Estimation of 3D reconstruction errors in a stereo-vision system  

Belhaoua A., Kohler S., Hirsch E.  

aLSIIT, Laboratoire des Sciences de l’Image, de l’Informatique et de la Télédétection UMR 7005, Pôle API, Boulevard Sébastien Brant, BP 10413 - 67412 Illkirch Cedex, France;  
bMIPS, Laboratoire Modélisation Intelligence Processus Systèmes EA 2332, 12, rue des frères Lumière 68093 - Mulhouse Cedex France  

ABSTRACT  

The paper presents an approach for error estimation for the various steps of an automated 3D vision-based reconstruction procedure of manufactured workpieces. The process is based on a priori planning of the task and built around a cognitive intelligent sensory system using so-called Situation Graph Trees (SGT) as a planning tool. Such an automated quality control system requires the coordination of a set of complex processes performing sequentially data acquisition, its quantitative evaluation and the comparison with a reference model (e.g., CAD object model) in order to evaluate quantitatively the object. To ensure efficient quality control, the aim is to be able to state if reconstruction results fulfill tolerance rules or not. Thus, the goal is to evaluate independently the error for each step of the stereo-vision based 3D reconstruction (e.g., for calibration, contour segmentation, matching and reconstruction) and then to estimate the error for the whole system. In this contribution, we analyze particularly the segmentation error due to localization errors for extracted edge points supposed to belong to lines and curves composing the outline of the workpiece under evaluation. The fitting parameters describing these geometric features are used as quality measure to determine confidence intervals and finally to estimate the segmentation errors. These errors are then propagated through the whole reconstruction procedure, enabling to evaluate their effect on the final 3D reconstruction result, specifically on position uncertainties. Lastly, analysis of these error estimates enables to evaluate the quality of the 3D reconstruction, as illustrated by the shown experimental results.  

Keywords: Quality Control, Stereo-vision, Segmentation, Error analysis, CAD Data, Planning, SGT  

1. INTRODUCTION  

In order to favor usage of industrial applications of computer vision, such as quality control and accurate measurement tasks leading to quantitative inspection of manufactured parts, it is necessary to develop fully automated tools for the accurate computation of 3D descriptions of the object of interest out of the image contents. The latter can then be compared with either ground truth or the CAD model of the object under investigation, in order to assess its quality.  

In our institution, an autonomous cognitive vision system is currently being developed for the optimal 3D reconstruction of manufactured parts, based on a priori planning of the task and built around a cognitive intelligent sensory system using so-called Situation Graph Trees as a planning / control tool. The planning system has been applied to structured light and stereo-vision based 3D reconstruction tasks, the aim being to develop an automated quality control system for manufactured parts evaluating quantitatively their geometry. This requires the coordination of a set of complex processes performing sequentially data acquisition, its quantitative evaluation (i.e., extraction of geometric features and their 3D reconstruction), and the comparison with a reference model (e.g., CAD model of the object) in order to evaluate quantitatively the object. Stereo-vision is recognized to be one of the techniques enabling to solve such tasks and is widely used for obtaining the required 3D information after having processed a pair of images. In this case, the outline of the object can be determined, based on the set of contours defining this outline. As a result, such a contour based approach provides good
(i.e., reasonable) reconstruction quality at low cost. The resulting accuracy of the three-dimensional position information is a crucial point for quality control tasks. Accordingly, accuracy has to be optimized and, in fact, the quality of the 3D reconstruction depends not only on the quality of the acquired images but also of their processing (e.g., segmentation, feature extraction, matching and reconstruction).

Over the last few years, some efforts have been spent on error analysis in stereo-vision based computer vision systems.\(^5\) As an example, Blostein and Huang\(^5\) have investigated the accuracy in obtaining 3D positional information based on triangulation using point correspondences derived using a stereoscopic camera setup. They have derived closed form expressions for the probability distribution of position errors along each direction (horizontal, vertical and range) of the coordinate system of the stereo rig. Also, a study of different types of error and their effects on 3D reconstruction results obtained using a structured light technique has been presented by Yang \textit{et al.}\(^6\) In their work, Yang \textit{et al.} have derived expressions for the errors observed for the 3D surface position, the orientation and the curvature measurements. Further, Ramakrishna \textit{et al.}\(^7\) proposed a new approach for estimating tight bounds on measurement errors, considering the inaccuracies introduced during calibration and triangulation. Balasuramanian \textit{et al.}\(^9\) analyzed the effect of noise (which is assumed to be independent and uniformly distributed) and of the geometry of the imaging setup on the reconstruction error for a straight line, their analysis being mainly based on simulation studies. Revira-Rios \textit{et al.}\(^10\) have analyzed the error when measuring dimensionally line entities, these errors being mostly due to localization errors in the image planes of the stereo setup. Consequently, in order to determine optimal camera poses, a non-linear program has been formulated, that minimizes the total MSE (Mean Square Error) for the line to be measured, while satisfying sensor related constraints. Lastly, the accuracy of 3D reconstructions has been evaluated through comparison with ground truth in contributions presented by Park \textit{et al.}\(^11\) and Albouy \textit{et al.}\(^12\) More recently, Jianxi \textit{et al.}\(^13\) have presented an error analysis for 3D reconstruction taking into account only the camera calibration parameter accuracy.

In all cases, the presence of noise in the acquired data affects the accuracy of the subsequent image processing and, thus, of the reconstruction process. In contrast to the above mentioned studies, our analysis focuses on error estimation for the segmentation process, the starting step for the whole reconstruction procedure. Using fitting technique, error bounds are established for each (geometric) feature that composes the object to be evaluated. The error bounds are then propagated up to the final 3D reconstruction step. This error analysis for stereo-vision based reconstruction tasks will then help to evaluate the quality of the 3D reconstruction. The resulting final error estimates will then enable to state if the reconstruction results fulfills \textit{a priori} defined criteria, including tolerance information.

The outline of the contribution is as follows. Section 2 describes the configuration of the stereo-vision system used in our work and identifies accordingly the sources of error. Section 3 introduces the related image processing and associated errors. In section 4, using the error models, experimental results are presented and discussed. Finally, Section 5 concludes the paper and provides a short outlook.

\section*{2. FRAMEWORK OF THE APPROACH}

In this section, firstly, we describe shortly the cognitive system built around two cooperative modules, an off-line planning procedure defining the whole application and an on-line module organizing and controlling the processing of the acquired data. Secondly, we point out various error sources likely to be observed in the various steps for the whole process.

\subsection*{2.1 System architecture}

Figure 1.a gives the block-diagram of the autonomous vision system built around an intelligent component using Situation Graph Trees (shortly SGTs)\(^1\) as a planning /control tool. This component is able to manage and to adjust dynamically and on-line the acquisition conditions (i.e., camera positions and lighting source parameters), leading eventually to re-planning of the process, in order to improve the quality of the image data required by the 3D reconstruction step. In particular, treatments and reasoning about local image contexts are favored when dynamic re-planning is required after failure of a given processing step. This planning tool has been applied to stereo-vision based 3D reconstruction tasks in order to extract optimally the contour data required by the reconstruction step. Figure 2 describes schematically the processing steps, from calibration of the stereo
The quality (e.g., accuracy) of the 3D reconstruction depends both on the quality of the acquired images and on the subsequent processing. Unfortunately, errors due to segmentation are propagated up to the reconstruction step. Therefore, these errors have to be estimated in order to evaluate the quality of the 3D reconstruction used to assess the workpiece under evaluation.

2.2 Measurement uncertainties and sources of error

Uncertainty is a critical issue when evaluating the quality of a 3D reconstruction. In our case, the errors in obtaining 3D information using stereo-vision are estimated using expressions defined in terms of system parameters and error sources likely to be observed. In addition to this quantitative determination of errors, other likely error sources affecting system modeling and calibration are also considered. In this section, as a result, we identify three major types of error sources for the system under investigation.

2.2.1 Camera model errors

As mentioned above, the sensor consists of two CCD cameras rigidly mounted on an arm. These latter are at the origin of three common types of noise, i.e., random noise, fixed pattern noise, and banding noise. Random noise can be characterized by intensity fluctuations in the acquired images and is always observed in the image.
Fixed pattern noise generally appears when using rather long exposure times and is usually exacerbated when temperature increases. Banding noise is highly camera-dependent and is introduced by the camera when reading data from the digital sensor. These latter two types of noise can usually be neglected. Lastly, in our case, an ideal pinhole camera model is used with optics of high quality. One can thus ignore camera lens distortion and other optical nonlinearities.

### 2.2.2 Image processing errors

Results provided by the stereo-vision system depend heavily on the quality of the 2D features extracted by edge detection algorithms and processing is affected by various kinds of errors. As far as the acquisition conditions are concerned, illumination conditions may decrease the quality of the edge detection procedure following image acquisition, due to shadow or glare effects. Another kind of error arises from the sampling process leading to the digital image. As a result, image point coordinates may differ from their true values because the sampling process restricts image pixels to lie on an integer grid. Furthermore, many edge detection schemes rely on thresholds to decide whether an edge or contour point exists or not. Therefore, localized edge detection algorithms intrinsically exhibit errors, as they can fail to detect an edge, which in fact is present in the image, or detect false edge.

### 2.2.3 Camera calibration errors

Camera calibration is a necessary step in stereo-vision in order to extract quantitative and metric 3D information from the set of two 2D images. Various types of error occur, due to inappropriate calibration of the imaging setup. These errors, which depend on both the camera set-up (e.g., distortion of the grid pattern, lens distortion, etc) and calibration algorithm, can affect the correspondence procedure (i.e., the matching step of 2D image points from the two images and belonging to the same 3D scene point) and thus the triangulation step (i.e., the determination of 3D coordinates using a matched pair of two image points, one from each image of the stereo pair).

### 3. EVALUATION OF THE SEGMENTATION STEP AND ASSOCIATED ERRORS

Edge or contour point detection relies in our case on a gradient based-method convolving the image with first derivatives of a Gaussian smoothing kernel. A decision is then made as to whether a pixel belongs to an edge or not, based on convolution results. Since the aim is to develop a fully automated system (i.e., with very restrictive human control), the parameters that control the edge detection process (e.g., the width \( \sigma \) of the Gaussian smoothing kernel and the threshold parameter values) are determined automatically. The \( \sigma \) parameter value is determined by the amount of camera noise observed in the image and by the fine-scale texture of specific object surfaces seen in the image. In our current implementation, a fixed value (\( \sigma = 1 \)) is used as a starting value. Chains of detected contour points are subsequently grouped together to form possibly closed contours. Each contour is then further subdivided either into straight-line segments or elliptical arcs using the method described in. These geometric features build the base of the outline of the imaged objects particularly or our test workpieces.

However, these simple geometric features do not contain necessarily only true edge points. This leads to uncertainties for the set of edge point positions belonging to a given geometric feature, which are further reflected in the parameter values of the fitting equation used to describe the data. Despite this, the line segment and curve descriptions can be determined using more or less standard fitting technique for the edge points supposed to be on these lines or curves. For that purpose, linear or non-linear least-squares fitting techniques are widely used. These latter minimize the sum of squared errors in predefined figures of merit. When using geometric fitting, named also the 'best fitting' technique, the error distance is defined by the orthogonal, or the shortest, distance of a given point to the geometric feature to be fitted. The geometric fitting for a line segment is a linear problem, whereas the geometric fitting of ellipses is a non-linear problem, which has to be solved iteratively. Gander et al. have proposed a geometric ellipse fitting algorithm in parametric form, which involves a large number of fitting parameters (\( m+5 \) unknowns in a set of \( 2m \) equations, where \( m \) is the number of measurement points), and where each measurement point carries an individual angular parameter to be estimated simultaneously, together with the five ellipse parameters.
Also, et al.\textsuperscript{16} discussed the relationship between the random perturbations of edge point positions and the variance of the least squares estimates of the corresponding line parameters. In their analysis, the noise is assumed to be independent and uniformly distributed. As the outline of our test workpieces is composed of simple geometric features such as lines and elliptical arcs, one can determine descriptors of these lines or arcs using a fitting technique similar to the ones described above, in order to obtain the parameters of these features. In this paper, we have extended Yi’s results to analyze the edge detection error for both straight line segments and elliptical arcs. The parameters of the fitted geometric features are finally used as a quality measure to determine confidence interval for the following matching / reconstruction procedure.

3.1 Fitting of straight line segments

It is well known that the least squares line parameters can be derived statistically as the line of maximum likelihood, assuming that the distribution of the points belonging to the line is normal. When this assumption is not fulfilled, one can apply a weighted least squares method.

A line in a plane $xy$ can be described by the following equation:

$$Ax + By + C = 0$$  \hspace{1cm} (1)

Equation 1 represents all geometric lines in the plane, including verticals ($B=0$) and horizontals ($A=0$). Also, it ensures finiteness of the moments of the estimates of the parameters, and helps to secure numerical stability in practical computations. The orthogonal fitting line can be found by minimizing the following criterion:

$$\mathcal{F}(a, b) = \frac{1}{A^2 + B^2} \sum_{i=1}^{n} (Ax_i + By_i + C)^2$$  \hspace{1cm} (2)

In our case, we have used the singular value decomposition (SVD) approach to find the parameters $A$, $B$ and $C$.

3.2 Fitting of elliptical arcs

To compute ellipse equations from given data point sets, various methods are discussed in.\textsuperscript{15,17,18} Gander et al.\textsuperscript{15} propose a geometric parametric ellipse fitting algorithm, as described above. Their method achieves a good compromise between accuracy of the results and necessary computing time. Solving this non-linear least squares problem relies here also on minimizing the sum of squares of the distance of the data points to the ellipse. In this paper, we have applied this method to our images.

4. RESULTS

One of the aims of this work is to quantify the errors for 3D reconstruction results. These latter depend on the quality of the acquired images and their processing. To validate our approach, we have used a quasi-polyhedral test object of known size (an example of which is shown on figure 3.a). The images of this object are composed of straight lines, arcs and ellipses (resulting from the perspective projection of circles from the object onto the image plane). Figures 3.a and 3.c show an example of a pair of images of the object acquired respectively with the right and the left camera of the stereo rig.

For each line or curve of the object, the corresponding contour points have been extracted using Canny’s edge operator, and, after determination / classification of contour point lists, we have applied the fitting technique described in sections 3.1 and 3.2 to these edge points supposed to belong to either a line or a curve. However, as stated above, the edge point positions are always affected by uncertainty, due to the image digitization process and noise in the system, and the non-ideal behavior of the image processing applied. Assuming that the errors in the images are independent and identically distributed, Yi et al.\textsuperscript{16} analyzed, for straight lines, how edge point position uncertainty is propagated through to the uncertainties of the fitting parameters. These uncertainties are statistically expressed as "confidence interval”.

A so-called confidence interval refers to the range of parameter values containing the limits or bounds of the parameter values, the interval being associated with a confidence level, which guarantees that the bounds are such
as to define an interval that contains surely the true parameter value. The parameter value limits define this way a range of values, lower bounded by a minimum parameter value (or lower confidence limit) and upper bounded by a maximum parameter value (or upper confidence limit), such that one can be confident (with a pre-specified level of confidence, e.g. such as 95% or 99%) that the true parameter value is within the confidence interval. The confidence intervals can be computed using the fitting parameters specifying the equations describing the geometric features making up the outline of the manufactured piece under evaluation. These confidence intervals are further propagated up to the triangulation procedure, achieving 3D reconstruction. The final 3D errors are then estimated applying an error propagation scheme.

Figures 3.b and 3.d show the results of the segmentation and classification steps, after having segmented the lists of contour points provided by the contour point extraction step. A same color is used in the two pictures to show the geometric features that have been associated, i.e. matched. These pairs of features are the input data for the reconstruction step. Figure 4 shows the computed 2D error distributions for different types of contours, composing the object of interest and extracted from both the right and left images. The error itself is defined as the distance between a given edge point of the contour point list defining a given feature and the corresponding geometric feature described using the feature parameters resulting from the fitting procedure. It has to be understood that the results should produce error distributions with zero mean values, as shown on figure 4.

The triangulation-based reconstruction method applied in this work relies on standard stereo-vision principles. After having determined pairs of corresponding points (belonging here to matched geometric primitives), they are mapped into the 3D world using the known geometry of the stereo system, as determined during the (off-line) calibration step. The correspondence problem is solved using matching functions based on the so-called epipolar constraint (see^2). As a result, if both the intrinsic and extrinsic parameters of the camera set-up are known, the reconstruction problem can be solved easily applying a simple procedure known as triangulation.

As a pre-requisite for applications (such as inspection or quality control tasks), the estimation of the error
Fig. 4. Error distributions for various types of matched geometric features in the right and left images of a representative test piece.

affecting the 3D reconstruction is essential in order to be able to evaluate quantitatively the dimensions of the object or to state if the reconstruction results fulfill tolerance rules. Further, these error estimates enable also to qualify the behavior of the vision system. With this objective, we firstly evaluate the accuracy / quality of the reconstructed contours using a linear orthogonal regression procedure in 3D-space based on a principal components analysis. Figure 5 illustrates the resulting 3D error distributions in the X, Y and Z directions of a reference system (here the camera coordinate system). The mean errors $\mu_x, \mu_y, \mu_z$ and variances $\sigma_x, \sigma_y, \sigma_z$ are also indicated, building the basic descriptors of the computed distributions. As can be seen on figure 5, these distributions are approximately normal, with a mean value of zero and a variance which varies in the range (0.02, 0.52) mm. As could be expected (the behavior being routinely observed for stereo-vision based reconstruction procedures), the error in the Z direction (i.e., range) is much higher than that in the X and Y directions (more or less, the directions of the coordinate axes of the image plane). This behavior has been very recently proved by Jianxi et al.13

Lastly, we have also propagated the 2D errors observed for edge points through to the 3D reconstruction computation step, in order to estimate the final 3D errors. In 3D space, these errors can be represented by ellipsoids bounding the reconstructed points. Figure 6 shows the 3D error distributions in the X, Y and Z directions after propagation of the 2D errors. Again, we can observe that the error in the Z direction is much higher than in the other two directions. Finally, Figure 7 shows examples of reconstructed 3D vertical contours after application of the triangulation procedure. The 3D error bounds, represented as bounding ellipses, are sketched in red.

5. CONCLUSION AND OUTLOOKS

In this contribution, the performance and the uncertainties of a cognitive stereo-vision system have been evaluated and quantified, specifically the uncertainty estimation related to the triangulation based 3D reconstruction problem. As a matter of fact, the triangulation process is sensitive to many kinds of error, a summary
Fig. 5. 3D error distributions in the X, Y and Z directions of a typical 3D reconstructed contour.

(a) 3D error distribution in the X direction
(b) 3D error distribution in the Y direction
(c) 3D error distribution in the Z direction

Fig. 6. Propagated 3D error distributions in the X, Y and Z directions of a typical 3D reconstructed contour.

(a) Propagated 3D error distribution (X direction)
(b) Propagated 3D error distribution (Y direction)
(c) Propagated 3D error distribution (Z direction)

Fig. 7. Examples of 3D reconstructed vertical contours (Perspective view) and associated error bounds represented as ellipses.
being provided in this paper. As a result, three major types of error have been identified, namely errors due to the image formation process (i.e., camera model used), errors related to image processing, and camera calibration errors. In this contribution, we mainly considered the source of errors due to the image segmentation procedure.

Assuming that the various errors, specifically those related to image processing, are independent and identically distributed, the errors of each processing step are evaluated independently. In particular, we have estimated the segmentation errors, using fitting results of contour point lists, in order to determine so-called confidence intervals. This latter are computed for each geometric feature that composes the object of interest. The confidence intervals have then been propagated through the whole chain of treatments of the images acquired with the stereo-vision system, using an error propagation scheme. First experimental results concerning this method of error estimation have been presented, specifically for the segmentation steps. The observed quantified error distributions are significant and, accordingly, validate the approach. Finally, the error of the stereo-vision system is estimated, taking, e.g., tolerance specifications into account.

With that objective, in order to evaluate the quality of the 3D reconstruction and to estimate its accuracy, we have applied a 3D fitting technique to each reconstructed geometric feature, using a linear orthogonal regression approach in 3D space based on principal components analysis. As a result, as expected, we have observed that the error in the Z direction is much higher than in the X and Y directions.

Finally, propagation of these various errors up to the final 3D reconstruction step, in order to compute a global and unique error value, taking into account all error sources, has been carried out. In particular, we have estimated the 3D reconstruction error, which can be represented as bounding ellipses positioned around the reconstructed 3D points. This helps to quantify on one hand the accuracy of the reconstruction and, on the other hand, to decide whether the reconstruction result fulfills tolerance rules or not. Related work is performed in order to reduce the 3D reconstruction errors as much as possible, by adjusting dynamically the image acquisition conditions, specifically the illumination conditions. As an extension to the work presented, all sources of errors, such as matching and calibration errors, will be taken into account.

REFERENCES


