L. Fredriksson, "Evaluation of 3D reconstructing based on visual hull algorithms," Bachelor Thesis, Dept. Comput. Sci. Faculty Eng. Sustain. Develop., Univ. Gävle, Sweden, Jun. 2011.
- [15] Y. Lu, J. Z. Zhang, Q. M. J. Wu, and Z.-N. Li, "A survey of motion-parallax-based 3-D reconstruction algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 34, no. 4, pp. 532–548, Oct. 2004.
- [16] T. S. Huang, "Modeling, analysis, and visualization of nonrigid object motion," in *Proc. Int. Conf. Pattern Recognit.*, vol. 1. Jun. 1990, pp. 361–364.
- [17] C. Leung, "Efficient methods for 3D reconstruction from multiple images," Ph.D. dissertation, Dept. Elect. Eng., Univ. Queensland, Queenstown, New Zealand, 2005.
- [18] T. Moons, L. V. Gool, and M. Vergauwen, "3D reconstruction from multiple images part 1: Principles," *Found. Trends Comput. Graph. Vis.*, vol. 4, no. 4, pp. 287–404, 2008.
- [19] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1. Jun. 2006, pp. 519–528.

- [20] J. Dornauer, G. Kotsis, C. Bernthalier, and M. Naderhirn, "A comparison of different computer vision methods for real time 3D reconstruction for the use in mobile robots," in *Proc. Int. Conf. Adv. Mobile Comput. Multimedia*, New York, NY, USA, Nov. 2008, pp. 136–141.
- [21] A. Koutsoudis, B. Vidmar, G. Ioannakis, F. Arnaoutoglou, G. Pavlidis, and C. Chamzas, "Multi-image 3D reconstruction data evaluation," *J. Cult. Heritage*, vol. 15, no. 1, pp. 73–79, 2014.
- [22] H. Fathi, F. Dai, and M. Lourakis, "Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges," *Adv. Eng. Inform.*, vol. 29, no. 2, pp. 149–161, 2015.
- [23] L. Gimenez, J.-L. Hippolyte, S. Robert, F. Suard, and K. Zreik, "Review: Reconstruction of 3D building information models from 2D scanned plans," *J. Building Eng.*, vol. 2, pp. 24–35, Jun. 2015.
- [24] S. M. Khan, P. Yan, and M. Shah, "A homographic framework for the fusion of multi-view silhouettes," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [25] S. M. Khan, "Multi-view approaches to tracking, 3D reconstruction and object class detection," Ph.D. dissertation, School Elect. Eng. Comput. Sci., Univ. Central Florida, Orlando, FL, USA, 2008.
- [26] B. Michoud, E. Guillou, and S. Bouakaz, "Real-time and markerless 3D human motion capture using multiple views," in *Human Motion—Understanding, Modeling, Capture and Animation*, vol. 4814, A. Elgammal, B. Rosenhahn, and R. Klette, Eds. Springer, Oct. 2007, pp. 88–103.
- [27] Q.-B. Zhang, H.-X. Wang, and S. Wei, "A new algorithm for 3D projective reconstruction based on infinite homography," in *Proc. Int. Conf. Mach. Learn. Cybern.*, Nov. 2003, pp. 2882–2886.
- [28] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, MA, USA: Cambridge Univ. Press, 2004.
- [29] C. Rother, "Multi-view reconstruction and camera recovery using a real or virtual reference plane," Ph.D. dissertation, Royal Inst. Technol., Stockholm, Sweden, 2003.
- [30] Z. Zhang and A. R. Hanson, "3D reconstruction based on homography mapping," in *Proc. ARPA Image Understand. Workshop*, 1996, pp. 1007–1012.
- [31] T. Wada, X. Wu, S. Tokai, and T. Matsuyama, "Homography based parallel volume intersection: Toward real-time volume reconstruction using active cameras," in *Proc. IEEE Int. Workshop Comput. Archit. Mach. Perception*, Sep. 2000, pp. 331–339.
- [32] P.-L. Lai and A. Yilmaz, "Projective reconstruction of building shape from silhouette images acquired from uncalibrated cameras," in *Proc. ISPRS Congr.*, 2008, pp. 103–108.
- [33] H. Lee and A. Yilmaz, "3D reconstruction using photo consistency from uncalibrated multiple views," in *Proc. Int. Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, vol. 1, 2010, pp. 484–487.
- [34] B. Zhang and Y. F. Li, "An efficient method for dynamic calibration and 3D reconstruction using homographic transformation," *Sens. Actuators A, Phys.*, vol. 119, no. 2, pp. 349–357, 2005.
- [35] C.-Y. Tang, Y.-L. Wu, P.-C. Hu, H.-C. Lin, and W.-C. Chen, "Self-calibration for metric 3D reconstruction using homography," in *Proc. IAPR Conf. Mach. Vis. Appl.*, 2007, pp. 86–89.
- [36] M. Sormann, C. Zach, J. Bauer, K. Karner, and H. Bishof, "Watertight multi-view reconstruction based on volumetric graph-cuts," in *Proc. 15th Scandian. Conf. Image Anal.*, vol. 4522, Jun. 2007, pp. 393–402.
- [37] R. Guerchouche, O. Bernier, and T. Zaharia, "Multiresolution volumetric 3D object reconstruction for collaborative interactions," *Pattern Recognit. Image Anal.*, vol. 18, no. 4, pp. 621–637, 2008.
- [38] T. C. S. Azevedo, M. R. S. Tavares, and M. A. P. Vaz, "3D object reconstruction from uncalibrated images using an off-the-shelf camera," *Adv. Comput. Vis. Med. Image Process.*, vol. 13, pp. 117–136, Oct. 2009.
- [39] T. C. Azevedo, J. M. R. Tavares, and M. A. Vaz, "Three-dimensional reconstruction and characterization of human external shapes from two-dimensional images using volumetric methods," *Comput. Methods Biomed. Eng.*, vol. 13, no. 3, pp. 359–369, 2010.
- [40] H. Aliakbarpour, K. Palaniappan, and G. Seetharaman, "Robust camera pose refinement and rapid SFM for multiview aerial imagery—Without RANSAC," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2203–2207, Nov. 2015.
- [41] H. AliAkbarpour, K. Palaniappan, and G. Seetharaman, "Fast structure from motion for sequential and wide area motion imagery," in *Proc. Int. Conf. Comput. Vis. Workshops*, Dec. 2015, pp. 1086–1093.
- [42] M. Jethwa, "Efficient volumetric reconstruction from multiple calibrated cameras," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2004.
- [43] S. M. Khan and M. Shah, "Reconstructing non-stationary articulated objects in monocular video using silhouette information," in *Proc. Comput. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [44] N. S. Sudipta, "Silhouettes for calibration and reconstruction from multiple views," Ph.D. dissertation, Univ. North Carolina, Chapel Hill, NC, USA, 2009.
- [45] F. Remondino, "3-D reconstruction of static human body shape from image sequence," *Comput. Vis. Image Understand.*, vol. 93, no. 1, pp. 65–85, 2004.
- [46] M. Maitre, Y. Shinagawa, and M. N. Do, "Symmetric multi-view stereo reconstruction from planar camera arrays," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [47] C. Ruwwe, B. Keck, O. Rusch, U. Zolzer, and X. Loison, "Image registration by means of 3D octree correlation," in *Proc. IEEE Workshop Multimedia Signal Process.*, Oct. 2008, pp. 515–519.
- [48] Y. Díez, F. Roure, X. Lladó, and J. Salvi, "A qualitative review on 3D coarse registration methods," *ACM Comput. Surv.*, vol. 47, no. 3, pp. 45:1–45:36, Feb. 2015.
- [49] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: A review," *Comput. Methods Biomed. Eng.*, vol. 17, no. 2, pp. 73–93, 2014.
- [50] A. Criminisi, I. Reid, and A. Zisserman, "A plane measuring device," *Image Vis. Comput.*, vol. 17, no. 8, pp. 625–634, 1999.
- [51] J. Meidow, C. Beder, and W. Förstner, "Reasoning with uncertain points, straight lines, and straight line segments in 2D," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 2, pp. 125–139, 2009.
- [52] S. Negahdaripour, R. Prados, and R. Garcia, "Planar homography: Accuracy analysis and applications," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Sep. 2005, pp. 1089–1092.
- [53] B. Ochoa and S. Belongie, "Covariance propagation for guided matching," in *Proc. Statist. Methods Multi-Image Video Process.*, May 2006, pp. 1–12.
- [54] O. Chum, T. Pajdla, and P. Sturm, "The geometric error for homographies," *Comput. Vis. Image Understand.*, vol. 97, no. 1, pp. 86–102, Jan. 2005.
- [55] S. Baker, A. Datta, and T. Kanade, "Parameterizing homographies," Dept. Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-06-11, Mar. 2006.
- [56] Y.-T. Chi, J. Ho, and M.-H. Yang, "A direct method for estimating planar projective transform," in *Proc. Asian Conf. Comput. Vis.*, 2011, pp. 268–281.
- [57] J. Dias, J. Lobo, and L. A. Almeida, "Cooperation between visual and inertial information for 3D vision," in *Proc. Medit. Conf. Control Autom.*, 2002, pp. 1–10.
- [58] J. Lobo, L. Almeida, J. Alves, and J. Dias, "Registration and segmentation for 3D map building—A solution based on stereo vision and inertial sensors," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 1, Sep. 2003, pp. 139–144.
- [59] J. Lobo and J. Dias, "Relative pose calibration between visual and inertial sensors," *Int. J. Robot. Res.*, vol. 26, no. 6, pp. 561–575, 2007.
- [60] J. Lobo, C. Queiroz, and J. Dias, "World feature detection and mapping using stereovision and inertial sensors," *Robot. Auto. Syst.*, vol. 44, no. 1, pp. 69–81, 2003.
- [61] L. G. B. Mirisola, J. Dias, and A. T. de Almeida, "Trajectory recovery and 3D mapping from rotation-compensated imagery for an airship," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, Nov. 2007, pp. 1908–1913.
- [62] G. Bleser, C. Wohlleber, M. Becker, and D. Stricker, "Fast and stable tracking for AR fusing video and inertial sensor data," in *Proc. Int. Conf. Central Eur. Comput. Graph., Vis. Comput.*, 2006, pp. 109–115.
- [63] G. Bleser and D. Stricker, "Advanced tracking through efficient image processing and visual-inertial sensor fusion," in *Proc. IEEE Virtual Reality Conf.*, Mar. 2008, pp. 137–144.
- [64] F. Ababsa, "Toward a real-time 3D reconstruction system for urban scenes using georeferenced and oriented images," in *Proc. Int. Conf. Comput. Electr. Eng.*, vol. 1, Dec. 2009, pp. 75–79.
- [65] I. M. Zendjebil, F. Ababsa, J. Didier, and M. Mallem, "A GPS-IMU-camera modelization and calibration for 3D localization dedicated to outdoor mobile applications," in *Proc. Int. Conf. Control Autom. Syst.*, Oct. 2010, pp. 1580–1585.
- [66] A. Hogue, A. German, and M. Jenkin, "Underwater environment reconstruction using stereo and inertial data," in *Proc. IEEE Int. Conf. Syst., Man, Cybren.*, Oct. 2007, pp. 2372–2377.
- [67] H.-W. Hsieh, H.-H. Yu, C.-C. Wu, and J.-S. Hu, "Concurrent multiple cameras calibration and robot localization from visual and 3D inertial measurements," in *Proc. SICE Annu. Conf.*, Aug. 2010, pp. 1914–1919.

- [68] R. R. Clark, M. H. Lin, and C. J. Taylor, "3D environment capture from monocular video and inertial data," *Proc. SPIE*, vol. 6056, p. 60560G, Feb. 2006.
- [69] Y. Gat, I. Kozintsev, and O. Nestares, "Fusing image data with location and orientation sensor data streams for consumer video applications," in *Proc. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2010, pp. 1–8.
- [70] E. Besdok, "3D vision by using calibration pattern with inertial sensor and RBF neural networks," *J. Sensors*, vol. 9, no. 6, pp. 4572–4585, 2009.
- [71] D. I. B. Randeniya, M. Gunaratne, S. Sarkar, and A. Nazef, "Calibration of inertial and vision systems as a prelude to multi-sensor fusion," *Transp. Res. C, Emerg. Technol.*, vol. 16, no. 2, pp. 255–274, 2008.
- [72] F. M. Mirzaei and S. I. Roumeliotis, "A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation," *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 1143–1156, Oct. 2008.
- [73] T. Okatani and K. Deguchi, "Robust estimation of camera translation between two images using a camera with a 3D orientation sensor," in *Proc. Int. Conf. Pattern Recognit.*, vol. 1, Aug. 2002, pp. 275–278.
- [74] M. Labrie and P. Hebert, "Efficient camera motion and 3D recovery using an inertial sensor," in *Proc. Can. Conf. Comput. Robot Vis.*, May 2007, pp. 55–62.
- [75] M. A. Brodie, A. Walmsley, and W. Page, "The static accuracy and calibration of inertial measurement units for 3D orientation," *Comput. Methods Biomed. Eng.*, vol. 11, pp. 641–648, Oct. 2008.
- [76] M. Kalantari, A. Hashemi, F. Jung, and J.-P. Guedon, "A new solution to the relative orientation problem using only 3 points and the vertical direction," *J. Math. Imag. Vis.*, vol. 39, no. 3, pp. 259–268, 2011.
- [77] P.-L. Lai and A. Yilmaz, "Efficient object shape recovery via slicing planes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–6.
- [78] H. Aliakbarpour and J. Dias, "Volumetric 3D reconstruction without planar ground assumption," in *Proc. 5th ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Aug. 2011, pp. 1–2.
- [79] H. Aliakbarpour and J. Dias, "Inertial-visual fusion for camera network calibration," in *Proc. IEEE Int. Conf. Indus. Informat.*, Jul. 2011, pp. 422–427.
- [80] H. Aliakbarpour, L. Almeida, P. Menezes, and J. Dias, "Multi-sensor 3D volumetric reconstruction using CUDA," *J. 3D Res.*, vol. 2, no. 4, pp. 1–14, Dec. 2011.
- [81] H. Aliakbarpour and J. Dias, "Three-dimensional reconstruction based on multiple virtual planes by using fusion-based camera network," *Comput. Vis. IET*, vol. 6, no. 4, pp. 355–369, Jul. 2012.
- [82] H. Aliakbarpour, K. Palaniappan, and J. Dias, "Geometric exploration of virtual planes in a fusion-based 3D data registration framework," *Proc. SPIE*, vol. 8747, p. 87470C, Jun. 2013.
- [83] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3D Vision*. New York, NY, USA: Springer, 2004.
- [84] H. Aliakbarpour and J. Dias, "IMU-aided 3D reconstruction based on multiple virtual planes," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl.*, Sydney, NSW, Australia, Dec. 2010, pp. 474–479.
- [85] W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, Jun. 1989.
- [86] A. Bjorck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA, USA: SIAM, 1996.
- [87] O. Faugeras, *Three-Dimensional Computer Vision*. Cambridge, MA, USA: MIT Press, 1993.
- [88] S. Heuel, *Uncertain Projective Geometry: Statistical Reasoning for Polyhedral Object Reconstruction*. Heidelberg, Germany: Springer, 2004.
- [89] E. Malis and M. Vargas, "Deeper understanding of the homography decomposition for vision-based control," INRIA, France, Tech. Rep. RR-6303, 2007.
- [90] H. Aliakbarpour, V. B. S. Prasath, and J. Dias, "On optimal multi-sensor network configuration for 3D registration," *J. Sens. Actuator Netw.*, vol. 4, no. 4, pp. 293–314, 2015.
- [91] B. Fassnutt-Mombot and J.-B. Choquel, "A new probabilistic and entropy fusion approach for management of information sources," *Inform. Fusion*, vol. 5, no. 1, pp. 35–47, 2004.
- [92] H. Aliakbarpour, P. Nunez, J. Prado, K. Khoshhal, and J. Dias, "An efficient algorithm for extrinsic calibration between a 3D laser range finder and a stereo camera for surveillance," in *Proc. Int. Conf. Adv. Robot.*, 2009.



HADI ALIAKBARPOUR (SM'07) is currently an assistant research professor in the Computer Science Department at the University of Missouri. He received his Ph.D. in Automation and Robotics, from University of Coimbra (Department of Electrical and Computer Engineering), Portugal, under supervision of Prof. Jorge Dias. He had been a researcher at the Institute of Systems and Robotic (ISR- Coimbra, Portugal), during September 2007 till March 2013. He was awarded a PhD fellowship from the European project BACS (Bayesian Approach to Cognitive Systems, within the 6th Framework Program of the European Commission) for 2007–2008. His was awarded a PhD fellowship from Portuguese Foundation for Science and Technology (FCT), for 2008–2012. He was actively involved in other European projects including PROMETHEUS (a European project on prediction and interpretation of human behavior based on probabilistic structures and heterogeneous sensors- within the 7th Framework Program of the European Commission) and MORFEU (the Multi-Objective Robot Fleet for improved communication, funded by Portuguese FCT). Hadi Aliakbarpour participated in several project/academic meetings and gave talks in different universities and research companies including, Technical University of Munich (Germany), University of Patras (Greece), Probayes Co. (France), University of Extremadura (Spain) and University of Malaga (Spain). He was a co-organizer and committee member of the Intelligent Space Monitoring (ISM) special session at the 2011 International Workshop in Image Analysis for Multimedia Interactive Services at Delft Technical University in The Netherlands. He has authored more than 45 peer-reviewed articles and has a U.S. Patent (pending) on the Structure-from-Motion. He practices in different research areas including vision-based robot control (visual-servoing), computer vision, remote sensing, video analytics, image-based multi-view 3D reconstruction, structure from motion, bundle adjustment, sensor fusion, sensor calibration, aerial image registration and computational human behavior analysis.



V. B. SURYA PRASATH received the Ph.D. degree in mathematics, under the supervision of Prof. A. Singh, from IIT Madras, Chennai, India, in 2009. He was a Post-Doctoral Fellow with the Department of Mathematics, University of Coimbra, Portugal. Since 2012, he has been with the Computational Imaging and VisAnalysis Laboratory, where he is involved in various image processing and computer vision problems. He had summer internships/visits with Kitware, The Fields Institute, Canada, and IPAM, UCLA. He is currently an Assistant Research Professor with the Computer Science Department, University of Missouri. His main research interests include regularization methods for inverse and ill-posed problems, optimization, variational, PDE based image processing, and computer vision with applications in remote sensing, bio-medical imaging, and biometrics domains.



KANNAPPAN PALANIAPPAN (S'84–M'90–SM'99) received the B.S. and M.S. degrees in systems design engineering from the University of Waterloo, Canada, and the Ph.D. degree from the University of Illinois at Urbana–Champaign. He is currently with the Department of Computer Science, University of Missouri, Columbia, where he directs the Computational Imaging and Visualization Analysis Laboratory and helped to establish the NASA Center of Excellence in Remote Sensing.

His research is at the synergistic intersection of image and video processing, computer vision, high performance computing, and machine learning to understand, quantify, and model physical processes with applications to biomedical, space, and defense imaging. His recent multidisciplinary contributions range across orders of scale from subcellular microscopy at the molecular level to aerial and satellite remote sensing imaging at the macro level. At NASA Goddard Space Flight Center and UMIACS, University of Maryland, he co-invented the Interactive Image SpreadSheet for visualizing large multispectral imagery and co-founded the Visualization and Analysis Laboratory that has produced a number of digital earth visualizations widely used by museums, media, and map search engines (i.e., BlueMarble). He is a member of the Editorial Board of the IEEE TRANSACTIONS ON IMAGE PROCESSING. He received the William T. Kemper Fellowship for Teaching Excellence at the University of Missouri, the ASEE Air Force and Boeing Welliver Summer Faculty Fellowships, the NASA Public Service Medal for pioneering contributions in data science for understanding petabyte-sized archives, and the U.S. National Academies Jefferson Science Fellowship.



GUNA SEETHARAMAN (S'82–M'88–SM'10–F'15) received the Ph.D. degree in electrical and computer engineering from the University of Miami in 1988. He was a Principal Engineer of Computer Vision and Video Exploitation with the Information Directorate, Air Force Research Laboratory, Rome, NY. He was an Associate Professor of Computer Science and Engineering with the Air Force Institute of Technology (AFIT) from 2003 to 2008 and the University of Louisiana, Lafayette

from 1988 to 2003. He co-founded the Team CajunBot that participated in the DARPA Grand Challenges in 2005 and 2007, and led the LIDAR data processing and obstacle detection efforts. He was a member of the Core AFIT Team that demonstrated and transitioned a wide area persistent imaging and surveillance system known as Angel Fire. He is currently a Senior Scientist with the Senior Executive Service For Advanced Computing Concepts and the Chief Scientist of Computational Sciences with the U.S. Naval Research Laboratory, Systems Directorate, Information Technology Division, DC. He has authored over 150 peer-reviewed articles in computer vision, low-altitude aerial imagery, parallel computing, VLSI signal processing, 3-D displays, nanotechnology, micro-optics, and 3-D video analysis. His research has been focused on high performance computing for video exploitation, such as computer vision, machine learning, content-based image retrieval, persistent surveillance, and computational science and engineering. He was a Guest Editor of the IEEE Computer Special Issue in 2006 on Unmanned Intelligent Autonomous Vehicles and the Special Issue of the EURASIP Journal on Embedded Computing on Intelligent Vehicles. He was the General Chair of the IEEE Workshop on Computer Architecture for Machine Perception in 2003, and the Co-Chair of the Technical Program Committee of the IEEE AIPR in 2014. He served as the Section Chair of the IEEE Mohawk Valley Section, NY, in 2013 and 2014. He serves as an Associate Editor of the Journal ACM Computing Surveys.



JORGE DIAS (M'96–SM'10) has a Ph.D. on Electrical Engineering by the University of Coimbra, Portugal, specialization in Control and Instrumentation. Jorge Dias have been professor at the Department of Electrical Engineering and Computers (www.deec.uc.pt) and researcher from the Institute of Systems and Robotics (ISR) (www.isr.uc.pt) from the University of Coimbra (UC) (www.uc.pt). Jorge Dias research is in the area of Computer Vision and Robotics and has contributions on the field since 1984. He has several publications in international journals, books, and conferences. Jorge Dias has been principal investigator from several research projects. Jorge Dias coordinates the research group for Artificial Perception for Intelligent Systems and Robotics (<http://ap.isr.uc.pt>) of Institute of Systems and Robotics from University of Coimbra and the Laboratory of Systems and Automation (<http://las.ipn.pt>) from the Instituto Pedro Nunes (IPN) (www.ipn.pt) - a technology transfer institute from the University of Coimbra, Portugal. Jorge Dias published several articles in the area of Computer Vision and Robotics that include more than 80 publications in international journals, 1 published book, 15 books chapters, and 279 articles in international conferences with referee. The research activities of Jorge Dias have been concentrated in the Artificial Perception Laboratory from Instituto of Systems and Robotics. Jorge Dias's coordination leveraged the Laboratory to be involved in over a dozen international cooperation projects, CyberMove - Cybernetic Transportation Systems for the Cities of Tomorrow (EC-RTD Project, EVK4 – 2001), VISOR - Visual Perception System for a Social Robot (EURON-European Robotics Research Network), IRPS - Intelligent Robotic Porter System (EU-IRPS FP6-IST-45048), BACS - Bayesian Approach to Cognitive Systems (FP6-IST-027140), PROMETHEUS - Prediction and interpretation of human behaviour based on probabilistic structures and heterogeneous sensors (FP7 - 214901) and HANDLE - Developmental pathway towards autonomy and dexterity in robot in-hand manipulation (FP7-2008-231640), Social Robot (FP7 Marie Curie - 285870) and in over 25 national projects: CHOPIN - Cooperation between Human and rObotic teams in catastrophes Incidents, TICE Mobilidade, TICE Healthy, DIVA – Instrumented Airship for Aerial Surveillance. Since July 2011, Jorge Dias is on leave of absence to setup the Robotics Institute and research activities on robotics at Khalifa University (Abu Dhabi, UAE). He has been acting as a Professor of ECE/Robotics at Khalifa University, Abu Dhabi.

• • •