Advanced Engineering Informatics 30 (2016) 643–659

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

# Semantic scan planning for indoor structural elements of buildings

B. Quintana *, S.A. Prieto, A. Adán, A.S. Vázquez

*3D Visual Computing and Robotics Lab, Universidad de Castilla-La Mancha (UCLM), Paseo de la Universidad, 4, Ciudad Real 13071, Spain*

ARTICLE INFO

*Article history:*
Received 30 March 2016
Received in revised form 16 August 2016
Accepted 22 August 2016

*Keywords:*
Large-scale 3D scanning
Automatic 3D acquisition
Dense 3D reconstruction
Next best view in 3D environments
Large point cloud processing

ABSTRACT

The objective of this paper is to propose a new semantic 3D data acquisition method which is focused on sensing data belonging to indoor structural elements of buildings. Our system uses and processes 3D information coming from a 3D laser scanner sensor. The presented approach deals with some essential key issues in the scanning world which are rarely dealt with in papers. These are: the final goal of the scanning process, the hypotheses about the scene, lack of dynamic spaces in the next-best-scan-based solutions and the quality evaluation of the data sensed. Whereas most of the Next Best Scan (NBS) based approaches do not discriminate between data and clutter, we propose a scanning process in which potential structural elements of building indoors are learned as a new scan arrives. Our workspace is not a priori hypothesized, but a dynamic space which is updated as a new scan is added. This allows us to deal with more complex shape scenarios (i.e. concave-shaped spaces). Through the so called Structural Element (SE) membership probability, we introduce the data-quality concept in the scanning process which highly reduces the point cloud to be processed. This system has been tested in inhabited indoors and has yielded promising results. An experimental comparison with three close techniques is presented in an extended and detailed experimental section. The results yielded from our experimental work demonstrate the quality and validity of the proposed method.



## 1. Introduction

The creation of building models using large range sensors, e.g. 3D scanners, is a research line that is very active nowadays in computer vision forums [1–10]. Obtaining accurate as-built 3D models of facilities is a time-consuming process that is commonly carried out by hand for 3D modelling in engineering and architectural companies. Nevertheless, automated strategies are increasingly required in this field because they can provide more precise, less time-consuming and effective solutions.

Among the most important stages in the automatic creation of building models, we clearly distinguish between two complementary processes: 3D scanning and 3D modeling. It is important to point out that this paper exclusively deals with 3D scanning.

Automatic generation of 3D models has been strongly being developed during the last five years and it is still an open topic in 3D vision conferences. The goal of this research line is to identify essential parts of buildings, such as walls, ceiling, floors and attached columns [1–4,7,8,10,11], or even whatever object of the scene, typically furniture [12–16], by using range data under complex circumstances, i.e. occlusion and clutter. Although this objective seems feasible to be reached, it becomes very difficult in inhabited buildings, in which we can find a wide variety of objects that occlude the structural elements of the building [8,17].

The second process, the automatic scanning of buildings, is a more recent and challenging topic in which fewer researchers are currently working on. The primary objective here is to accumulate the maximum amount of data of the scene (indoor/outdoor) to be further modeled. As in the modeling stage, 3D digitization of buildings has been carried out manually by an operator for years. However, there are some serious problems mainly related with: the arbitrary number of scans taken, the time spent on the process, the data redundancy, the accuracy of the data and the completeness of the sensed scene. Such problems may also be reduced by automating the scanning process. It is important to point out that the quality of the final 3D models highly depends on the quality of the data, that is, the scanning process.

In summary, this paper presents a novel method for automatic scanning of structural elements of building indoors as an essential part of an automatic 3D model creation system. Thus, the scanning method presented in this document is just a component of a more complex system, in our case composed of a 3D laser scanner on board a mobile robot. The objective is that this setup is able to

* Corresponding author.
*E-mail addresses:* Blanca.Quintana@uclm.es (B. Quintana), Samuel.Prieto@uclm.es (S.A. Prieto), Antonio.Adan@uclm.es (A. Adán), Andress.Vazquez@uclm.es (A.S. Vázquez).

http://dx.doi.org/10.1016/j.aei.2016.08.003
1474-0346/

generate an accurate 3D model of the building structure by itself. Our system does not scan continuously, however it scans from specific positions called Next Best Positions. Therefore, the method here explained is exclusively addressed to provide the point cloud of structural elements, so that an accurate geometrical 3D model can be obtained in further stages.

## 2. Automatic scanning approaches: related works and weaknesses

To begin with, manual and semiautomatic scanning techniques have to be separated from automatic ones. Thus, approaches that perform permanent scan on cars, on board humans or even with commanded mobile robots, perform just semiautomatic scanning tasks, most of those being SLAM (Simultaneous Localization and Mapping) systems manually commanded. Here, a human decides where to go and how to perform a complete scanning in a large scenario. For example, in Toschi et al. [18] a laser scanning system on board a chauffeur-driven car digitizes the downtown of Trento. Xiao et al. [19] reconstruct indoors by means of a RGB-D camera that is handled by a person. The individual who carries the system takes care of avoiding data redundancy while he moves. A commanded robot with a 3D scanner is used in [20] and micro aerial vehicles (MAV) extract 3D information on interior scenes in [21]. All these related works can be considered as non autonomous/automatic methods.

One of the keys in automatic scanning is the one of the good selection of the next scanner position. The decision of the next best position should lead us to a complete, high-quality and non-time-consuming digitization of the scene. This is known in the literature as the Next Best View problem (NBV). In our context, it could be renamed as Next Best Scan (NBS). The original NBV's concept was discussed by Connolly in [22] and applied for object reconstruction.

In two dimensions, the NBS problem can be approached with traditional geometric solutions such as in Xie et al. [23] or with randomized methods such as in González-Baños [24]. The Art Gallery Problem (AGP) is set as a 2D optimization problem over a polygon representing the outline of an art gallery. The problem consists in finding the smallest set of positions where the guards visually cover the entire gallery. Many variations and solutions of this problem have been studied in the literature, but all of them usually assume some initial model of the scene.

Couto et al. [25] present an exact algorithm for the AGPVG (AGP with vertex guards). The algorithm iteratively discretizes a witness set creating a sequence of AGP instances which are then modeled as set cover problem (SCP). Kröller et al. [26] develops another approach aiming at solving the AGP in an exact way. The idea of their algorithm is again to discretize not only the witness set but also the guard set and to model the restricted AGP as an SCP. Lately, Tozoni et al. [27] present a practical iterative algorithm for the AGP with point guards, which finds a sequence of decreasing upper bounds and increasing lower bounds for the optimal value. They prove the effectiveness of the method after testing it for 1400 instances of polygons. Finally, Borrmann et al. [28] present a video in which an integer programming approach for general point guards is presented. The method is based on an analysis of possible guard and witness positions in the arrangement of visibility polygons.

The extension of AGP to a 3D context is usually based on planning a set of initial views by using two-dimensional maps of the region. These are "offline" problems in which the complete instance is known. All scanning locations in this initial phase are planned in advance, before any 3D data acquisition occurs [28,29]. Of course, this resembles again the art gallery problem. Therefore, this view planning strategy makes the assumption that if we can see the 2-D footprint of a wall then we can see the wall in 3D, but this is not always true. In addition, the occlusion problem is not dealt with in these approaches. Therefore, 3D planning solutions from 2D maps might be inefficient in indoor areas with high occlusion.

The underlying problem in this paper is of the nature "online", i.e., with incomplete information of the scene and under a 3D context. This means that we solve the art gallery problem without a priori knowing the complete instance.

Our NBS method is developed with 3D information. Several works have been dedicated to introducing NBS solutions in buildings and facilities using 3D data taken from range sensors and laser scanners [30–33]. Nevertheless, there are a good number of underlying questions and weaknesses in this field that are rarely debated in papers and that determine the goodness of the automatic scanning method. In order to argue the validity and contributions of our approach in this context, we will make a critical discussion around five key issues.

### 2.1. The objective of the scanning process is imprecise

The majority of the current approaches can be inefficient in building model creation tasks because the scanning is just seen as a process in which the objective is to accumulate as much data as possible. Thus, the objective is to scan everything which lies inside [34,31] or outside [29,35] the building, but the approach does not usually deal with the meaning of the data. Redundancy and cluttering are frequently ignored, so that a huge amount of redundant 3D data is processed in order to generate a building model.

In this sense, those methods are inefficient in the scanning stage because a great part of the stored 3D data could be irrelevant. For example, if the objective is to recognize the set of openings within the room, the data belonging to the furniture is irrelevant information.

In contrast with these methods, our proposal is addressed from the beginning to capture the data belonging to the structural elements (SE) of the scene, which essentially are floor, walls, attached columns and ceiling. Thus, our NBS algorithm is based on the current knowledge of structural elements and, consequently, we highly reduce the volume of data and alleviate the algorithmic complexity in further processes.

### 2.2. The a priori knowledge and strong hypotheses imposed

Many scan planning methods work due to the fact that important and essential knowledge of the scene is a priori assumed. For example, one of the inputs of the approach proposed in [36] to generate automatic scanning plan is the same 3D model of the facility. Normally, the existing NBS algorithms assume bounding boxes or convex-hulls that determine the borderlines of the scene from the beginning of the scanning process [34,37,31] or, at least, they set it as a result of the initial scan of a single room [38], or of a preliminary complete point cloud of the environment [29]. For example, if NBS depends on the number of occluded voxels viewed from a particular position of the scanner, the method a priori assumes an enveloping workspace.

Contrary to these methods, the boundaries of our workspace are not hypothesized but they are calculated and updated as a new scan is added.

### 2.3. The absence of dynamic workspaces

An essential point to be considered is how to cope with large point clouds which are generated from the scanner after a few