Online Inspection of 3D Parts via a Locally Overlapping Camera Network

Tolga Birdal
Technical University of Munich
tolga.birdal@tum.de

Emrah Bala, Tolga Eren
Gravi Ltd
{emrah,tolgaeren}@gravi.com.tr

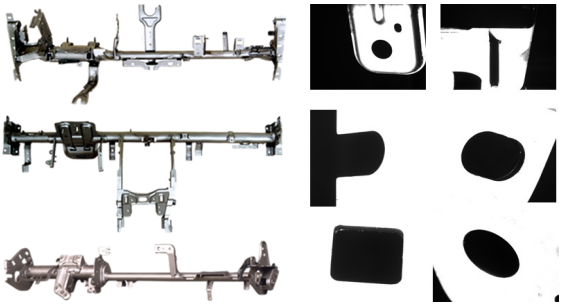
Slobodan Ilic
Siemens AG
slobodan.ilic@siemens.com

Abstract

The raising standards in manufacturing demands reliable and fast industrial quality control mechanisms. This paper proposes an accurate, yet easy to install multi-view, close range optical metrology system, which is suited to online operation. The system is composed of multiple static, locally overlapping cameras forming a network. Initially, these cameras are calibrated to obtain a global coordinate frame. During run-time, the measurements are performed via a novel geometry extraction techniques coupled with an elegant projective registration framework, where 3D to 2D fitting energies are minimized. Finally, a non-linear regression is carried out to compensate for the uncontrollable errors. We apply our pipeline to inspect various geometrical structures found on automobile parts. While presenting the implementation of an involved 3D metrology system, we also demonstrate that the resulting inspection is as accurate as 0.2 *mm, repeatable and much faster, compared to the existing methods such as coordinate measurement machines (CMM) or ATOS.*

1. Introduction

Dimensional monitoring has become an important part of manufacturing processes due to the gradually advancing standards in many sectors, such as automotive (chasis), flight (wings) or energy (turbines). The welded components found on the backbone of all these high-end products are sensitive and require careful manufacturing and thus, careful quality control. The de-facto solution to inspect these critical parts relies on the contact-based CMMs (coordinate measuring machines), which are expensive, hard to maintain and slow to operate. These machines can neither be installed on the production lines nor provide 100% statistics on the manufactured parts. Such lack of online inspection coerces many assembly lines to revert to mechanical fixtures and apparatus to verify the quality of manufacturing. This has many drawbacks. First of all, notwithstanding the cost of assembling mechanical control fixtures, sustaining the precision of such assembly over the long term requires significant amount of maintenance effort. Yet, no information on errors or statistics could be provided by the fixed control equipment and the inspections cannot be documented. Last but not least, such inspection necessitates an operator constantly engaging in the production process.



(a) (b)

Figure 1. Samples from the parts, our system could inspect.

The bottleneck caused by the current quality control procedures is addressed by Tuominen as the lack of reliable, accurate and generic metrology / photogrammetry techniques, which are amenable to be installed on production lines [30]. Tuominen further shows that the ultimate quality control performance is achieved when the parts are either fully inspected or no inspection is carried out at all. This makes CMMs an unviable option. While the industry suffers such drawbacks, the quality requirements imposed by the consumer market keep demanding more and more, stimulating the research on online 3D inspection [17].

This paper presents a thorough methodology on the research and development of an inspection unit, which addresses all of the aforementioned downsides. Our system is composed of multiple static cameras, calibrated to a global coordinate frame, being able to measure arbitrary geometrical structures. Our algorithms are novel and carefully designed to fit the accuracy requirements. The resulting system is accurate within acceptable error bounds and flexible, in the sense that the design can be adopted to the inspection of different parts with minimal modifications to the software. The entire implementation cost is lower than a CMM, while the capability of inspection is 100%, saving us from

the sampling requirement. We present quantitative comparisons with the industry-strength commercial metrology systems by evaluating our approach against an optical metrology unit (ATOS) and a standard CMM (Hexagon) both for calibration and measurement. We also present our repeatability values along with the actual measurements. Finally, we demonstrate that our timings clearly outperform the state of the art, making the system applicable to online inspection. See Figure 1 for the parts we can inspect as well as the images of the measured geometries.

The contributions of this paper are three-fold: First, we propose an affordable and effective way to calibrate a mulicamera systetm, which is composed of local and global camera networks. We base our calibration on bundle adjustment and global registration. Secondly, we develop a novel, robust, multi-view projective CAD registration procedure along with a multiview edge detection scheme. We show that this new approach is real-time capable, while not compromising accuracy. We review in detail the implementation aspects and draw a complete picture in the design and realization of such a multi-view metrology unit. Finally, we carry out extensive experimental evaluation, in comparisons to existing, state of the art metrology methodologies.

2. Prior Work

Optical metrology enjoys a history of over 30 years. The different aspects of the issue such as calibration, triangulation, stereo reconstruction or pose estimation have been tackled many times in computer vision literature.

Analyzing the colinearity relations, photogrammetrists were the first to use image data to conduct measurements [25, 16]. Computer vision tried to improve the lower level sub-problems such as calibration [29, 35], pose estimation [24, 28] or SLAM [2]. In spite of the vast amount of literature in computer vision, the care for accuracy and precision is not very well established in such works.

Many commercial metrology software exist with different application areas [1,2]. While being highly accurate, some of these software are not capable of measuring generic CAD geometries, but rather rely either on feature points or predefined markers. In contrast, the tools which could recover arbitrary 3D geometries cannot operate on generic prior 3D models. Even today, a well described, truly capable machine vision technique to automate the process control remains to be intact [17].

Unlike the commercial arena, development of accurate and reliable close range machine vision systems is rather unexplored in academia [17]. Jyrkinen et.al. [14] designed a system to inspect sheet metal parts, but their approach cannot be generalized to arbitrary geometries. Mostofi et. al. [21] target a similar problem like ours. Their technique is dependent on human intervention and utilize a single moving camera. Such choice is far from realtime concerns. Yet, authors report the visibility of measurement points from 6 cameras, which drastically increases the number of views and the effort for measurement. Moreover, they report an accuracy of 0.5mm even when the measurement points are visible in 6 views. Such accuracy is well below what is achievable today [31]. Bergamasco et. al. [4] developed a novel method to precisely locate ellipses in images, but their method involves perfect overlap of views and doesn't generalize to measurement of arbitrary 3D shapes. Similar to our work, Malassiotis and Strintzis developed a stereo system to measure holes defined by CAD geometries [19]. While posing the CAD fitting problem as an optimization procedure, just like ours, they make use of explicit primitive modeling, which restricts the measurement capabilities. The 3D primitives are re-generated at each step of the optimization and this comes at the expense of computational complexity, degrading the real-time (or online) capabilities. Their approach is also not applicable to triangulate arbitrary 3D geometries. Moreover, all of these approaches assume a perfect calibration and lack a well established methodology to calibrate the camera networks.

Our system is uniquely positioned in the application field of non-contact multi-view measurement, which was shown to be one of the most accurate metrology methods. It is tuned for inspection of points of interest [18] and retrieves its power from the developments in low level and geometric 3d computer vision e.g. [9, 5].

3. Proposed Approach

System Setup Our system consists of 48 *The Imaging Source* cameras providing 1280x960 pixels at 30 fps. Depending on the installation distance, we choose either 25mm, 35mm or 50mm industrial lenses. The scene is illuminated via 10 white global LED lights, installed on the external skeleton. The cameras are positioned such that a point of interest is visible to at least 3 cameras. We will refer to these points of interests as the *measurement points* $\mathbf{P}_i \in \mathbb{R}^3$ and to such a local camera group as the *local camera network* $\mathbf{C}^i = \{[\mathbf{C}^i_1, \mathbf{C}^i_2, ..., \mathbf{C}^i_k] : k < N\}$. For the sake of accuracy, it is highly unlikely that a single camera would capture different measurement points and thus the local camera groups remain independent (no overlap and thus no direct calibration exists between camera groups). The set of these K local camera groups form the *global camera network* $\mathbf{C} = \{\mathbf{C}^1, \mathbf{C}^2, ..., \mathbf{C}^K\}$, in total composed of N cameras. Eventually, our system consists of multiple local camera networks, which are calibrated to form $\mathbf{C}$, the global camera network. In this sense, the entire network is arranged in a hierarchy, as shown in Figure 2.

Target Parts For demonstration purposes, and specific to

[1] http://www.photomodeler.com/

[2] http://www.gom.com/3d-software/atos-professional.html

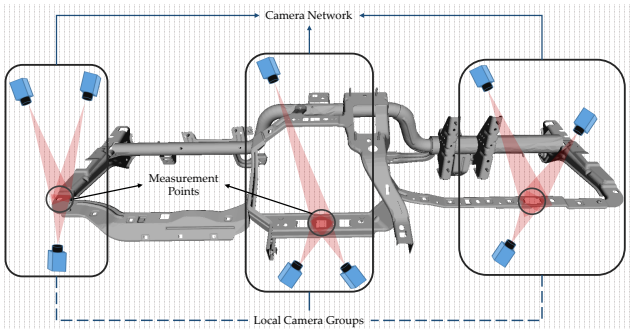


Figure 2. System Setup. Figure depicts the locally overlapping camera groups forming a global camera network. The shown CAD model belongs to the actual part.

our application, we aim to inspect 15 critical points ($\mathbf{P}_i$) on a $2m$ long and $80cm$ wide metal industrial automotive chasis. The part is shown in Figure 2. Even though our setup is capable of inspecting arbitrary CAD models from different manufacturing areas, for the purpose of clarity, we will stick to the automotive application. Typically, the geometries to be measured, as well as the CAD models are known beforehand and the accuracy is limited to 1-1.5%.

Image Acquisition We control all cameras over a GigE Network ($1000Mbit/s$). The lights are strobed to be in synch with the exposure. This prevents the multi-threaded capturing. We initialize each local network group per each measurement, and utilize the full Gig-E bandwidth (using 30fps/camera). This optimizes the speed and durability. Some images of the measured points, acquired by our system during a single runtime are shown in Figure 1(a).

3.1. Calibration

We calibrate all cameras intrinsically and extrinsically to the same global coordinate frame. Intrinsic calibration is carried out prior to installation using the standard methods [10, 6]. We use calibration plates made up of 3mm circular reflective dot stains. On the other hand, considering the extremely non-overlapping nature of our cameras, and narrow depth of fields, the task of extrinsic calibration is significantly difficult and crucial.

Extrinsic Calibration We use a pre-manufactured calibration object (reference body), *The Calibrator* to calibrate the absolute poses. Counter intuitively, this object is not precisely manufactured and does not use expensive material or components, but only acts as a mounting apparatus for circular dots, forming the relation over different camera groups. It is produced at a CNC machine such that when it is imaged by the multiview system, enough 3D reference points (calibration dots) would be visible on each camera. The shape resembles the rough approximation of the target object (part to be measured). We then manually attach random circular markers so that enough overlap is created. The calibrator (virtually and physically) is shown in Fig. 3.

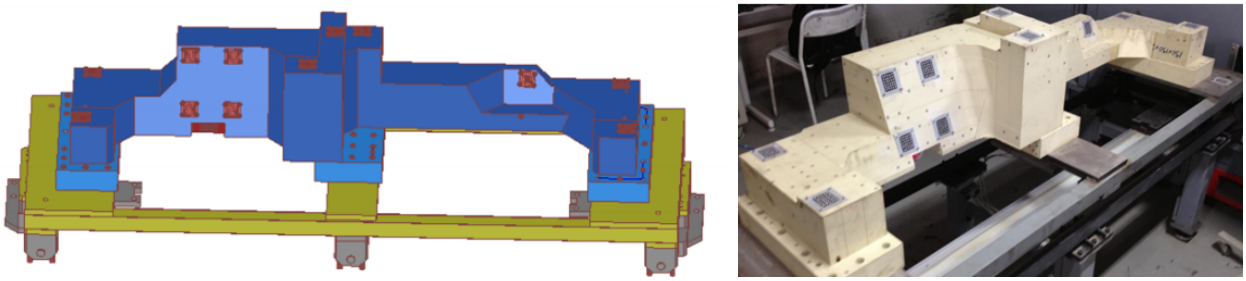
The calibration grids and random dots attached on the calibrator are first reconstructed and bundle adjusted by taking multiple overlapping shots with a high-resolution SLR camera. We take around 128 images, and compute the pose graph (a graph, in which each edge depicts an overlapping view) semi-automatically. The global optimization is solved with Google Ceres [1].

After creating The Calibrator, we extrinsically calibrate the cameras by imaging the calibration plates multiple times for each view. During this repetitive stage, at each inspection, the positions of the calibration plates are computed and subsequent measurement errors are recorded. Once enough measurements are collected and repeatability values become acceptable, the images are gathered to be used in multiview calibration. This way, we make sure that enough variance of positioning is covered. Such extrinsic calibration is performed as a succeeding offline bundle adjustment (BA), where the poses and 3D structure are refined simultaneously [11, 15].

Learning Systematic Errors Finally, the so-far neglected non-rigidness effects, mis-manufacturing and other systematic errors are compensated. To achieve that, we employ an error correction (bias reduction) procedure, in principle similar to [23, 22, 12]. While these methods were designed to perform on CMMs, we generalize the technology to optical inspection via a learning approach. Initially, we sample a set of carefully-manufactured 3D parts. We measure each of those on various CMM devices, repeatedly (5 times for each) to cancel out the biases. If the desired precision is verified, the median measurement per point is taken as the ground truth. We then measure the same parts with our system, repetitively, using the technique described in Sec. 3.2. Based on the differences (errors) of two methodologies (CMM and ours), we train both a Nearest Neighbour and a Support Vector Machine [7] classifier. Our features are the concatenated $3D$ point coordinates. The responses (outputs) are then the differences of these measured coordinates to the collected training data i.e. coordinate-wise regressed offset from ground truth. Thanks to the acceptable mechanical precision, we could easily cover almost all the entire space of 3D variations.

3.2. Measurement

Our measurement stage follows 4 steps: Acquisition, extraction of geometries of interest, fitting 3D models and triangulation. The part is taken with with a linear axis conveyor belt, giving us a good initial condition. This section will focus on extraction of region of interests and fitting CAD models. Note that, in our setting, we are interested only in measuring the 3D center coordinates, but not the



(a) (b)

Figure 3. The Calibrator a) 3D model b) Manufactured. Sparse random dots (inter-reference points) exist in the real object, whereas they are absent in the model. They do not contribute to the extrinsic calibration, but serve as a guide for the optimization of the global reference frame. On the target part, instead of the calibration plates, lies the measurement points.

dimensions. Following the centroid determination, we always triangulate the spatial interest points to obtain the real 3D coordinates.

3.2.1 Geometry Extraction

As the goal is to measure both the geometric primitives and arbitrary edges, we design a generic geometry extractor, which simultaneously determines the edges of interest in multiple views. We will refer to every connected edge fragment, which cannot be represented by a single geometric primitive, as an *arbitrary contour*. With arbitrary contours, one cannot utilize prior information of geometry. We rather rely directly on the edges to obtain the matching structures across multiple views. Our procedure starts with a spatial sub-pixel accurate edge detection performed individually in all views of a local camera network, using a 3^{rd} order contour extraction technique [27]. The resulting edges are smooth, continuous and accurate up to $1/20^{th}$ pixel. Due to edge detection in individual frames, we do not have a representation of correspondence information. In general, obtaining correspondences across multiple images involve stereo matching algorithms, which are computationally expensive. Luckily, our system is calibrated and we are not interested in exact correspondences but rather in sets of consistent points across views (correspondence candidacy). In other words, we seek to find a set of segmented edges per each view. The real correspondences are then generated through the fitting algorithm in Section 3.3.

Multi-view Edge Segmentation Given a multiview setup, the essential matrix, relating view i to view j can be computed by $\mathbf{E}_{ij} = [\mathbf{t}_{ij}]_x \mathbf{R}_{ij}$, where $\mathbf{R}_{ij}$ is the relative rotation and $[\mathbf{t}_{ij}]_x$ is the relative translation between views i and j. Then $\forall \mathbf{p}^i, \exists \mathbf{l}^{i->j}_{p^i_k} : \mathbf{l}^{i->j}_{p^i_k} = \mathbf{E}_{ij} p^i_k$, where $\mathbf{l}^{i->j}_{p^i_k}$ is the epipolar line in view j, obtained by back-projecting k^{th} point in i^{th} view (p^i_k). From the epipolar constraint, $\mathbf{l}^{i->j} p^j_k = 0$ holds if p^j_k lies on $\mathbf{l}^{i->j}$. We will now derive an algorithm for a correspondence search in multiview images.

The idea is to iteratively transfer the edges from the first view to the last one, using the essential matrix. Each transfer step involves the intersection of the target frame's edges with the epipolar line, generated from the source edge point in the previous image. The third view possesses a unique intersection. This is often referred as the trinocular constraint [3]. For the rest of the views, the distance between the edge pixels and the optimum intersection is minimized. More formally, we start by obtaining the view with the least number of edge pixels and denote this view $C_0 = 0$. Then, let $\mathbf{p}^0 = \{p^0_i\}$ be all the subpixel edge pixels in this view. Then for a given pixel in view 0, p^0_i, it is possible to find a single correspondence hypothesis in view n in the following fashion:

$$x^n_i = \begin{cases} \arg\min_{p^n_j} \| (\mathbf{l}^{0\to n}_{p^0_i} \cap \mathbf{l}^{1\to n}_{p^1_k}) - p^n_j \|, & n = 2 \\ \arg\min_{p^n_j} \sum_{k=2}^{n-1} d(p^n_j, \mathbf{l}^{k\to n}_{x^k_i}), & n > 2 \end{cases} \quad (1)$$

x^n_i denotes the optimum intersection for p^0_i in view n. $d(\cdot)$ is the point to line distance in 2D space. This is indeed a recursive formulation, where $x^0_i = p^0_i$ and $\forall x^0_i, \exists \mathbf{x}^1_i : \mathbf{x}^1_i = p^1_j \in (l^{0->1}_{p^0_i} \cap \mathbf{p}^1)$. Multiple hypotheses, generated by any given stage of the algorithm are pruned by retaining only the closest intersection points. Similarly, only the intersecting edge points are transferred to the next view. Still, computing the intersection of all the edges with an epipolar line increases the search space and can result in undesired matches. Thus, we only match and transfer the subpixel contours respecting the consistency of intensity gradient in the direction of the epipolar line. Such consistency is achieved via the *epipolar gradient feature* [34]. Eventually, for each point p^j_0 a tracked trajectory is obtained if the point is visible in at least 3 cameras. This procedure is summarized in Fig. 4.

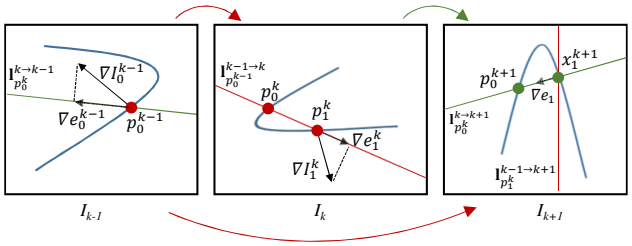


Figure 4. Edge pruning. The figure demonstrates a 3-view scenario: *The edges in the first image I_{k-1} (e.g. red point) are transferred to view I_k. The edge points of I_k lying on the epipolar line (red points) and which have consistent epipolar gradient features are selected and transferred to the I_{k+1}. The intersections of this new epipolar line with the edges of the I_{k+1} are selected in a similar fashion. In parallel, the epipolar line from I_{k-1} to I_{k+1} ($\mathbf{l_1^{k-1->k+1}}$) is also generated. At this point, the edges lying closest to the intersection are found to correspond in all 3 views (intersection of red and green lines). All the other correspondence hypotheses are discarded.* $\nabla\mathbf{e}$ *depict the epipolar gradients features.*

Edge Selection on CAD Models Typically, CAD geometry involves mesh edges, which do not correspond to physical structure. Therefore, we also process the CAD models to obtain image consistent edges. As in [32], synthetic views of CAD model are generated by projecting & rendering the model onto a 3-channel image, where each channel encodes one direction of the normal information, establishing the angle-magnitude relations. This image is then processed to select the image-visible-CAD-edges, using standard edge extraction techniques, e.g. Canny. Finally, a noise cleaning algorithm is conducted in the form of a connected edge segment filtering. The goal is to retain only the longest edge segment. Steger founded an edge linking algorithm suited for sub-pixel contours, which we directly use in our scenario [26]. Succeeding the sub-pixel edge linking, we filter the short edge segments, while merging the collinear ones to form the final curvilinear structures.

The steps as described in this entire section are depicted in Fig. 5. The output of this section is a set of 2D contours in distinct views, that are to be measured via CAD fitting. An image consistent 3D point cloud is also computed to be used in model fitting.

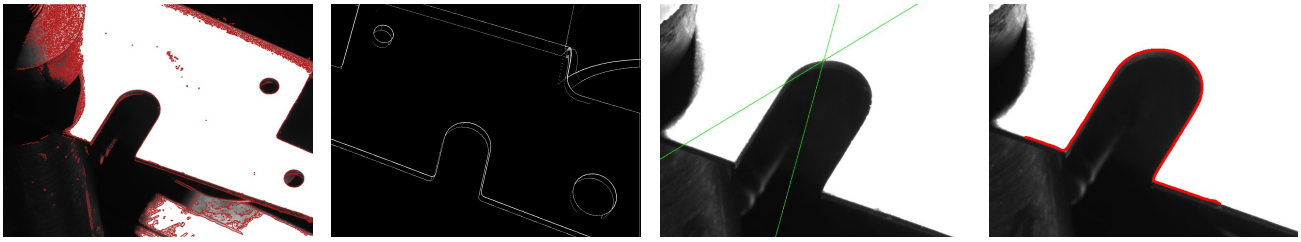
3.3. CAD Model Fitting and Triangulation

Even though the subpixel edge segmentation is reliable, the parts are almost always subject to severe noise due to cracks, crusts, surface defects, mis-manufacturing and calibration errors. Moreover, even if the edges found on multiple views constitute a reasonable cue, there is no clear method on how to reconstruct arbitrary contours (as the multiple-view correspondences are not apparent).

Therefore, we seek to find a robust method to relate the 2D geometric information with the CAD model, under the constraint that the registration respects the inspected 3D shape. To do that, we propose a novel projective variant of robust ICP. For the moment, we start by assuming that the inspected part goes through a rigid transformation, i.e. we handle non-rigidity in error compensation. Let $\mathbf{X}_k = \{X_{k1}^T, X_{k2}^T, ..., X_{kM_k}^T\}$ denote the points sampled around the measurement point k, on the partial 3D CAD model. $\mathbf{X}_k$ exists only for the cameras containing this partial model in their FOV. Thus, $\mathbf{X}^i = \{\mathbf{X}_k^i\}$ is a vector of edges of the measurement points being visible per each camera view i when projected. These points are obtained via sampling on the CAD model to desired resolution. The generated model point cloud is further subject to visible edge selection. Thus, number of CAD points differs per each view e.g. $\|\mathbf{X}_k^1\|_L \neq \|\mathbf{X}_k^2\|_L$. Similarly, let $\mathbf{p}^k = \{p_1^k, p_2^k, ..., p_{N_k}^k\}$ be the selected image edges in view k. We have $\mathbf{P} = \{\mathbf{p}^1, \mathbf{p}^2, ..., \mathbf{p}^K\}$, where K refers to the number of cameras capturing X_k. Due to the positioning on the conveyor belt, we assume that initial pose is within reasonable error bounds, i.e. the inspected part should remain fully within the view of the camera. We then formulate the CAD matching as the minimization of multiview re-projection error, as follows:

$$E(\theta) = \sum_{k=1}^{K} \sum_{i=1}^{M_k} w_{ki} \|\Omega(\mathbf{x}_{ki}, \mathbf{p}^k) - \mathbf{x}_{ki}\| \qquad (2)$$
$$\mathbf{x}_{ki} = Q(\theta_i, \mathbf{X}_i^k)$$

where $\theta = \{f_x, f_y, p_x, p_y, \mathbf{q}, \mathbf{t}\}$ are the standard pinhole projection parameters (intrinsic and extrinsic). We parametrize the rotations by the quaternion $\mathbf{q}$. Q denotes a perspective projection according to the pinhole camera model. For this step we avoid self-calibration of intrinsic orientation and only update the exterior pose parameters. Since fixing the intrinsics frees us from FOV adjustment, we could simply back-project the image points beforehand and work purely on normalized coordinates. This significantly decreases the computational load and speeds up the convergence as we only optimize over 7 parameters per camera. Ω is a closest point function, used to find the point p_j^k in view k, which has the shortest Euclidean distance to the projection of the 3D point X_i^k. w_k^j are the robust weights associated with the residuals and $\|\cdot\|$ is the Euclidean norm. Note that the visible CAD edges, and thus the synthesised model are unique per each rendered view. Naturally the multi-view correspondences are not available preventing us from simply referring to the bundle adjustment literature. For these reasons, we exploit an iteratively re-weighted least squares optimization where at each iteration we obtain the correspondences in the projected domain, while updating the parameters (and solve for the transformation) in full $3D$. The weighting is introduced to wane the effect of the outliers. This is already well studied in robust computer vision literature [20]. For our application, we select the weights w_k^j to



(a) Original View with Edges (b) Visible Edges of CAD Model (c) Epipolar Lines on Edge Point (d) Epipolar Edge Selection

Figure 5. Edge prunning. a) 1^{st} image of multi-view system. Subpixel edges are overlayed. b) View synthesized by edge selection on CAD model. c) Epipolar lines in 1^{st} image given the correspondents on other two. d) Selected edges due to multi-view epipolar intersection.

be Tukey's bi-weights as follows:

$$w(\hat{r}_i) = \begin{cases} \hat{r}_i(1-(\frac{\hat{r}_i}{c})^2)^2 & \text{if } |\hat{r}_i| \leq c \\ 0 & otherwise \end{cases} \quad (3)$$

where $\hat{r}_i = \frac{r_i}{\hat{s}}$ with r_i being the current residual and $c = 4.685$. $\hat{s} = \frac{mad(r_i)}{0.6745}$ is the estimated scale parameter, using $mad(r_i)$, the median absolute deviations. Such formulation inherently introduces outlier rejection, when the weights are set to 0. This problem becomes non-linear and we benefit from Levenberg Mardquardt solvers to directly minimize $E(\theta)$. A similar idea is already presented in the context of ICP by [8] and named *LM-ICP*. Thus, we refer to our solver as the *Robust Projective LM-ICP*. Our formulation can also be thought as an alternating minimization between the closest point assignment step and bundle adjustment.

Malassiotis *et al*. [19] uses a similar method for the case of stereo. However, they take into account the parametric nature of the CAD model and re-generate the model points in each iteration. This requires a parametric CAD geometry, decreases the well-posedness of the problem and increases the number of parameters to optimize. Also, it is not apparent in their method, whether the scale or anisotropic changes in the model points are due to the registration error or really stem from the model manufacturing. We find such setting restrictive. In contrast, by using less parameters, our method is efficient, could well operate on non-analytical CAD models and doesn't suffer from scale ambiguities.

To improve the convergence speed, we employ a coarse-to-fine scheme, in which the pose is refined in a hierarchy of down-samplings. We also use the idea of Picky-ICP [36], which establishes unique correspondences at each iteration. Finally, 3D measurement is concluded by a simple look up of the center position, of the registered and transformed CAD part. The computed 3D positions are corrected via error compensation procedure as described in Section 3.1.

4. Experiments and Results

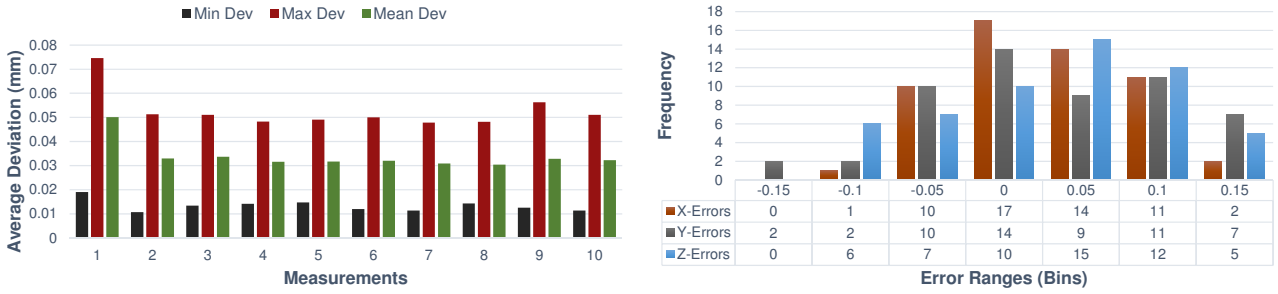
To demonstrate the validity and effectiveness of our approach, we run a series of experiments concerning calibration and measurement.

4.1. Calibration Accuracy

Intrinsic Calibration In all experiments, the intrinsic calibration errors (re-projection errors) were below 0.125 pixels. Considering the image resolution, such accuracy is enough for capturing as fine details as 10μ. Intrinsic orientation is one of the least significant sources of error in our system and is kept constant throughout all the experiments.

Extrinsic Calibration Our system relies not on the accurate production of the calibrator but on the measurement of reference points, placed on it. As described in section 3.1, 3 Canon 6D cameras are assembled and the calibration plates, as well as the overlapping inter-reference points (thanks to random dots) are measured with multiple shots and bundle adjustment. Due to the presence of optical inspection points, the manufactured calibrator cannot be measured with a CMM and we do not have any ground truth to compare against. Thus, we compare our measurements with ATOS, an industrial structured light triangulation system. As ATOS relies on fringe projection, it cannot recover the tiny details over complex structures (such as subpixel edges) accurately. But, it is very accurate in measuring reference points found on planar/smooth surfaces, such as the calibrator.

Fig. 6(a) presents a comparison between our system and ATOS in measuring the reference field composed of 180 reference points, on the calibrator. To be compatible with ATOS's precision requirements and to minimize the distortion effect, we use circular markers with a radius of 3 mm. The measurements are repeated 10 times, resulting in a total of 1800 measurements. We register ATOS and our system to the same coordinate frame. We accomplish this by a well known transformation estimation proposed by [13] using all the measured points. To this end, we assume that because both systems are sufficiently accurate, the errors in registration would have minimal disturbance to the coordinate frame alignment. It is noticed from the results that the registration error is small enough and the root mean squared error (RMSE) is 0.03mm. Every residual per coordinate contributes equally to RMSE. We accept this error



(a) (b)

Figure 6. **a)** 180 feature points on the calibration plates are measured with ATOS and with our system, 10 times. Using all the measurements, a 3D transformation is computed between the coordinate systems and the coordinate-wise difference in the common frame is recorded. It is shown that our system shows resemblance with ATOS system in the measurement of reference points ($\varepsilon_{max} < 0.1mm$). Note that, ATOS is good at measuring such points, as the fiducials lie on planar surfaces. Yet, it will later become ineffective in measuring complicated structures due to fringe projection drawbacks. **b) Accuracy Evaluations** Histogram of registration errors to the CMM coordinate space. Bins indicate the error amount, while frequencies correspond to number of registrations. The data is taken over 55 measurements. Note that, the errors are recorded after coordinate space registration between our system and common CMM frame. In this respect, the figure plots the accuracies relative to a CMM, which is the main measure of error in our system. Values are in milimeters.

and conclude that the calibrator is measured accurately and precisely to achieve the desired accuracy.

4.2. Measurement Analysis

CAD Registration The iterations of refinement for a local camera group are visualized in Fig. 7. Regarding the fitting of CAD model to arbitrary subpixel contours, the final RMSE registration error is typically in the range of 0.25 pixels. To quantify the numerical accuracy, we conducted a set of experiments on different types of model parts. Fig. 8(a) plots the convergence of our algorithm, while Fig. 8(b) shows the inliers retained after each iteration.

Accuracy (Errors in Comparison to CMM) Error compensation step involves repeated measurement and registration to CMM space in order to reduce the effect of systematic errors. CMMs report an error of 0.01mm for the points of interest and stand out to be the ground truth reference for our system [33]. The fact that we accept CMM results as a baseline for our measurements, make us interested in the deviations from this pseudo-ground truth. For this reason, we sample 55 random measurements of 5 different parts, which are measured both with our system and with Hexagon CMM. We then register each part from our coordinate space to CMM and record the errors. As this experiment is performed on the measured part but not on the calibrator, we end up with an evaluation of the accuracy of our system, with respect to the CMM. Fig. 6(b) shows a histogram over the collected data. It is shown that the maximum registration error is well below 0.2mm. This is both within the industry standards and within the tolerances we require.

Precision We will again refer to CMM comparisons, this time for precision. Fig. 9 enlists a series of experiments conducted on 5 different measurement points. In total 25 random measurements are displayed and the inter-measurement errors, which we will be referring as the repeatability measures are tabulated. Note that the maximum error remains significantly under 0.15mm, while appearing below 0.05mm most of the times.

The reported accuracy in the previous section is less than the precision values due to the undesired compensation er-

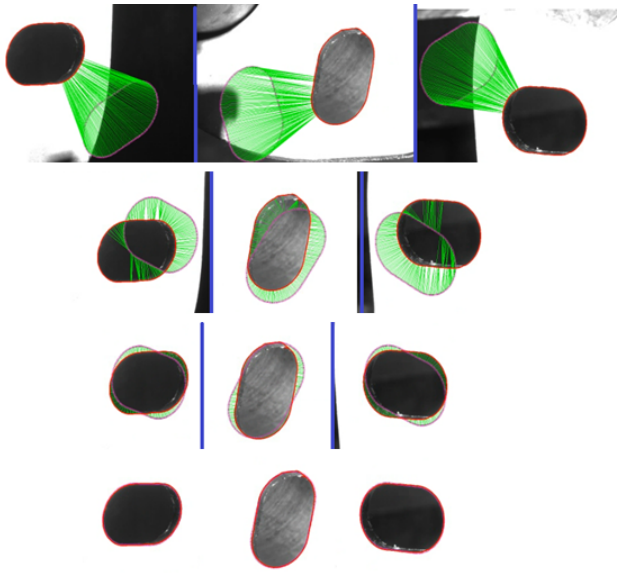
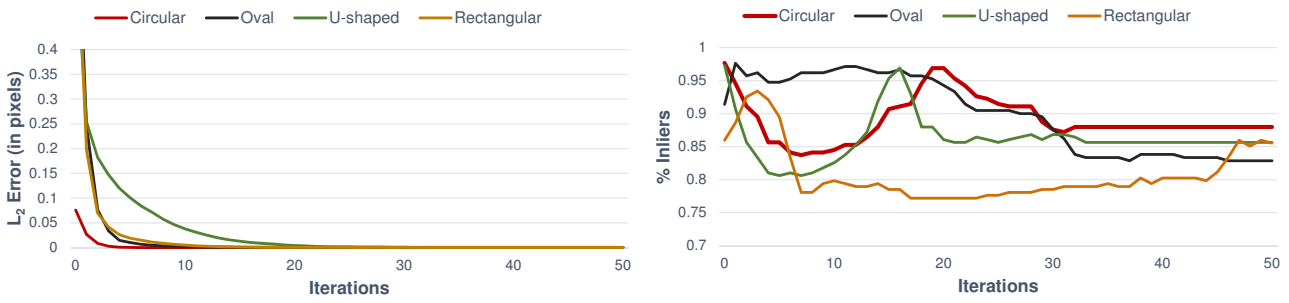


Figure 7. The process of CAD fitting: Red points indicate the found edges, purple points are the model projections, while the green points are the extracted edge pixels. Each column is a view and each row shows an iteration.



(a) Registration errors for different types of measurement points

(b) Inliers per iteration

Figure 8. Evaluation of Registration Performance: **a)** Each run is started with 2.5cm translational shift and 10°rotational offset. The registration error between the inliers is recorded at each iteration. Note that the convergence is super-linear in all types of geometrical structures. **b)** Ratio of Inliers over iterations.

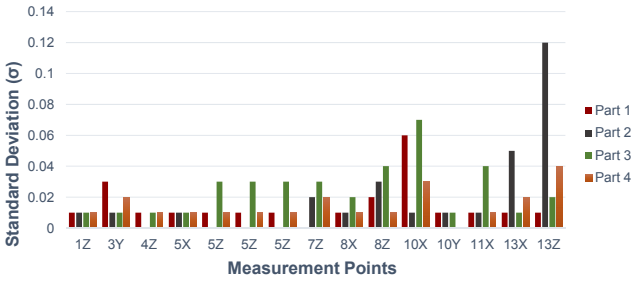


Figure 9. Repeatability results for 4 different parts, manufactured from the identical CAD model. For each part, 15 out of 26 points are measured, 5 times each (randomly taken out of 41). The standard deviation per point is plotted.

rors. In other words, the bias learned in Section 3.1 does not always represent the systematic components and is, to a certain extent, subject to random/unstructured noise. For this reason, while we are able to get measurements with maximum 0.15mm, we introduce certain errors in the registration to final coordinate space. Nevertheless, this error is not more than 0.05mm and is tolerable in our scenario.

4.2.1 Runtime Performance

Our system aims online operations, where the complete unit is installed on an operating production line. Therefore speed plays a key role in our design. Table 1 tabulates the timings for different stages of the runtime on a machine with 3GhZ Intel i7 CPU and 8GB of RAM. Note that image acquisition (30fps) step is responsible for most of the delays due to synchronous capture and bandwidth. Respectively, each camera captures the individual frames with a 0.5sec of strobe timing (until LEDs reach the desired intensity level). No processing starts until all the cameras finish the acquisition. Triangulation is only applied after every feature point is successfully extracted.

Our system is incomparably fast w.r.t. ATOS or CMM. It is, in that sense, not a replacement for a CMM device but a much more effective online competitor of the existing mechanical fixtures. Inspecting a single part took 1h when measured with ATOS and 45min on CMM. Both of these results are far from online 100% inspection.

Table 1. Average timings in measurement stage. Timings are reported as an average of 10 runs. MP refers to *measurement point*. Even though processing of the measurement points differ, their contribution to final runtime is averaged.

	Per MP (sec)	For All MPs (sec)
Image Acquisition	5,00	75,00
Feature Extraction	0,30	4,50
CAD Registration	0,15	2,25
Bundle Adjustment	0,10	1,50
	Total	**83,25**

5. Conclusions

We presented an in depth analysis of an accurate optical metrology system, designed to operate on industrial settings, where realtime performance and durability are of concern. To the best of our knowledge, this was the first work to address jointly the calibration and measurement processes, while at the same time providing novel algorithms targeting the fast, robust and accurate triangulation of different types of geometries. In this regard, even though the software requires care in implementation, the set-up of the system is not very complicated and can be customized without altering the software. It can be configured to operate on different parts with minimal effort. It is invariant to the physical system as much as possible and accounts for imprecise manufacturing and mechanical errors at the software level. We compared our system against the industry standard measurement devices, one mechanical and one optical. Our measurement results are well within the accepted tolerances. In the future work, we will evaluate self calibration methods and research on the error handling mechanisms, allowing the system to recover from severe misuses.

References

[1] S. Agarwal and K. Mierle. Ceres solver: Tutorial & reference. *Google Inc*, 2, 2012.

[2] J. Aulinas, Y. Petillot, J. Salvi, and X. Lladó. The slam problem: A survey. In *Proceedings of the 2008 Conference on Artificial Intelligence Research and Development: Proceedings of the 11th International Conference of the Catalan Association for Artificial Intelligence*, pages 363–371, Amsterdam, The Netherlands, The Netherlands, 2008. IOS Press.

[3] N. Ayache and F. Lustman. Trinocular stereo vision for robotics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1):73–85, 1991.

[4] F. Bergamasco, L. Cosmo, A. Albarelli, and A. Torsello. A robust multi-camera 3d ellipse fitting for contactless measurements. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pages 168–175. IEEE, 2012.

[5] P. J. Besl and N. D. McKay. Method for registration of 3-d shapes. In *Robotics-DL tentative*, pages 586–606. International Society for Optics and Photonics, 1992.

[6] J.-Y. Bouguet. Camera calibration toolbox for m atlab. 2004.

[7] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[8] A. W. Fitzgibbon. Robust registration of 2d and 3d point sets. *Image and Vision Computing*, 21(13):1145–1153, 2003.

[9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[10] J. Heikkila. Geometric camera calibration using circular control points. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1066–1077, 2000.

[11] L. Heng, B. Li, and M. Pollefeys. Camodocal: Automatic intrinsic and extrinsic calibration of a rig with multiple generic cameras and odometry. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 1793–1800. IEEE, 2013.

[12] G. Hermann. Geometric error correction in coordinate measurement. In *Acta Polytechnica Hungarica*, volume 4, pages 47–62, 2007.

[13] B. K. Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.

[14] K. Jyrkinen, M. Ollikainen, V. Kyrki, J. P. Varis, and H. Kälviäinen. Optical 3d measurement in the quality assurance of formed sheet metal parts. In *Proceedings of International Conference on Pattern Recognition*, volume 120, page 12, 2003.

[15] F. Kahlesz, C. Lilge, and R. Klein. Easy-to-use calibration of multiple-camera setups. In *Workshop on Camera Calibration Methods for Computer Vision Systems (CCMVS2007)*, Mar. 2007.

[16] R. Loser, T. Luhmann, and L. H. PMU. The programmable optical 3d measuring system pom-applications and performance. *International Archives of Photogrammetry and Remote Sensing*, 29:533–533, 1993.

[17] T. Luhmann. Close range photogrammetry for industrial applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6):558–569, 2010.

[18] T. Luhmann, F. Bethmann, B. Herd, and J. Ohm. Comparison and verification of optical 3-d surface measurement systems. *The international archives of the photogrammetry, remote sensing and spatial information sciences*, 37:51–56, 2008.

[19] S. Malassiotis and M. Strintzis. Stereo vision system for precision dimensional inspection of 3d holes. *Machine Vision and Applications*, 15(2):101–113, 2003.

[20] P. Meer. Robust techniques for computer vision. In *Emerging topics in computer vision*, pages 107–190, 2004.

[21] N. Mostofi, F. Samadzadegan, S. Roohy, and M. Nozari. Using vision metrology system for quality control in automotive industries. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1:33–37, 2012.

[22] S. Phillips, B. Borchardt, A. Abackerli, C. Shakarji, D. Sawyer, P. Murray, B. Rasnick, K. Summerhays, J. Baldwin, R. Henke, et al. The validation of cmm task specific measurement uncertainty software. In *Proc. of the ASPE 2003 summer topical meeting Coordinate Measuring Machines*, 2003.

[23] S. Sartori and G. Zhang". "geometric error measurement and compensation of machines". volume "44", pages "599 – 609", "1995".

[24] S. Savarese and L. Fei-Fei. Multi-view object categorization and pose estimation. In *Computer Vision*, pages 205–231. Springer, 2010.

[25] C.-T. Schneider and K. Sinnreich. Optical 3-d measurement systems for quality control in industry. *International Archives of Photogrammetry and Remote Sensing*, 29:56–56, 1993.

[26] C. Steger. An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(2):113–125, Feb. 1998.

[27] A. Tamrakar and B. Kimia. No grouping left behind: From edges to curve fragments. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8, Oct 2007.

[28] A. Tejani, D. Tang, R. Kouskouridas, and T.-K. Kim. Latent-class hough forests for 3d object detection and pose estimation. In *Computer Vision ECCV 2014*, volume 8694 of *Lecture Notes in Computer Science*, pages 462–477. Springer International Publishing, 2014.

[29] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *Robotics and Automation, IEEE Journal of*, 3(4):323–344, 1987.

[30] V. Tuominen. Cost modeling of inspection strategies in automotive quality control. *Engineering Management Research*, 1(2):p33, 2012.

[31] V. Uffenkamp. State of the art of high precision industrial photogrammetry. In *Third international workshop on accelerator alignment.*, 1993.

[32] M. Ulrich, C. Wiedemann, and C. Steger. Combining scale-space and similarity-based aspect graphs for fast 3d object recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(10):1902–1914, Oct 2012.

[33] N. Van Gestel. *Determining Measurement Uncertainties of Feature Measurements on CMMs (Bepalen van meetonzekerheden bij het meten van vormelementen met CMMs)*. 2011.

[34] E. Vincent and R. Laganiere. Matching with epipolar gradient features and edge transfer. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–277. IEEE, 2003.

[35] Z. Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334, 2000.

[36] T. Zinßer, J. Schmidt, and H. Niemann. A refined icp algorithm for robust 3-d correspondence estimation. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 2, pages II–695. IEEE, 2003.