

An Industrial Augmented Reality Solution For Discrepancy Check

Pierre Georgel₁*

Pierre Schroeder₁

Mirko Appel₂†

Selim Benhimane₁

Nassir Navab₁

Stefan Hinterstoisser₁

₁Chair for Computer Aided Medical Procedures and Augmented Reality, TUM Munich

₂Siemens CT, Munich, Germany

ABSTRACT

Construction companies employ CAD software during the planning phase, but what is finally built often does not match the original plan. The procedure of validating the model is called “discrepancy check”. The system proposed here allows the user to easily obtain an augmentation in order to find differences between the planned 3D model and the built items. The main difference to previous body of work in this field is the emphasis on usability and acceptance of the solution. While standard image-based solutions use markers or rely on a “perfect” 3D model to find the pose of the camera, our software uses Anchor-Plates. Anchor-Plates are rectangular structures installed on walls and ceiling in the majority of industrial edifices. We are using them as landmarks because they are the most reliable components often used as reference coordinates by constructors. Furthermore, for real industrial applications, they are the most suitable solutions in terms of general applicability. Unfortunately, they have not been designed with Computer Vision applications in mind. On the contrary, they are often made or painted in such way that they are not easily popping out. They are therefore difficult targets to segment and to track.

This paper proposes a solution to extract and match them to their 3D counterparts. We created a software that uses the detected structures for pose estimation and image augmentation. The software has been successfully employed to find discrepancies in several rooms of two industrial plants.

Index Terms: H.5.1 [Multimedia Information Systems]H.5.2 [User Interfaces] I.4.6 [Segmentation] I.4.9 [Applications]

1 INTRODUCTION

Nowadays, industrial plant designers plan projects using CAD software. These systems give them the ability to design extremely precise models for the items to be constructed. Additionally, they provide the possibility to share, upgrade, and archive the designed model easily. These models are then given to builders as a guideline to erect the plant. During the construction, even if the planning is done perfectly, mistakes occur: pipes are swapped, handles are shifted, parts are missing, etc. This creates a gap in the digital workflow. The designed 3D model cannot be trusted anymore. If it needs to be used again for maintenance, it will have to be controlled and possibly updated.

In order to check the built items, several alternatives are used in industry. The standard solution is to take manual measurements of the room. This is time demanding and error prone. Nevertheless, these techniques are used to verify certain components (e.g. concrete structures). For complex systems such as pipes, this approach may lead to a low precision. To overcome this problem, industry is trying to use laser scanning technology. This promising technique

uses a laser beam, which measures the distance between the apparatus and all surrounding opaque objects. This returns a dense and accurate point cloud. But there is no automated solution yet to obtain a parametric representation from the data, which could be used to validate the model. Thus it is complicated to use this technology.

In this paper, a new approach for solving the “discrepancy check” task is proposed. The idea is to allow the engineers to superimpose images from the plant with the CAD model developed during the planning phase. In order to recover the imaging geometry relative to the 3D model, several techniques are available. The common solution is to use optical markers [7, 3]. This is used in several projects to find the pose of images [15, 4] and has been successful in the past due to the automatic localization and decoding of markers. Unfortunately, the “discrepancy check” can be seen as a wide area tracking problem, where markers proved to be complex to operate with [17]. Additionally, industry is not much interested using markers as they require installation and alter the plant.

Another approach would be to use feature base tracking. For examples [8, 2] use lines extracted from the 3D model. Also tracking of image features with *keyframes* [16] has proven to be successful [14]. But these techniques rely on a perfect model which is clearly not available in this situation. Furthermore, the visual environment consists mainly of pipes, which are complex to detect. The texture is sparse and repetitive so standard feature point based methods even with strong descriptors like SIFT [9] are delicate to handle.

To overcome these problems, the use of trusted components is necessary. These components should have been built as planned. In addition, they should consist of simple primitives to ease the detection and be widespread to make the process reproducible. After discussion with plant designers, we decided to exploit Anchor-Plates¹ because they are positioned with lots of care. In fact they are the most reliable components in the factory. Anchor-Plates are metallic structures embedded into the concrete walls. They are mounting points to fix other components: pipes, supports, cable-racks, control boxes, etc. After site construction they are often partially occluded. And since they are not designed to be distinguishable (i.e. they are painted like the walls), their detection is complex. We developed an interactive method to segment them combining Canny edge detector and Hough transform. The advantage of Anchor-Plates is that their positions are almost unique within every single wall. This leads to a robust and automatic 2D-3D matching procedure, performed with a homography estimation method.

The solution presented in this paper proposes an alternative to the use of laser scans or other techniques to verify the correctness of CAD offering a simple, fast and intuitive way to check the built items. The efficiency of the interactive registration algorithm enables the user to obtain quickly a superimposition of the plant’s image by the 3D model.

*e-mail:georgel@in.tum.de

†e-mail:mirko.appel@siemens.com

¹In some facilities, the so-called Dowel-Plates are used instead. They are similar to Anchor-Plates in shape and function.

2 RELATED WORK

There has been a lot of work on superimposing CAD model and images. Noelle et al. [13, 12] focused on developing good visualization techniques to display augmentation. They show the results of crash-test simulations overlaid on top of real images. However, since this prototypical system is marker-based, it is not suitable for our application.

The Cylicon project [11, 10, 1] is targeting the problem of as-built reconstruction of industrial systems from pictures. They superimpose industrial drawings (e.g. floor-maps) onto pictures of the plant by performing several steps: placing markers, measuring some of them, and acquiring photos. This system targets the 3D reconstruction problem and not the “discrepancy check” task. Access3D [18] is capable of overlaying the planned CAD model on pictures, using marker technologies. It was published as a test-bench for point-based registration, but was not studied as an application to check the planned model against the actual built item. All of these methods require a long off-line process to equip the environment with markers and register their position to the CAD reference frame.

Our approach minimizes the preparation step by allowing to directly use images of the object of interest. Also, no markers are used because the 3D data is directly extracted from the CAD model.

3 “DISCREPANCY CHECK” PROCEDURE

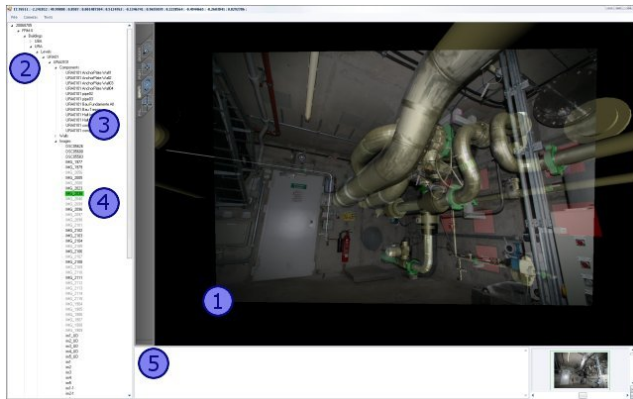


Figure 1: The main user interface where the 3D viewer (1) takes the majority of the screen. The hierarchy tree (2) gives access to the CAD components(3) and to the stored images(4), any additional text information is displayed in the text box(5).

The system was designed in close collaboration with industrial partners. The goal is to create an application for a fast discrepancy check. It is a complete software that guides the user during every stage (see figure 1). This gives the opportunity to someone, with no knowledge related to Computer Vision or Augmented Reality techniques, to use this tool because no expert knowledge is needed. All complicated procedures (camera calibration and image registration) are wizard driven, so the users do not have to understand the underlying algorithms to use the software.

Instead of going on site to measure and document a room, our procedure lets the controller inspect the site with a calibrated² camera and acquire pictures from the items to study. Back to the office, using the AR software the controller selects the room to inspect and uploads the pictures to a database. The photos are undistorted automatically. Based on the software, the user now registers the image

²Our software have an integrated off-line automatic calibration method that computes the intrinsic parameters and the distortion coefficients.

to the CAD coordinate system. This is done in two steps: first, segmentation of Anchor-Plates in the images, and second, matching them with the 3D Anchor-plate models extracted from the CAD.

When this is done, the user can start the inspection. He can now select the 3D components (retrieved from the database) to be superimposed. The user can change the transparency or the depth of the frustum to avail possible discrepancies. Finally, he decides whether an item is well built or if it needs reconstruction. This decision comes from the user by visualizing all planned CAD components against actual pictures of the factory. The item is considered well built when there is no visible misalignment between the model and its picture.

Since the software is designed to be used in large facilities by numerous users, the application is attached to a database server. The software is a client that will store/retrieve all the data into/from the server (building organization, levels, rooms, camera information, images, poses, etc). The original CAD model was converted into VRML using a third-party software. The model was simplified to facilitate the data transfer and its browsing. The software integrates a Cortona viewer, which takes care of the rendering.

4 REGISTRATION PROCESS

The registration is based on a specific components: the Anchor-Plates which, as mentioned earlier, are complex to segment. In consequence, we decided to implement an interactive segmentation techniques where the user will select the area where an Anchor-Plate lays and the software will return possible candidates letting the final choice to the user. Once two Anchor-Plates from the same wall are selected, the software automatically detects their 3D counterparts. Then the pose can be computed and the position of the image in 3D is determined.

4.1 Landmark Extraction

Despite their simple geometry, performing the automatic extraction of the Anchor-Plates is far from being easy. In fact, Anchor-Plates are used as mounting points for many kinds of plant components. Consequently they are, in most of the cases, partially occluded. Therefore, in addition to their wide variety, their appearance is highly dependent on the attached object, on the viewpoint and on the illumination conditions. The techniques that work well for marker detection, such as adaptive threshold followed by an exhaustive search of closed quadrangles, do not permit extracting the Anchor-Plates since their borders are often incomplete and not well contrasted.

We evaluated different approaches to automatically segment the Anchor-Plates and determine their borders but few methods provided acceptable results. The best results were obtained using a semi-assisted approach, where the user selects an area around an Anchor-Plate and then validates a choice among a set of candidates suggested by the system. The candidates are obtained in four steps: 1. A Canny edge detector is performed in the selected area; 2. Using the obtained edges, a line detection based on a Hough transform makes it possible to reconstruct incomplete borders; 3. All the quadrangles that can be obtained using the detected lines are then formed; 4. Constraints on the angles, the size, the relative length of the sides and the orientation of the quadrangles are used to reduce their number to a small set that contains the correct Anchor-Plate borders.

This algorithm gives very good results due to the combination of the Canny edge detector and the Hough transform that permits to have a strong robustness to partial occlusions and to various illumination conditions. Once the sides are reconstructed, the constraints based on the general geometry of the Anchor-Plates allow us to obtain the desired extraction as it can be seen in the Figure 2. The proposed algorithm is a good compromise between precision and speed. It satisfies the targeted objective of user-friendly software

since the users only need to validate a choice among a small set of possible extractions.

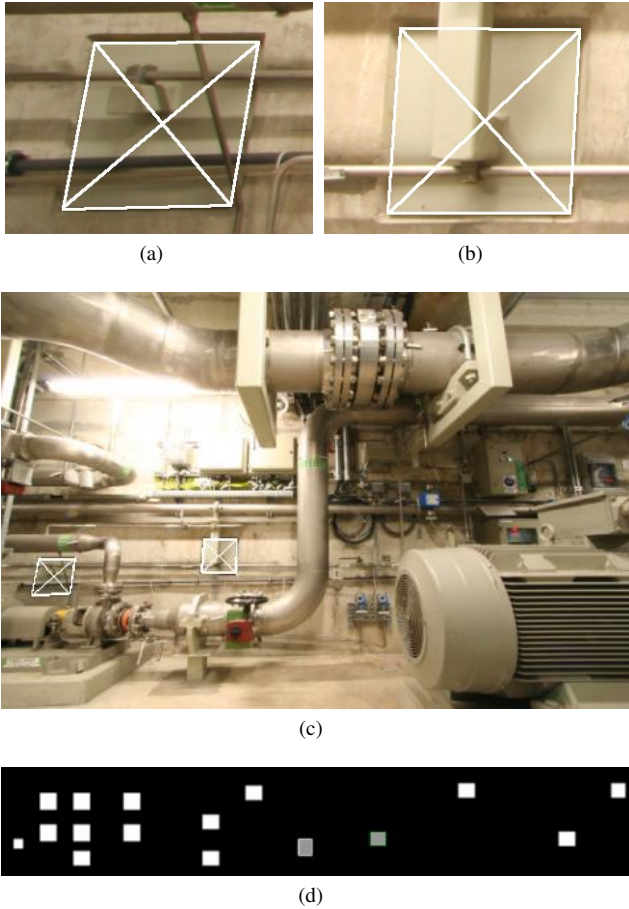


Figure 2: (a) and (b) show examples of extracted Anchor-Plates; (c) presents typical layouts of the extracted Anchor-Plates and (d) shows the corresponding walls; in gray are the matched Anchor-Plates.

4.2 Matching and Pose Estimation

Once the Anchor-Plates are detected, they have to be matched with the 3D data in order to find the pose. The layout of the Anchor-Plates on a wall is mostly unique, there are rarely two similar organization in a room. So it is possible to find the correspondences between the Anchor-Plates extracted from the image and multiple 3D candidates automatically. Since the detected Anchor-Plates all belong to the same plane in 3D, one could find a homography to validate the matching. Supposing that the correspondences between the Anchor-Plates in the image (AP^{2d}) and the 3D ones (AP^{3d}) is known, there is a homography that defines this geometrical relationship.

In both the image and the 3D world, Anchor-Plates are represented by their four corners. Given one 3D and one 2D anchor-plate, the correspondence between corners cannot be found. For a rectangular shape, there are 4 solutions (mirroring not included). For the moment, this local orientation will be supposed given by the EXIF (Exchangeable Image File Format) rotation tag. This is provided by most of modern cameras to determine whether an image is a portrait or a landscape. Usually, there are few detectable Anchor-Plates, most of them are not in the field of view or too cluttered to be extracted. For each putative correspondences between

the elements of AP^{3d} and AP^{2d} , we compute a homography H_k . This is done using the Direct Linear Transform (DLT) [5] with all four correspondences.

In order for the results to be as stable as possible, we normalized the points (2D and 3D), as suggested by Hartley in [6]. Empirically, we found out that using a non-isotropic normalization was more robust. That is maybe because our points are often spread in one dimension as shown in 2. During the normalization the dimensionality of the 3D points is reduced to 2. This is done by projecting the 3D points onto the wall plane.

We use a mean reprojection error in order to grade the quality of each homography. The right correlation between the 2D set and the 3D should minimize the following formula:

$$e(H_k) = \frac{1}{4n} \sum_{i=1}^n \sum_{j=1}^4 \left\| AP_{ij}^{2d} - H_k * AP_{ij}^{3d} \right\|^2, \quad (1)$$

where AP_{ij}^{2d} is the j^{th} corner of the i^{th} 2D anchor-plate, AP_{ij}^{3d} is the 3D point related to the j^{th} 2D corner of the corresponding i^{th} anchor-plate and n is the number of Anchor-Plates considered.

To refine the result, we re-normalize the 3D points once some putative matching have been obtained. We normalize these points only using the matched points. This gives more comparable reprojection error. In case of multiple answers (i.e. several putative matches which have $e(H_k)$ under a threshold) the user will have to decide which one is correct within a subset of 5 possibilities. Since the same layout of the Anchor-Plates on a wall is rarely reproduced, Anchor-Plates from all walls are used as input. This eases the workflow since no walls have to be selected. For pictures of the ceiling where the EXIF rotation tag can not be used to determine the local orientation (landscape or portrait), we supply additional inputs to the matching algorithm. These additional inputs are generated by rotation of the order within each Anchor-Plate leading to 4 additional inputs for the algorithm. Typically, there are 4 walls and 4 ceilings with different orientations as input for the matching algorithm for a 4-corners room.

When all the points are matched, the pose computation is executed. This process returns the camera pose in the coordinate system of the 3D model. This allows us to create an augmented CAD by positioning the image into the 3D view. These transformations are stored in the database and linked to the image.

5 RESULTS AND DISCUSSION

This software was successfully used to verify two rooms of two different factories. The 3D model and the coordinates of the 3D components used to do the registration (i.e. Anchor-Plates) were transferred to a database server. Two different cameras were used to test the software, one Sony Cybershot with a fixed lens and one Canon EOS 350D mounted with wide-angle lens. These cameras were set in manual mode for the experiments. Pictures from all the walls were acquired, where Anchor-Plates were visible in order to test the proposed procedure. All the images were automatically undistorted and added to the database.

The procedure has been tested on more than 50 images. The segmentation performed well with respect to the condition (lightning, occlusions, etc). We tried to segment 45 different Anchor-Plates. A third of them has been correctly segmented. In order to have a maximum number of images to test the registration process, the user manually selected the 4 corners, whenever the segmentation failed. The matching has performed extremely well, more than 90% of proposed matches were correct. It failed mainly because of wrong 3D data often coming from an outdated CAD model. Using only two Anchor-Plates has shown to be enough to obtain a correct augmentation. When a wrong pose is computed, it is obvious for the user because of the massive misalignment.

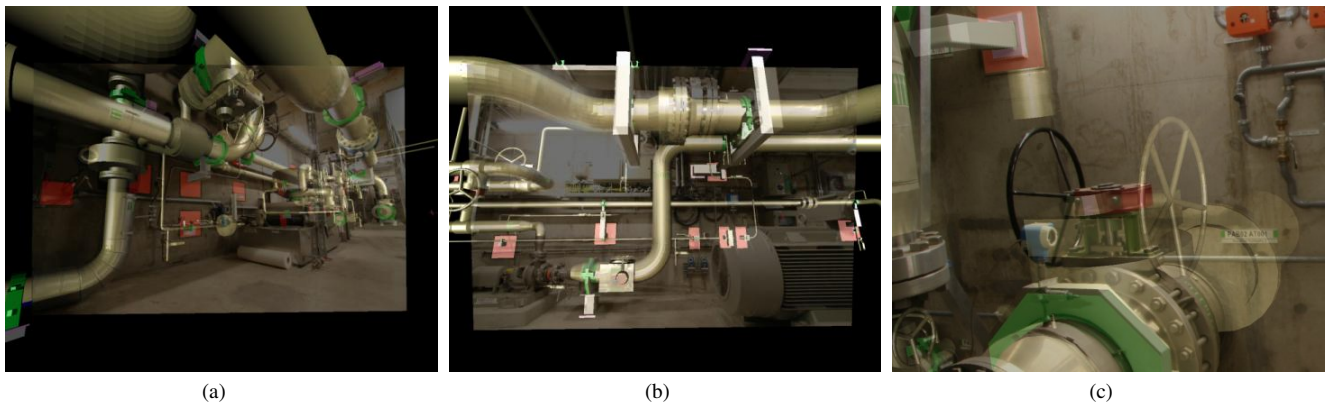


Figure 3: (a) (b) display typical augmentation realized with our system; (c) presents a close-up view of one of this augmentation we can see that a valve in the CAD (rendered in gold) was designed to be oriented to the right while in the current installation is mounted to the left.

By offering a seamless mean to quickly obtain an augmentation, it enables the use of AR for a “discrepancy check” task. This software, as shown in figure 3, gives the possibility to do a detailed inspection of the built items by changing the depth or transparency of the frustum.

In addition, our solution provides the CAD designers with registered images of the factory in its current state. This is a great advantage because even with an up-to-date model, all objects are not represented but they may be very well visible in the images. The designer’s workflow is eased during the update and the maintenance because they have access to all present items in the plant.

6 CONCLUSION

In this paper, we presented a system that guides the user from image acquisition to CAD augmented visualization for discrepancy check. The proposed solution gives the possibility to register pictures to the associated CAD model for inspection processes. Contrarily to standard approaches that make use of markers, here, we use actual reference Anchor-Plates that are commonly employed in industrial environments. The proposed solution includes a step-by-step segmentation and registration process. For the discrepancy evaluation, different interactions are used to modify the augmentation and make the differences visible. The presented software has proven to be useful for discrepancy check and also for upgrade planning in order to inform designers about the presence of non-planned objects that should be taken into account.

Future Work: In order to register images that do not contain Anchor-Plates, we plan to use a feature-based matching with already registered images. This would imply a robust feature extractor, for poorly textured environment common in industrial application. In parallel, we will implement some 3D image browsing techniques to offer a better navigation system. So it would be easier to intuitively navigate within a dense set of registered images. Finally a study on the error propagation will have to be performed to evaluate whether the augmentation is reliable enough for its large scale application.

ACKNOWLEDGEMENTS

The authors would like to thank our partners at Areva NP : Werner Biemann, Ralf Keller, Martin Neuberger and Erwin Rusitschka for their time and advices. They would also like to express their gratitude toward Harald Türck of TUM for his help and patience. Finally, they appreciate the commitment from the CAMPers especially Ivan Borisov.

REFERENCES

- [1] M. Appel and N. Navab. 3D Reconstruction from Co-Registered Orthographic and Perspective Images: Theoretical Framework and Applications. *ICPR*, 04:40021, 2002.
- [2] C.-C. Chang and W.-H. Tsai. Reliable Determination of Object Pose from Line Features by Hypothesis Testing. *IEEE Trans. PAMI*, 21(11):1235–1241, 1999.
- [3] M. Fiala. ARTag, a Fiducial Marker System Using Digital Techniques. *CVPR*, 2:590–596, 2005.
- [4] S. Goose, S. Sudarsky, X. Zhang, and N. Navab. Speech-Enabled Augmented Reality Supporting Mobile Industrial Maintenance. In *IEEE Pervasive Computing*, volume 2, Issue 1, pages 65–70, 2003.
- [5] R. Hartley and A. Zisserman. *Multiple View Geometry*. Cambridge University Press, 2000.
- [6] R. I. Hartley. In defense of the eight-point algorithm. *IEEE Trans. PAMI*, 19(6):580–593, 1997.
- [7] H. Kato and M. Billinghurst. Developing AR Applications with AR-ToolKit. In *ISMAR*, pages 305–305, USA, 2004.
- [8] Y. Liu, T. S. Huang, and O. D. Faugeras. Determination of Camera Location from 2-D to 3-D Line and Point Correspondences. *IEEE Trans. PAMI*, 12(1):28–37, 1990.
- [9] D. G. Lowe. Distinctive Image Features from Scale-Invariant Key-points. *IJCV*, 60(2):91–110, 2004.
- [10] N. Navab, B. Bascle, M. Appel, and E. Cubillo. Scene augmentation via the fusion of industrial drawings and uncalibrated images with a view to marker-less calibration. In *IWAR*, pages 125–133, 1999.
- [11] N. Navab, N. Craft, S. Bauer, and A. Bani-Hashemi. CyliCon: Software package for 3D reconstruction of industrial pipelines. In *WACV*, page 278, USA, 1998.
- [12] S. Noelle. Stereo Augmentation of Simulation Results on a Projection Wall by Combining two Basic ARVIKA Systems. In *ISMAR*, pages 271–272, Germany, 2002.
- [13] S. Noelle and G. Klinker. Augmented Reality as a Comparison Tool in Automotive Industry. In *ISMAR*, pages 249–250, USA, 2006.
- [14] J. Platonov, H. Heibel, P. Meier, and B. Grollmann. A mobil marker-less AR system for maintenance and repair. In *ISMAR*, USA, 2006.
- [15] D. Reiners, D. Stricker, G. Klinker, and S. Mueller. Augmented reality for construction tasks: Doorlock assembly. In *IWAR*, pages 31–46, USA, 1998. AK Peters.
- [16] L. Vacchetti and V. Lepetit. Stable Real-Time 3D Tracking Using Online and Offline Information. *IEEE Trans. PAMI*, 26(10):1385–1391, 2004.
- [17] M. Wagner. Building Wide-Area Applications with the AR Toolkit. In *The First IEEE Int. Augmented Reality Toolkit Workshop*, 2002.
- [18] X. Zhang. Projection Matrix Decomposition in AR - A Study with Access3D. In *ISMAR*, pages 258–259, USA, 2004.