



Facoltà di Ingegneria

Dipartimento di Ingegneria Elettronica ed Ingegneria Informatica

Dottorato di Ricerca in Ingegneria dell'Informazione
X Ciclo – Nuova Serie

TESI DI DOTTORATO

Novel models and methods for structured light 3D Scanners

CANDIDATO: **MAURIZIO BEVILACQUA**

TUTOR: **ING. ALFREDO PAOLILLO**

COORDINATORE: **PROF. ANGELO MARCELLI**

Anno Accademico 2010 – 2011

To my grandparents,

and my uncle.

Thanks to all who were close to me during these three years. Thank you for believing in my abilities and for supporting me in difficult times.

Indice

INTRODUCTION.....	9
CHAPTER 1.....	11
INTRODUCTION TO THE 3D SCANNING SYSTEMS AND APPLICATION	11
1.1 INTRODUCTION.....	11
1.1.1 Contact systems	12
1.1.2 Contactless systems	15
1.2 APPLICATIONS.....	18
1.2.1 Reverse engineering	18
1.2.2 The Digital Michelangelo Project.....	19
1.2.3 Digital Hammurabi: scanning of cuneiform tablets	21
CHAPTER 2.....	23
CLASSICAL MODEL FOR THE CALIBRATION OF 3D STRUCTURED LIGHT SCANNING SYSTEMS.....	23
2.1 INTRODUCTION.....	23
2.1.1 Structured light profilometer	23
2.1.2 Calibration.....	24
2.2 DIGITAL PATTERN PROJECTION SCANNER	25
2.2.1 Rusinkiewicz implementation	25
2.2.2 ICP.....	31
2.2.3 Evaluation of the implemented scanner	33
2.2.4 Drawback.....	37
2.3 ANALOG PATTERN PROJECTION SCANNER	37
2.3.1 Fringe pattern projection and phase shift	37
2.3.2 Classical calibration techniques	40
2.3.3 Drawback.....	52
2.4 LIMITATIONS OF CLASSICAL SYSTEMS	57
2.4.1 Calibration one step	57
2.4.2 Insensitivity to environmental noise	58
2.4.3 Reduction of uncertainty and faster recalibration	58
CHAPTER 3.....	61

NOVEL CALIBRATION TECHNIQUES FOR 3D STRUCTURED LIGHT	
SCANNING SYSTEMS.....	61
3.1 INTRODUCTION.....	61
3.1.1 <i>Pin hole camera model</i>	61
3.1.2 <i>Limits of the application to the projector</i>	65
3.2 STEREO MODEL FOR CAMERA-PROJECTOR PAIR	66
3.2.1 <i>Projector model</i>	67
3.2.2 <i>Calibration algorithm</i>	69
3.2.3 <i>DLT</i>	75
3.3 STEREO CALIBRATION IN A FRINGE PATTERN PROFILOMETER	77
3.3.1 <i>Simulated phase maps</i>	77
3.3.2 <i>Insensitivity to environmental noise</i>	80
CHAPTER 4.....	83
ADVANCED MODELS FOR RECONFIGURABLE SCANNER	
83	
4.1 INTRODUCTION.....	83
4.1.1 <i>How to reduce the uncertainty?</i>	84
4.1.2 <i>Introduce the lens distortion parameters</i>	84
4.1.3 <i>Faster recalibration</i>	84
4.2 LENS DISTORTION PARAMETERS	84
4.2.1 <i>Classical approach to the problem</i>	84
4.2.2 <i>Model for the variation of lens distortion parameter</i>	89
4.3 ADVANCEND GEOMETRIC CAMERA MODEL	91
4.3.1 <i>Pinhole model</i>	94
4.3.2 <i>Axial model</i>	94
4.3.3 <i>General Model</i>	95
4.3.3 <i>Modeling of distortion parameters trend by the Axial model</i>	96
4.4 SELF CALIBRATION	98
4.4.1 <i>Self-calibration procedure</i>	98
4.4.2 <i>The set of parameters</i>	99
4.4.3 <i>The objective function: the average epipolar error</i>	101
CHAPTER 5.....	105
TEST FOR THE VALIDATION OF THE NEW CALIBRATION	
ALGORITHMS.....	
105	
5.1 INTRODUCTION.....	105
5.1.1 <i>Assumptions about the model of the projector</i>	106
5.2 CHARACTERIZATION OF THE CALIBRATION	108
5.2.1 <i>Test station</i>	108
5.2.2 <i>Comparison of calibration techniques</i>	109
5.2.3 <i>Re-projection error</i>	111
5.2.4 <i>Real object reconstruction</i>	113

5.3	ADVANCED GEOMETRIC CAMERA MODEL VALIDATION.....	115
5.3.1	<i>Synthetic images creation</i>	<i>116</i>
5.3.2	<i>Comparison of distortion parameters trends</i>	<i>116</i>
5.4	SELF CALIBRATION	119
5.4.1	<i>The epipolar error surfaces.....</i>	<i>119</i>
5.4.2	<i>Simulation results.....</i>	<i>121</i>
5.4.3	<i>Comparison with Camera-Projector calibration</i>	<i>125</i>
5.4.4	<i>Comparison with the 8-points algorithm</i>	<i>127</i>
CONCLUSION		129
REFERENCES.....		131

Introduction

Introduction

The work made during the PhD course in Information Engineering, was focused on the possibility to find out novel techniques for the quick calibration of a cheap 3D Scanner. It is based on a simple camera and a commercial projector, in order to develop low-cost devices with high reliability capable to quickly acquire large areas.

Many systems based on this configuration exist, which have benefits and disadvantages. They can acquire objects with a large surface in a few seconds and with an adequate accuracy. On the other hand, they need a lengthy calibration and they are very sensitive to the noise due to the flicker of the light source. Considering these problems, I tried to find new robust calibration techniques in order to reduce the sensitivity to noise, and, in this way, to have high-performance low-cost 3D scanners with short-time calibration and reconfiguration.

There are many calibration techniques available for these systems. First, it is necessary to calibrate the camera and then the overall system for projecting analog encoded patterns, typically sinusoidal or digital, such as Gray codes. These techniques are very time-consuming because they require a prior camera calibration phase separate from the calibration of the whole system and also disturbing factors are introduced by the ambient light noise. Indeed, a lot of projection patterns, used in mapping the calibration volume, are required to be projected.

In order to achieve our goal, different types of structured light scanner have been studied and implemented, according to the schemes proposed in literature. For example, there exist scanners based on sinusoidal patterns and others based on digital patterns, which also allowed the implementation in real time mode. On these systems classical techniques of calibration were implemented and performance were evaluated as a compromise between time and accuracy of the system.

Introduction

Classical calibration involves the acquisition of phase maps in the volume calibration following a pre-calibration of the camera. At first, an algorithm that allows calibration through the acquisition of only two views has been implemented, including camera calibration, modeled by pin-hole model, in the calibration algorithm. To do this, we have assumed a geometric model for the projector which has been verified by the evaluation of experimental data. The projector is then modeled as a second camera, also using the pin-hole model, and we proceeded with the calibration of camera-projector pair as a pair of stereo cameras, using a *DLT* calibration. Thanks to the acquisition of two views of the target volume in the calibration, it is possible to extract the parameters of the two devices through which the projected pattern can be generated. Thus the acquisition by the camera can be performed overpassing the problem of noise introduced by the ambient light.

This system result to be a good compromise between calibration time and uncertainty. The former is reduced from half an hour to a couple of minutes, whereas uncertainty is reduced according to one percentage point of calibration volumes (that were chosen of a depth of 10 centimeters).

The percentage errors could be reduced by considering the lens distortion. During the period spent as a visiting student at the Machine Vision Group of Oulu in Finland, under the supervision of Prof. J. Heikkilä, problems related to the change of parameters of distortion in a pin-hole model as a function of distance between target and camera have been studied. After several experimental tests, first in simulation and later on real images, it has been concluded that the parameter variations can be justified with the use of a simple model such as pin-hole. The use of advanced geometric models like Axis model and subsequently the Generic model, can incorporate these variables within the model itself, to make the camera and projector calibration more accurate when extended to a larger volume range.

Finally, Self-Calibration for Stereo system has been used for the 3D scanner system in order to reconfigure on-site, i.e. outside the laboratory, a scanner previously calibrated in the laboratory.

Chapter 1

Introduction to the 3D scanning systems and application

1.1 Introduction

It is important in many and various applications have digitalized tridimensional models that represent the shape and the color features of real objects. The classical techniques of computer-aided design (CAD) were been accompanied, in the last years, by a methodology called *3D scanning* [1].

The automatic object shape acquisition technique, used in a lot of environments, are as multiple and varied as their applications are.

A conceptual scheme is illustrated in Figure 1.1

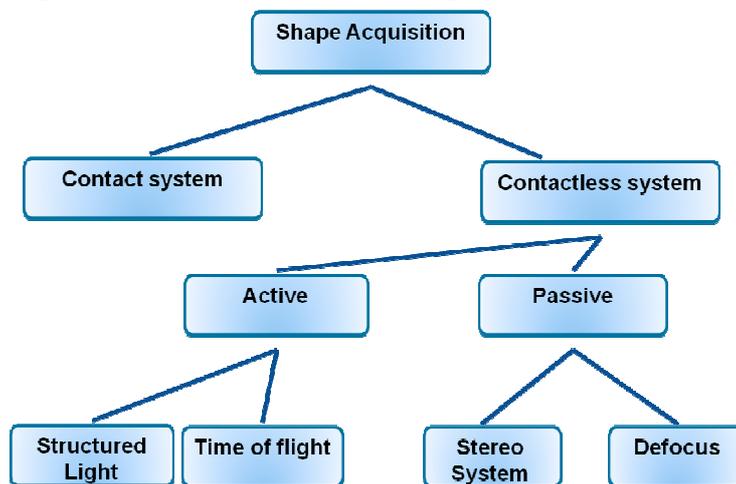


Figure. 1.1 : Classification of the major profilometer systems

1.1.1 Contact systems

The coordinate measuring machines (CMM) are the most important systems that allow the acquisition of 3D shape by touching the surface which is to be acquired.

They work using analogical solid probe of accurately known diameter that, acquire and store the coordinate of touched points based on the deflection of the same.

Due to their extreme accuracy, they are used very commonly for the detection of the surface profiles in the industry. On the contrary, however, the CMM have some disadvantages such as the slowness and the high cost, the limitations in the acquisition of entire surfaces, and mainly of object of very irregular shapes.

Less than 60 minutes are required for the installation of machinery, however these times are frustrated by the long time of acquisition, which grows in relation to the size and complexity of the shape of the object under measurement.

Furthermore, the application of a force on the surface of the object to be studied could ruin the part in the case of soft materials or, in the case of flexible materials, alter the measurement. It is clear how this type of technique can be invasive in the case where the object to be detected should be an ancient find difficult to handle because of its fragility.

The forces applied in these cases range from 1 to 3 N, the tolerances and repeatability are typically about ± 10 micrometers.

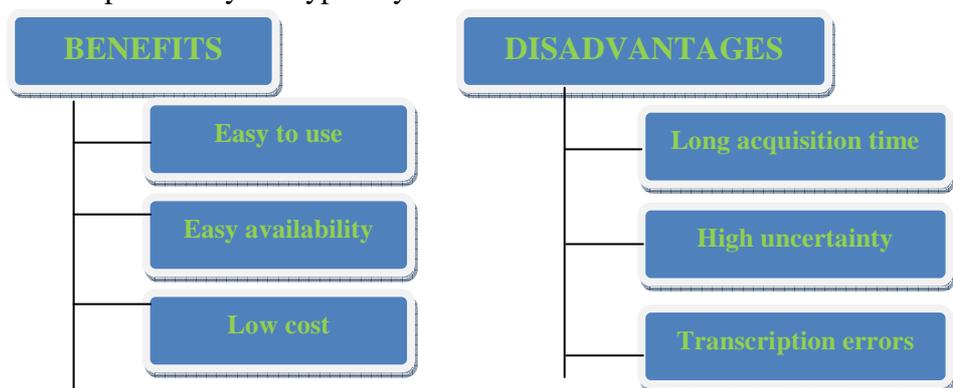


Figure. 1.2 : Benefits and disadvantages of contact systems

13 Chapter 1 Introduction to the 3D scanning systems and application

An instrument widely used in the reliefs and measurement with contact methods is represented in Figure 1.3.

A measuring arm is a mechanical digitizer capable of detecting the 3D coordinates of any form contained in a hemispherical volume with different sizes and with different accuracies. The mechanics is based on an articulated arm balanced with pointer able to move freely through 360 ° about any axis with extreme simplicity and precision



Figure. 1.3 : Measuring arm

When we use this tool, the joints are designed to fit any movement in space without any obstacle and they have the necessary fluidity. To make a measurement, the end of the arm is carried through to the point to be measured, and a series of sensors provide to decode the position of the end of the arm in real time transferring it to a computer.

14 Chapter 1 Introduction to the 3D scanning systems and application



Figure. 1.4 : Real application

The acquisition of the physical model can be made through the probing of the desired points on the model using the most common CAD programs that run directly measuring arm, using the proper tools of the program as lines, splines, NURBS, also having the ability to view the result of a scan as a real-time three-dimensional video.

This tools incorporate a pointing system that continuously identifies an absolute the position of the probe tip relative to the origin of an absolute reference system. Thus the movements of the arm in the measurement step always refer to the absolute reference system identified as the beginning of the relief during the calibration of the instrument itself.

This allows to perform all the necessary movements of the instrument and to acquire parts of an object in separate sessions even of a relief, permitting the construction of a cloud of points also composed of several parts, but always related to the same reference system.

1.1.2 Contactless systems

These acquisition systems today are those that find application in various industrial sectors: from aerospace to automotive and marine industry.

In this case are used tools with which the relief takes place without that these come into contact with the objects to be detected. These machines are very expensive and extremely sensitive.

One of the advantages is certainly the high precision. Are also used in the field of precision mechanics, allowing a quick assessment of any changes by acting on CAD systems "model" point clouds acquired, reaching the final result quickly.

These systems allow the acquisition of complex shapes. In fact, you can scan complex objects such as statues, monuments and objects, which generally have very irregular shapes. Unfortunately, they present high cost. One of the latest generation of laser scanners can cost around 90,000 euros, which is why, very often, in areas where high accuracy is not required, it falls back to other solutions. In this case the cost may be a discriminating factor in the choice of method to be used in the relief campaign. The laser scanners project a beam of light on the object, but, since they are sensitive to ambient light, in some conditions it may be difficult, if not impossible, to the relief. Even the colors of the object can create difficulties in such devices, since the dark colors or some shiny surface finishes can cause the absorption of light or refractions, interfering with the used instrumentation.

This high quantity of data to be manipulated can not sometimes be an advantage, but may actually represent a redundancy that in some way must be managed and often simplified.

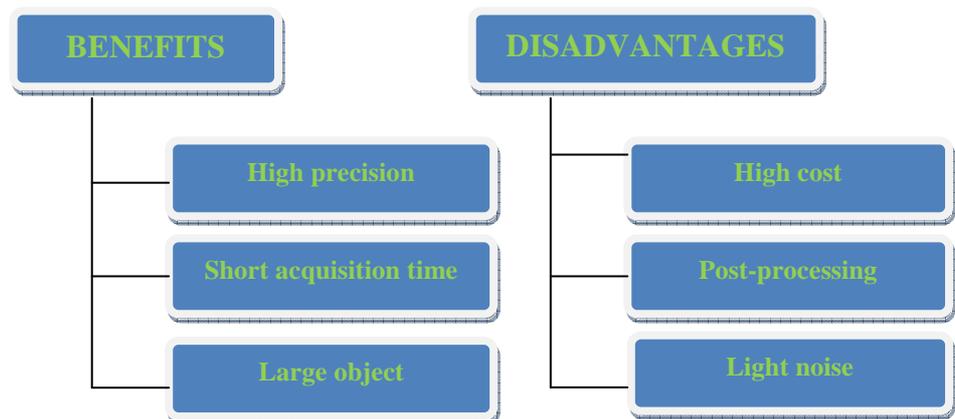


Figure. 1.5 : Benefits and disadvantages of contactless systems

One popular tool in the reliefs and measurements with non-contact methods is the MicroScan, powerful and above all very cheap. It uses a system of scanning based on the optical triangulation; the shape of objects is reproduced by sending a laser beam directly onto the surface to be scanned.



Figure. 1.6 : Microscan

17 Chapter 1 Introduction to the 3D scanning systems and application

The acquired data are then managed by software.

Among the multiple 3D scanning technologies particular importance is represented by systems not in contact, ie the optical systems, which can be classified into two main categories [2]:

1. **Passive optical systems:** generally based on acquisition of many images (in color) taken from different views, on the reconstruction of the contours of the object taken and finally on the integration of these boundaries for the reconstruction of the 3D model. Is not used any light source to derive the position of the points: it is obtained by triangulation from analysis of the images obtained by a pair of cameras positioned in a different poses compared to the same scene, as shown in figure 1.7:
These systems are extremely cheap and simple to use. Against the quality and the accuracy of the product model are generally too low for most applications. However, they are not used to the great difficulty of correlating the pairs of images in an automatic way, so for this reason it is necessary to bring a number of references on the object (some markers or grids), or manually identify the correlations, a process that is slow and difficult.
2. **Active optical systems:** are generally constituted by a pair source-sensor, where the source emits appropriate electromagnetic radiation (light patterns, laser light, etc.) and the sensor acquires the return signal, reflected from the surface of the object. The light source marks the space on a regular basis and the system returns a 2D array encoding these points, usually called range map, which shows the spatial information about the surface part of the visible object from the scan tool. Usually to get the automatic acquisition of 3D profile for an object is not enough to simply acquire a range map, but is required a simple acquisition of a set of range map defined so as to cover completely the entire surface of object with a certain overlap between the different range map. Such systems in

18 Chapter 1 Introduction to the 3D scanning systems and application

their turn are differentiated according to additional criteria: based on the light source used there may be laser systems, [3], [4] systems Moiré [5], [6], [7], (projection of interference fringes), structured light systems (emitted by projector), but these types are expensive and not always applicable even if they ensure high resolutions [8], [9].

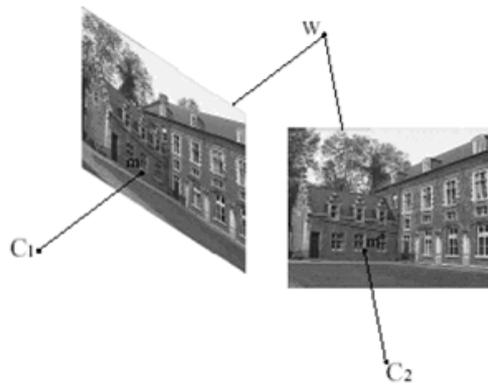


Figure 1.7 Example of a passive optical system

1.2 Applications

Now will be described some application examples of these systems that show the variety and effectiveness of their use in various fields. This type of contactless machine is used in medical, textile, movie industry and in game development.

1.2.1 Reverse engineering

The 3D scanners are used in the entertainment industry to create tridimensional digital models for movies and video games [10]. In the cases where there is an equivalent real object model is much easier to

19 Chapter 1 Introduction to the 3D scanning systems and application

scan the real object that implementing a model from scratch using solid modeling software [11]. Some artists prefer to sculpt the desired subject and then digitize it with a scanner rather than create the 3D digital model[12].

The reverse engineering in industrial field requires an accurate digital model of the object to be reproduced. It's needed a model, rather than a cloud of points, constituted by polygons, or by NURBS curves and surfaces, or from a CAD model of the object. The 3D scanners can digitize surfaces of any shape, whose point clouds can be processed with suitable software for reverse engineering.

1.2.2 The Digital Michelangelo Project

An important study on contactless technology has been made by the Department of Stanford University with the work "The Digital Michelangelo Project" [13] published on the occasion of the Second International Conference on 3D Digital Imaging and Modeling in Canada in 1999, under a project on the relief with a 3D laser scanner of Michelangelo's David. From the image below you can see the tower 3D laser scanner positioned in front of the sculpture.

20 Chapter 1 Introduction to the 3D scanning systems and application



Figure. 1.8 : Tower 3D laser scanner

The phase of acquisition of an object of this type is divided into several phases precisely because of the complexity of the shapes to be detected. In fact it is necessary to perform scans from different angles to prevent the creation of areas of shadow ruin the measurement.

The lattice of the reconstructed surface contains 8 million polygons, each approximately 2.0 mm in size. The cloud of points from which the loop was built contains 2 billion acquisitions, at the sampling distance of 0.25 mm on the surface of the statue.

1.2.3 Digital Hammurabi: scanning of cuneiform tablets

Cuneiform tablets are witnesses of an ancient form of writing in which wooden pipes were used to record the signs of forms of wet clay tablets.

Once dried, the clay tablets preserved the written text with remarkable precision and durability. There are currently hundreds of thousands of cuneiform tablets spread around the world in museums and in private collections. The study and manipulation of these artifacts presents several problems for scholars. It may be difficult or impossible to gain access to a given library. In addition, the photographic records of the tablets can be insufficient for the correct examination. The photographs do not allow you to alter the lighting conditions and the point of view. As a solution to these problems, at the Johns Hopkins University, Baltimore, they used a 3D scanner able to acquire the shape of a tablet in the three dimensions [14]. This data set can then be stored in an online library and manipulated by users by modifying the point of view and the lighting conditions. The scanner uses a camera capable of capturing the images of the tablet under different lighting conditions controlled. The image data is processed with a photometric stereo and structured light technique to determine the shape of the object. The surface is sampled with a lateral spatial distance of 26.8 μm and the depth is calculated on a scale even less. Scans of several adjacent sides are assembled together to form a 3D surface model.



Figure. 1.9 : Cuneiform tablet

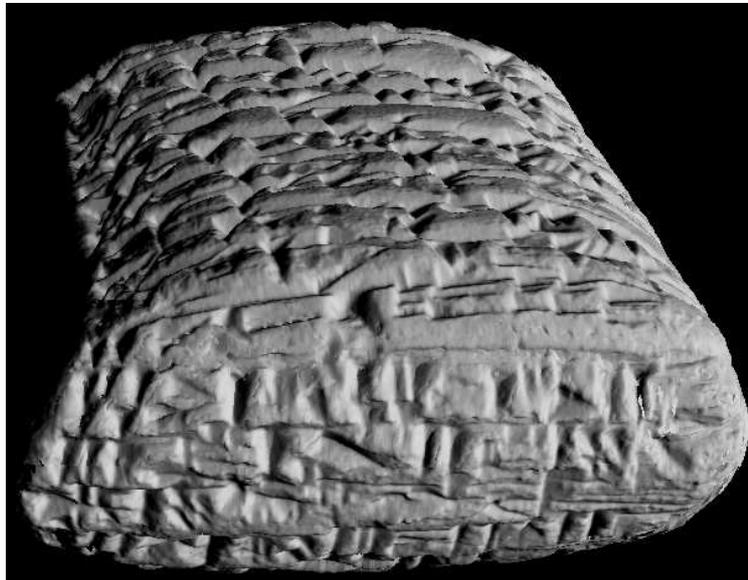


Figure. 1.10 : Rendering of a cuneiform tablet

Chapter 2

Classical model for the calibration of 3D structured light scanning systems

2.1 Introduction

This chapter gives an overview of structured light 3D scanners in the literature. These can be divided into two categories, those using the projection of digital patterns and using triangulation techniques for the reconstruction of the profile, and those that project analog pattern.

On the second scanner of this type have focused my studies, such as scanners allow increased resolution of the measured profiles, it's also possible to improve the performance of these scanners adopting new techniques of calibration

2.1.1 Structured light profilometer

In general a profilometer works in the following way, see Figure 2.1: light patterns are generated and projected, they can be, as we have said, digital (i.e. Gray encoding), or analog (usually using sinusoidal pattern), these patterns projected on the object, which we want to scan the profile, is deformed, this deformation is captured by a camera and a computer performs the post-processing that links these deformations with the profile of the object.

How to associate the entity of the deformation with the depth value is the problem that the calibration has to solve, thanks to which you can map the phase values of the deformation to the depth values, in order to obtain a cloud of points of the profile that you are capturing.

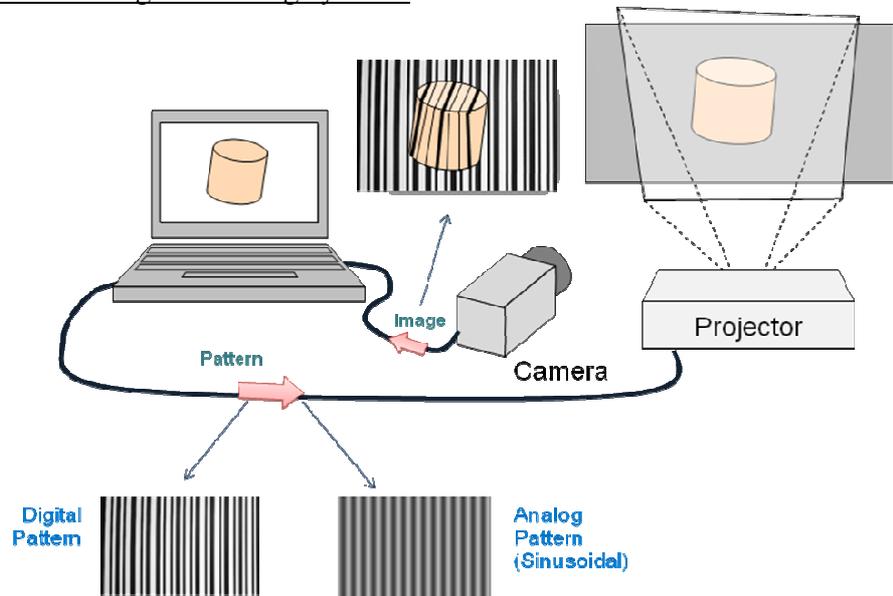


Figure. 2.1 : Structured light profilometer

2.1.2 Calibration

The calibration serves to bind the trend of the variation of light intensity of the projected pattern to the deep of the profile you are illuminating.

In the classical schemes of calibration[15],[16],[17], as shown in Figure 2.2, the depth Z is mapped in a look-up table, generally formed by N phase maps as there are the acquisitions that you performed during the calibration at regular steps inside the calibration volume

Under certain hypotheses, the phase maps can be interpolated, in this way could be aquired a minor number of reference plane.

The number of planes, that you need to acquire, is the factor of compromise which plays an important role in the calibration, it must achieve a proper trade-off between a good accuracy, a large number of plans and great flexibility, a few number of planes then the calibration should be faster and easier to implement, since a high number of phase maps involves a considerable computational cost and storage space for the large data files.

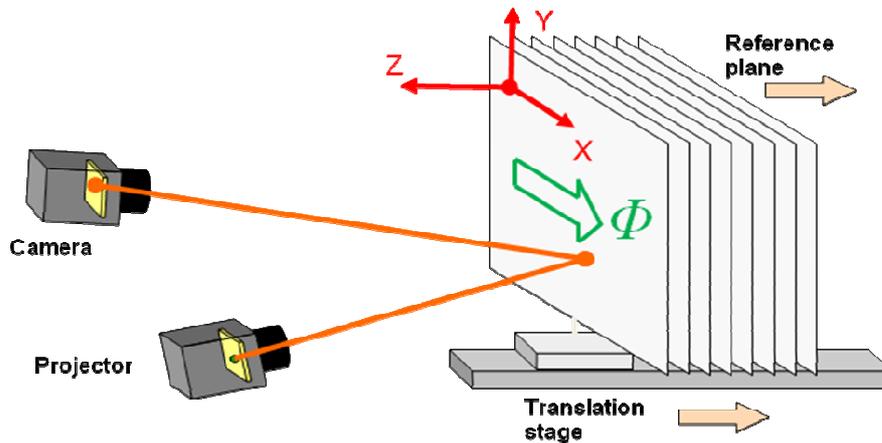


Figure. 2.2 : Common calibration schemes for 3D Scanners

To understand the limitations of these classical techniques I have been implemented some different 3D scanners based on both digital and analog pattern and I have been tested with different calibration techniques. Understood the limitations and shortcomings of these systems I have been thinking about how to build a more robust scanner while keeping low costs and timing of acquisitions, ensuring comparable uncertainties if not lower than the classical methods

2.2 Digital pattern projection scanner

2.2.1 Rusinkiewicz implementation

Rusinkiewicz algorithm for depth extraction consists of segmenting each video field into illuminated and unilluminated regions (i.e., black and white), finding stripe boundaries, matching these boundaries to those on the previous field, and using information about the illumination history of each boundary to determine the plane in space to which it corresponds. Depth is then obtained via ray-plane triangulation.

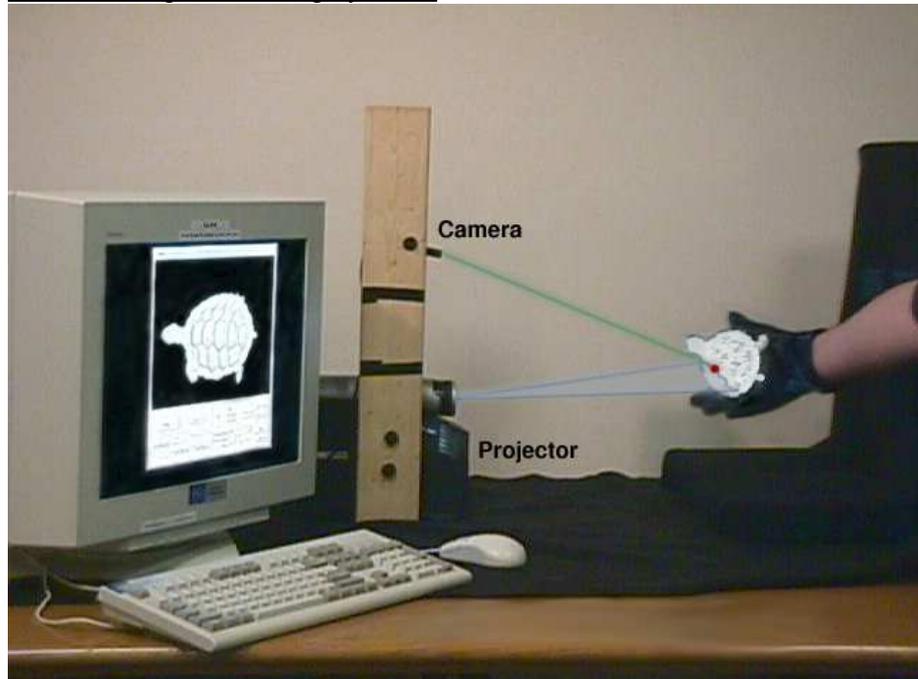


Figure. 2.3 : Rusinkiewicz implementation

This implementation makes possible the real-time acquisition, it will be discussed the tradeoffs made in implementing each stage of this pipeline (Figure 2.3) on current hardware, as well as possible extensions to these algorithms to make them more robust as hardware capabilities increase.

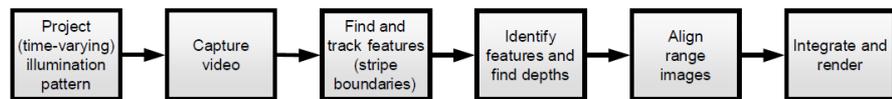


Figure. 2.4 : Real-time model acquisition pipeline

Pattern Projection and Video Capture: As mentioned above, the patterns are projected using a projector based on digital light processing (DLP) technology [18]. These projectors have the advantage of being relatively inexpensive, and have very short transition times between patterns. Because projector and camera must

structured light scanning systems

be synchronized (so that we capture exactly one video frame for each projected frame), has been chosen to drive the projector with an S-video signal and to genlock the video camera to this signal. In addition, to prevent interpolation between projector pixels, we orient the pattern such that the stripes run along the scanlines of the projector. For this reason, we are currently limited to 240 projected stripes, as compared to the 1024 potentially available from the projector. Since we currently use a 4-frame sequence consisting of 111 stripes, this limitation is not significant. However, expanding to a larger number of stripes (to increase the working volume) would require driving the projector with a VGA or DVI signal, thus requiring a different method of projector-camera synchronization. Since we use a standard video camera to capture frames, our captured video fields are interlaced. This results in a slight shift in the position of stripe boundaries from field to field. Since the effect is small, we currently do not correct for it in the processing pipeline, but because the effect of interlacing is completely known it would be possible to compensate for it. Note that any translation in the 3D model resulting from not considering the interlacing is corrected by frame-to-frame alignment.

Segmentation Algorithm: The problem of finding the stripes (and hence the stripe boundaries) in a captured video frame may be considered a special case of the general segmentation and edge detection problems. Both of these problems have been studied extensively in the computer vision community and many sophisticated algorithms are available [19].

In this application, however, it is needed a method that is robust and runs in real time, while taking advantage of the known features of the projected illumination. In particular, given the assumption of local reflectance coherence, it may assume that the highest-frequency variations in the captured frames are due to illumination, not texture. Moreover, we may assume that the projected stripes (and hence the edges we wish to find) are roughly perpendicular to the camera scanlines. Therefore, we process each scanline independently, looking for local maxima and minima along each row, and assume that these correspond to white and black projected stripes, respectively. Between each adjacent local maximum and minimum, we look for a pixel with intensity halfway between that of the minimum and maximum

structured light scanning systems

(optionally using subpixel interpolation), and use this as the location of a stripe boundary.

For scenes without high-frequency textures, it was found this method to be effective and robust, while still running in real time. In particular, it was found this algorithm less sensitive to variations in reflectivity and changes in ambient illumination than both thresholdbased segmentation methods and derivative-based edge detectors.

Stripe Matching Algorithm: Since this approach relies on time-coding the boundaries between stripes, a critical part of our algorithm is matching the boundaries visible in each frame to those in previous frames. This is a nontrivial problem for two reasons. First, the boundaries move from frame to frame, potentially with large velocities. Second, the fact that our code contains .ghost. boundaries means that not all boundaries are visible in each frame.

It is the presence of ghosts (i.e., the inferred black-black and white-white stripe .boundaries) that distinguishes this stripe matching problem from the traditional feature tracking literature. To make the problem tractable, it must used the constraints that there may be at most one ghost between each pair of visible stripe boundaries, and that ghost must match to a visible stripe boundary in the previous and following frames.

These conditions limit the possible matches and allow to determine, in many cases, whether certain boundaries should match to other visible boundaries or to ghosts. Even these conditions, however, are not enough to disambiguate the situation shown in Figure 2.5. The two possibilities of having the center stripes match to each other and having them match to ghosts in the other frame are both allowed by the constraints mentioned above.

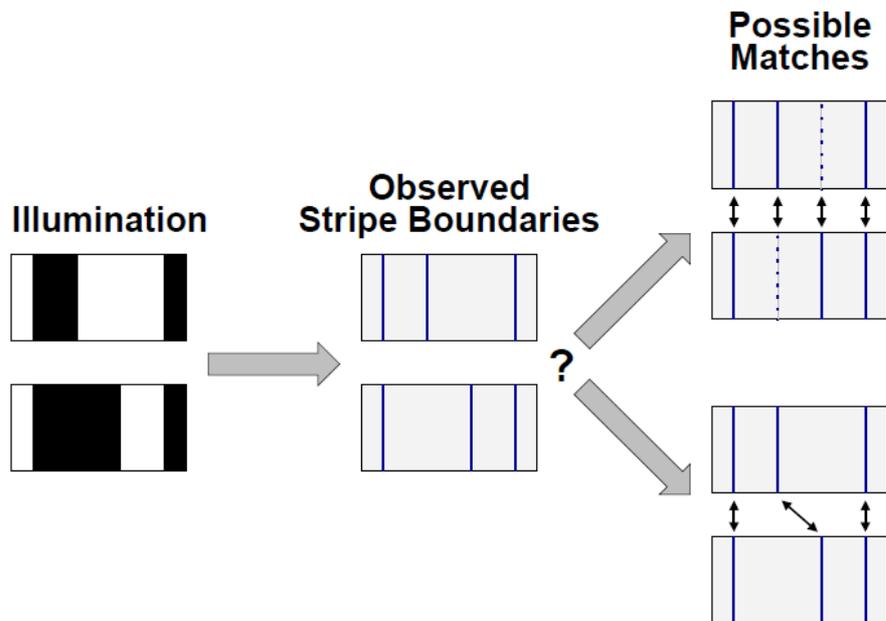


Figure. 2.5 : Matching stripe boundaries becomes difficult in the presence of “ghosts”

Although there is a large literature on tracking algorithms that could potentially be adapted to this application, including multiple-hypothesis methods [20] and methods that use velocities [21], most of these approaches are too slow for real-time implementation.

Therefore, it is currently implemented only a simple matching algorithm that hypothesizes all possible locations of ghosts and matches each visible boundary to the closest stripe or hypothesized ghost in the previous frame. As discussed later, this places a constraint on the maximum allowable velocity of stripes, hence limiting the speed at which objects in the scene can move.

Future systems may incorporate better matching heuristics, permitting correct stripe matching in the presence of greater frame-to-frame motion.

Decoding Algorithm: Once the stripe boundaries in the frame have been matched to those in the previous frame, the illumination history has been propagated (i.e., the color of the stripes on either side of the

structured light scanning systems

boundary over the past four frames) from the old boundaries to the new ones. If this boundary have seen and successfully tracked for at least four frames, this history identifies it uniquely. Note that the boundary remains identified at every frame thereafter, since the four-frame illumination history contains all four patterns.

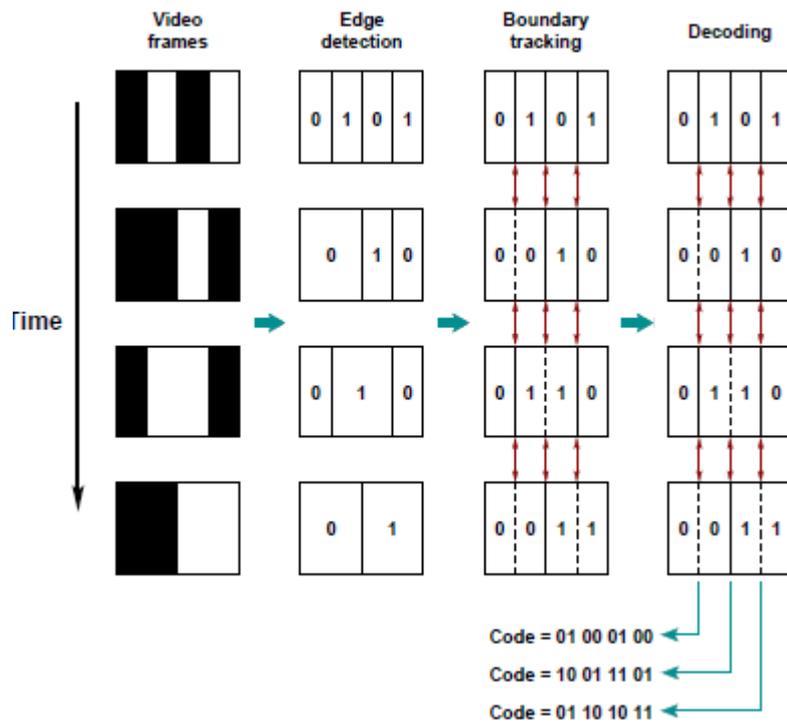


Figure. 2.6 : Digital pattern codecs, time boundary decoding

Triangulation: Given a stripe boundary identification, the plane in space to which the boundary corresponds is determined. Then the intersection of that plane, with the ray corresponding to the camera position at which the boundary was observed; is found. This determines the 3D location of a point on the object being scanned. An important difference between this approach and traditional projected-stripe systems based on Gray codes is that this scheme only gives us depth values at stripe boundaries. These depths, however, are very

accurate: we triangulate with an exact plane (the stripe boundary), rather than a wedge formed by two planes (the stripe itself). For smooth surfaces without high-frequency texture, we may perform sub-pixel estimation of the location of the stripe boundaries to further reduce depth errors.

This triangulation process requires the knowledge of the internal parameters for both the camera and projector, as well as their relative pose. In order to calibrate intrinsics, we currently use the method of [22]. This can be followed by moving a target to known 3D positions and optimizing to find the relative pose of the camera and projector.

One could imagine an automatic calibration method that would permit a calibration target to be moved around (by hand), then simultaneously solve for the scanner calibration and the positions to which the target was moved (a similar approach was demonstrated by [23]).

2.2.2 ICP

Iterative Closest Point (ICP) is an algorithm employed to minimize the difference between two clouds of points. ICP is often used to reconstruct 2D or 3D surfaces from different scans, to localize robots and achieve optimal path planning (especially when wheel odometry is unreliable due to slippery terrain), to co-register bone models, etc.

The algorithm is conceptually simple and is commonly used in real-time. It iteratively revises the transformation (translation, rotation) needed to minimize the distance between the points of two raw scans.

This algorithm use as inputs the points from two raw scans, initial estimation of the transformation and criteria for stopping the iteration.

It's return as output the refined transformation.

Essentially the algorithm steps are the follows:

- 1) Associate points by the nearest neighbor criteria.
- 2) Estimate transformation parameters using a mean square cost function.
- 3) Transform the points using the estimated parameters.
- 4) Iterate (re-associate the points and so on).

The main algorithm drawback is that it is prone to accumulative errors, which can lead to the mapping algorithm failure.

In order to apply this algorithm to a real-time 3D scanner it was chosen a high-speed ICP algorithm by combining some different algorithm known. Like Blais and Levine, we propose using a projection-based algorithm to generate point correspondences. Like Neugebauer, it was combined this matching algorithm with a point-to-plane error metric and the standard “select-match-minimize” ICP iteration. The other stages of the ICP process appear to have little effect on convergence rate, so it was chosen the simplest ones, namely random sampling, constant weighting, and a distance threshold for rejecting pairs. Also, because of the potential for overshoot, it was avoided extrapolation of transforms.

All of the performance measurements presented so far have been made using a generic ICP implementation that includes all of the variants described in literature.

It is, however, possible to make an optimized implementation of the recommended high-speed algorithm, incorporating only the features of the particular variants used. When this algorithm is applied to a “fractal” testcase, it reaches the correct alignment in approximately 30 milliseconds. This is considerably faster than the baseline algorithm (based on [24]), which takes over one second to align the same scene. It is also faster than previous systems that used the constant-time projection strategy for generating correspondences; these used computationally expensive simulated annealing [25] or Levenberg-Marquardt [26] algorithms, and were not able to take advantage of the speed of projection-based matching.

It's shown in Figure 2.7 an example of the algorithm on real-world data: two scanned meshes of an elephant figurine were aligned in approximately 30 ms.



Figure 2.7: High-speed ICP algorithm applied to scanned data

There are other dissertation that propose a high-speed ICP algorithm suitable for realtime use. David Simon, in his Ph. D. dissertation [27], demonstrated a system capable of aligning meshes in 100-300 ms. for 256 point pairs (one-eighth of the number of pairs considered throughout this chapter). His system used closest-point matching and a point-to-point error metric, and obtained much of its speed from a closest-point cache that reduced the number of necessary $k-d$ tree lookups. As we have seen, however, the point-to-point error metric has substantially slower convergence than the point-to-plane metric that Rusinkiewicz uses. As a result, this system appears to converge almost an order of magnitude faster, even allowing for increase in processor speeds. In addition, this system does not require preprocessing to generate a $k-d$ tree.

2.2.3 Evaluation of the implemented scanner

In order to evaluate the performance of the implemented system it was performed acquisitions of objects of known dimensions and performance was evaluated in terms of time and of resolution of the profile obtained. As mentioned above to make sure that the

structured light scanning systems

triangulation has a match in terms of measurable depth we need to have available the parameters of both devices, camera and projector, as these parameters were calculated is explained in the next section and represents the innovation on which is based this work, since before an overall and exhaustive calibration of both devices did not exist in the literature with a so precise modeling

Please note that the results of a calibration procedure are intrinsic and extrinsic parameters of the calibrated system. For the evaluation of the results will be presented with two MPP matrix for the camera and the projector and the relative rotation matrix (R), translation vector (T) and intrinsic parameters matrix (A).

For the calibration method that will be shown in next chapter, the following results were obtained:

Camera MPP:

$$\begin{bmatrix} 1302.253 & -31.190 & -487.987 & 399991.664 \\ -79.292 & 1183.532 & -595.285 & 352296.830 \\ 0.007 & -0.226 & -0.974 & 1510.109 \end{bmatrix}$$

Projector MPP:

$$\begin{bmatrix} 2043.431 & -131.753 & -1031.049 & 175856.587 \\ -220.120 & 2092.220 & -1027.466 & 169814.865 \\ -0.252 & -0.140 & -0.958 & 1440.352 \end{bmatrix}$$

and reprojection errors relating of that calibration are the following:

camera reprojection error: 0.586 pixel;

projector reprojection error: 1.25 pixel;

The values of some parameters of these matrices are an indicator of the results obtained. In fact, if we analyze the R is expected that the values along the main diagonal are very close to unity while the others are close to zero. For the vector T is expected that the last value (t_{13}) that relating to the translation along the Z axis is equal to the distance (in mm) of the device from the reference plane for the acquisition. The

35 Chapter 2 Classical model for the calibration of 3D structured light scanning systems

matrix A is expected to find the pixel coordinates of the focal center (A_{13}, A_{23})

Then this scanner was used for the acquisition of a mask of known dimensions. The probe mask is shown in Figure 2.8, and the results of the acquisition is shown in Figure 2.9



Figure 2.8:Probe mask

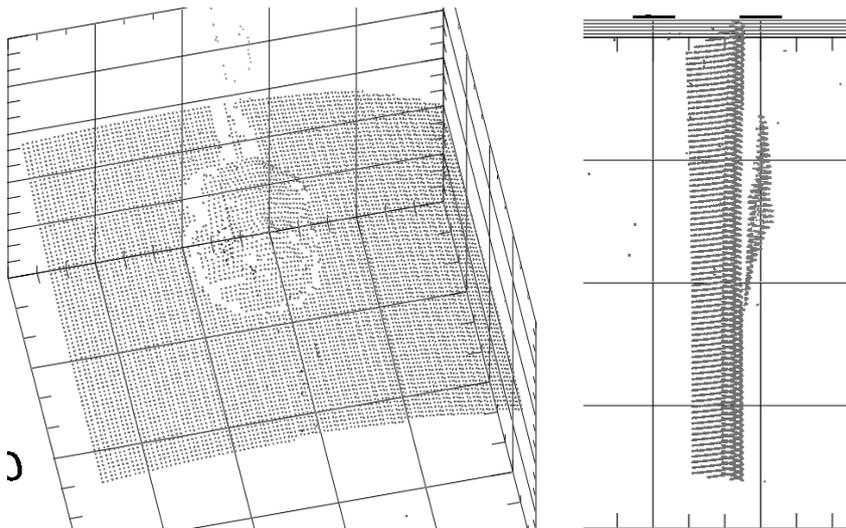


Figure 2.9:Results of acquisition

The results of the process of triangulation are the real points coordinates (world) acquired.

In the figure above you can see the details of the profile of the object acquired. The details seem not perfectly delineated, actually in the shown reconstruction is not rendered, which would allow a better view.

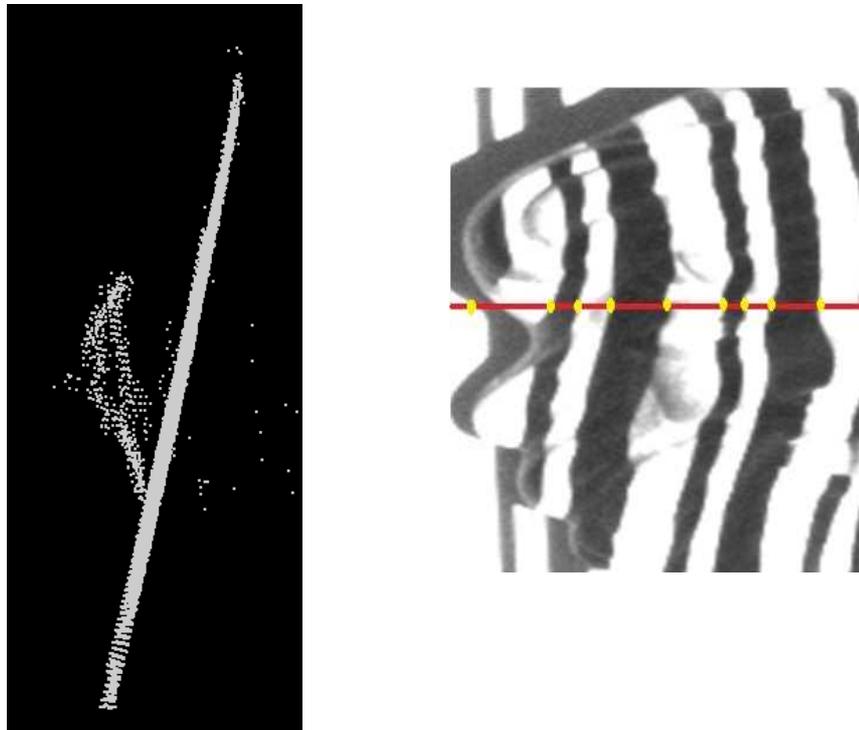


Figure 2.10: Cloud Points and a frame acquisition

We have achieved performance of about 10 frames per second as having single-frame acquisition times close to 100 ms. In fact the algorithm uses about 60 ms for the acquisition of the image from the camera. The acquisition time is a parameter which depends only on the type of camera used and how the software manages the communication with it, so that it's not possible to act on it if not using different tools.

2.2.4 Drawback

The limitations encountered in the use of this scanner are *a)* the low resolution of the acquisitions, since an increase in resolution does not make it possible to reconstruct it in real time, and *b)* the problem (thus far not common in the literature) of the precise knowledge of the intrinsic and extrinsic projector parameters, that is able to calibrate both devices in a short time and with a high precision. For this reason, errors, in the reconstruction phase of this scanner, remain high and unfeasible in applications that require high resolutions such as the reverse engineering. Thus the adoption of the scanner is limited to applications of a qualitative nature such as the creation of models for video games or augmented Reality

2.3 Analog pattern projection scanner

When a fringes pattern, assuming sinusoidal, is projected onto the surface of an object it undergoes a deformation with respect to the case where it is projected onto a flat surface, considered as reference. The deformation depends on the deviation of the surface of the object from the reference surface (difference of optical path between the reference plane and the measured one) which is found encoded in the phase distribution pattern as deformed and acquired from the camera.

2.3.1 Fringe pattern projection and phase shift

To be able to trace three-dimensional topographical information of object, then, you must do:

- numerically demodulating the distribution of spatial phase by an automated method of analyzing the fringes pattern;
- find the relationship of conversion phase-depth.

The acquisition of the phase map is obtained through a “local” approach, said phase-stepping or phase-shifting (PSP), which determines the phase information from the local intensity. The

structured light scanning systems

procedures that may follow are two-phase temporal stepping and spatial phase-stepping.

In the temporal phase-stepping, the phase information is obtained by making k images, obtained by acquiring a periodic grating supposed to sinusoidal intensity, projected on the object and traversed k times of a quantity $2\pi / k$. Since the process of image acquisition requires a certain time, can not be studied moving objects using this approach. It is resistant to ambient light and to the variations of reflections.

In the spatial phase-stepping is required a carrier so that a given number N of contiguous pixels should correspond to the period of the signal intensity (condition of "tuning"). In this case, the recorded values from adjacent pixels play the role of intensity shifted in phase and the phase can be calculated from a single frame. However, the condition of "tuning" is quite compelling, and often the method is not applicable

Phase Shift technique.

The Phase-Shift technique is characterized by the subsequent projection of N sinusoidal pattern shifted in phase of $2\pi/N$ on a reference plane and on the object under test, and by the corresponding acquisition, by a camera, of N images, in which the sinusoidal pattern is distorted because of the prospect and the trend space of the illuminated surface [5].

The intensity of the generic pixel (i, j) of the 'k-th image' can be expressed by the following formula:

$$I_k(i, j) = a(i, j) + b(i, j) \cos[\phi(i, j) + \delta_k]$$

where

(i, j) are the pixel coordinates;

$a(i, j)$ is the mean value of the gray pixel (average brightness or light background);

$b(i, j)$ is the amplitude modulation (the variation of (i, j) location);

in other terms represent the distortion of the light intensities and depend on non-uniform lighting, the reflectivity of the object surface and the non-uniformity in the response of the camera sensor;

δ_k represent the different shift of the phase;

ϕ is the phase value you want to measure.

39 Chapter 2 Classical model for the calibration of 3D structured light scanning systems

We can rewrite the equation describing the intensity, by simple mathematical manipulations, as follows:

$$I_k(i, j) = c_0(i, j) + c_1(i, j) \cos \delta_k + c_2(i, j) \sin \delta_k$$

where:

$$\begin{aligned} c_0(i, j) &= a(i, j) \\ c_1(i, j) &= b(i, j) \cos \phi(i, j) \\ c_2(i, j) &= -b(i, j) \sin \phi(i, j) \end{aligned}$$

As seen from the above equation, to derive the phase you must first know the coefficients and as the phase information is contained in them. To this regard, it applies the method of least squares, that allows to calculate the coefficients desired in order to minimize the sum of the squares of the differences between the empirical values and the theoretical ones. The result obtained, for simplicity, it can be written in matrix form as follows:

$$\underline{\underline{P}}c(i, j) = q(i, j)$$

Where:

$$P = \begin{bmatrix} p_1 & p_2 & p_3 \\ p_2 & p_4 & p_5 \\ p_3 & p_5 & p_6 \end{bmatrix} = \begin{bmatrix} K & \sum_{k=0}^{K-1} \cos \delta_k & \sum_{k=0}^{K-1} \sin \delta_k \\ \sum_{k=0}^{K-1} \cos \delta_k & \sum_{k=0}^{K-1} \cos^2 \delta_k & \sum_{k=0}^{K-1} \cos \delta_k \sin \delta_k \\ \sum_{k=0}^{K-1} \sin \delta_k & \sum_{k=0}^{K-1} \cos \delta_k \sin \delta_k & \sum_{k=0}^{K-1} \sin^2 \delta_k \end{bmatrix}$$

And it's only function of δ_k shift in the fringe system, while

$$q(i, j) = \begin{bmatrix} q_0(i, j) \\ q_1(i, j) \\ q_2(i, j) \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{K-1} I_k(i, j) \\ \sum_{k=0}^{K-1} I_k(i, j) \cos \delta_k \\ \sum_{k=0}^{K-1} I_k(i, j) \sin \delta_k \end{bmatrix}$$

is composed of the weighted sum of the intensities of the images acquired.

Now this report is obtained by reversing the vector and finally the phase will be:

$$\phi(i, j) = \text{arctg} \left(\frac{-c_2(i, j)}{c_1(i, j)} \right)$$

Following a similar procedure is also obtained the phase only on the reference plane.

The phase distribution calculated in this way has a continuous pattern within each fringe with a value in the range $[0, 2\pi]$ in each period, that contains a discontinuity in the form of 2π as the arctangent function is extended to four radians. To have, therefore, a unique correspondence between each point and the value of the phase is necessary to make the operation of "unwrapping" or "unrolling" of the phase, with the aim of returning a unique distribution of the continuous phase without the presence of steps or discontinuity.

2.3.2 Classical calibration techniques

In order to derive the relation-phase depth is necessary an initial stage of calibration of the system. It allows to calculate the parameters that are a function of geometry and disposition of the various components of the scanner, such as camera, projector and the reference plane of the volume inside which will return the object to be scanned.

Then will be shown in order the various steps of the calibration performed according to classical techniques known in the literature, starting from the camera calibration for the calculation of the intrinsic and extrinsic parameters, to arrive at the calibration of the entire system through the calculation of the parameters of longitudinal and transverse of whole system.

In all scanners you perform a pre-calibration of the camera usually using a well-known algorithm, proposed by Zhang and Tsai [28],[29] who model the camera through the pinhole model, then you calibrate the profilometer for the calculation of the phase relation-depth, longitudinal calibration, and the scaling relations on the reference plane, trasversal calibration. These two calibrations typically take place through the projection of N patterns within the volume

calibration. In the following it will be shown a couple of recent techniques for the calculation of these relations through the acquisition of a limited number of phase maps.

Camera calibration

To better understand what is and what information you can get to calibrate a camera, suppose you have a camera that sees a scene and fix in it a three-dimensional Cartesian reference system, XYZ , with a random origin. In this way each point of the scene is detected by the camera from a set of three Cartesian coordinates (x, y, z) .

The image produced by the camera is two-dimensional for which we consider a two-dimensional Cartesian reference system uv , whose coordinates are expressed in pixels. Each point in the image will be so identified by a pair (u, v) [30].

P is a point in the interior of the scene and P' is the corresponding point in the image produced by the camera, as shown in figure 2.12. In general, every point of the scene will have its corresponding point within the image. The points scene and the image points are linked by a kind of transformation that is operated by the camera. The calibration process is to find this transformation, or derive the parameters that govern the way in which a scene is reflected in a point image, as in Figure 2.11.

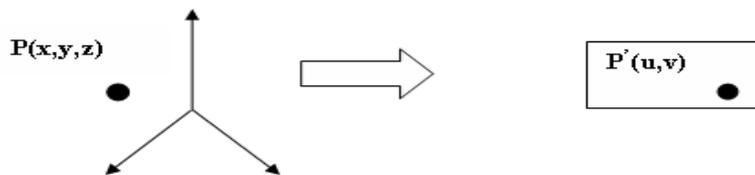


Figure 2.11: Correspondence between points in space and on the sensor

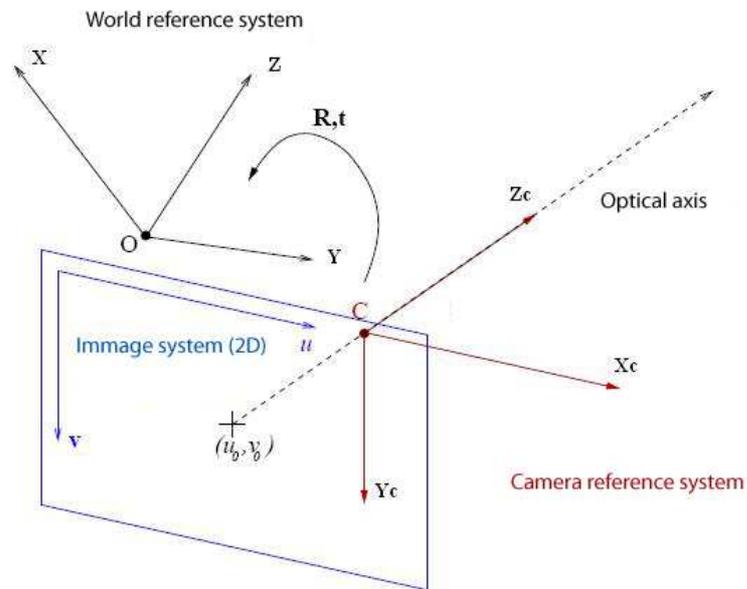


Figure 2.12: Reference systems

In other words, the calibration allows to calculate the parameters that govern the projection operated by a camera on the points in space. Therefore you are able to know in what way the 3D points are projected on the plane of image from a camera whose parameters are known.

The question arises whether it is possible to operate a reverse transformation, i.e. from the points of an image, extract information about the corresponding points of space. It is actually possible, but doing some clarifications. Note that the projection is operated by a camera implies an inevitable loss of information: the passage from a point of three-dimensional space to a point two-dimensional image, provides for the loss of a coordinate, the one that takes into account the depth. Thus in a hypothetical inverse passage there is the problem of recovering the information related to this "lost" coordinate. This problem is overcome by a technique known as stereopsis. It consists in the recovery of the third coordinate, using two or more views of the same scene from different angles and thus in this way it is able to make the inversion of the transformation: this is the process known by the name of 3D reconstruction.

43 Chapter 2 Classical model for the calibration of 3D structured light scanning systems

The reference system centered in the world is the optical center C, which is also the origin of the reference system of the camera, but in general these two systems do not coincide, then we introduce three different reference systems:

1. the reference system, also called 3D world system;
2. the standard reference system 3D camera, centered in C;
3. reference system for the 2D image.

Given any point P of the scene, it can be identified in each of these three different systems, in particular it is indicated with:

$X = (X, Y, Z)$ coordinates of the point P in the world system;

$X_C = (X_C, Y_C, Z_C)$, the coordinates of the camera system;

$x = (x, y)$ coordinates of the point P in the system of the image;

$w = (u, v)$ coordinates of the point P in the system of the digital image, taking into account the discretization made during the image conversion from analog to digital.

Whereas the two systems are not identical, they are linked by an isometric transformation composed by a traslation T and a rotation R, which are unknown.

The world system and the camera system are related by making use of homogeneous coordinates as follows:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-f}{Z_C} & 0 & 0 & 0 \\ 0 & \frac{-f}{Z_C} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix}$$

One aspect that we must not neglect is that you are working with digital images, i.e. formed by a number of pixels and the relation

structured light scanning systems

between the image coordinates (x, y) to the coordinates of the pixels (u, v) is:

$$u = u_0 + k_u x$$

$$v = v_0 + k_v y$$

where

(u_0, v_0) are the coordinates of the principal point;

(k_u, k_v) are the inverse of the effective size of the pixel, respectively along the direction u and v .

Through simple matrix manipulations we can write the relation between the coordinates in the world to those pixels in the image:

$$w = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} X$$

It defines:

$$\begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as a perspective projection matrix (MPP). This matrix represents the geometric model of the camera.

Calibration of the fringe pattern profilometer

Figure 2.13 shows the coordinate systems in a typical system of measurement of the profile projection of fringes based on the technique of phase shift (PSPFP) [31].

There are the following reference systems:

- world or absolute reference system XYZ;
- reference system relative to the RC photo sensor;
- UVW reference system on the lens;
- XgYgZg reference system relative to the grid to be projected;
- Reference system XpYpZp on the projection lens.

The coordinate system XYZ is the reference system fixed to represent the shape of the object under test. The coordinate system RC is placed

in the plane of the photosensor with the axes R and C parallel to the direction of the rows and columns of the sensor, respectively. The origin of the coordinate system RC is placed in the center of the first pixel in the upper left. The origin of the coordinate system UVW is the main point of the lens shooting. The W-axis coincides with the main axis of the lens. The origin of the reference system relative to the grating generated by the projector is a point where the absolute phase appears to be zero. The grid lies in the plane $X_g Y_g$ with the fringe orthogonal to the axis X_g . The coordinate system relative to the lens of the projector is defined in a similar way to the coordinate system relative to the lens.

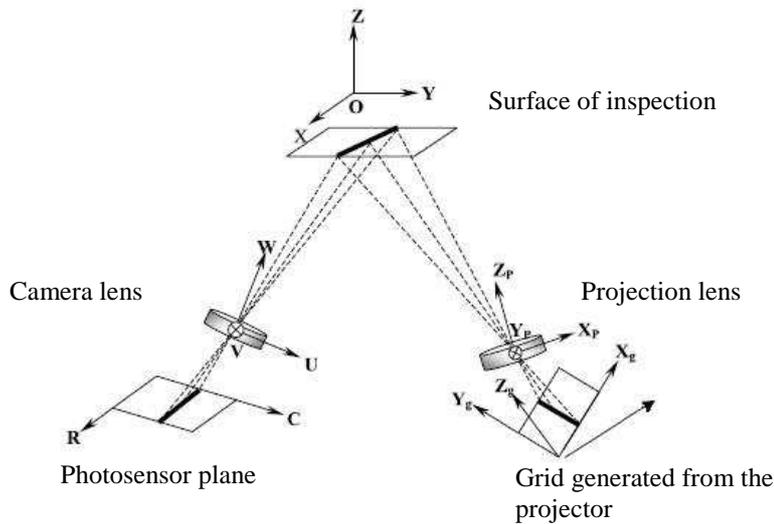


Figure 2.13: Geometric system of the profilometer

A point can be transported from a reference system to another using a rotation matrix and a translation vector. For example, if the point of the plane of the photodetector corresponds to the point of the reference system of the lens, then the coordinates are related by

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} r_{11}^{DC} & r_{12}^{DC} & r_{13}^{DC} \\ r_{21}^{DC} & r_{22}^{DC} & r_{23}^{DC} \\ r_{31}^{DC} & r_{32}^{DC} & r_{33}^{DC} \end{bmatrix} \begin{bmatrix} r \\ c \\ 0 \end{bmatrix} + \begin{bmatrix} t_x^{DC} \\ t_y^{DC} \\ t_z^{DC} \end{bmatrix}$$

structured light scanning systems

where the symbols D and C present in the apices represent the coordinate systems associated with the plan and the lens, respectively. The order of these symbols indicates the direction of the coordinate transformation. This convention is also adopted in the hereafter.

Consider a generic point in the absolute system of the surface to be reconstructed in which the absolute phase is φ . The absolute phase can be obtained from its image point in the reference plan. The quantity r, c and φ there can be written as:

$$r = f_r(x, y, z)$$

$$c = f_c(x, y, z)$$

$$\varphi = f_\varphi(x, y, z)$$

Where $r = f_r(x, y, z)$, $c = f_c(x, y, z)$ and $\varphi = f_\varphi(x, y, z)$ are non-linear functions that depend on the measuring system as a whole. When these functions are determined by the system calibration can be determined by simultaneously solving the system of nonlinear previous equations. The equations contain all the subsequent information to reconstruct the profile of an object. To implement this approach, we must make explicit the connections between the quantities to be measured and the unknown. Assuming negligible effects of distortion introduced by the projection lens and the lens recovery, perspective transformations can be written as follows:

$$\frac{u}{w} = \frac{r_{11}^{WC} x + r_{12}^{WC} y + r_{13}^{WC} z + t_x^{WC}}{r_{31}^{WC} x + r_{32}^{WC} y + r_{33}^{WC} z + t_z^{WC}}$$

$$\frac{v}{w} = \frac{r_{21}^{WC} x + r_{22}^{WC} y + r_{23}^{WC} z + t_y^{WC}}{r_{31}^{WC} x + r_{32}^{WC} y + r_{33}^{WC} z + t_z^{WC}}$$

Where (u, v, w) indicates a point in the reference system of the lens, (x, y, z) indicates the corresponding point in the absolute system, and the apex WC indicate the coordinate transformation from the absolute system to that of the lens of the camera.

Set the point of observation, the words to the second member of previous equations become constants. Solving these equations simultaneously, we can express x, y as functions of z , namely:

$$x = a_1 z + a_0$$

$$y = b_1 z + b_0$$

Where the coefficients a_0 , a_1 , b_0 and b_1 can be derived by simple algebraic calculations, this also shows that all the points observed by a given pixel lie on a straight line called line of sight of the pixel.

Combining equations that is obtained:

$$\varphi = \Phi'(z)$$

This equation shows that, for a fixed observation point, the measured phase depends only on the depth z . In a properly designed system, the function $\Phi'(z)$ is monotonic and invertible. This allows to represent the depth z as a function of the measured phase φ :

$$z = Z(\varphi)$$

Where $Z(\varphi)$ is the inverse of $\Phi'(z)$ and is also a reversible monotone function.

The profilometric system can be described analytically by the following equations:

$$x = a_1 z + a_0$$

$$y = b_1 z + b_0$$

$$z = Z(\varphi)$$

Longitudinal calibration

Using an ideal projector, the equiphase lines in the projected sinusoidal grid produce a set of equiphase planes in the space image of the projector.

When the line of sight of a sequentially pixel meets these equiphase planes, the phase value observed by the pixel changes monotonously. For a given value of phase φ , the corresponding depth of the surface is given by the z coordinate of the point at which the line of sight intersects the plane of the pixel of the phase φ .

The equations of equiphase planes can be obtained by recalling that they pass through the corresponding equiphase lines and the center of the projection lens. Assuming use of fringes parallel to the Y axis, consider for a generic fringe a line L parallel to the axis Y . For this straight line L which has constant phase φ , its plane grid has the following expression:

$$x_G = \frac{\varphi}{K}$$

$$z_G = 0$$

where K represents the wave number of the sinusoidal fringe pattern. The y -coordinate does not appear because it was considered a pattern of sinusoidal fringes parallel to the axis Y .

In the reference system of the projector, the reticular plane is represented by:

$$r_{11}^{PG} x_p + r_{12}^{PG} y_p + r_{13}^{PG} z_p + t_x^{PG} = \frac{\varphi}{K}$$

$$r_{31}^{PG} x_p + r_{32}^{PG} y_p + r_{33}^{PG} z_p + t_z^{PG} = 0$$

The equiphase plane produced by the product by the projection of the line L is given by:

$$(d_1 x_p + d_2 y_p + d_3 z_p) \varphi = e_1 x_p + e_2 y_p + e_3 z_p$$

where

$$d_i \equiv \frac{r_{3i}^{PG}}{K}$$

$$e_i \equiv r_{1i}^{PG} t_z^{PG} - r_{3i}^{PG} t_x^{PG}$$

The line of sight of the pixel can be obtained by transforming the final equation, which represents the line in the absolute reference system, in the projector system:

$$x_p = a'_1 z + a'_0,$$

$$y_p = b'_1 z + b'_0,$$

$$z_p = c_1 z + c_0$$

Substituting the previous equations gives the ideal phase-depth relation

$$z = \frac{m_1 \varphi + m_0}{n_1 \varphi + n_0}$$

where

$$m_0 = a'_0 e_1 + b'_0 e_2 + c_0 e_3$$

$$m_1 = -a'_0 d_1 - b'_0 d_2 - c_0 d_3$$

$$n_0 = -a'_1 e_1 + b'_1 e_2 - c_1 e_3$$

$$n_1 = a'_1 d_1 + b'_1 d_2 + c_1 d_3$$

To carry out the calibration phase-depth, a sinusoidal pattern is projected onto a flat surface placed perpendicularly to the z-axis direction. The resulting phase distribution on the flat surface can be measured using a phase-shifting algorithm. The phase measurement is repeated when the flat surface is subsequently translated in positions at different depths, until the limit position is reached. The measured phase maps are unrolled along the first transverse direction and then along the longitudinal direction in order to recover the information on the continuity of the phase. After this procedure, you get a series of steps associated with the absolute positions of the plane at different depths for each pixel. Starting from this information (phase values and relative depth for each pixel), by means of an appropriate minimization process, one can determine the values of the unknown coefficients present in the relationship phase-depth.

Divide numerator and denominator by n_0 to obtain:

$$z = \frac{m'_1 \varphi + m'_0}{n'_1 \varphi + 1}$$

Then

$$m'_1 \varphi + n'_1 \varphi z + m'_0 = z$$

Selecting φ , $z\varphi$ and 1 as functions of the base, the estimation of the parameters can be done transformed into a linear problem. When more than three pairs of values phase-depth becomes available for the interpolation, it is possible to use, for the estimation of the parameters, the least squares minimization.

The least squares problem can be formalized as follows:

$$\mathbf{Ax} = \mathbf{b}$$

Is a linear system where the matrix of coefficients $\mathbf{A} \in \mathbf{R}^{n \times m}$ is such that $n \geq m$. If the system is oversized and has no solution, fixed the vector norm $\| \dots \|_2$, the problem becomes that of determining a vector such that

$$\| \mathbf{Ax} - \mathbf{b} \|_2 = \min_{x \in \mathbf{R}^m} \| \mathbf{Ax} - \mathbf{b} \|_2$$

Transverse calibration

It's now presented the other type of calibration, the transversal one, which allows to obtain the other two coordinates, x and y, and be able to fully reconstruct the object under test.

The calibration procedure of the camera provides information that can be exploited for the calculation of the parameters required for determining the transverse coordinates. We know that the coordinates x and y can be calculated using the z coordinate note the following relations

$$x = a_1 z + a_0$$

$$y = b_1 z + b_0$$

In which the parameters a_0 , a_1 , b_0 and b_1 are to be determined during calibration of the system. The number of parameters to be determined is four, which is why we need at least four equations. The expression of a line of sight of a pixel is formed by two equations. Thus in order to determine the unknown parameters is necessary to know the couple for at least two different values of z. To solve this problem is posted on a chessboard floor used for calibration. The plan is put first in correspondence of the reference position and then in the bottom of the volume of calibration,

So we can write the following equation system

$$\begin{cases} x_1 = a_1 z_1 + a_0 \\ y_1 = b_1 z_1 + b_0 \\ x_2 = a_1 z_2 + a_0 \\ y_2 = b_1 z_2 + b_0 \end{cases}$$

where z_1 corresponds to the depth of the reference plane, so that it is $z_1 = 0$, while z_2 corresponds to the depth of the plan when this is placed in the bottom of the volume of calibration, so it is $z_2 = z_{\max}$.

This is valid for each pixel and the system equation becomes:

$$\begin{cases} x_1 = a_0 \\ y_1 = b_0 \\ x_2 = a_1 z_{\max} + a_0 \\ y_2 = b_1 z_{\max} + b_0 \end{cases}$$

Substituting into the previous equations we obtain the result:

$$\begin{cases} a_0 = x_1 \\ b_0 = y_1 \\ a_1 = \frac{x_2 - a_0}{z_{\max}} \\ b_1 = \frac{y_2 - b_0}{z_{\max}} \end{cases}$$

This type of calibration can be performed using the parameters previously calculated in the calibration of the camera, so as to reduce the time and calibrate the scanner at the same time the camera, in this way, if we remember that:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} -fk_u & 0 & u_0 & 0 \\ 0 & -fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

So we can write the following relations, reformulating them in such a way to write two equations in which X and Y are a function of the unknowns u, v, Z:

$$x = \frac{(E_y F_z - F_y E_z)z + E_y F_0 - E_0 F_y}{F_y E_x - F_x E_y}$$

$$y = \frac{(E_z F_x - F_z E_x)z + E_0 F_x - E_x F_0}{F_y E_x - F_x E_y}$$

where

$$\begin{aligned} E_x &= (u - u_0)r_{31} - fk_u r_{11} & F_x &= (v - v_0)r_{31} - fk_v r_{21} \\ E_y &= (u - u_0)r_{32} - fk_u r_{12} & F_y &= (v - v_0)r_{32} - fk_v r_{22} \\ E_z &= (u - u_0)r_{33} - fk_u r_{13} & F_z &= (v - v_0)r_{33} - fk_v r_{23} \\ E_0 &= (u - u_0)t_3 - fk_u t_1 & F_0 &= (v - v_0)t_3 - fk_v t_2 \end{aligned}$$

If we compare the equations with the previous one:

$$x = a_1(u, v)z + a_0(u, v)$$

$$y = b_1(u, v)z + b_0(u, v)$$

We finally obtain:

$$\begin{aligned}
 a_1 &= \frac{E_y F_z - F_y E_z}{F_y E_x - F_x E_y} & b_1 &= \frac{E_z F_x - F_z E_x}{F_y E_x - F_x E_y} \\
 a_0 &= \frac{E_y F_0 - E_0 F_y}{F_y E_x - F_x E_y} & b_0 &= \frac{E_0 F_x - E_x F_0}{F_y E_x - F_x E_y}
 \end{aligned}$$

The transverse parameters thus defined, may be calculated using the intrinsic and extrinsic calibration data of the camera using data of the roto-translation of the reference plane of the volume of calibration, which provides for a reduction of the times of acquisition and processing of the previously expected from the classical technique.

2.3.3 Drawback

This classical technique has some problems, that we talk about following.

Higher order harmonics

The implementation of the classical phase-shift introduces some errors due to simplification of the model used for the modeling of the pattern, this simplification does not take into account the higher order harmonics arising from the use of a digital projector for the generation of fringe patterns.

These noises will dirty the projected patterns with infinite harmonics, but the main error is introduced by the second harmonic signal that is significantly stronger than the other.

A solution to this drawback, known in literature, has been studied in [32] and is introduced into the solution to the problem that goes for modeling more detailed patterns. This technique takes the name of I3PSP (improved three-step phase shifting profilometry).

First, filtering is performed by a low-pass filter. This procedure allows the elimination of higher harmonics to the first, but given the non-ideality of the second harmonic filter will not disappear completely. Is introduced at this point a term which takes into account the residual noise of the filtered second harmonic, as such term precisely the residue with the previous model was not taken into account.

The resolution of the method assumes a different form as the function that describes the pattern shown is the function that describes the deformed patterns on the object and captured by the camera feature

53 Chapter 2 Classical model for the calibration of 3D structured light scanning systems

in terms of frequency that is twice the resolution of the PSP method (step phase shifting profilometry) above. Then extracted from these functions is the relative phase of the object that has distorted the patterns obtained. To do this, therefore, introduces a discourse on the powers of the signals for the solution of these equations, which are trigonometric functions with different frequencies, and proceed with the calculation of the phase.

Starting from the assumptions above we can write the model as it changes;

$$\begin{aligned}\hat{s}_n(x) &= a_0 + a_1 \cos\left(\theta(x) + \frac{2\pi(n-1)}{3}\right) + b * \cos\left(2\theta(x) + 2 * \frac{2\pi(n-1)}{3}\right) \\ \hat{d}_n(x) &= a_0 + a_1 \cos\left(\theta(x) + \frac{2\pi(n-1)}{3} + \phi(x)\right) \\ &+ b * \cos\left(2\theta(x) + 2 * \frac{2\pi(n-1)}{3} + 2\phi(x)\right)\end{aligned}$$

$\hat{s}_n(x)$ is the projected signal;

$\hat{d}_n(x)$ is the acquired signal that include the phase variation to measure $\theta(x) = 2\pi f_0 x$;

b is the amplitude of the second harmonic that goes out from the low-pass filter

$\phi(x)$ is the phase that we want to estimate.

To solve we procede considering \bar{S} and \bar{D} as following:

$$\bar{S}_A = -\sum_{n=1}^3 s_n(x) \sin\left(\frac{2\pi(n-1)}{3}\right)$$

$$\bar{S}_B = \sum_{n=1}^3 s_n(x) \cos\left(\frac{2\pi(n-1)}{3}\right)$$

$$D_A = \frac{3}{2} [a_1 \sin(\theta + \phi) - b \sin(2\theta + 2\phi)]$$

$$D_B = \frac{3}{2} [a_1 \sin(\theta + \phi) + b \sin(2\theta + 2\phi)]$$

It continues with the calculation of these quantities that are obtained as shown below:

$$\hat{P} = \left(\hat{S}_A \right)^2 + \left(\hat{S}_B \right)^2 = \frac{9}{2} a_1 b \cos(3\theta)$$

$$\hat{Q} = \left(\hat{D}_A \right)^2 + \left(\hat{D}_B \right)^2 = \frac{9}{2} a_1 b \cos(3\theta + 3\phi(x))$$

Finally the last step to be performed is that of extracting the phase that is executed in this way;

$$\phi(x) = \frac{\text{Im} \left[\ln \left(\hat{Q}^* \hat{P} \right) \right]}{3}$$

This led to a solution that allows a reconstruction of the phase in a more accurate and robust immune to errors that are introduced in a deterministic manner by the measuring instrument, and in particular from the projector of the patterns.

It can be seen in the images shown in Fig.2.14 and Fig.2.15 the improvements obtained by this technique against the classical technique that does not account for this phenomenon.

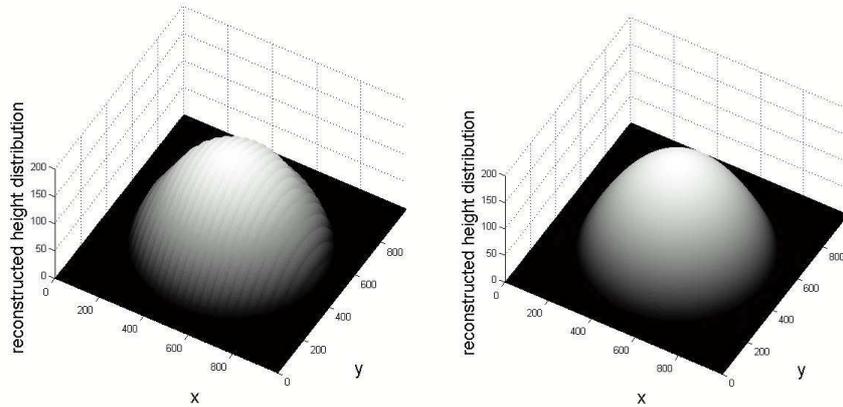


Figure 2.14: Comparison between PSP and I3PSP reconstruction

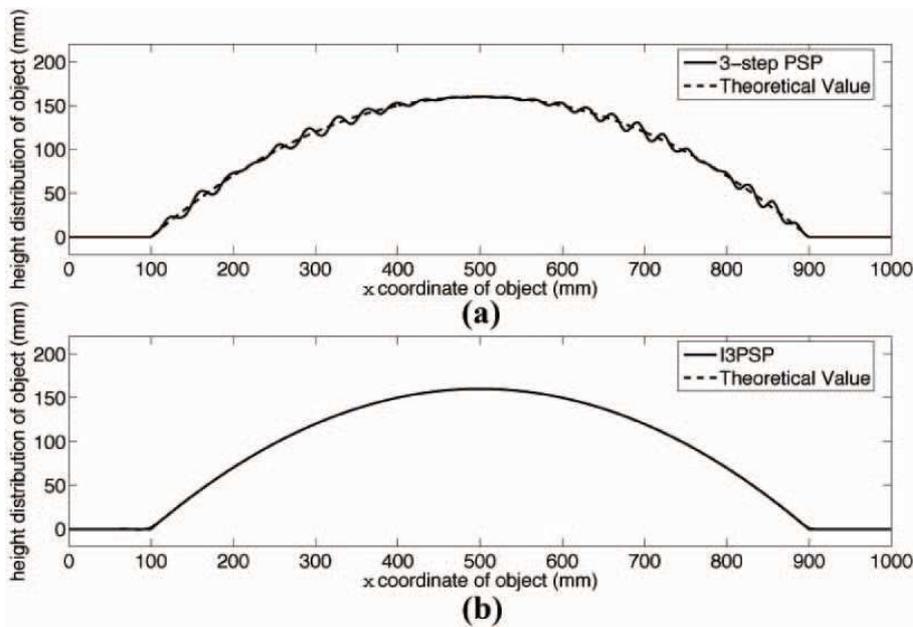


Figure 2.15: Comparison of a row profile between PSP and I3PSP

Saturation problems

The technique of the phase shift used to derive the map of the phase actually provides phase value in the range $[-\pi, \pi]$, i.e. the phase is "rolled". Comes the need to unroll the phase. In literature there are

structured light scanning systems

various techniques to solve such a problem, but all, in order to operate properly, need to be checked the condition that the variation of the true phase between two adjacent pixels is less than π . This implies that we must avoid periods of extremely small, for example, avoid the situation where it goes from black to white or vice versa going from one pixel to its adjacent inevitable if the period of the fringe is equal to only two pixels. Then the number of fringes in fact must be smaller than the half of M , horizontal resolution of the CCD camera.

The considerations made so far appear to be true in the case where the projected and acquired image contain the same number of pixels and the scanned image covers exactly the area illuminated by the projector. Unfortunately this does not happen in reality: we must not forget that the projector and the camera sees the object from two different directions and so it is virtually impossible to achieve this condition. What is done is to place only the central part of the illuminated trying to converge on the center of the illuminated area with the center of the scanned image, when you are acquiring the reference plane.

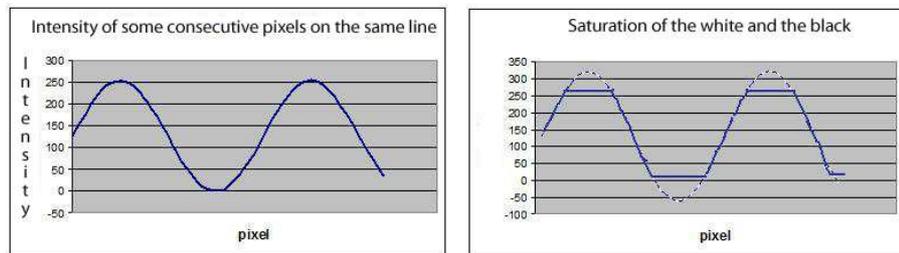


Figure 2.16: Saturation problem

As mentioned now suggests that, in reality, what you acquire is a larger version of the projected image and limited only to its central part. This implies that for a given size of projected image and scanned image, the period of the fringes acquired is greater than the period of the fringes projected and this results in a decrease in sensitivity compared to the expected value.

Remains to be done, however, some consideration on the range of intensity values to be used.

Convenient to use a resolution no more than eight bits, i.e. intensity values between 1 and 255. Increase over this is of no advantage

considering the little time available. Actually should not cover the full range available as it incurs into problems of saturation introduced by the projector and the camera.

In fact what you want to project and then capture is a sinusoidal signal, that is, if we fix a row of the image that we project, it is desired that the intensity values have a projected trend of the type in Figure 2.16 (a).

Because of the gain of the projector and camera and other parameters involved in the optics of the camera, what is acquired may be a sinusoid truncated as shown in figure 2.16 (b). To avoid these drawbacks can also be reduced by 25% the range occupied by the intensity values of the fringes so that they go from 25 to 217.

2.4 Limitations of classical systems

On the basis of what has been described so far, my focus was to improve some features of classical techniques, as they sinned before in some points of view for the applicability in some contexts.

I'll make a simple list of these features:

- The ability to speed up the calibration of a 3D scanner in a single step,
- Make these scanners robust and insensitive to environmental noise and interference caused by bright lights,
- Keep the uncertainty comparable, if not less than traditional systems,
- At least to perform a quick calibration on-site of a scanner pre-calibrated previously in the laboratory.

2.4.1 Calibration one step

Classical calibration procedures include, as mentioned above, that the camera has been calibrated previously respect to the overall system. To obtain a unique calibration procedure for the system, there has been the idea to calibrate the camera at the moment when the trasversal calibration is performed via two planes placed at two different depths of the volume of calibration. In this way we get

structured light scanning systems

intrinsic and extrinsic parameters of the camera that are used at the same time for the transversal calibration of the system. This procedure optimizes a classical technique, to try to change structurally the calibration and then also speed it up, it was being considered an innovative solution that models the camera-projector pair as a pair of stereo cameras, the basic problem is that the projector can not see the calibration targets, this problem will be described and resolved in the next chapter.

2.4.2 Insensitivity to environmental noise

Another problem that is addressed is to reduce problems related to light interference and to higher order harmonics of quantization of the projector. A solution can be found through the generation of simulate phase maps through the parameters obtained by the projector calibration, which previously had never been tested because of the lack of a model and because there was no calibration procedure of the projector before. In this way it is possible to simulate the longitudinal calibration of the projection by simulated sinusoidal pattern, in this way there will be no problem in the quantization of the pixels that it produces by the projector light interference caused by lamps present in the environment. The critical point is the modeling of the projector, but after testing it was concluded that the hypothesized model is correct. A further reduction of errors can be made considering the lens distortion of camera and projector which, as discussed below.

2.4.3 Reduction of uncertainty and faster recalibration

Final issue, that was addressed, was the distortion introduced by the optics. Fundamental problem is that the optical distortion parameters show a variation as a function of target distance, this variation is modeled using empirical equations. Introducing these parameters in the camera-projector model encounters a problem that these parameters vary with the plan that takes into account within the volume of calibration. Since we perform the calculation of these parameters on the two planes at the border of this volume, we got different values of distortion parameters. It was studied the possible

structured light scanning systems

modeling of the variation of this parameter by using advanced models i.e., we have moved beyond the classic pin-hole model and we have evolved into a new model, the axial type. This work was conducted under the supervision of Professor Janne Heikkilä at University of Oulu in Finland. What has been achieved it was a calibration tool that could, through the axial model, include the drift phenomenon of distortion parameters. To conclude the whole work is applied a self-calibration procedure for stereo cameras pairs, to the camera-projector pair in order to make fast recalibration and quickly reconfiguration on site of a pre-calibrated scanner in the laboratory.

Chapter 3

Novel calibration techniques for 3D structured light scanning systems

3.1 Introduction

This chapter will address the problem of calibration in a single step. Will be introduced an innovative model for the projector and camera system and it will be showed the calibration algorithm for a system so modeled. We are showing the main advantages of this innovation consisting, in addition to the fast calibration in a single step, to make the system insensitive to ambient light noise.

To do this we recall, in this introduction, the pinhole camera model which will be applied also to the projector.

3.1.1 Pin hole camera model

The pinhole camera model defines the geometric relationship between a 3D point and its 2D corresponding projection onto the image plane. When using a pinhole camera model, this geometric mapping from 3D to 2D is called a perspective projection. We denote the center of the perspective projection (the point in which all the rays intersect) as the optical center or camera center and the line perpendicular to the image plane passing through the optical center as the optical axis (see Figure 3.1). Additionally, the intersection point of the image plane with the optical axis is called the principal point. The pinhole camera that models a perspective projection of 3D points onto the image plane can be described as follows.

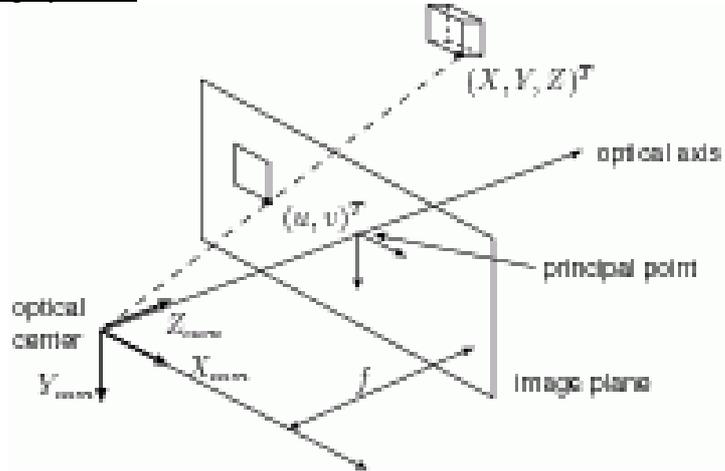


Figure 3.1: Pinhole model, intrinsic parameters

Let us consider a camera with the optical axis being collinear to the Z_{cam} -axis and the optical center being located at the origin of a 3D coordinate system.

The ideal pinhole camera model describes the relationship between a 3D point $(X, Y, Z)^T$ and its corresponding 2D projection (u, v) onto the image plane.

The projection of a 3D world point $(X, Y, Z)^T$ onto the image plane at pixel position $(u, v)^T$ can be written as:

$$u = \frac{Xf}{Z} \text{ and } v = \frac{Yf}{Z},$$

where f denotes the focal length. To avoid such a non-linear division operation, the previous relation can be reformulated using the projective geometry framework, as:

$$(\lambda u, \lambda v, \lambda)^T = (Xf, Yf, Z)^T.$$

This relation can be expressed in matrix notation by:

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix},$$

where $\lambda = Z$ is the homogeneous scaling factor.

Most of the current imaging systems define the origin of the pixel coordinate system at the top-left pixel of the image. However, it was previously assumed that the origin of the pixel coordinate system corresponds to the principal point $(o_x, o_y)^T$, located at the center of the image. A conversion of coordinate systems is thus necessary. Using homogeneous coordinates, the principal-point position can be readily integrated into the projection matrix. The perspective projection equation becomes now:

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & o_x & 0 \\ 0 & f & o_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}.$$

To derive the relation described by the last matrix equation, it was implicitly assumed that the pixels of the image sensor are square, i.e., aspect ratio is 1:1 and pixels are not skewed. However, both assumptions may not always be valid. First, for example, an NTSC TV system defines non-square pixels with an aspect ratio of 10:11. In practice, the pixel aspect ratio is often provided by the image-sensor manufacturer. Second, pixels can potentially be skewed, especially in the case that the image is acquired by a frame grabber. In this particular case, the pixel grid may be skewed due to an inaccurate synchronization of the pixel-sampling process. Both previously mentioned imperfections of the imaging system can be taken into account in the camera model, using the parameters η and τ , which model the pixel aspect ratio and skew of the pixels, respectively. The projection mapping can be now updated as:

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & \tau & o_x & 0 \\ 0 & \eta f & o_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = [\mathbf{K} \quad \mathbf{0}_3] \mathbf{P},$$

with $P = (X,Y,Z,1)^T$ being a 3D point defined with homogeneous coordinates. In practice, when employing recent digital cameras, it can be safely assumed that pixels are square ($\eta = 1$) and non-skewed ($\tau = 0$). The projection matrix that incorporates the intrinsic parameters is denoted as K throughout this thesis. The all zero element vector is denoted by 0 .

As opposed to the intrinsic parameters that describe internal parameters of the camera (focal distance, radial lens parameters), the extrinsic parameters indicate the external position and orientation of the camera in the 3D world. Mathematically, the position and orientation of the camera is defined by a 3×1 vector C and by a 3×3 rotation matrix R (see Figure 3.2).

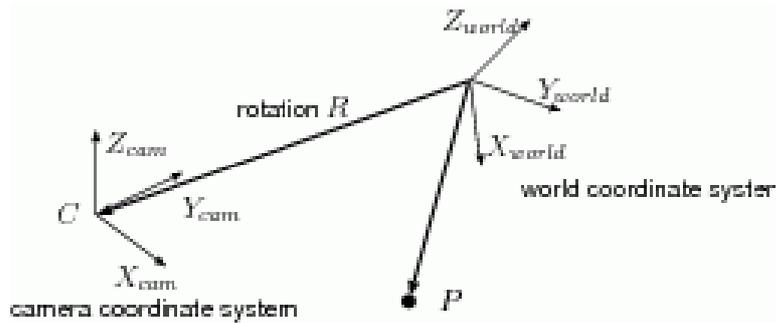


Figure 3.2: Pinhole model, extrinsic parameters

The relationship between the camera and world coordinate system is defined by the camera center C and the rotation R of the camera.

To obtain the pixel position $p = (x,y,1)^T$ of a 3D-world homogeneous point P , the camera should be first translated to the world coordinate origin and second, rotated. This can be mathematically written as:

$$\lambda p = [K | \mathbf{0}_3] \begin{bmatrix} R & 0 \\ \mathbf{0}_3 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{0}_3^T & -C \\ 0 & 1 \end{bmatrix} P.$$

Alternatively, when combining matrices, previous equation can be reformulated as

$$\lambda \mathbf{p} = [\mathbf{K} | \mathbf{0}_3] \begin{bmatrix} \mathbf{R} & -\mathbf{RC} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \mathbf{P} = \mathbf{KR} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \mathbf{KRC}.$$

Previously, the process of projecting a 3D point onto the 2D image plane was described. We now present how a 2D point can be back-projected to the 3D space and derive the corresponding coordinates. Considering a 2D point \mathbf{p} in an image, there exists a collection of 3D points that are mapped and projected onto the same point \mathbf{p} . This collection of 3D points constitutes a ray connecting the camera center $\mathbf{C} = (C_x, C_y, C_z)^T$ and $\mathbf{p} = (x, y, 1)^T$.

From previous equation, the ray $\mathbf{P}(\lambda)$ associated to a pixel $\mathbf{p} = (x, y, 1)^T$ can be defined as:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \underbrace{\mathbf{C} + \lambda \mathbf{R}^{-1} \mathbf{K}^{-1} \mathbf{p}}_{\text{ray } \mathbf{P}(\lambda)},$$

where λ is the positive scaling factor defining the position of the 3D point on the ray. In the case Z is known, it is possible to obtain the coordinates X and Y by calculating λ using the relation

$$\lambda = \frac{Z - C_z}{z_3} \text{ where } (z_1, z_2, z_3)^T = \mathbf{R}^{-1} \mathbf{K}^{-1} \mathbf{p}.$$

The back-projection operation is important for depth estimation and image rendering. For depth estimation, this would mean that an assumption is made for the value of Z and the corresponding 3D point is calculated. With an iterative procedure, an appropriate depth value is selected from a set of assumed depth candidates.

3.1.2 Limits of the application to the projector

The camera calibration takes place as shown in the previous chapter through the acquisition of a target and the solution of a system of equations to obtain the intrinsic and extrinsic parameters of the model just described. The fundamental problem that limits the applicability of this calibration and the model to the projector is due to the fact that

the projector can not identify the aims of the target as it can not acquire the projected scene. As a solution is expected to project a target known in projector coordinates and reverse the process with the proper conventions as discussed below.

3.2 Stereo model for camera-projector pair

The extension of a pin-hole model to a desktop projector is theoretically straightforward, its practical application cannot neglect that it is impossible to know directly the projector coordinates (u_p, v_p) under which the projector illuminates the real point P.

However, the potential advantages of such an approach have urged researchers into designing indirect procedures for stereo-like calibrations of camera-projector pairs. One of the most interesting solutions is described in [33], where the transformation between camera coordinates and projector coordinates is achieved by projecting a time-coded pattern composed of sinusoidal fringes, which is analyzed with the known phase-shifting algorithm [34]. With this technique, the spatial phase of the projected sinusoidal pattern as viewed by the camera at the pixel (u, v) is directly related to one of the projector coordinate (e.g. the up-direction if the fringes are vertical). A complete correspondence between camera coordinates and projector coordinates can be estimated by repeating the procedure both with horizontal and vertical fringes. The proposed algorithm assures a good accuracy, but the procedure requires subsequent steps and the phase unwrapping, necessary for the estimation of the spatial phase, is a time-demanding algorithm. In [35], a time sequence of Gray-coded digital patterns is projected on a target, with both vertical and horizontal stripes, and both positive and negative, in order to estimate a correspondence between projector coordinates and world coordinates. The procedure estimates the homography between projector coordinates and two world coordinates on the plane of the target. The discussion about the accuracy in the estimation of pixel coordinates and the causes of error is very interesting. A drawback of this procedure can be seen in that it requires some steps, and thus a relevant amount of time, for the processing. In [36] the epipolar geometry is exploited in order to calibrate the camera and projector

pair. The proposed procedure requires a preliminary calibration of the camera, and then the relationship between the camera view and the projector is estimated with the normalized 8-point algorithm [37]. Also this procedure is composed of subsequent steps, and the camera has to be calibrated in a separate task.

3.2.1 Projector model

The common pin-hole model for cameras is adopted both for camera and projector. Figure 3.3 shows the reference systems and the basic operation. The projector illuminates a point P in the real world with a ray originating from the projector grid at projector pixel coordinates (u_p, v_p) . The same point P generates an image on the camera sensor plane at camera pixel coordinates (u_c, v_c) . The camera (projector) reference system x_c, y_c, z_c (x_p, y_p, z_p) is centered at the camera (projector) optical centre, respectively, and an absolute “world” coordinate system is located at arbitrary position and orientation in space depending on the specific application. Both the line of sight and the line of projection can be geometrically represented by a perspective projection. If P is at (x, y, z) world coordinates:

$$s \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} k_{uc}f_c & 0 & u_{oc} & 0 \\ 0 & k_{vc}f_c & v_{oc} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Camera intrinsic parameters}} \underbrace{\begin{bmatrix} \underline{R}_c & \underline{t}_c \\ \underline{0} & 1 \end{bmatrix}}_{\text{Camera extrinsic parameters}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Camera matrix \underline{P}_c

The arbitrary parameter s takes into account the indetermination of the projection. The parameter k_{ufc} (k_{vfc}) is the horizontal (vertical) focal length expressed in pixels, respectively, the principal point (u_{oc}, v_{oc}) is the pixel where the optical axis z_c of the camera intersects the sensor plane, the 3x3 orthogonal matrix \underline{R}_c and the 3x1 vector \underline{t}_c are the rotation and the translation, respectively, relating the camera reference system x_c, y_c, z_c and the world reference system x, y, z .

Definitions for the corresponding parameters of projector are similar, and the relationship is as follows:

$$l \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} k_{up}f_p & 0 & u_{op} & 0 \\ 0 & k_{vp}f_p & v_{op} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Projector intrinsic parameters}} \underbrace{\begin{bmatrix} R_p & t_p \\ \underline{0} & 1 \end{bmatrix}}_{\text{Projector extrinsic parameters}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Projector matrix P_p

If previous equations described an actual pair of cameras, intrinsic and extrinsic parameters of both devices could be estimated through some calibration algorithm.

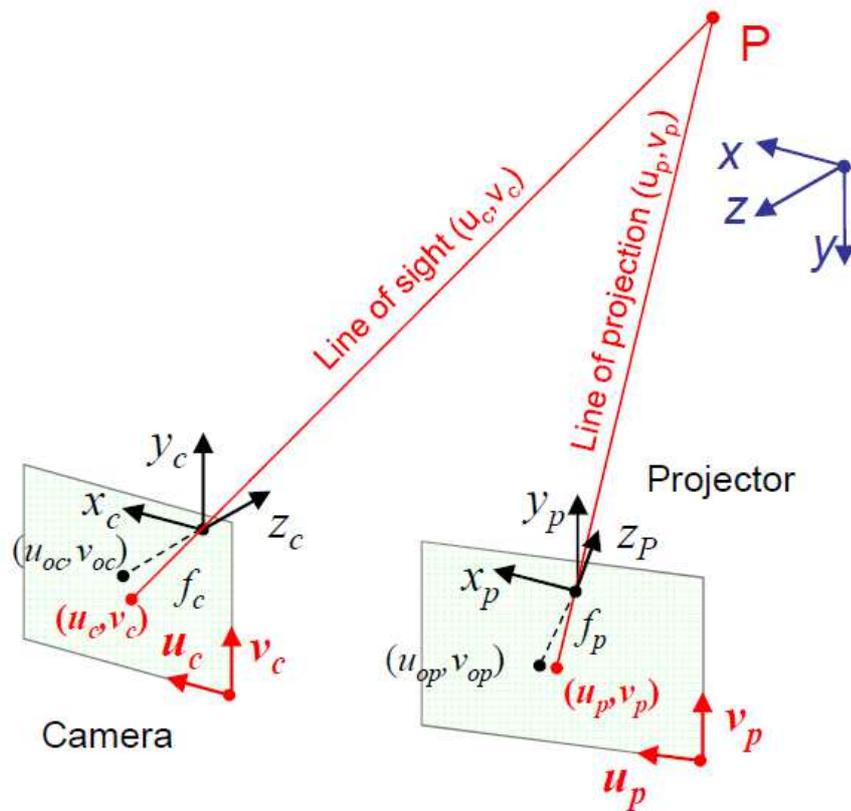


Figure 3.3: Model for camera and projector

One of the most common calibration algorithms is the Direct Linear Transformation (DLT) [1], which basically requires the estimation of pixel coordinates (u_{ci}, v_{ci}) and (u_{pi}, v_{pi}) of a set of calibration points at world coordinates (x_i, y_i, z_i) and the simultaneous solution for the unknown parameters of an overdetermined linear system built by stacking four equations composed of a pair of both matricial equations for each calibration point.

Depending on the algorithm, subsequent elaborations could refine the estimations of camera parameters, which are nonoptimal when input data are noisy, and possibly correct them for non-linearity of lenses. On the other hand, a triangulation algorithm basically measures the world coordinates of P by solving the matricial equations for x, y, z, being known (u, v) , (u_p, v_p) and all the parameters of the pin-hole models.

3.2.2 Calibration algorithm

The novel pre-calibration procedure we describe yields three paired sets of coordinates of the same array of calibration points: the world coordinate set (x_i, y_i, z_i) , the camera-pixel coordinates (u_{ci}, v_{ci}) and the projector pixel coordinates (u_{pi}, v_{pi}) . Since these are the data required by most stereo-vision calibration algorithms, the subsequent estimation of intrinsic and extrinsic parameters of camera and projector through any camera calibration algorithm is straightforward, even though results will be shown for a DLT calibration.

One of the advantages of this pre-calibration procedure was that the camera and the projector are calibrated at the same time.

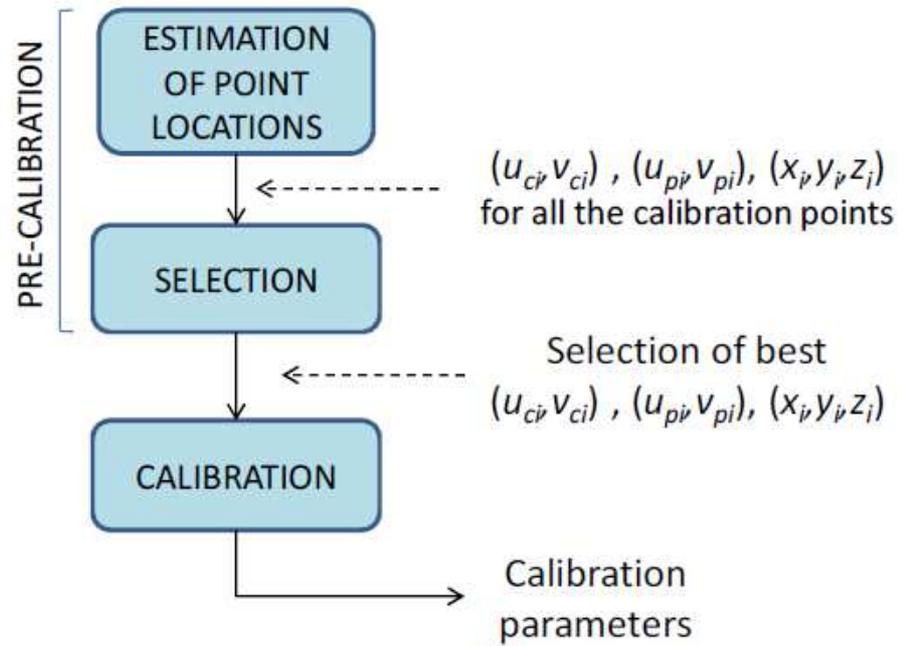


Figure 3.4: The proposed calibration procedure

The overall structure of the proposed method of calibration is reported in Figure 3.4. The pre-calibration module yields the three paired sets of coordinates of calibration points. The precalibration includes a selection module which eliminates possible outliers, namely those points which may have been localized with poor accuracy, and then the actual calibration algorithm is run in order to evaluate calibration parameters.

These modules will be described in this section.

The pre-calibration procedure

A novelty of the procedure is that it was designed in order to estimate the three sets of coordinates for the centers of an array of projected calibration points instead of an array of real target points. In fact, a planar calibration target with printed circles at known world coordinates is adopted during the procedure, but the coordinates of real target points are used only for intermediate processing steps. Since the calibration algorithm chosen to be run after the pre-calibration, as most calibration algorithms, requires a set of non-

coplanar calibration points, two parallel poses of the planar target will be processed in order to collect a valid set of calibration points. Thus the pre-calibration procedure described in the following has to be executed twice actually, once for each pose of the target. The steps of the proposed pre-calibration procedure are reported correspondently in Figure 3.5.

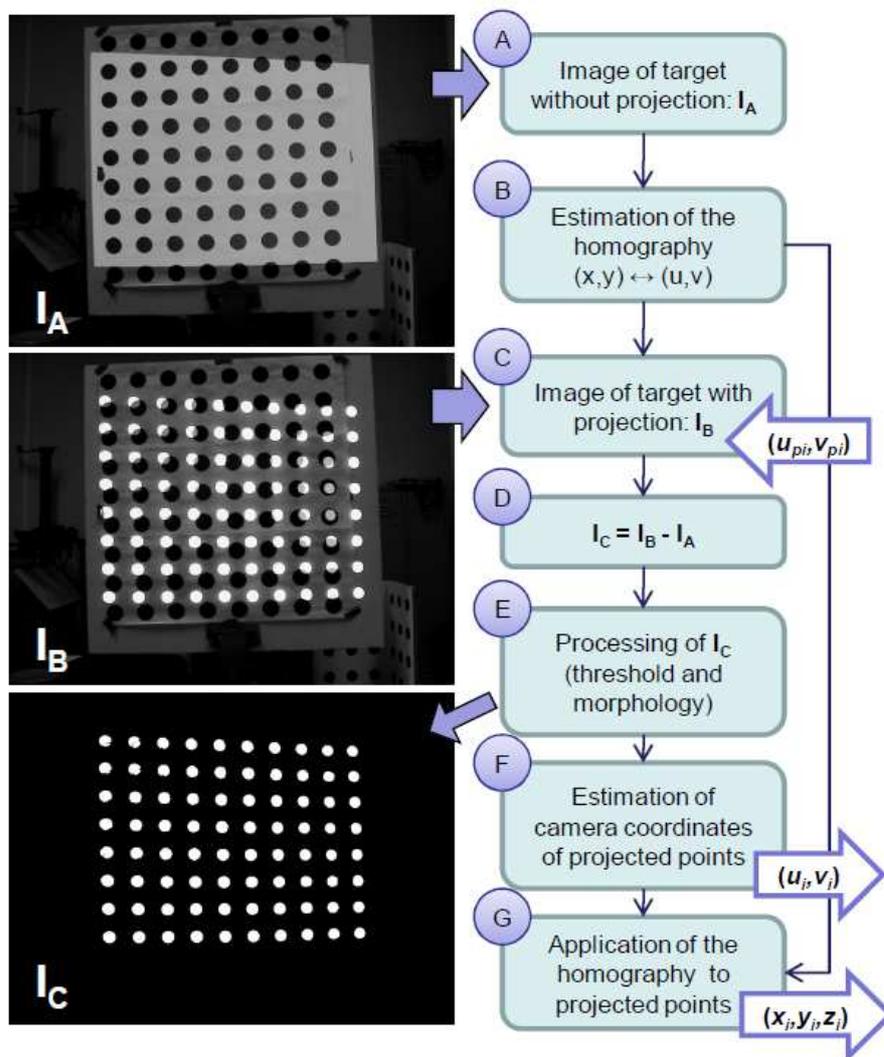


Figure 3.5: The proposed pre-calibration procedure

A. The image I_A of the planar target with real accurately printed circles is acquired while the projector is switched off. The transformation between the world target coordinates on the target plane and the camera pixel coordinates is estimated from this image. Said (x_{ti}, y_{ti}, z_{ti}) the world coordinates of the target points and (u_{ti}, v_{ti}) the camera pixel coordinates of the target points as seen from the camera, respectively, on the target plane at a known $z_{ti} = \text{constant}$, the 2-D projective transformation can be described as a homography:

$$[u_{ti} \quad v_{ti} \quad 1]^T = \underline{H}[x_{ti} \quad y_{ti} \quad 1]^T$$

B. The 3x3 matrix \underline{H} is estimated with the DLT technique since a linear system with two equations per each calibration point can be written from previous relation and solved for the unknown elements of \underline{H} .

C. A pattern of white circles at known projector pixel coordinates (u_{pi}, v_{pi}) is projected onto the target and a new image, I_B , is acquired.

D. Since projected circles may superimpose onto printed circles in I_B , centroids of projected circles cannot be estimated directly on I_B . In order to isolate the projected circles from the printed ones, a third image is calculated with pixel-by-pixel subtraction:

$$I_C = I_B - I_A$$

E. A thresholding and a morphological opening is applied in order to regularize the images of circles, which are expected to have elliptical shape.

F. The image pixel coordinates (u_{ci}, v_{ci}) of the projected circles are estimated on the resulting image I_C (see Figure 3.5) as centroids of the elliptical shapes.

G. The world coordinates (x_i, y_i, z_i) of projected target points are estimated by applying the inverse homography given by the first one and evaluated in step B to the image pixel coordinates (u_{ci}, v_{ci}) :

$$[x_i \quad y_i \quad 1]^T = \underline{H}^{-1}[u_i \quad v_i \quad 1]^T$$

for each calibration point. The three arrows of Figure 3.5 show where the three sets of coordinates of the same calibration points appear in the procedure.

The selection of target points and the calibration

The images of projected circles in the difference image I_C should appear as ellipses in ideal conditions, but in some cases their shapes may be altered by artifacts mainly when projected circles were partially superimposed on printed circles, as shown in Fig. 3.6.

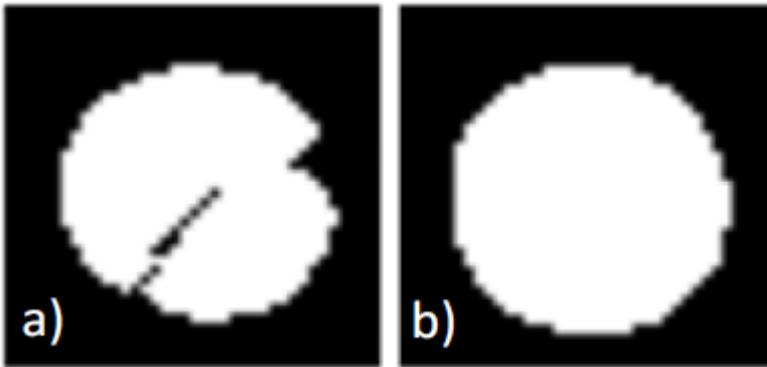


Figure 3.5: Calibration circles in image I_C a)with artifacts; b)correct

Since the reference point of a calibration circle is estimated as the centroid of the image of the circle, the artifacts may cause errors in the localization of target points. To overcome this problem, the subsequent calibration can be performed on a subset of those target points minimizing a given error function. After the calibration, given the world coordinates (x_i, y_i, z_i) of a target points, we can calculate the “re-projected” estimations (\hat{u}_i, \hat{v}_i) and $(\hat{u}_{pi}, \hat{v}_{pi})$ of camera and projector pixel coordinates, respectively. Then the average distance between measured and re-projected pixel coordinates, called the re-projection error, can be used as the error function to be minimized:

$$e = \frac{1}{2N} \sum_{i=1}^N \left[\sqrt{(u_i - \hat{u}_i)^2 + (v_i - \hat{v}_i)^2} + \sqrt{(u_{pi} - \hat{u}_{pi})^2 + (v_{pi} - \hat{v}_{pi})^2} \right]$$

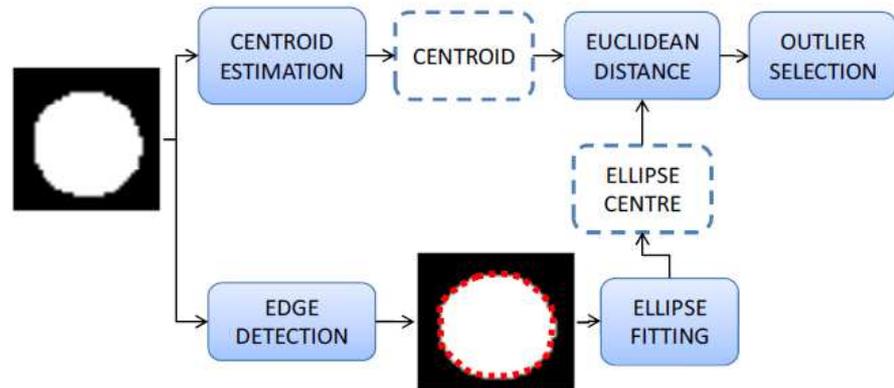


Figure 3.6: The algorithm for the selection of target points

A common approach in computer vision problems is to select target points by using RANSAC (RANDOM Sample Consensus) [38] algorithms, which randomly build subsets of target points, selecting the subset minimizing an error function.

The disadvantage of these algorithms is that the required elaboration time is very high for typical numbers N of points, and they cannot be adopted when the elaboration time is a constraint. For this reason, an alternative selection algorithm has been applied, whose block diagram is depicted in Figure 3.6.

The idea is to compare the centroid with an alternative estimation of the centre of the circle image, and discard the target point if the distance between them is statistically significant. In this case, such an alternative estimation is the centre of the ellipse fitted to the edge points sampled along the contour of the white region [39]. As an example, a plot of the observed distances between centroids and fitting ellipse centres is reported in Figure 3.7, where three points show distances significantly greater than the threshold t_s determined with a t-Student test, and then are discarded. The subsequent estimation of camera and projector intrinsic and extrinsic parameters is performed by a DLT algorithm, as previously said.

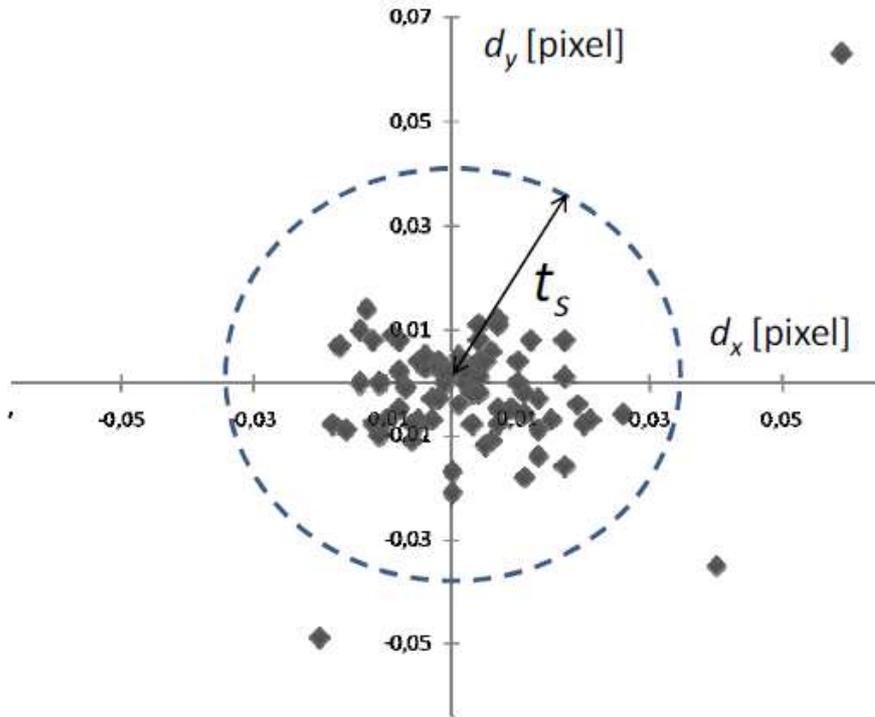


Figure 3.6: Distances between fitting ellipse centres and centroids

3.2.3 DLT

Direct linear transformation (DLT) is an algorithm which solves a set of variables from a set of similarity relations:

$$\mathbf{x}_k \propto \mathbf{A} \mathbf{y}_k \text{ for } k = 1, \dots, N$$

In our case we have to solve this relationship

$$\lambda \tilde{\mathbf{m}}_i = \underline{\mathbf{P}} \tilde{\mathbf{M}}_i$$

Where $\underline{\mathbf{M}}_i = (x_i, y_i, z_i)$ are the coordinates of the target points estimated by the pre-calibration procedure previous described, and $\underline{\mathbf{m}}_i = (u_i, v_i)$ are the coordinates of the pixel located on the image sensor and that are evaluated through the image processing, this is shown in the Figure 3.7.

The camera matrix $\underline{\mathbf{P}}$ is the least-square solution of the linear system that is obtained as follows.

The final linear system that we obtain from all the aim correspondances are:

$$\begin{cases} \dots & \dots & \dots \\ \underline{p}_1^T \underline{M}_i - u_i \underline{p}_3^T \underline{M}_i = 0 \\ \underline{p}_2^T \underline{M}_i - v_i \underline{p}_3^T \underline{M}_i = 0 \\ \dots & \dots & \dots \end{cases}$$

Then \underline{P} can be factorized and the model parameters can be calculated from:

$$\underline{A} \underline{p} = \underline{0}$$

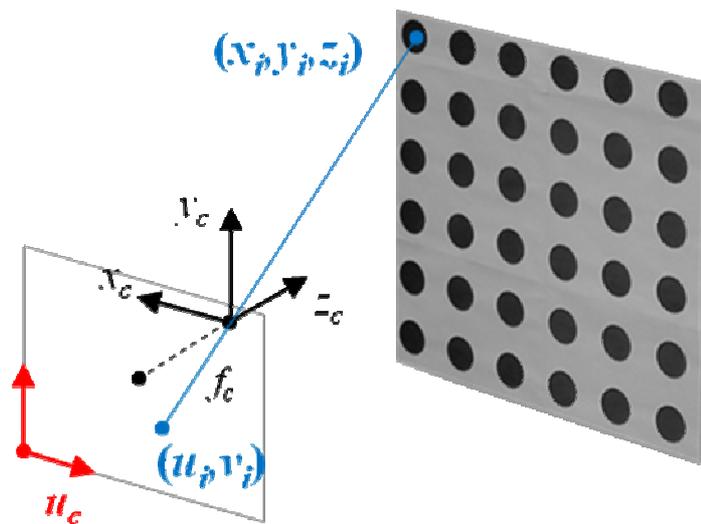


Figure 3.7: DTL in Stereo-like system

We can find the solution, then the parameters, as last-mean square solution of:

$$\min_p \|\underline{A} \underline{p}\|_2$$

It solves the least squares decomposing the matrix A using SVD decomposition

$$\underline{A} = \underline{U} \underline{S} \underline{V}^T \quad \begin{cases} \underline{U} \text{ orthogonal } 2N \times 2N, \\ \underline{S} \text{ diagonal } 2N \times 12 \\ \underline{V} \text{ orthogonal } 12 \times 12 \end{cases}$$

3.3 Stereo calibration in a fringe pattern profilometer

Once obtained the parameters of both devices, one can replace the classic longitudinal calibration procedure with an innovative one which provides the calculation of simulated phase maps in order to overcome the problems of the light interference and harmonics quantizer generates in output from the projector. And we proceeded as follows

3.3.1 Simulated phase maps

Starting from the knowledge of the perspective projection matrix \underline{P}_c and \underline{P}_p of camera and projector is possible to calculate the phase maps $\phi(u,v)$ to a generic Z .

If we impose the perspective equations of camera (C) and projector (P) as:

$$\begin{aligned} s \underline{m}_C &= \underline{P}_C \underline{M} \\ b \underline{m}_P &= \underline{P}_P \underline{M} \end{aligned}$$

where \underline{M} is a generic point on the plane where we want to calculate the phase map.

From the second equation

$$\underline{M} = \underline{P}_C^{-1} (s \underline{m}_C)$$

And replacing into the first one

$$b \underline{m}_P = \underline{P}_P \underline{P}_C^{-1} (s \underline{m}_C)$$

The two scale factors can incorporate into one $g = s/l$:

So:

$$\underline{m}_P = \underline{P}_P \underline{P}_C^{-1} (g \underline{m}_C)$$

We can now calculate, using a ray tracing procedure, the beam from the projector to the left point M and then simulate the image acquired by the camera.

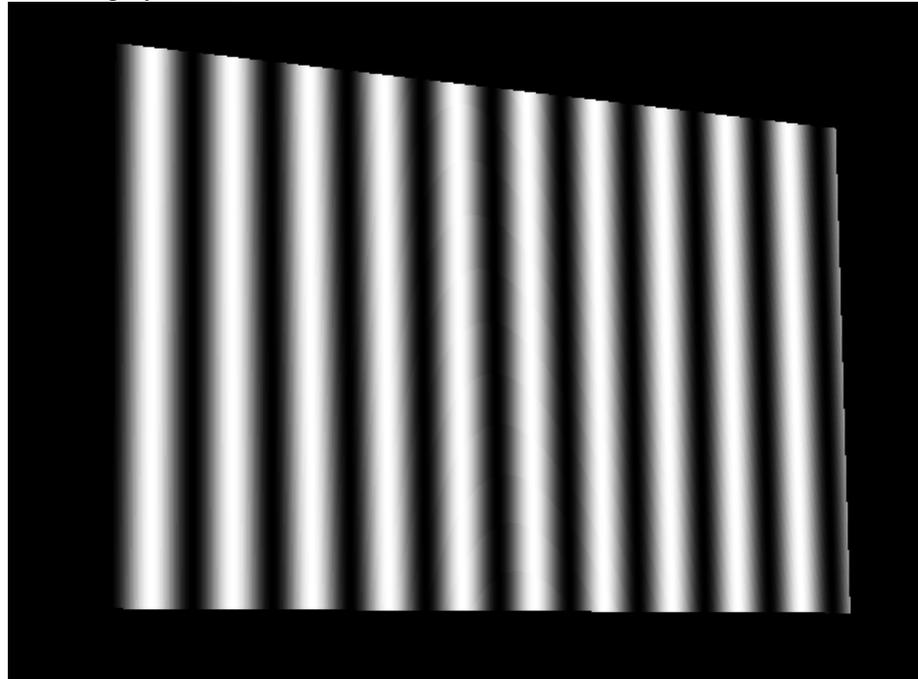


Figure 3.8: Simulated image acquired for the longitudinal calibration

The objective is to assign a value of gray, or color phase to all the pixels of an image to be synthesized (or simulated) of fixed size, using a ray tracing procedure.

The procedure of ray tracing scans, with two loops For, of all the $m_K(u, v)$ of the image, and for each m_K :

- calculation of the pixel-projector m_P from which started the beam illuminated the point M, then framed by the camera pixel coordinates m_K .

- the pixel of coordinates m_K of the image under construction is assigned the value of gray, color, or phase that the pattern projected had to the coordinate m_P .

To do this we have to consider the follow relations:

Camera model, with the real point M on the plane at $z=z_K$

$$s \begin{bmatrix} u_k \\ v_k \\ 1 \end{bmatrix} = \begin{bmatrix} q_{11}^k & q_{12}^k & q_{13}^k & q_{14}^k \\ q_{21}^k & q_{22}^k & q_{23}^k & q_{24}^k \\ q_{31}^k & q_{32}^k & q_{33}^k & q_{34}^k \end{bmatrix} \begin{bmatrix} x \\ y \\ z_k \\ 1 \end{bmatrix}$$

Projector model, with the real point M on the plane at $z=z_K$

$$b \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} q_{11}^b & q_{12}^b & q_{13}^b & q_{14}^b \\ q_{21}^b & q_{22}^b & q_{23}^b & q_{24}^b \\ q_{31}^b & q_{32}^b & q_{33}^b & q_{34}^b \end{bmatrix} \begin{bmatrix} x \\ y \\ z_k \\ 1 \end{bmatrix}$$

The values of the transversal coordinates of $M=(x,y,z_K)$ can be found solving the following system, that we can obtain rewriting the first system into the variables x, y, s (the z_K is known and the s doesn't need to be evaluated)

$$\begin{bmatrix} q_{11}^k & q_{12}^k & -u_k \\ q_{21}^k & q_{22}^k & -v_k \\ q_{31}^k & q_{32}^k & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ s \end{bmatrix} = \begin{bmatrix} -q_{13}^k z_k - q_{14}^k \\ -q_{23}^k z_k - q_{24}^k \\ -q_{33}^k z_k - q_{34}^k \end{bmatrix}$$

From the second system the projector coordinates can be found:

$$\frac{bu_p}{b} = u_p = \frac{q_{11}^p x + q_{12}^p y + q_{13}^p z_k + q_{14}^p}{q_{31}^p x + q_{32}^p y + q_{33}^p z_k + q_{34}^p}$$

$$\frac{bv_p}{b} = v_p = \frac{q_{21}^p x + q_{22}^p y + q_{23}^p z_k + q_{24}^p}{q_{31}^p x + q_{32}^p y + q_{33}^p z_k + q_{34}^p}$$

Once the coordinates of the projector are found (u_p, v_p), it's possible to use as index in the projected pattern image, to determine the value of the intensity projected on M and that the camera seen at (u_c, v_c). Then, to simulate the acquisition of an image, the intensity, that the projected pattern has at (u_p, v_p) coordinates, must be write at (u_c, v_c) coordinates of the image.

3.3.2 Insensitivity to environmental noise

The preceding leads to the calibration of structured light 3D systems, which become unresponsive to light noise like the environment in which the calibration is done. Implement phase maps simulated from the parameters of both the devices that were not available before, leads to implementation of the longitudinal calibration without projecting any real pattern, the only thing you need are the parameters of the two devices depending on the model chosen and the reference plane on which you then run the scan. This calibration can be longitudinal, generating as many acquisitions are to be within the volume of calibration to then proceed to the calculation of longitudinal parameters as follows.

As seen above, the phase a pattern which can be well approximated as linear as a function of depth along the lines of sight, and for this reason we hypothesize to be able to write the following linear relationship:

$$z_m = m_1(u, v)\varphi_m(u, v) + m_0(u, v)$$

where:

z_m is the depth to be measured

φ_m is the measured phase in u,v coordinates, along the lines of sight of the camera.

What should be done to calibrate the system longitudinally is to find the coefficients m_1 and m_0 that bind the phase to the depth along the lines of sight.

The simplified model thus allows the calculation of these parameters due to the longitudinal resolution of the following system of equations

$$\begin{cases} z_0 = m_1(u, v)\varphi_0(u, v) + m_0(u, v) \\ z_1 = m_1(u, v)\varphi_1(u, v) + m_0(u, v) \end{cases}$$

Imposing $z_0 = 0$ and $z_1 = d$ reach the following results for the longitudinal parameters:

$$\begin{cases} m_1(u, v) = \frac{d}{\varphi_1(u, v) - \varphi_0(u, v)} \\ m_0(u, v) = \frac{-d\varphi_0(u, v)}{\varphi_1(u, v) - \varphi_0(u, v)} \end{cases}$$

To calculate the parameters it's necessary to simulate at least two longitudinal phase maps at two different depths at a known distance and then solve the system for each pixel of the image and save the parameters obtained from a text file.

Regarding the transversal calibration, it is performed as mentioned in the previous chapter from the knowledge of the camera parameters obtained in this case by the stereo calibration.

Chapter 4

Advanced models for reconfigurable scanner

4.1 Introduction

Now that we have moved the classical problems in the calibration only on the choice of projector and camera parameters, the critical point becomes to estimate these parameters from the models. So far we have reasoned with ideal models that do not bring into account factors such as lens distortion and all the secondary factors such as chromatic aberrations and effects that usually introduce negligible nonlinearity in classical models.

At this point we have to consider and then introduce these effects so that the modeling of objects such as camera and projector is to get closer to reality, so as to be able to reduce the uncertainties associated with non-idealities previously considered.

Which will be treated in this chapter, is the distortion introduced by optics such as for the real camera or projector. After a review on the models of distortion will be seen as such models have the disadvantage that distortion parameters of an optics show a non-linear variation with different distances from the target, this variation is taken into account in the literature by empirical models. What I intended, in this chapter, is to search for the cause of this trend that led me to consider models for advanced optics, which after some testing I have confirmed that they can be used to model this variation. Finally, as the last issue, is addressed the choice to apply the self-calibration to recalibrate a pre-calibrated scanner rigidly mounted on site that was previously calibrated in the laboratory.

4.1.1 How to reduce the uncertainty?

Reducing uncertainty is a key point, as ideal models using non-negligible errors are introduced during the reconstruction of point clouds.

In order to do this, I focused my attention on the distortion problem and the possibility to refine the calibration parameters with a procedure of selfcalibration that gets more robustness to the final and innovative calibration

4.1.2 Introduce the lens distortion parameters

Introduce the effects of distortion of the lens approaching the model to reality, in this way the fast calibration technique proposed can become more precise offering high accuracy and a short time.

It will be shown the classical distortion model and the drawback applied to the Scanner calibration problems. It will be introduced a new type of modelling the camera and projector optic to fix the problem that the classical distortion model has.

4.1.3 Faster recalibration

At the end of this chapter will be introduced a procedure of self-calibration on the stereo system of camera-projector pair.

In this way, using an iterative optimization that starts from the previous calculated parameters, we can obtain a refinement that is based on the minimization of the mean epipolar error by the Neadel Mead simplex method.

4.2 Lens distortion parameters

We show first the classical model of lens distortion and then proceed to the problem addressed in the calibration of the scanner due to drift of these parameters within the volume of calibration.

4.2.1 Classical approach to the problem

Real camera lenses typically suffer from non-linear lens distortion. In practice, radial lens distortion causes straight lines to be mapped as curved lines. As seen in Figure 4.1, the radial lens distortion appears more visible at the image edges, where the radial distance is high. A standard technique to model the radial lens can be described as follows.

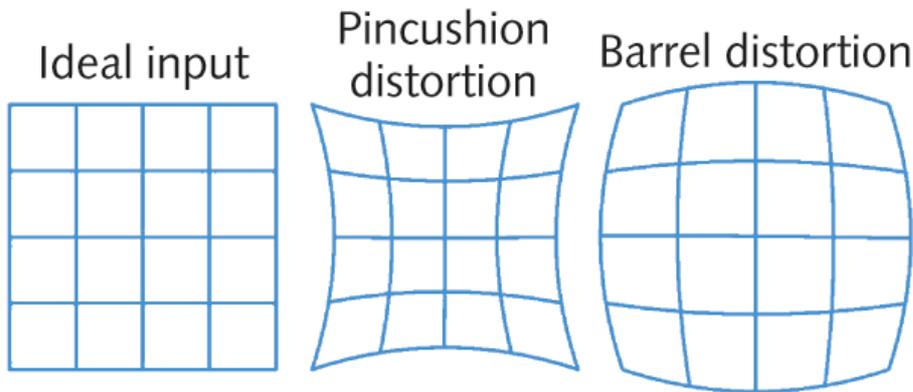


Figure 4.1: Examples of lens distortions

Let $(x_u, y_u)^T$ and $(x_d, y_d)^T$ be the corrected and the measured distorted pixel positions, respectively. The relation between an undistorted and distorted pixel can be modeled with a polynomial function and can be written as

$$\begin{pmatrix} x_u - o_x \\ y_u - o_y \end{pmatrix} = L(r_d) \begin{pmatrix} x_d - o_x \\ y_d - o_y \end{pmatrix},$$

where

$$L(r_d) = 1 + k_1 r_d^2 \quad \text{and} \quad r_d^2 = (x_d - o_x)^2 + (y_d - o_y)^2.$$

In the case $k_1 = 0$, it can be noted that $x_u = x_d$ and $y_u = y_d$, which corresponds to the absence of radial lens distortion.

It should be noted that the first equation provides the correct pixel position using a function of the *distorted pixel position*. However, to generate an undistorted image, it would be more convenient to base the function $L(r)$ on the *undistorted pixel position*. This technique is usually known as the *inverse mapping* method. The inverse mapping technique consists of scanning each pixel in the output image and re-

sampling and interpolating the correct pixel from the input image. To perform an inverse mapping, the inversion of the radial lens distortion model is necessary and can be described as follows. First, similar to the second part of second equation, we define

$$r_u^2 = (x_u - o_x)^2 + (y_u - o_y)^2.$$

Then, taking the norm of first equation it can be derived that

$$(x_u - o_x)^2 + (y_u - o_y)^2 = L(r_d) \cdot ((x_d - o_x)^2 + (y_d - o_y)^2),$$

which is equivalent to

$$r_u = L(r_d) \cdot r_d.$$

this equation can be rewritten as a cubic polynomial:

$$r_d^3 + \frac{1}{k_1} r_d - \frac{r_u}{k_1} = 0.$$

The inverted lens distortion function can be derived by substituting last equation into the first one and developing it from the right-hand side:

$$\begin{pmatrix} x_d - o_x \\ y_d - o_y \end{pmatrix} = \frac{r_d}{r_u} \begin{pmatrix} x_u - o_x \\ y_u - o_y \end{pmatrix},$$

where r_d can be calculated by solving the cubic polynomial function. This polynomial can be solved using Cardano's method, by first calculating the discriminant Δ defined as $\Delta = q^2 + 427p^3$ where $p = 1/k_1$ and $q = -r_u/k_1$. Depending on the sign of the discriminant, three sets of solutions are possible.

If $\Delta > 0$, then the equation has one real root r_{d1} defined as

$$r_{d1} = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} + \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}}.$$

If $\Delta < 0$, then the equation has three real roots r_{dk} defined by

$$r_{dk} = 2\sqrt{\frac{-p}{3}} \cos\left(\frac{\arccos\left(\frac{-q}{2}\sqrt{\frac{27}{-p^3}}\right) + 2k\pi}{3}\right),$$

for $k = \{0,1,2\}$, where an appropriate solution r_{dk} should be selected such that $r_{dk} > 0$ and $r_{dk} < r_{uk}$. However, only one single radius corresponds to the practical solution. Therefore, the second case $\Delta < 0$ should not be encountered. The third case with $\Delta = 0$ is also impractical. In practice, we have noticed that, indeed, these second and third cases never occur.

As an example, Figure 4.2 depicts a distorted image and the corresponding corrected image using the inverted mapping method, with $\Delta > 0$.

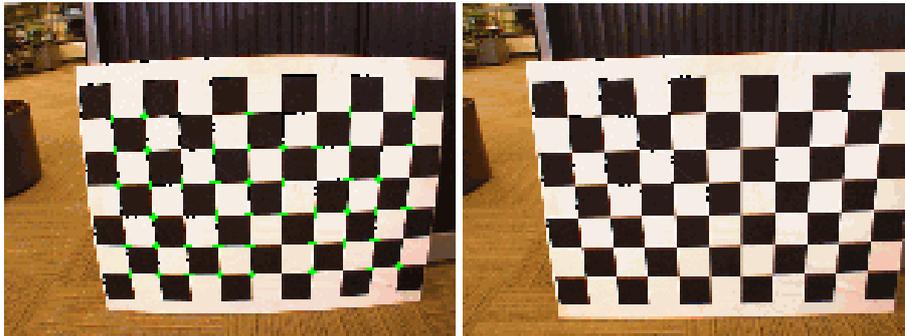


Figure 4.2: Distorted and corrected images

Estimation of the distortion parameters

The discussed lens-distortion correction method requires knowledge of the lens parameters, i.e., k_1 and $(o_x, o_y)^T$. The estimation of the distortion parameters can be performed by minimizing a cost function that measures the curvature of lines in the distorted image. To measure this curvature, a practical solution is to detect feature points belonging to the same line on a calibration rig, e.g., a checkerboard calibration pattern (see Figure 4.2). Each point belonging to the same line in the distorted image forms a bended line instead of a straight line. By

comparing the deviation of the bended line from the theoretical straight line model, the distortion parameters can be calculated.

Unusually modeling

Real lenses may deviate from precise radial symmetry and therefore we supplement our model with an asymmetric part. For instance, the lens elements may be inaccurately aligned causing that the projection is not exactly radially symmetric. With conventional lenses this kind of distortion is called decentering distortion. However, there are also other possible sources of imperfections in the optical system and some of them may be difficult to model. For example, the image plane may be tilted with respect to the principal axis or the individual lens elements may not be precisely radially symmetric. Therefore, instead of trying to model all different physical phenomena in the optical system individually, it is usually proposed a flexible mathematical distortion model that is just fitted to agree with the observations.

To obtain a widely applicable, flexible model, it proposes to use two distortion terms as follows. One distortion term acts in the radial direction

$$\Delta_r(\theta, \varphi) = (l_1\theta + l_2\theta^3 + l_3\theta^5)(i_1 \cos \varphi + i_2 \sin \varphi + i_3 \cos 2\varphi + i_4 \sin 2\varphi)$$

and the other in the tangential direction

$$\Delta_t(\theta, \varphi) = (m_1\theta + m_2\theta^3 + m_3\theta^5)(j_1 \cos \varphi + j_2 \sin \varphi + j_3 \cos 2\varphi + j_4 \sin 2\varphi)$$

where the distortion functions are separable in the variables θ and Φ . Because the Fourier series of any 2π -periodic continuous function converges in the L2-norm and any continuous odd function can be represented by a series of odd polynomials we could, in principle, model any kind of continuous distortion by simply adding more terms to previous equations, as they both now have seven parameters.

By adding the distortion terms, we obtain the distorted coordinates $\mathbf{x}_d = (x_d; y_d)^T$ by

$$\mathbf{x}_d = r(\theta)\mathbf{u}_r(\varphi) + \Delta_r(\theta, \varphi)\mathbf{u}_r(\varphi) + \Delta_t(\theta, \varphi)\mathbf{u}_\varphi(\varphi)$$

This model is usually not used because the tangential distortion terms are as much as two orders of magnitude lower than radial distortion terms and they are generally neglected.

4.2.2 Model for the variation of lens distortion parameter

The knowledge of a distance or magnification dependence of lens distortion [40] is older than 50 years. In 1955 Magill published his work about variation in distortion with magnification [41], where he mentioned the phenomenon of a changing amount of distortion depending on the working distance to the viewed object.

Let the magnification m_s for a distance s be defined as

$$m_s = \frac{f}{(s - f)}$$

where f is the focal length and denote by $M_s = 1/m_s$ the inverse magnification. In the following distortion should mean radial distortion if not otherwise said.

Magill developed a formula [41] in order to calculate the distortion at any arbitrary distance or magnification which was extended by Brown [42] who stated a formula for the situation of known two radial distortion values Δr_{s_1} and Δr_{s_2} for the two distances s_1 and s_2 in order to predict the distortion value at any arbitrary focus distance s :

$$\Delta r_s = \alpha_s \Delta r_{s_1} + (1 - \alpha_s) \Delta r_{s_2}$$

with

$$\alpha_s = \frac{s_2 - s}{s_2 - s_1} \cdot \frac{s_1 - f}{s - f}$$

This formula was further modified and has been verified experimentally for radial and decentering distortion for certain conventional film camera. Brown developed an extended model in order to describe the radial distortion variation outside the plane of best focus by a scaling factor γ_{ss} .

$$\Delta r_{ss'} = \frac{1}{\gamma_{ss'}} \Delta r_{s'}$$

where $\Delta r_{ss'}$ is the radial distortion at an object distance s' for a lens focussed at an object distance s and Δr_s the radial distortion at an object distance s' for a lens focussed at an object distance s' . The scaling factor $\Delta \gamma_{ss'}$ is given by

$$\gamma_{ss'} = \frac{s'(s-f)}{s(s'-f)}$$

Fraser and Shortis [43] suggest the introduction of an empirically determined correction factor:

$$\Delta r_{ss'} = \Delta r_s + g_{ss'}(\Delta r_{s'} - \Delta r_s)$$

where $g_{ss'}$ is an empirically derived constant value and Δr_s the radial distortion at an object distance s for a lens focussed at an object distance s . Another suggestion is given by Dold [44] who suggests a set of parameters which may be completely determined within a bundle adjustment process:

$$\Delta r_{dist} = \frac{1}{Z^*} \left[D_1 r' (r'^2 - r_0^2) + D_2 r' (r'^4 - r_0^4) + D_3 r' (r'^6 - r_0^6) \right]$$

where $Z^* \sim s$. For more details see [44, 45]. Brakhage introduces a method for the consideration of distance dependence of distortion effects for fringe projection systems with telecentric projection lenses using Zernike-polynomials [46].

In principal, for the description of lens distortion several different models are possible. The key point is that the chosen model fits sufficiently the actual occurring distortion.

4.3 Advancend Geometric Camera model

After a study of the geometric models for cameras[47],[48] that bases its modeling on the path of the rays that from the real world through the lens is impacting on the camera sensor, has been able to verify that the use of such advanced models can take account of changes of the parameters of distorting the previously described

I will briefly reported the geometric models developed for camera and then explain how it was possible to confirm the hypothesis that was supposed.

We have to introduce the Plücker Coordinates to understande the geometrical route of the rays passing through the optical system.

Plücker Coordinates

We represent projection rays as 3D lines, via Plücker coordinates. Several definitions exist for them; we use the following. Let \mathbf{A} and \mathbf{B} be the homogeneous coordinates of 3D points defining a line. The line can be represented by the skew-symmetric 4 x 4 Plücker matrix $\mathbf{L} = \mathbf{A}\mathbf{B}^T - \mathbf{B}\mathbf{A}^T$.

It is independent (up to scale) of the points used to represent the line. An alternative representation for the line is its Plücker coordinate vector of length 6:

$$\mathbf{L} = \begin{pmatrix} A_4B_1 - A_1B_4 \\ A_4B_2 - A_2B_4 \\ A_4B_3 - A_3B_4 \\ A_3B_2 - A_2B_3 \\ A_1B_3 - A_3B_1 \\ A_2B_1 - A_1B_2 \end{pmatrix}$$

We sometimes split it in two 3-vectors \mathbf{a} and \mathbf{b} ,

$$\mathbf{a}^T = (L_1 \quad L_2 \quad L_3) \quad \mathbf{b}^T = (L_4 \quad L_5 \quad L_6)$$

which satisfy the so-called Plücker constraint: $\mathbf{a}^T\mathbf{b} = 0$. Consider a metric transformation defined by a rotation matrix \mathbf{R} and a translation vector \mathbf{t} , acting on points via:

$$\mathbf{C} \rightarrow \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{pmatrix} \mathbf{C}$$

Plücker coordinates are then transformed according to

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \rightarrow \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ -[\mathbf{t}]_\times \mathbf{R} & \mathbf{R} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$$

Two lines intersect if the following relation holds:

$$\mathbf{L}_2^\top \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \mathbf{L}_1 = \mathbf{a}_2^\top \mathbf{b}_1 + \mathbf{b}_2^\top \mathbf{a}_1 = 0$$

Points/lines cutting rays	Description
None	Non-central camera
1 point	Central camera
2 points	Camera with a single ray
1 line	Axial camera
1 point, 1 line	Central 1D camera
2 skew lines	X-slit camera
2 coplanar lines	Union of a non-central 1D camera and a central camera
3 coplanar lines without a common point	Non-central 1D camera

Figure 4.3: Camera models, defined by 3D points and lines that have intersection with all projection rays

These three classes of camera models may be defined as: existence of a linear space of d dimensions that has an intersection with all projection rays: $d = 0$ defines central, $d = 1$ axial and $d = 2$ general non-central cameras.

Intermediate classes do exist. X-slit cameras are a special case of axial cameras: there actually exist 2 lines in space that both cut all projection rays. Similarly, central 1D cameras (cameras with a single row of pixels) can be defined by a point and a line in 3D. Camera models, some of which without much practical importance, are summarized in Figure 4.3.

Camera model	Central		Axial		X-slit			
	finite	infinite	finite	infinite	finite+finite		finite+infinite	
Parameterization of projection rays	$\begin{pmatrix} a \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ a_3 \\ b_1 \\ b_2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} a \\ b_1 \\ b_2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ a_2 \\ a_3 \\ b \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ W & 0 & -Y & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & W & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} a_1 \\ a_3 \\ b_1 \\ b_2 \end{pmatrix}$

Figure 4.4: Parameterization of projection rays for different camera models

It is worthwhile to consider different classes due to the following observation: the usual calibration and motion estimation algorithms proceed by first estimating a matrix or tensor by solving linear equation systems (e.g. the calibration tensors or the essential matrix). Then, the parameters that are searched for (usually, motion parameters), are extracted from these. However, when estimating for example the 6 x 6 essential matrix of non-central cameras based on image correspondences obtained from central or axial cameras, then the associated linear equation system does not give a unique solution (much like when estimating a fundamental matrix from correspondences coming from coplanar 3D points). Consequently, the algorithms for extracting the actual motion parameters, can not be applied without modification. In the following, we deal with central, axial, x-slit and fully non-central cameras.

Multi-view geometry are formulated in terms of the Plücker coordinates of camera rays. For other models than the fully non-central one, camera rays belong to constrained sets, as explained in the previous section. We may thus choose the cameras' local coordinate systems such as to obtain "simpler" coordinate vectors for camera rays, and in turn simpler matching constraints. Since we deal with calibrated cameras, rays are given in metric coordinate systems, and we may apply rotations and translations to fix local coordinate

94 Chapter 4 Advanced models for reconfigurable scanner systems. Appropriate parameterizations for different models are explained in the following.

4.3.1 Pinhole model

All rays go through a single point, the optical center, as you can see in the Figure 4.5. We distinguish the cases of a finite and infinite optical center.

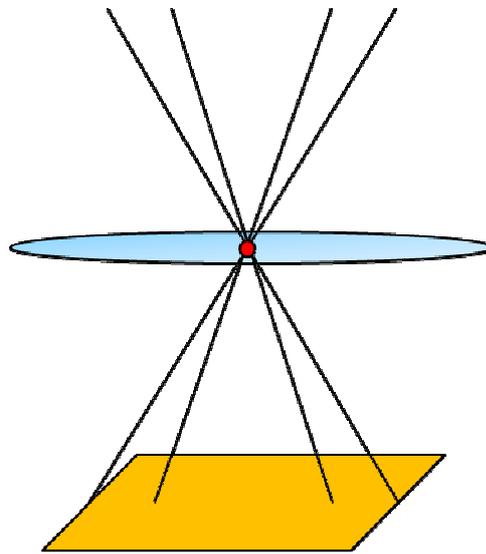


Figure 4.5: $d=0$ Central camera (Pinhole)

Finite optical center. We choose a local coordinate system with the optical center as origin. This leads to projection rays whose Plücker sub-vector b is zero, cf. figure 4.4. This is one reason why the multi-focal tensors, e.g. the fundamental matrix, can be written with a "base size" of 3.

Infinite optical center (e.g. affine camera). We can not adopt the optical center as origin, thus choose a coordinate system where it has coordinates $(0; 0; 1; 0)^T$. Projection rays are then of the form given in the 3rd column of table 2.

4.3.2 Axial model

All rays touch a line, the camera axis. Again, by choosing local coordinate systems appropriately, the formulation of the multi-view relations may be simplified. We distinguish the cases of a finite and an infinite camera axis.

Finite axis. Assume that the camera axis is the Z-axis. Then, all projection rays have Plücker coordinates with $L_6 = b_3 = 0$, cf. the 4th column of Figure 4.4.

Infinite axis. We choose a local coordinate system where the axis is the line at infinity with coordinates $(1; 0; 0)^T$ (line coordinates on plane at infinity). The camera axis' Plücker coordinates are then given by $(0; 0; 0; 1; 0; 0)^T$. Projection rays thus have coefficients with $L_1 = a_1 = 0$, cf. the 5th column of Figure 4.4.

Multi-view relations for axial cameras, with finite or infinite axis, can thus be formulated via tensors of "base size" 5, e.g. the essential matrix will be of size 5×5

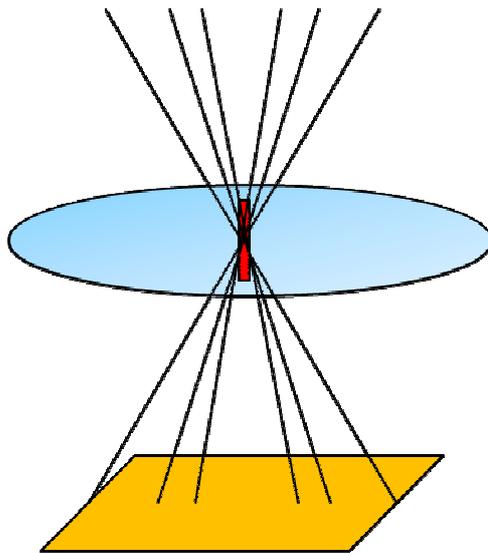


Figure 4.6: $d = 1$ Axial camera

As you can see in the Figure 4.6 all the incoming rays pass through a monodimensional site that is the axis of the optic system.

4.3.3 General Model

In this case, the rays pass through a bidimensional surface, as can be seen in Figure 4.7 and is the more general case that leads into account the complete model of the tensor of Plücker coordinates.

So no such simplification occurs, and multi-view tensors will have .base size. 6.

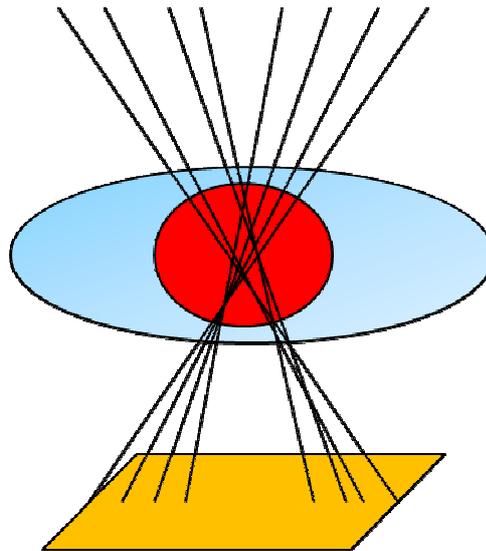


Figure 4.7: $d = 2$ General camera

4.3.3 Modeling of distortion parameters trend by the Axial model

It's possible, under geometric arguments previously made, written in matricial form the relations for the axial model. They are similar to those for the pinhole model except for the focal length f , which becomes a function of the incident rays which passes through the optics.

$$\begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \begin{bmatrix} -f(r)k_u & 0 & u_0 \\ 0 & -f(r)k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ \underline{\underline{0}} \\ 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

If it is assumed a variable trend of the focal length as a function of the distance r from the optical axis of the incident ray like that, (also we can see that in the Figure 4.8):

$$f(r) = f_0 + f_1 r + f_2 r^2 + \dots$$

In general f_0 is the nominal focal length of the pinhole model and $f_1 \in [-1; 0]$ while r is the distance of a generic pixel from (u_0, v_0) . Then it can be assumed that the more external rays suffer of a distortion at the same focal distance.

This makes us hope that this model can be used just to model the drift phenomenon of distortion parameters discussed above

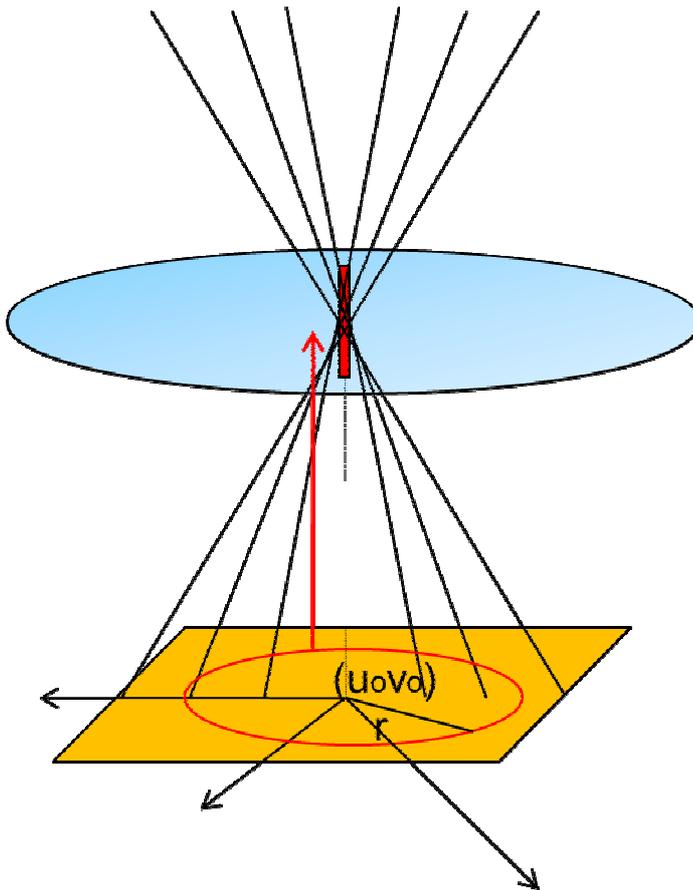


Figure 4.8: Modelling of the variable focal length

The next chapter will take the verifications carried out to validate the thesis that the variation of distortion parameters can be modeled by the axial model.

4.4 Self calibration

Now we introduce an algorithm for self-calibration of the camera-projector pair system, in this way we can have a fast reconfiguration of the scanner system on site. Many techniques exist [49][50][51].

4.4.1 Self-calibration procedure

The aim of this task was to design a self-calibration procedure for a user-configurable structured-light 3-D scanner which could be easy to be applied in short time, without the need of precision target, and in places as close as possible to the objects to be measured. The geometry of the scanning system [52],[53] is still that of Figure 4.9, and the block diagram of the proposed procedure is reported in Figure 4.10. During the “point acquisition” small circular reference points are projected onto the same surface that will be measured afterwards.

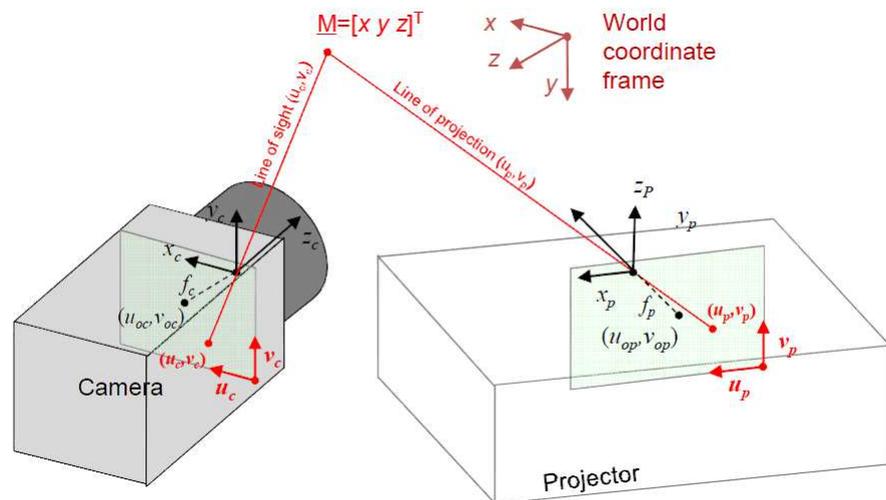


Figure 4.9: Model for camera-projector pair

In the first setup the reference points are projected one by one, but patterns of more than one point could be introduced in order to speed up this step, if necessary. Then corresponding pairs are established between projector pixel coordinates of projected points (they are known since they are generated within the software-generated image sent to the projector) and camera pixel coordinates of projected points, estimated as centroids of the observed shapes from the acquired images.

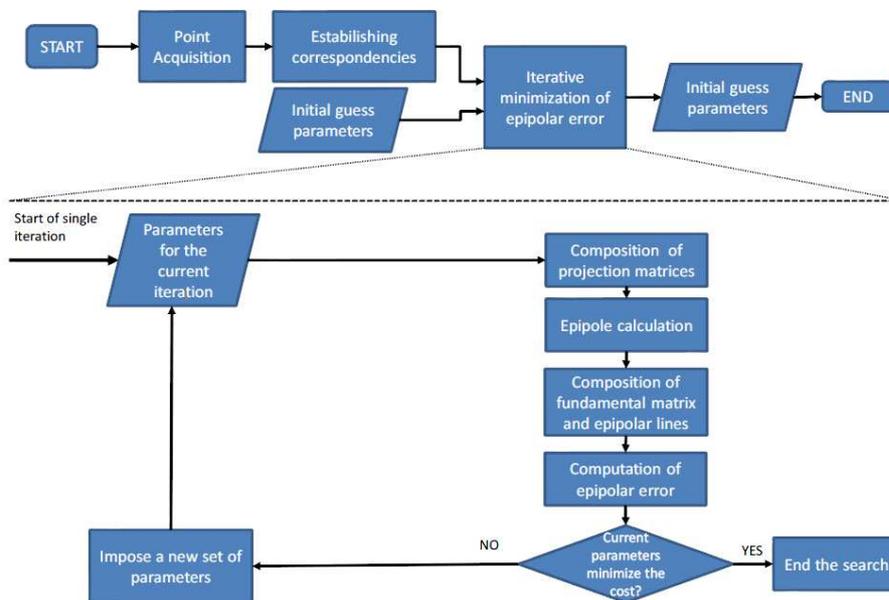


Figure 4.10: Block diagram of the proposed procedure

The corresponding pairs are inputs for the calibration step, which is based on a numerical optimization aiming to find the set of camera and projector parameters minimizing the epipolar error. Hence, the following issues have to be considered: i) the composition of the set of parameters to be optimized; ii) the starting guess values of the set of parameters; iii) the objective function to be minimized.

4.4.2 The set of parameters

The overall model of the camera–projector pair, represented by stereo model, has 20 independent parameters:

$$k_{ucf_c}, k_{vcf_c}, u_{oc}, v_{oc}, \varphi_c, \vartheta_c, \psi_c, t_{xc}, t_{yc}, t_{zc}, \\ k_{upf_p}, k_{vpf_p}, u_{op}, v_{op}, \varphi_p, \vartheta_p, \psi_p, t_{xc}, t_{yc}, t_{zc}$$

where each one of the two rotation matrices have been represented by three Euler angles. Some hypotheses can be made in order to reduce the number of parameters. At first, since the world coordinates of target points are not known, then extrinsic parameters of camera can be considered constant and posed equal to arbitrary values. Furthermore, horizontal and vertical pixel sizes can be assumed equal, so that horizontal and vertical pixel focals can be considered equal:

$k_{ucf_c} = k_{vcf_c}$ and $k_{upf_p} = k_{vpf_p}$ Eventually, after these hypotheses the set of parameters is the following:

$$k_{ucf_c}, u_{oc}, v_{oc}, k_{upf_p}, u_{op}, v_{op}, \varphi_p, \vartheta_p, \psi_p, t_{xc}, t_{yc}, t_{zc}$$

Preliminary numerical simulations have shown that: a) iterative minimization algorithms have a relevant sensitivity with respect to starting guess values, and generally non-global minimum solutions cannot be avoided; b) the variations of intrinsic parameters due to the changes of focus and iris during the normal operation of the scanner are small, of the order of few unit percent of their values. The position and orientation of the projector with respect to the camera has to comply some constraints, since camera and projector have to coarsely share a common field of view. Then the proposed procedure performs a number of calibrations, each one started with a different vector of starting values, and eventually the overall solution is the one giving the lowest value of the objective function. The different vectors of starting values have intrinsic parameters equal to those measured during a preliminary one-off traditional calibration with the DLT procedure of Chapter 3, and projector extrinsic values such that the projector center covers a box (in our tests, with x in 150-1000 mm, y in 400-600 mm, z in 2000-2600 mm, and the orientation angles such that camera and projector axis point approximately to a common center of view).

4.4.3 The objective function: the average epipolar error

Given the vector of the parameters to be optimized and the other known values (the paired set of camera pixel coordinates and projector pixel coordinates, and the remaining fixed parameters of the two pin-hole models), the objective function evaluates the average of the epipolar errors associated to the corresponding pairs, according to the lower part of Figure 4.10. At first, the two camera matrices \mathbf{P}_c and \mathbf{P}_p are composed from camera and projector model parameters as in the following matrix relationships:

for the camera;

$$s \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} k_{uc}f_c & 0 & u_{oc} & 0 \\ 0 & k_{vc}f_c & v_{oc} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Camera intrinsic parameters}} \underbrace{\begin{bmatrix} R_c & t_c \\ 0 & 1 \end{bmatrix}}_{\text{Camera extrinsic parameters}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Camera matrix \mathbf{P}_c

and for the projector;

$$l \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} k_{up}f_p & 0 & u_{op} & 0 \\ 0 & k_{vp}f_p & v_{op} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{Projector intrinsic parameters}} \underbrace{\begin{bmatrix} R_p & t_p \\ 0 & 1 \end{bmatrix}}_{\text{Projector extrinsic parameters}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Projector matrix \mathbf{P}_p

Then the fundamental matrix is calculated as

$$\mathbf{F} = [\tilde{\mathbf{e}}_p]_{\times} \mathbf{P}_p \mathbf{P}_c^{\nabla}$$

where \mathbf{P}_c^{∇} is the pseudoinverse matrix of \mathbf{P}_c and:

$$[\tilde{\mathbf{e}}_P]_{\times} = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}$$

is the skew symmetric matrix associated to the projector-side epipole $\tilde{\mathbf{e}}_p$, namely the intersection of the baseline (i.e. the line through the camera and projector centers) and the projector grid plane, expressed in homogeneous projector-pixel coordinates. The epipole can be evaluated as the camera center as viewed by the projector:

$$\tilde{\mathbf{e}}_P = \mathbf{P}_P \tilde{\mathbf{C}}_C = \mathbf{P}_P \cdot (-\mathbf{R}_p \mathbf{t}_p)$$

The well-known correspondence relationship $\mathbf{x}_L^T \mathbf{F} \mathbf{x}_R = 0$ can be written for each corresponding pair $\tilde{\mathbf{m}}_{C_i}$ and $\tilde{\mathbf{m}}_{P_i}$:

$$\tilde{\mathbf{m}}_{P_i}^T \mathbf{F} \tilde{\mathbf{m}}_{C_i} = 0$$

This eq. means that the point $\tilde{\mathbf{m}}_{P_i}$ must lie on the so-called epipolar line $\mathbf{F} \tilde{\mathbf{m}}_{C_i}$ on the projector grid plane, and vice versa, the point $\tilde{\mathbf{m}}_{C_i}$ must lie on the epipolar line $\mathbf{F}^T \tilde{\mathbf{m}}_{P_i}$ on the camera sensor plane. Then the average epipolar error can be evaluated as follows:

$$e_{epi} = \frac{1}{2N} \sum_{i=1}^N \{d(\mathbf{m}_{C_i}, \mathbf{F}^T \tilde{\mathbf{m}}_{P_i}) + d(\mathbf{m}_{P_i}, \mathbf{F} \tilde{\mathbf{m}}_{C_i})\}$$

where $d(\mathbf{x}, \mathbf{l})$ is the distance between a point \mathbf{x} and a line \mathbf{l} in pixel Cartesian coordinates. The algorithm chosen for the minimization is the Nelder-Mead simplex method.

The epipolar geometry that was used is shown in the Figure 4.11 below. You can understand how the epipolar error is composed and what it geometrically represents.

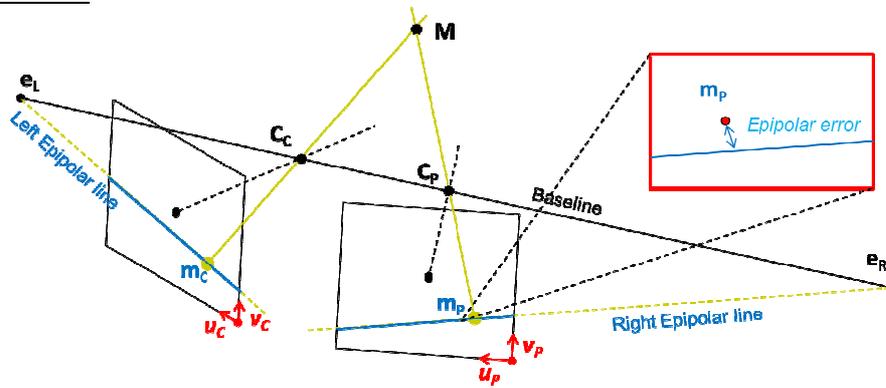


Figure 4.11: Epipolar geometry

Chapter 5

Test for the validation of the new calibration algorithms

5.1 Introduction

Task of the latter chapter is to show the tests carried out during the comparison of the new calibration technique with classical techniques and also the testing of hypotheses on the new geometric model which takes into account the variation of the distortion parameters.

It will be shown by first the comparison of the new procedure for calibrating camera-projector pair with some classical techniques, will be then shown some reconstructions of known objects to assess the quality of the reconstruction.

Then will be shown the tests conducted and the procedure followed to check that the variation of the parameters of distorting the variation of the target is rewritable in terms of axial model. And finally, I will take the tests carried out by self-calibrating the overall scanner and the comparison with the first stereo-like technique.

Will be shown for each test the main results, Labview[®] was used for the stereo calibration steps since it was more easy and quick to use in Firewire communication with the cameras and an optimization in terms of time for the extraction of the features was guaranteed by the images manipulations functions, even for the creation of the simulated images and for the comparison of the model axial was easier to use the Labview. Later Matlab[®] was also used in the stages of the software optimization , i.e. those minimization algorithms minimization for the self-calibration. The two programs, however, can be used interchangeably with the use of the correct libraries for analysis of

106 Chapter 5 Test for the validation of the new calibration algorithms

images and communication drivers with cameras. This is the why will be shown only the conceptual schemes of the tests without giving details of the software implementation because it would be redundant.

5.1.1 Assumptions about the model of the projector

Starting from the first results, shown in the Figure 5.1, obtained from the calibration of the scanner as a pair of stereo cameras we can qualitatively evaluate the parameters obtained for the projector since it is not possible to compare them with other techniques because there are not in the literature. We can only make assessments based on the geometric model and the geometry of the projector lens.

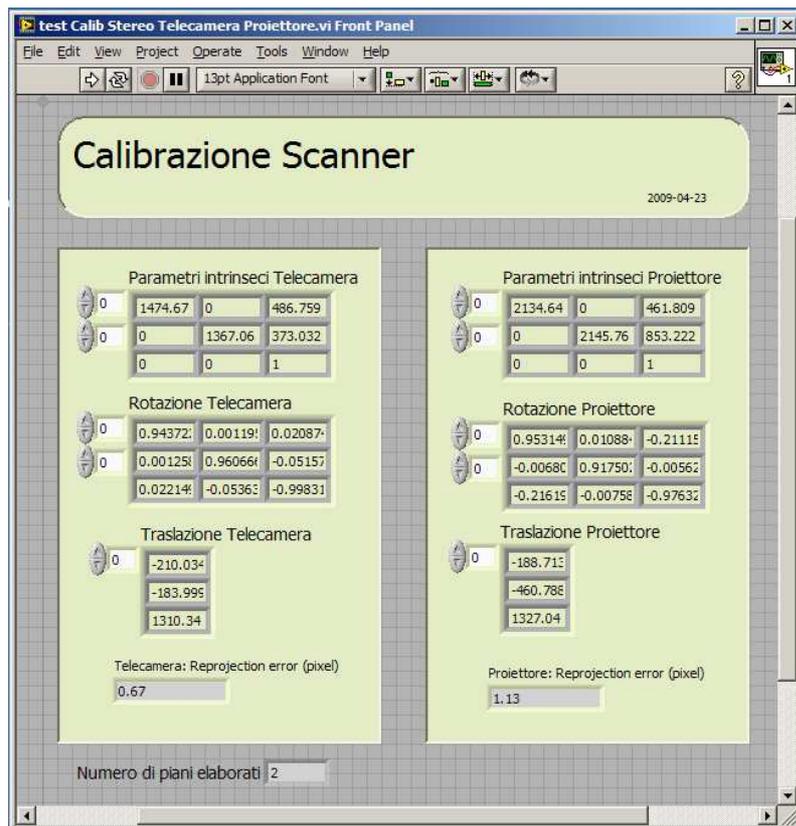


Figure 5.1: Front panel of Stereo like calibration software

107 Chapter 5 Test for the validation of the new calibration algorithms

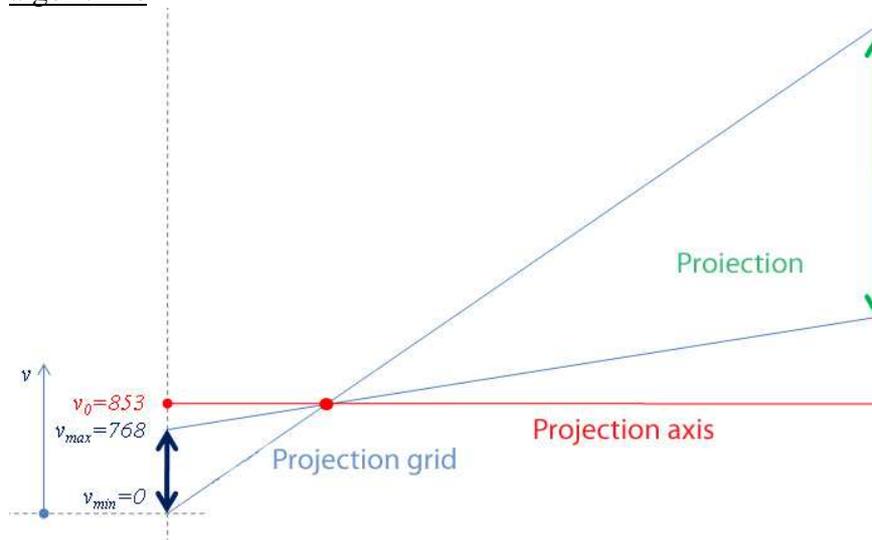


Figure 5.2: Projector perspective modeling

What we can evaluate is the focal length in pixels, knowing the nominal optics focal length, between [20.4 mm, 24.5 mm] and the size of the pixels of the CCD of the projector, that is $10\mu\text{m}$, we obtain a focal length of around 2000 pixels.

Also we can evaluate the coordinates of the center of projection that usually fall in the neighborhood of half the resolution of the sensor. In this case this is true for the horizontal coordinate as the value of the calibration is 461 which is about half the horizontal resolution of 1024 pixels. Is not observed due to the vertical coordinate, since it is a value of 853 pixels in relation to a vertical resolution of 768 pixels. This can be explained looking at the Figure 5.2 shows that the projector has a shift of the optical axis upwards to allow a projection non-symmetrical. Evaluated that we may consider extrinsic parameters of both devices, making measurements of the reference systems world, camera and projector and compare with the values of the translation vector that are congruent to the geometry of the scanner.

5.2 Characterization of the calibration

The new calibration technique was compared with a classical calibration but it will only return the parameters of the camera because as said previously there are not techniques in the literary that allow us to calibrate the projector modeled as a camera.

5.2.1 Test station

The Figure 5.3 shows the configuration of the test station. The geometry and the distance between the various devices are important to be able to assess the goodness of the extrinsic parameters calculated from the calibration.

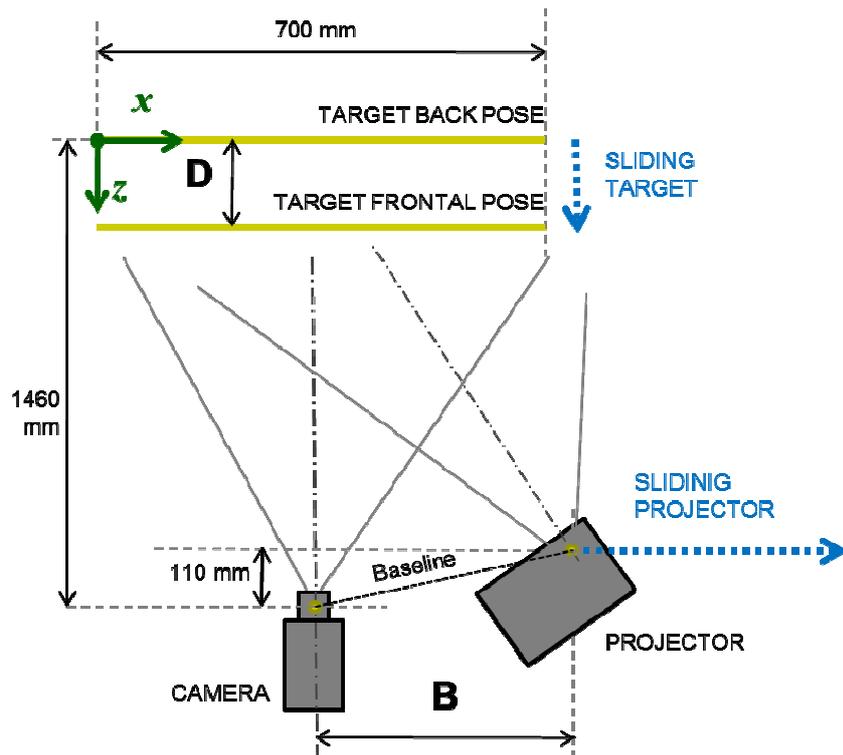


Figure 5.3: Geometrical configuration of the scanner

Our scanner is composed by the following devices:

109 Chapter 5 Test for the validation of the new calibration algorithms

Camera = ImagingSource DFK31F03, 1024x768, 30 fps, 6 mm lens.

Projector = NEC NP62, 1024x768, 60 Hz, focal in [20.4 mm, 24.5 mm]

Slider = M-415 CG by Physik Instrumente (PI) GmbH

And we can see them in the Figure 5.4.



Figure 5.4: Scanner 3D

5.2.2 Comparison of calibration techniques

To make a comparison with the calibration technique of Zhang were performed $N = 90$ calibrations with a volume of calibration $D = 50\text{mm}$ and a baseline variable from 310 to 470 mm. In this way was calculated the uncertainty of type A relative to each parameter. The intrinsic parameters of the camera are comparable with those obtained via Zhang algorithm, as we can see in the Figure 5.5, but improved in terms of uncertainty that is reduced. For the intrinsic parameters of the projector is valid the arguments made above where they are evaluated geometrically.

110 Chapter 5 Test for the validation of the new calibration algorithms

Parameters	k_{uofc} [pixel]	k_{vofc} [pixel]	u_{oe} [pixel]	v_{oe} [pixel]	k_{uofp} [pixel]	k_{vofp} [pixel]	u_{op} [pixel]	v_{op} [pixel]
Proposed → mean	1396	1329	507	298	2257	2315	503	754
Proposed → uncertainty	2	1	3	2	15	19	12	3
Zhang → mean	1378	1308	488	326	---	---	---	---
Zhang → uncertainty	26	24	10	20	---	---	---	---
	Camera				Projector			

Figure 5.5: Intrinsic parameters comparison

With regards to the extrinsic parameters we can see their variation as a function of the baseline that is varied by moving the projector towards the right and away from the camera. Figure 5.6 shows the extrinsic camera parameters to vary the baseline, since the camera remains stationary with respect to the world system such parameters remain the same, they are just subject of the oscillations due to uncertainty, but contained below the percentage point.

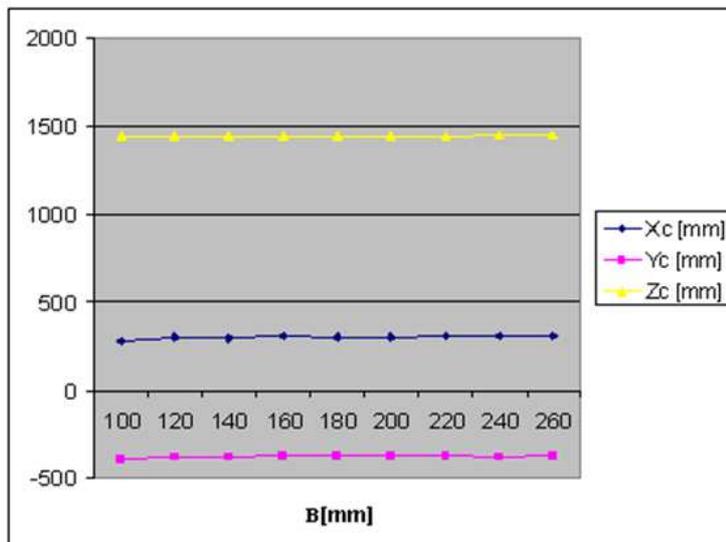


Figure 5.6: Camera extrinsic parameter evaluation at Baseline variation

111 Chapter 5 Test for the validation of the new calibration algorithms

The extrinsic parameters of the projector, instead, show a linear variation of the X coordinate, as shown in Figure 5.7, that corresponds to the distance of the center of the projector system from the world system to the right along the X coordinate, and a slight decrease of the Z coordinate because the projector is slightly rotated to illuminate the same target during the moving to the right side and this causes a convergence of the two reference systems along the Z axis, the Y-axis instead remains unchanged because the projector is fixed with the base along which it moves.

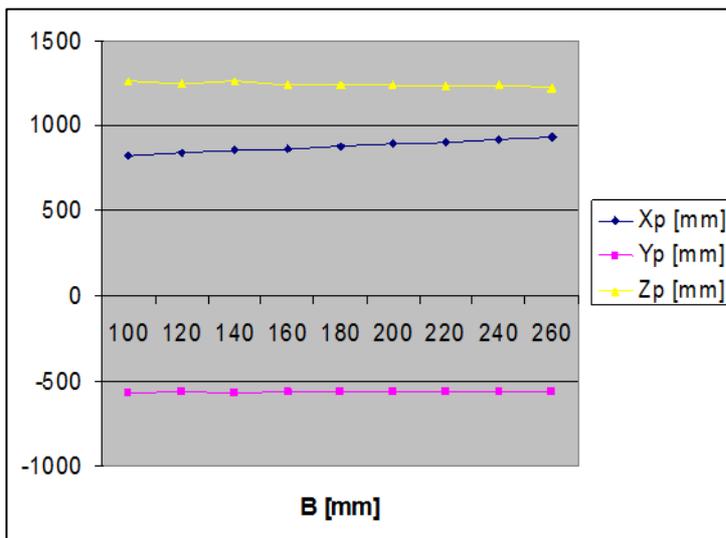


Figure 5.7: Projector extrinsic parameter evaluation at Baseline variation

5.2.3 Re-projection error

Another index that has been evaluated both as a function of baseline variation and variation of the depth of the volume of calibration is the error of reprojection. The error of reprojection is defined as follows.

It's the average of the N distances of the observed aims m_i from the reprojected points obtained as the product of the points on the sensor for the matrix of the perspective projection.

112 Chapter 5 Test for the validation of the new calibration algorithms

$$e = \frac{1}{N} \sum_{i=1}^N \text{dist}(\underbrace{\tilde{m}_i}_{\text{Observed}}, \underbrace{P\tilde{M}_i}_{\text{Estimated simulating a new projection}})$$

As shown in Figure 5.8 we can see that this error is reduced with increasing Baseline and the increase of the volume of calibration. This is a normal feature of the stereo type systems that improve their performance with the increase of the baseline, what is new even is the knowledge of an increase in the precision in growing of the volume of calibration. The error remains always below the pixel, this is a very good result compared with the classical techniques that they approach to these benefits with difficulty.

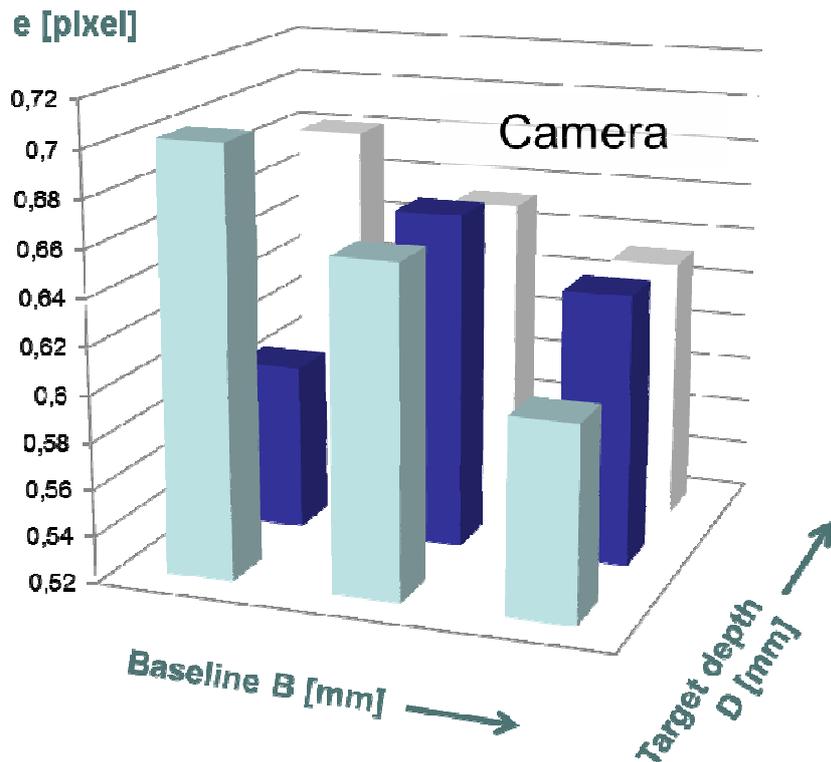


Figure 5.8: Camera re-projection error

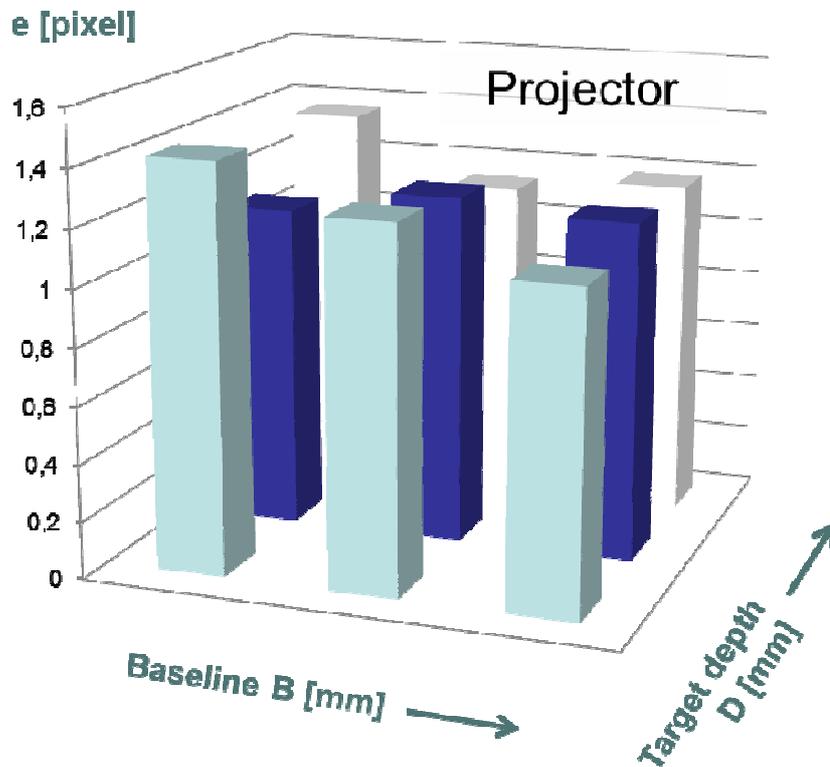


Figure 5.9: Projector re-projection error

The same result can be shown to the projector in Figure 5.9. Here, unfortunately, the reprojection error, it's still above the pixel, but never exceeds 1.5 pixels. This may be due to effects to having modeled the projector using the model of a camera. It's still remain a very good approximation considering the fact that there are no direct models for the projector.

5.2.4 Real object reconstruction

Finally, in order to assess the quality of the innovative procedure introduced, acquisitions of known objects have been made and the measures were compared with acquisitions performed by much more accurate 3D laser scanner that can be used as a reference.

114 Chapter 5 Test for the validation of the new calibration algorithms

In Figure 5.10 is shown the acquisition of a face of clay by a 3D laser scanner like the FaroArm.

The mesh of the points cloud were performed thanks to a software for the manipulation of 3D clouds , to be able to perform measurements with the digital gauges.

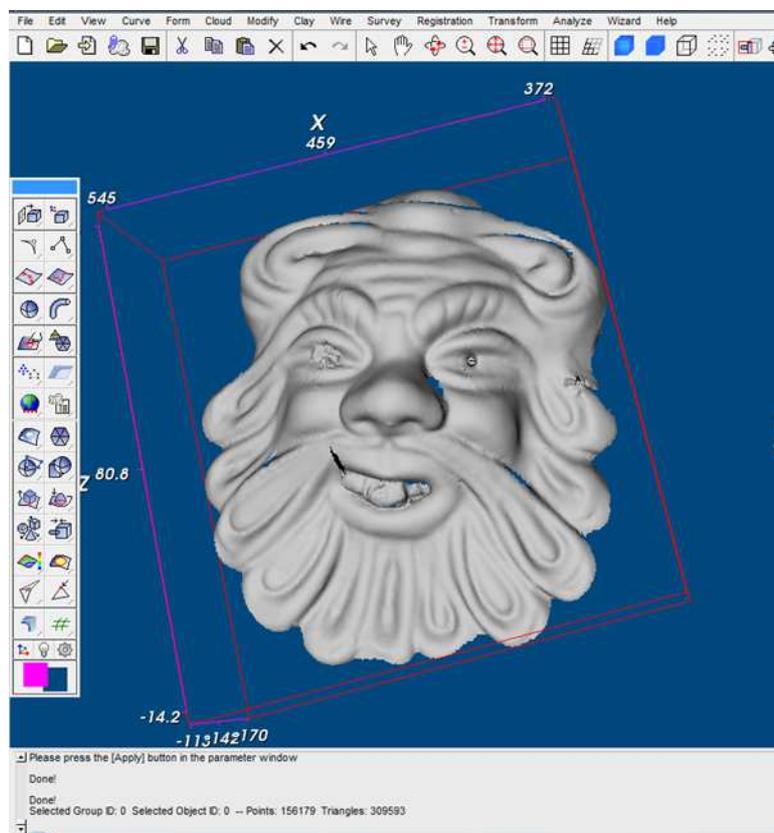


Figure 5.10: Laser scanner acquisition

In Figure 5.11 is shown the same mask acquired with our scanner, they have been carried out the same measurements with the digital caliper and it has come to the differences along the X and Y axes that is less than 1% instead it's of about 3% the difference between measures along the Z axis.

115 Chapter 5 Test for the validation of the new calibration algorithms

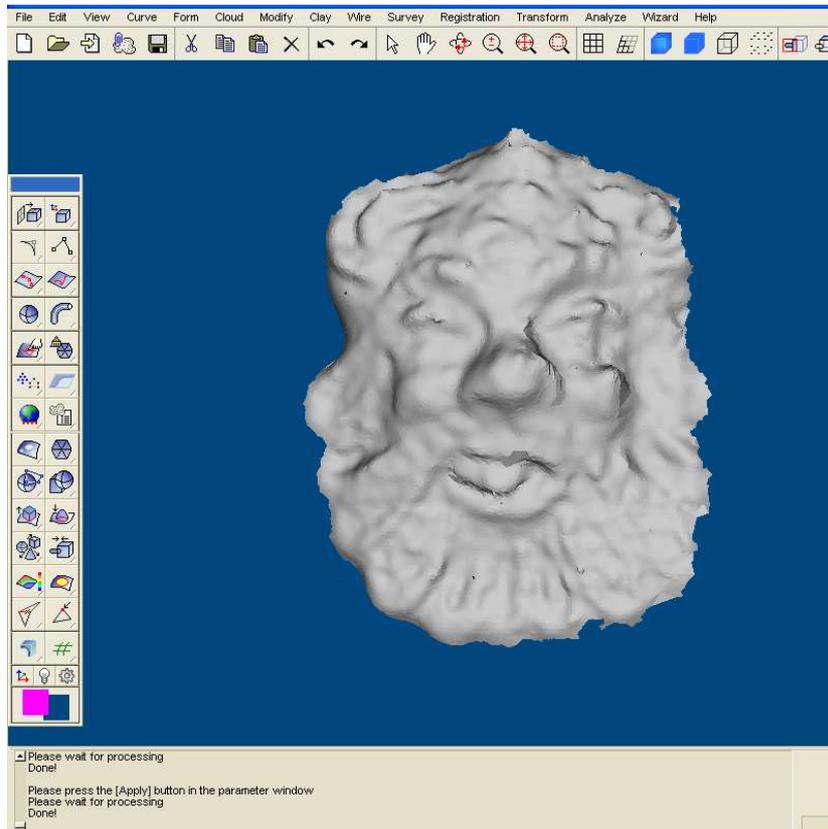


Figure 5.11: 3D Fringe profilometer with Stereo Calibration acquisition

5.3 Advanced Geometric Camera model Validation

To test the hypothesis that the dependence of the variation of distortion parameters as a function of target distance was linked to the type of model used, namely the pin-hole, repeated calibrations of a camera were carried out, with different target distances, using the Camera Toolbox Calibration of J.Heikkilä that implements the pin-hole model, and it was plotted the trend of the distortion parameters at the variation of this distance. Simulated images of a target were made, Figure 5.n, that they were generated as if they were captured by a

116 Chapter 5 Test for the validation of the new calibration algorithms

camera modeled by Axial model, without distortion, they were calibrated with the same toolbox and the trends of the parameters were plotted.

5.3.1 Synthetic images creation

The procedure can be summarized in the following steps:

- Creation of synthetic images of a target at different distances using the Axial model by Labview functions.
- Calibration with synthetic images using the Camera Calibration Toolbox for Matlab with the pinhole model, so we can verify the variation of the parameters to vary the distance of the target.
- Acquisition of real images of the same target.
- Calibration with real images

Examples of both of these images are shown in Figure 5.12.

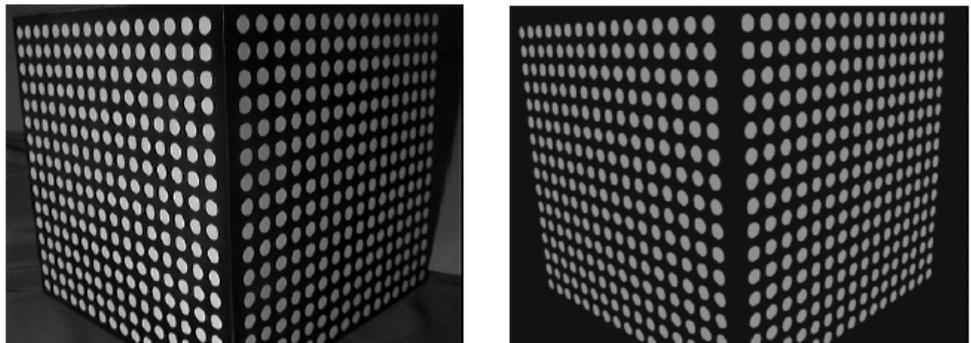


Figure 5.12: Calibration image, real on the left side and synthetic on the right one

5.3.2 Comparison of distortion parameters trends

From the trends of the parameters of distortion thus obtained, shown in the Figures 5.13-16 both in simulation and on real images it could be concluded that the variation of the coefficients of radial distortion to vary the distance from the target is linked to the use of pin-hole model, incomplete in describing this phenomenon. Next step is to

117 Chapter 5 Test for the validation of the new calibration algorithms

implement a calibration tool for the axial model of the camera, thanks to which even calculate f_1 , which is not taken into account in the model pin-hole.

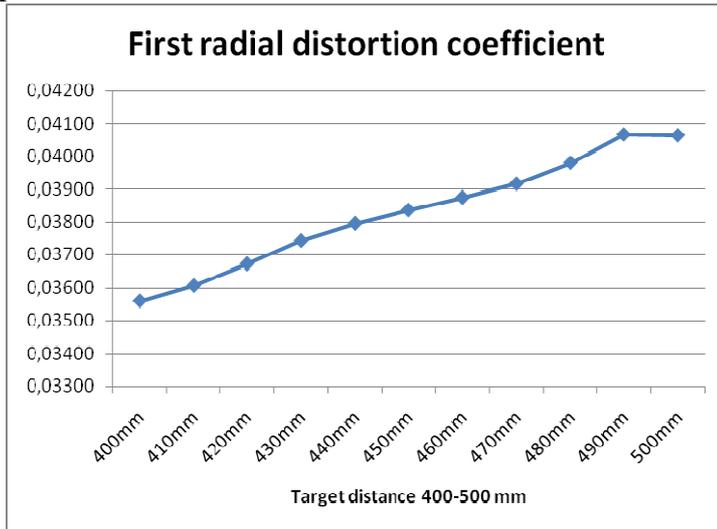


Figure 5.13: Trend of the first radial distortion parameter in the synthetic images

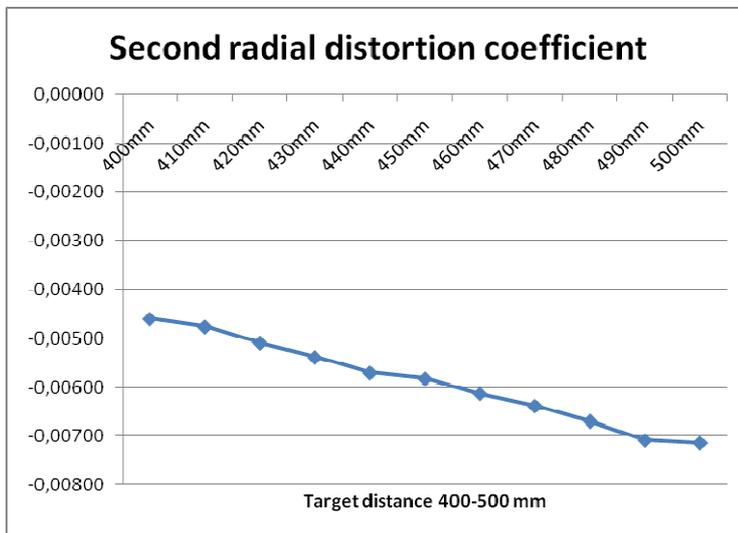


Figure 5.14: Trend of the second radial distortion parameter in the synthetic images

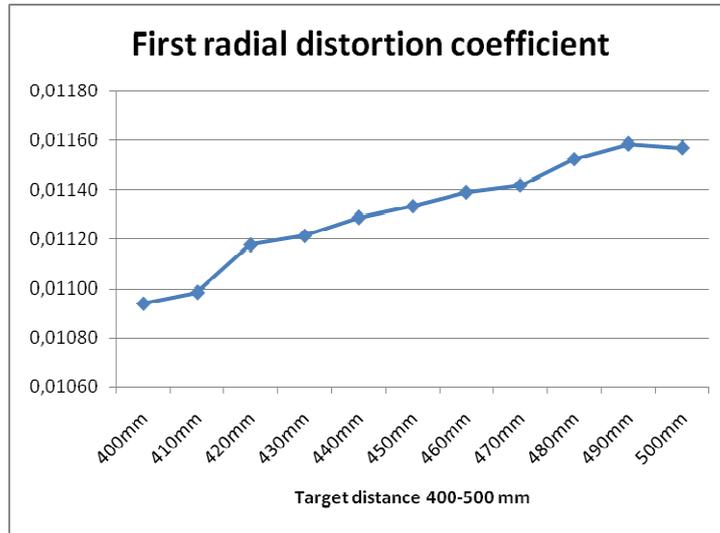


Figure 5.15: Trend of the first radial distortion parameter in the real images

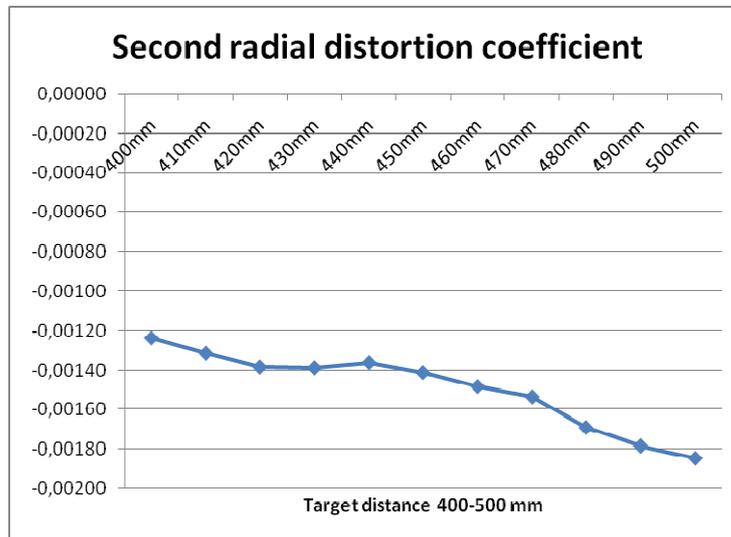


Figure 5.16: Trend of the second radial distortion parameter in the real images

119 Chapter 5 Test for the validation of the new calibration algorithms

The trend of the distortion parameters is the same, so we can conclude that this variation can be modeled with the Axial Model extension.

5.4 Self calibration

The proposed procedure has been implemented and tested both with numerical tests and experimentally for the calibration of a 3-D scanner composed of a desktop NEC NP62 projector and an Imaging Source DMK 21F04 monochrome camera. Numerical and experimental results have been also compared with those obtained with the 8-point algorithm and the subsequent decomposition of the fundamental matrices into extrinsic and intrinsic parameters. Moreover, the self-calibration method can be used for the quick tune up in loco of a system previously characterized in laboratory: under this point of view it is possible to narrow the search interval for the parameters and adopting a bounded numerical minimization algorithm.

5.4.1 The epipolar error surfaces

The function to be minimized is a N-dimensional function whose trend has been studied, as the simplex algorithm may bring the search to fall into local minima that are not correct for the purpose of calibration. The trend of the error surface was studied while maintaining fixed N-2 variables and making the other two variably, so as to build an error surface. It was noted that in many configurations these surfaces have heights cusps, as you can see in the Figure 5.17-18, and make the calculation of the absolute minimum almost impossible.

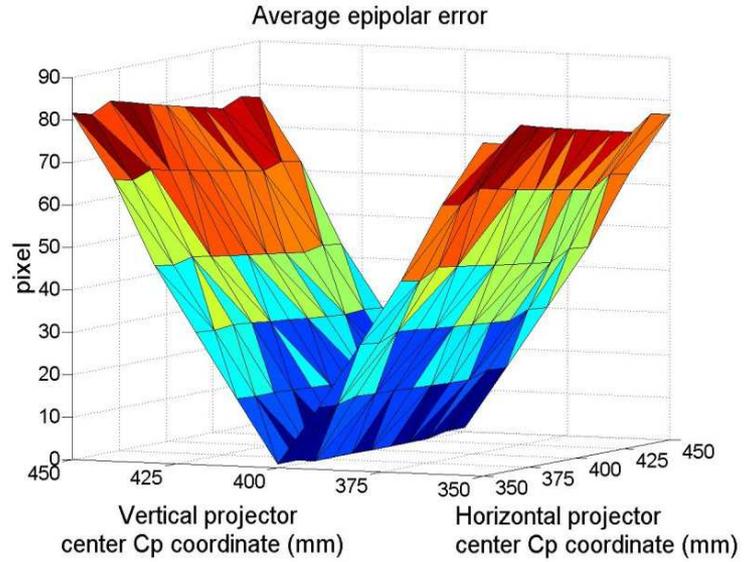


Figure 5.17: Average epipolar error surface

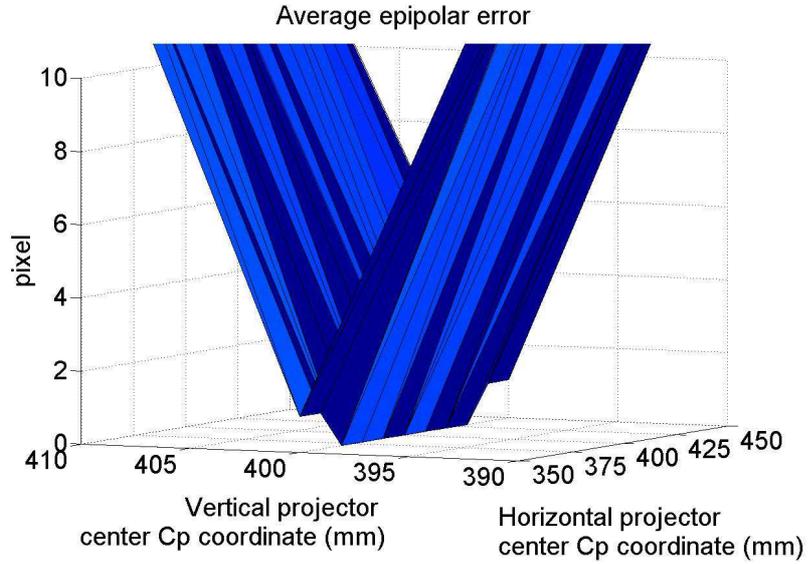


Figure 5.18: Zoom of the cusp of the average epipolar error surface

To obviate this, geometric constraints were introduced to the cost function which made possible the deformation of that surface as shown in Figure 5.19.

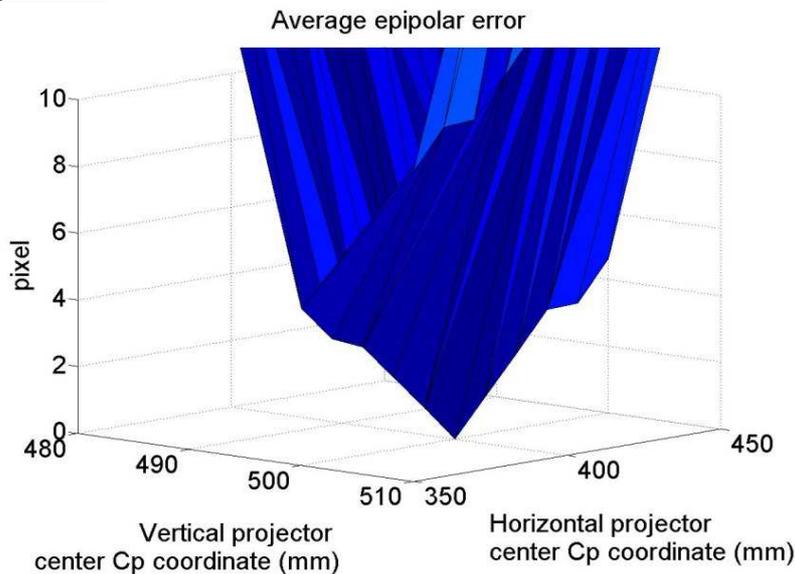


Figure 5.19: Deformation of the cusp by additional constraints

Some of the constraints can be used, for example, should be the fact that if the baseline increase then the projector Yaw angle must rotate towards the camera, so as to exclude geometric configurations outside the range of use of the two devices. So we proceeded to subtract in the error function all scanner configurations not geometrically feasible.

5.4.2 Simulation results

The simulations have been conducted using correspondent sets of points on the two sensors, upon which the selfcalibration is to be based. Starting from points in 3-D space, the image acquisition (generation) on the camera sensor (from the projector grid) has been simulated, respectively, implementing the previous models. The geometry of the simulated target was chosen based on the principle that calibration points must not belong to a same plane, for the selfcalibration to be effective. The adopted targets are shown in Figures 5.21. The number of points does influence the results of the selfcalibration: it can be seen that the highest the number of targeting points, the more accurate the results of the algorithm, as might be expected, although after a given number of points the improvements are less relevant. In the tests in the following, $N = 23$ was chosen. To

122 Chapter 5 Test for the validation of the new calibration algorithms

retain the chance to compare simulation results with the experimental ones achieved with the real system which will be shown further on, during the simulations the resolution of the camera sensor (projector grid), respectively, have been assumed as 640x480 pixels (1024x768 pixels). As for the values of the characteristic parameters of the system, please refer to Figure 5.20.

CAMERA		PROJECTOR	
Intrinsic Parameters $k_{vf_{camera}} = k_{uf_{camera}} = 1000$ pixels $u0_{camera} = 320$ pixels $v0_{camera} = 240$ pixels		Intrinsic Parameters $k_{vf_{projector}} = k_{uf_{projector}} = 2000$ pixels $u0_{projector} = 512$ pixels $v0_{projector} = 800$ pixels	
Extrinsic Parameters <i>Center of projection (C_c)</i>		Extrinsic Parameters <i>Center of projection (C_p)</i>	
$x_{C_c} = -250$ mm $y_{C_c} = 400$ mm $z_{C_c} = 2300$ mm	<i>Euler Angles</i> $\varphi = 0^\circ$ $\theta = 200^\circ$ $\psi = 0^\circ$	$x_{C_p} = 900$ mm $y_{C_p} = 500$ mm $z_{C_p} = 2400$ mm	<i>Euler Angles</i> $\varphi = 0^\circ$ $\theta = 165^\circ$ $\psi = -6^\circ$

Figure 5.20: Values of the parameters of the simulated system

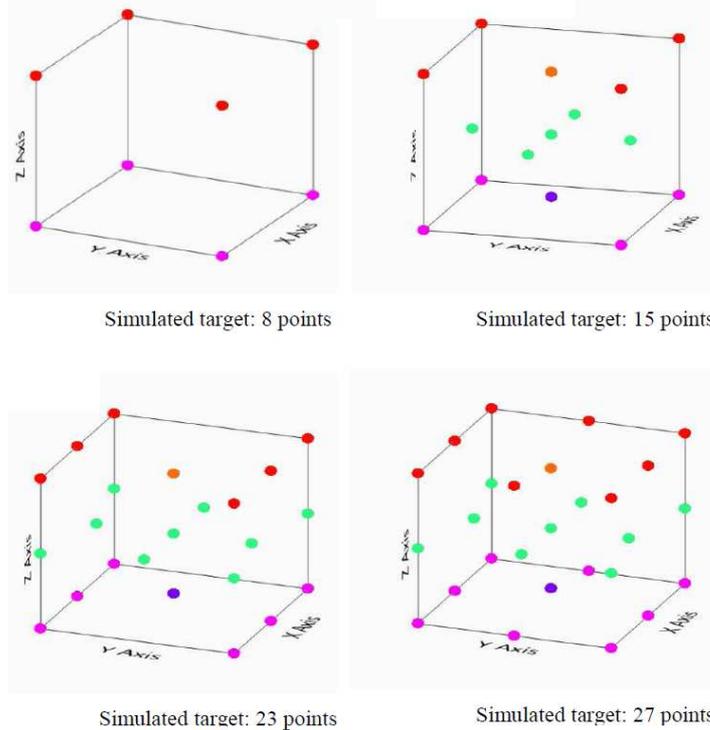


Figure 5.21: Values of characteristic parameters of the simulated system

123 Chapter 5 Test for the validation of the new calibration algorithms

Since in the real applications the images upon which the self-calibration has to be based are noisy, the immunity of the self-calibration algorithm to noise has been verified. Gaussian noise with standard deviation from 0.1 to 2 pixels has been superimposed to the image points coordinates, executing repeated self-calibrations, and the mean and the dispersion in the resulting values of the model parameters have been considered. Figures 5.22-25 show the behavior of the algorithm for some of the parameters with respect to the noise standard deviation. As shown in the graphs, the average value differs from the true one as the noise increases, and also the 3σ interval becomes greater.

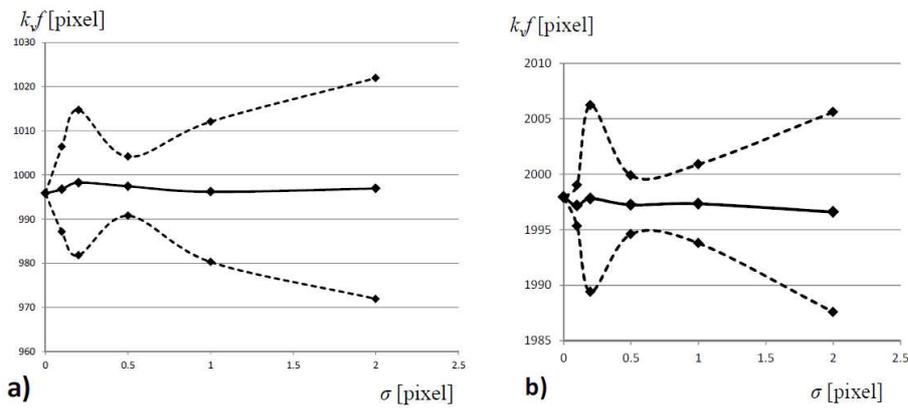


Figure 5.22: Intrinsic a) camera and b) projector parameters $k_v f$,

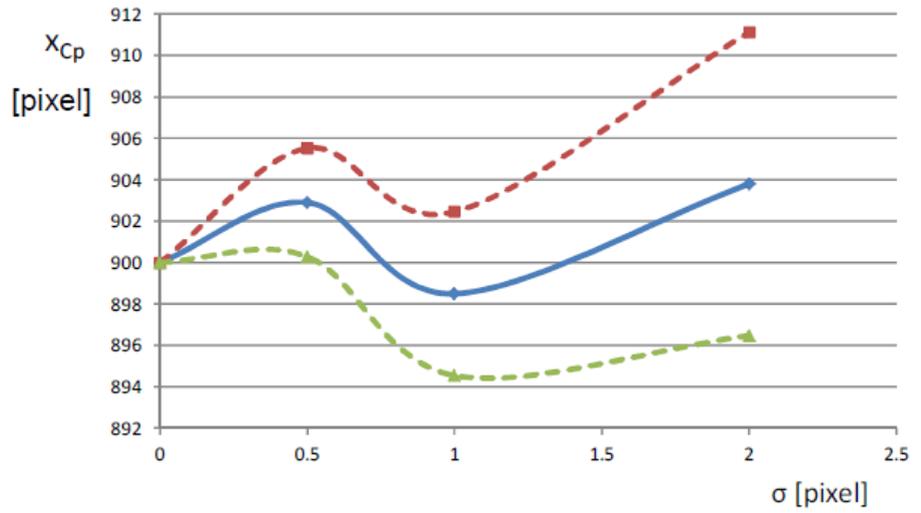


Figure 5.23: Behavior of extrinsic parameter x_{Cp} of the projector

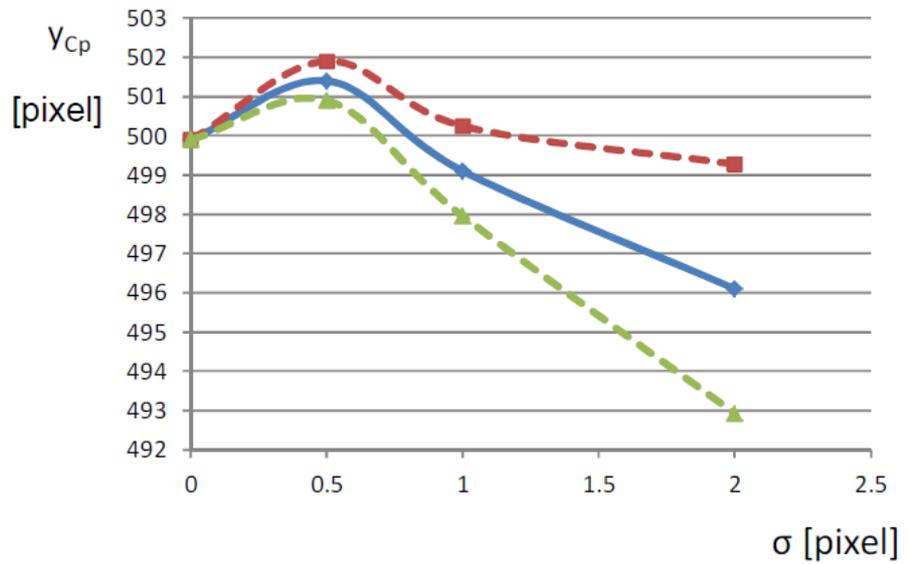


Figure 5.24: Behavior of extrinsic parameter y_{Cp} of the projector

125 Chapter 5 Test for the validation of the new calibration algorithms

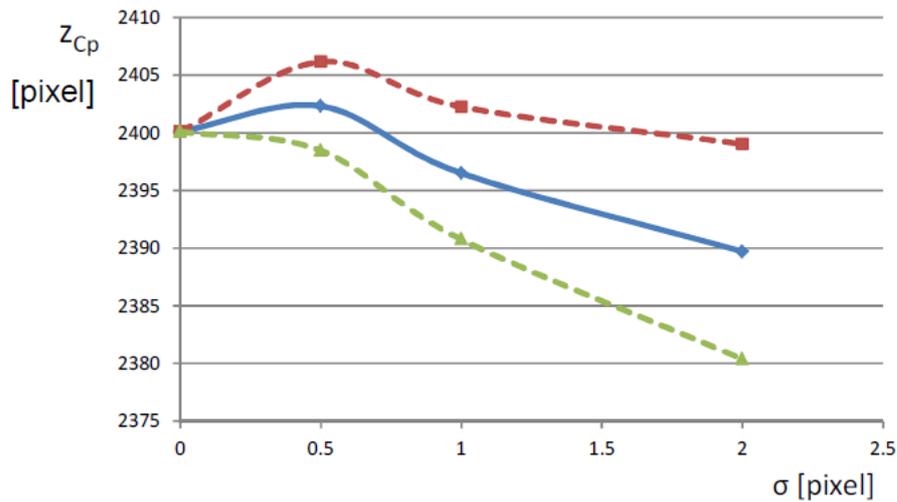


Figure 5.25: Behavior of extrinsic parameter z_{Cp} of the projector

5.4.3 Comparison with Camera-Projector calibration

The validation of the self-calibration method has been done on a real case. The procedure involves the projection of features onto the scene that has to be acquired, which allow to find correspondences between image points in the two sensors. The projector is treated as a camera, and its acquired image corresponds to the image that the projector itself projects onto the scene. The experimental results have been obtained using a helmet as a non-planar object onto which the features are projected. The results obtained with this procedure were compared with the Calibration procedure for Stereo Camera-Projector pair.



Figure 5.26: Two images acquired by the camera during the self-calibration procedure

Intrinsic Parameter	Stereo Camera-Projector Pair Calibration	Self-Calibration Epipolar minimization bounded
k_u camera	1123.7	1050.0
k_v camera	1113.1	1149.9
u_0 camera	290.9	329.0
v_0 camera	264.3	248.4
k_u projector	2242.5	2221.4
k_v projector	2333.7	2341.5
u_0 projector	391.9	449.7
v_0 projector	739.9	745.6

Figure 5.27: Comparison of intrinsic parameters. All values are in pixels

Extrinsic Parameter	Stereo Camera-Projector Pair Calibration	Self-Calibration Epipolar minimization bounded
Φ_{camera}	-9.498	-9.500
θ_{camera}	-0.802	-0.800
Ψ_{camera}	-9.789	-9.800
X_{camera}	171.5	171.5
Y_{camera}	342.8	342.8
Z_{camera}	1322.0	1322.0
$\Phi_{\text{projector}}$	-4.459	-2.200
$\theta_{\text{projector}}$	20.121	24.900
$\Psi_{\text{projector}}$	-3.969	-3.800
$X_{\text{projector}}$	625.2	550.0
$Y_{\text{projector}}$	435.8	401.8
$Z_{\text{projector}}$	1222.1	1170.3

Figure 5.28: Comparison of extrinsic parameters. All lengths are in mm and all angles are in degree.

5.4.4 Comparison with the 8-points algorithm

A self-calibration algorithm well known in literature is the algorithm of the 8 points. It has several limitations, the first one being the need of 3 views to operate. Moreover, it is extremely sensible to noise, even in its normalized version. To provide a reference for the proposed method, the algorithm has been applied in the case of 3 views, and the results have been compared to the ones provided by the 8 point algorithm. In both cases, two versions of the algorithms have been considered: one, unbounded, finds the best set of parameters without restrictions on the values to assign to each parameter; the bounded version limits the space of research to a domain that can be defined with an initial calibration (in laboratory, using classical calibration procedures) of the system. Figure 5.29 shows the results in the different cases.

Parameter	8-points unbounded	8-points bounded	Epipolar minimization unbounded	Epipolar minimization bounded
k, f camera	1211.4	1070	1014.2	1070
u_0 camera	319.8	326.3	319.1	319.3
v_0 camera	239.6	227	239.6	223.4
k, f projector	2453.2	2053	2020.2	1999.4
u_0 projector	600.2	501	511.8	510.2
v_0 projector	841.2	799.5	847.9	810.4

Figure 5.29: Comparison of experimental results. All values are in pixels.

128 Chapter 5 Test for the validation of the new calibration algorithms

Conclusion

In this work an original method of calibration was proposed, at the conclusion of a doctoral thesis, which concerned the study of fringe patterns systems and their metrological characteristics.

The main issues on which attention was focused during this work were the search for an innovative model that allowed, by the design and calibration of projector parameters, to eliminate problems involved in the projection of the patterns and go over the problems created by the light noise present in the calibration environment.

This goal has led to the implementation of a stereo-type model for the camera-projector pair that has no precedent in the literature. This modeling is a good approximation as has been verified experimentally, it also benefits from all the advantages of stereo techniques established in the literature, since we can obtain adequately low reprojection errors and acceptable reconstructions, whereas the time to calibrate the total 3D scanner system went from 40 minutes that the classical techniques of projection patterns and phase-shift require to about 3-4 minutes for this new technique. Without excessive loss in terms of accuracy as shown by the comparison tests.

Then, to be able to further reduce such uncertainty, during the visit to the laboratories of the Machine Vision Group of Oulu in Finland, thanks to the professor J.Heikkilä the problems related to the distortion of the optics were then introduced into the study. The main problem that is faced when introducing these parameters in the model is that they vary within the volume of calibration/measuring. This problem has been studied and it has come to formulate a thesis about the possibility of including such variation within a model developed for the optics, i.e. the axis model.

This is a remarkable achievement for the cases which in literature used to be modeled using empirical formulas, and can now be taken into account with geometric formulas which extend the previous pin-hole model. In conclusion of the work of this thesis, a self-calibration algorithm has been implemented that allowed a faster reconfiguration

of a scanner pre-calibrated in the laboratory when it is moved on the field for the measurements.

The techniques and models introduced make, therefore, possible to use both fast calibration and reliable low uncertainties for a low-cost 3D scanner, since it deals with a camera and a projector that cost few hundred euros and a PC, compared to systems such as laser based ones whose costs are many tens of thousands of euros.

In the future it is possible to think of applications of this calibration also in the techniques of real-time scanning, or even one can develop an application that can be self-calibrated during its operation time in order to always update the data calibration according to changes in environmental factors during the exercise period.

Models for generic optics could also be studied, trying to understand what other effects can be taken into account with these extensions of the model.

An idea for future research is the ability to further improve the search algorithm in selfcalibration being the error function the multivariable simplex algorithm has still some cases where returns as the solution of local minima.

References

- [1] Roberto Scopigno, “3D scanning: potenzialità e limiti delle tecnologie di acquisizione automatica”, DDD rivista trimestrale di Disegno Digitale e Design.
- [2] Andrea Fusiello, “Visione Computazionale: appunti delle lezioni”, Dipartimento di Informatica, Università di Verona, 1 Marzo 2003.
- [3] Halioua M., Liu H.C., “Optical Three-Dimensional Sensing by Phase Measuring Profilometry”, Optics and Lasers in Engineering, Vol. 11, pp. 185-215, 1989.
- [4] Takasaki H., “Moiré Topografy”, Applied Optics, Vol.9, Issue 6, June 1970.
- [5] Post D., Han B., Ifju P., “Moiré Methods for Engineering and Science-Moiré Interferometry and Shadow Moiré”, Photomechanics for Engineers, Pramod Rastogi ed., Springer-Verlag, 2000.
- [6] Sirat G., Psaltis D., "Conoscopic Holography", SPIE vol. 523 Application of Holography, 1985.
- [7] Conti P., Nigrelli V., Petrucci G., “Computer aided holographic investigation of the deformation of cylindrical shells”, Proceedings of the International Conference on Non-destructive testing & stress-strain measurement, FENDT '92, Tokyo, 1992.
- [8] Lombardo E., Martorelli M., Nigrelli V., “Non-Contact Roughness Measurement in Rapid Prototypes by Conoscopic Holography”, XII ADM International Conference, Rimini, Italy, 5-7 september 2001.

- [9] Schirripa Spagnolo G., Guattari G., Sapia C., Ambrosini D., Paoletti D., Accardo G., "Contouring of artwork surface by fringe projection and FFT analysis", *Optics and Lasers in Engineering*, vol. 33, 2000.
- [10] Várady T., Martin R.R., Cox J., "Reverse engineering of geometric models – an introduction", *Computer Aided Design*, vol.29, n.4, 1997, pp.255-268.
- [11] Humbert D., Cunin D., Barlier C., "Les Différentes Technologies de Numérisation et Capteurs Associés", in *Proc. of Conference Numerisation 3D*, pp.1-14, Paris May 1997.
- [12] Chen F., Brown G.M., Song M., "Overview of three-dimensional shape measurement using optical methods", *Optical Engineering*, vol. 39, n. 1, pp. 10-22, 2000
- [13] "The Digital Michelangelo Project" *Department of Stanford University Second International Conference on 3D Digital Imaging and Modeling in Canada*, 1999.
- [14] Subodh Kumar, Dean Snyder, Donald Duncan, Jonathan Cohen, Jerry Cooper, "Digital Preservation of Ancient Cuneiform Tablets Using 3D-Scanning," *Fourth International Conference on 3-D Digital Imaging and Modeling*, 2003, pp.326-333.
- [15] Zhang Z., Zhang D., Peng X., Performance analysis of a 3D full-field sensor based on fringe projection. (2004) *Optics and Lasers in Engineering*, 42 (3), pp. 341-353.
- [16] H. Liu, W. H. Su, K. Reichard, S. Yin, "Calibration based phase shifting projected fringe profilometry for accurate absolute 3D surface profile measurement", *Science Direct* November 2002.

- [17] P. J. Tavares, M. A. Vaz, "Linear calibration procedure for the phase to height relationship in phase measurement profilometry", Science Direct February 2007.
- [18] Hornbeck, L. J. and Nelson. W. E. "Bistable Deformable Mirror Device," OSA Technical Digest, Vol. 8, Spatial Light Modulators and Applications, 1988.
- [19] Faugeras, O. Three-Dimensional Computer Vision: A Geometric Viewpoint, MIT Press, 1993.
- [20] Reid, D. "An Algorithm for Tracking Multiple Targets," Trans. Auto. Control, Vol. 24, No. 6, 1979
- [21]. Brown, R. G. and Hwang, P. Y. C. Introduction to Random Signals and Applied Kalman Filtering, 3 ed., John Wiley & Sons, 1997.
- [22] Heikkilä, J. and Silven, O. "A Four-Step Camera Calibration Procedure with Implicit Image Correction," Proc. CVPR, 1997.
- [23] Pollefeys, M. "Self-Calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences," Ph. D. Dissertation, Katholieke Universiteit Leuven, 1999.
- [24] Pulli, K. "Multiview Registration for Large Data Sets," Proc. 3DIM, 1999.
- [25] Blais, G. and Levine, M. "Registering Multiview Range Data to Create 3D Computer Objects," Trans. PAMI, Vol. 17, No. 8, 1995.
- [26] Neugebauer, P. "Geometrical Cloning of 3D Objects via Simultaneous Registration of Multiple Range Images," Proc. SMA, 1997.
- [27] Simon, Fast and Accurate Shape-Based Registration, Ph. D. Dissertation, Carnegie Mellon University, 1996.

- [28] Z. Zhang. "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations" Microsoft Research, December 1998.
- [29] Roger Y. Tsai "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses" IEEE Journal of Robotics and automation, Vol. RA-3, No. 4, August 1987
- [30] R. C. Gonzalez e P. Wintz, "Digital Image Processing", MA: Addison-Wesley, 1987.
- [31] H. Liu, W. H. Su, K. Reichard, S. Yin, "Calibration based phase shifting projected fringe profilometry for accurate absolute 3D surface profile measurement", Science Direct November 2002.
- [32] "Improved Three-step Phase Shifting Profilometry Using Digital Fringe Pattern Projection" Yingsong Hu, Jiangtao Xi, Joe Chicharo and Zongkai Yang.
- [33] S. Zhang, P. S. Huang, "Novel method for structured light system calibration," Optical Engineering, 45(8), 083601, August 2006.
- [34] H. Liu, W. Su, K. Reichard, S. Yin, "Calibration based phase-shifting projected fringe profilometry for accurate absolute 3D surface profile measurement," Optics communications 216, pp. 65-80, 2003.
- [35]M. Drouin, G. Godin, S. Roy, "An Energy Formulation for Establishing the Correspondence Used in Projector Calibration," The Proceedings of the Fourth International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT). Atlanta, Georgia, USA. June 18-20, 2008. NRC 50357.

- [36] W. Gao, L. Wang, Z. Hu, "Flexible Calibration of a Portable Structured Light System through Surface Plane," *Acta Automatica Sinica*, Vol. 34, No. 11, November 2008, pp. 1358 – 1362.
- [37] Hartley, Zisserman, "Multiple View Geometry in Computer Vision," 2nd Ed, Cambridge University Press.
- [38] A. J. Lacey, N. Pinitkarn, N. A. Thacke, "An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration," *Proceedings of the British Machine Vision Conference (BMVC)2000*.
- [39] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct Least Square Fitting of Ellipses," *IEEE Trans. on PAMI*, vol. 21, no. 5, May 1999, pp. 476-480.
- [40] C. Bräuer-Burchardt, M. Heinze, C. Munkelt, P. Kühmstedt, and G. Notni, "Distance Dependent Lens Distortion Variation in 3D Measuring Systems Using Fringe Projection", *British Machine Vision Conference 2006, BMVC06. Proceedings. Vol. 1: The 17th BMVC was held 4-7 September 2006 in Edinburgh. Edinburgh: BMVA, 2006, pp. 327-336*
- [41] A.A. Magill. Variation in distortion with magnification. *Journal of Research of the National Bureau of Standards*, 54(3):153–142, 1955.
- [42] D.C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.
- [43] C.S. Fraser and M.R. Shortis. Variation of distortion within the photographic field. *Photogrammetric Engineering and Remote Sensing*, 58(6):851–855, 1992.
- [44] J. Dold. Ein hybrides photogrammetrisches Industriemesssystemhöchster Genauigkeit und seiner Überprüfung. PhD thesis, Universität der Bundeswehr

- [45] E. Kruck. Lösung großer Gleichungssysteme für photogrammetrische Blockausgleichungen mit erweitertem funktionalem Modell. PhD thesis, Universität Hannover, 1983. München, 1997.
- [46] P. Brakhage, G. Notni, and R. Kowarschik. Image aberrations in optical threedimensional measurement systems with fringe projection. *Applied Optics*, pages 3217–3223, 2004.
- [47] P. Sturm, S. Ramalingam. A generic concept for camera calibration. *ECCV*, 2004.
- [48] P. Sturm, INRIA Rh^{one}-Alpes, 38330 Montbonnot, France “Multi-View Geometry for General Camera Models”
- [49] J. Tian, Y. Ding, X. Peng. “Self-calibration of a fringe projection system using epipolar constraint”, *Optics & Laser Technology*, Volume 40, Issue 3, April 2008, Pages 538-544.
- [50] R. Hartley, “In defence of the 8-point algorithm”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, June 1997.
- [51] P. R. S. Mendonça, R. Cipolla, “A simple technique for Self-Calibration”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Fort Collins, Colorado, USA (June), volume I, pages 500-505, 1999.
- [52] M. Bevilacqua, C. Liguori, A. Paolillo, “Stereo Calibration for a Camera – Projector Pair”, *Proceedings of IEEE International Instrumentation and Measurement Technology Conference (I²MTC)*, Austin, TX, USA, 3-6 May 2010, pp. 492-497, ISBN 978-1-4244-2833-5, ISSN 1091-5281.
- [53] R. Hartley, A. Zisserman, “Multiple View Geometry in Computer Vision”, 2nd. Ed., Cambridge University Press, 2003