3D ACTIVE STEREO MEASUREMENT IN A REGULAR MESH WITH RANDOM PATTERN AND LASER SPECKLE PROJECTION

Tiago L. F. da Costa Pinto , Dr. Eng.
Armando Albertazzi G. Jr., Dr. Eng.
Labmetro - EMC - UFSC, CP 5053, Florianópolis - SC, Brasil, 88040-970
tlp@labmetro.ufsc.br
albertazzi@labmetro.ufsc.br

Abstract. *Optical 3D surface measurement technologies have achieved tremendous advances in research and development in last decades. Optical measurement systems using stereo vision with sinusoidal fringe projection are one of the most used methods to measure free form surfaces. Recent developments in this field include fringe code reduction, novel encoding methods and projection of laser speckles with temporal correlation methods. In the classical triangulation approach, 3D points are calculated from homologous image point pairs previously determined. In the inverse triangulation method, previously developed by the authors, several candidate 3D tested points are mathematically projected in both images planes and the information associated with these points are compared until a valid 3D point is reached, resulting in a regular and organized mesh of points. The objective of this work is to compare the inverse triangulation method performance for three kinds of projected patterns: sinusoidal fringe with Gray code projection, random pattern projection and laser speckle projection with temporal correlation. The comparison includes measurements of a reference plane as described in the VDI-VDE 2634 guidelines. The results show the suitability of the three tested projected patterns and the influence of random pattern and laser speckle number of acquired images. Sinusoidal fringe and Gray code projection, followed by random pattern and laser speckle projection, shows the best performance.*

Keywords: *3D measurement, stereo vision, random projection, laser speckle.*

1. INTRODUCTION

Increasing commercialization and use of three-dimensional surface measurement technologies in industrial inspection was caused by tremendous advances in research and development in the last decades resulting in improved accuracy, high density of measured points and short measurement times (Geng J., 2011) (Bräuer-Burchardt, C., *et al* 2011).

Optical measurement systems based on sinusoidal fringe projection are one of the most commonly used methods to measure full-field free form 3D surfaces being one of the most active research areas in optical metrology (Liu, Y., *et al*., 2011). The range of applications of these systems is large and includes measurement of mechanical components and assemblies, cultural heritage, reverse engineering and human body shape (Gorthi, S. and Rastogi, P., 2010).

Recent developments include fringe code reduction and parallel data processing (Bräuer-Burchardt, C., *et al* 2011), novel encoding methods (Liu, Y., *et al*., 2011), defocused projection of binary patterns (Lei, S. and Zhang, S., 2009), projection of laser speckles (Schaffer, M., Grosse, M. and Kowarschik R., 2010) (Schaffer, M, *et al.*, 2011) and projection of band limited random patterns (Wiegmann, A., *et al.*, 2006). Though stereo vision triangulation methods showed improvements, they are more or less evolutions (Lindstrom, P. 2010) of already established methods (Hartley R. and Zisserman A., 2003).

Anyway all of these approaches, using conventional triangulation methods, have in general the disadvantage that the acquired data is an unordered and non structured cloud of points in 3D space. These non-structured clouds of points are acceptable for visualization purposes, but not very handy to use for calculations and feature extraction. Data sets organized in a regular mesh of points are also a lot easier to visualize, unify and compare.

The algorithm previously developed by the authors (Pinto T. *et al*, 2012), named here as Inverse Triangulation, overcome some of these disadvantages. Inverse Triangulation change from a sensor oriented organization of the point cloud to an object of interest naturally oriented organization of the point cloud, with user defined density.

The objective of this work is to analyze and compare the Inverse Triangulation method performance for three different kinds of projected patterns: (a) sinusoidal fringe and Gray code projection, (b) band limited random pattern projection with temporal correlation and (c) laser speckle projection with temporal correlation. The comparison includes measurements of a reference plane as described in the VDI-VDE 2634 guidelines.

2. MEASUREMENT PRINCIPLES

This section describes the measure principles related to this work based on the pinhole camera model, as the classical triangulation approach and explain the developed inverse triangulation approach. It also shows how to detect the

homologous points using temporal correlation for projection of band limited random patterns and projection of laser speckles.

2.1 Pinhole camera model

The pinhole camera model considers that the projection of a 3D scene on an image plane through a lens can be described by projecting 3D points in a plane through a central point named the center of projection.

For any point $\boldsymbol{M}$ in 3D space, its representation in the image $\boldsymbol{m}$ is located where the line connecting $\boldsymbol{M}$ with the center of projection $\boldsymbol{C}$ intersects the image plane π. The projection $\boldsymbol{m} = (x.w, y.w, w)^T$ of a 3D point $\boldsymbol{M} = (X, Y, Z, 1)^T$ in the plane π can be described by the equation (Hartley and Zisserman, 2003):

$$\boldsymbol{m = P.M = A.[R\ t].M} \tag{1}$$

The projection matrix **P** is a 3x4 matrix containing a combination of extrinsic ($\boldsymbol{R}$, $\boldsymbol{t}$) and intrinsic ($\boldsymbol{A}$) parameters and not consider distortions introduced by imperfections of the lenses used, which can be quite significant. An usual model is consider the introduction of radial and tangential distortion in the projected point, better reproducing the actual light rays paths when they pass through the lens. The projection matrix and distortion coefficients should be determined by calibration (Heikkilä J. and Silvén O., 1997) (Zhang, S and Huang, P, 2006). In this work, the calibration is realized using the "Camera Calibration Toolbox for Matlab" available on the internet (Bouguet, J., 2010).

2.2 Triangulation

Stereo vision is a method for the three-dimensional reconstruction of a scene from corresponding points by triangulation with two cameras (Kanatani *et al*, 2008). In this case, the intrinsic camera parameters and the relative position between the cameras are not changed during the measurement allowing calibration as a preliminary step (Sünderhauf and Prötzeler, 2006).

To perform the triangulation and determine the three dimensional position of a point it is necessary to determine its homologous points, or its corresponding position in each of the two images, as a preliminary step. For this, the scene must contain heterogeneous texture to allow the use of simple digital image correlation. For the measurement of parts that have a homogeneous texture, auxiliary systems for the projection of structured light may be used. The epipolar geometry can be used to facilitate the determination of corresponding or homologous points between image pairs (Hartley and Zisserman, 2003).

From the pinhole model, the previously collected calibration data and the corresponding positions of homologous points in each image, it is possible to determine the lines in 3D-space that may contain these points. The intersection of these lines determines the position of point $\boldsymbol{M}$ in space by the triangulation process.

2.3 Inverse Triangulation

In the inverse triangulation method (Pinto T. *et al*, 2012), two spatial coordinates of an unknown 3D point are selected by the user and the third coordinate is scanned. The resulting 3D point is mathematically projected in both images planes and a test quantity, associated with these projected points, is compared until a valid 3D point is reached. Since the user can freely determine two coordinates, a regular and organized mesh of points can be obtained. The test quantity associated with these points can be phase values, spatial or temporal correlation coefficients.

To obtain a regular 3D point cloud by inverse triangulation, a regular grid in X and Y direction in the world coordinate system is first defined with a user defined point spacing. From the X and Y coordinates of a given point on the regular grid, the search for the corresponding Z coordinate is performed numerically. The value of Z is changed and the resulting 3D point is projected onto each image plane of the cameras using the projection equations (1) with lens distortion. The test quantity corresponding to the calculated camera pixel is determined with subpixel resolution and compared for both cameras. The correct value of the Z coordinate is achieved when the test quantity retrieved from the image planes are equal for all cameras.

The test quantity obtained from the image planescan be: (a) absolute phase difference if sinusoidal fringe projection is used, (b) spatial correlation if natural surface texture or projected texture is used and (c) temporal correlation if a sequence of structured patterns (e.g. random patterns) is projected on the object surface. In practice, the smallest absolute difference between the phase values or maximum correlation must be found. Thus, the estimate of the Z coordinate associated with each grid point (X, Y) is determined.

This principle can be seen in Fig. 1: for different Z coordinates, for a grid point (X, Y), their projected points on the image planes of the cameras are calculated. The algorithm can be summarized as:

1. Define the grid spacing and ranges in the X and Y direction to be measured;
2. Define the limits of the Z value variation: $Zmin$ and $Zmax$ (e.g. $Zmin$ = -100 mm, $Zmax$ = 100 mm);

3. Set ΔZ (e.g. ΔZ = 0,01 mm);
4. Choose a pair of *X, Y* coordinates from the defined grid;
5. Project each point ***M*** formed by the coordinates *X, Y* and (*Zmin* + ΔZ * *i*) ,with *i* = 0,1,...,(*Zmax*–*Zmin*)/ΔZ, in the image planes using equation (1) and distortion for each camera.
6. Determine the projected point (and *Z*) with the smallest phase value difference or largest correlation;
7. Set the *Z* coordinate for this *X*, *Y* coordinates in the grid;
8. Return to step 4 to measure another point of the predefined grid, i.e. different *X* and *Y*, until the *Z* values for all points on the grid are calculated.

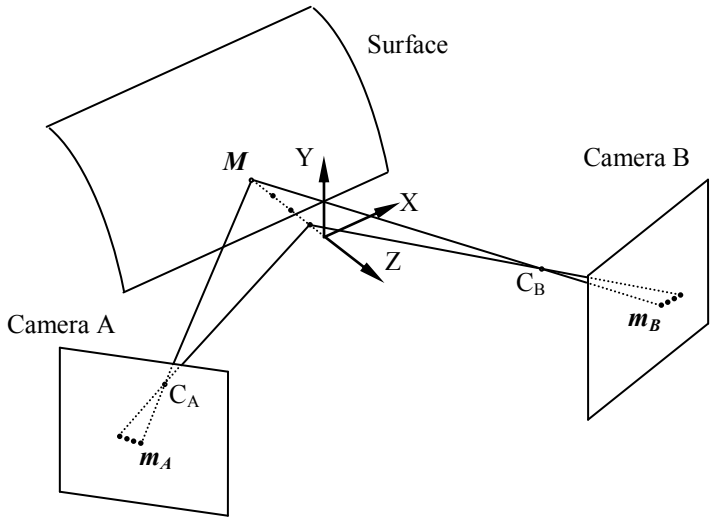

Figure 1: *Z* coordinate determination for an object point with pre-defined *X, Y* coordinates (Pinto T. *et al*, 2012).

The Inverse Triangulation technique, exposed in this section, allow measurement of a dense point cloud (which can be millions of points), organized in a regular grid in a predefined coordinate system defined in the camera calibration process. This approach allows points to be naturally structured and organized in the object space and the 3D points are determine simultaneously with the homologous points, intrinsically respecting the epipolar geometry. Author's previous work on the subject, with sinusoidal and gray code projection, include (Pinto, T. *et al*, 2011) (Fantin, A., *et al*, 2007) (Pinto T. *et al*, 2012).

In Classical Triangulation the 3D point are determined by straight lines leaving from homologue points in the cameras, in contrast, in the Inverse Triangulation the 3D point is determined by projecting it to the cameras image planes. The inverse triangulation approach can be naturally extended for multiple cameras triangulation.

2.4 Temporal correlation

In the Inverse Triangulation, the phase-value retrieved from cameras can be compared to determine the best Z for each grid point (Pinto T. *et al*, 2012). Instead of using the phase-values, it is also possible to use temporal correlation with several successively acquired images to calculate a normalized cross-correlation coefficient for each projected point in both image planes. For each projected point, the information (gray value) is retrieved along the successively acquired images, or in another words, through time. Temporal correlation is already used with the classical triangulation method (Schaffer, M. *et al*, 2011).

For a tested 3D point in the inverse triangulation, each projected point (*i,j*) in one camera can be correlated with each projected point (*i',j'*) from a second camera using the temporal normalized cross-correlation coefficient equation (Schaffer, M. *et al*, 2011):

$$\rho(i,j,i',j') = \frac{\sum_{t=1}^{N}(g(i,j,t)-\bar{g})(g'(i',j',t)-\bar{g}')}{\sqrt{\sum_{t=1}^{N}(g(i,j,t)-\bar{g})^2}\sqrt{\sum_{t=1}^{N}(g'(i',j',t)-\bar{g}')^2}} \quad (2)$$

For each projected point, $\bar{g}$ and $\bar{g}'$ are the mean gray value for the grey value vectors ***g*** and ***g'*** at each pixel along time *t,* for each camera. The tested *Z* resulting the pixel pair (*i, j, i', j'*) with highest correlation value ρ define the *Z* for

this (*X,Y*) node in the grid. A threshold for the correlation value ρ can be used to avoid outliers. The temporal vector is defined with subpixel interpolated gray levels for every projected point on image planes. Homologous pixels corresponding to the same object point share a similar gray value vector as shown in figure 2.

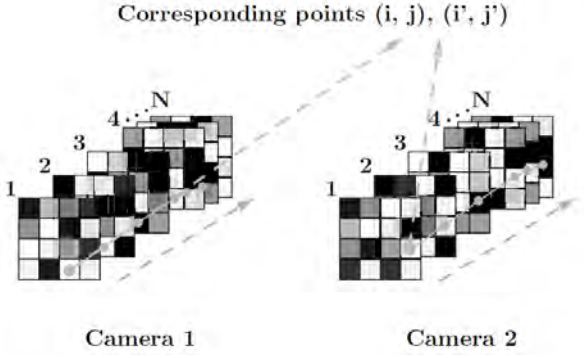

Figure 2: Corresponding pixels from two views with similar gray value sequence (Schaffer, M. *et al*, 2011).

Sequentially, the projection of a random pattern onto the object and image acquisition for each camera is realized several times, defining the temporal sequence *t* of *N* images. In this work, the projected patterns were images with band limited random patterns (BLP) projected by a multimedia projector as described by A. Wiegmann, *et al.*, 2006 or laser speckle projection (LS) by a specially constructed laser projector similar as described by Schaffer, M. *et al*, 2011. Figure 3 shows the images acquired from these projected patterns with (a) band limited random pattern projected by a multimedia projector and (b) blue laser speckle projection.

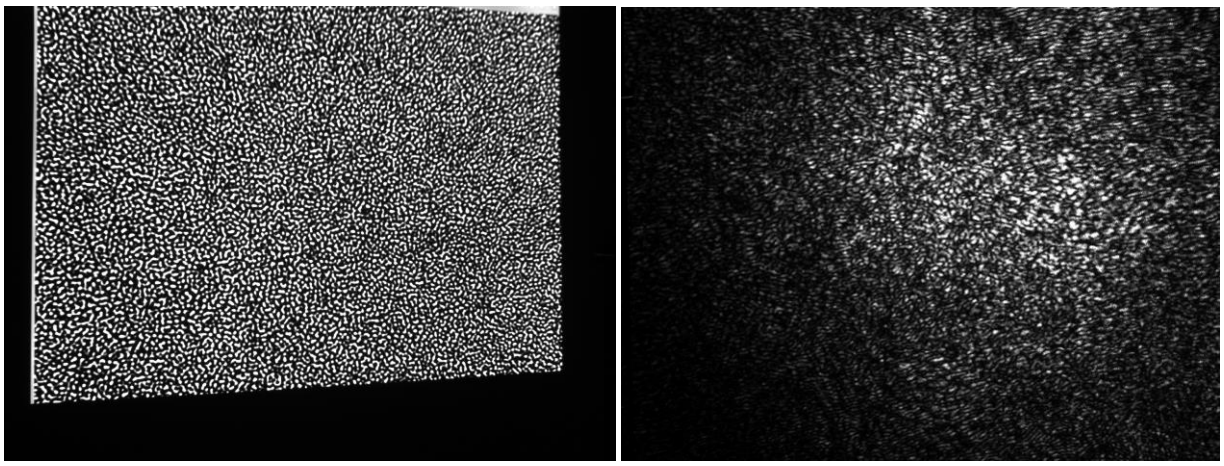

Figure 3: Acquired images from (a) band-limited random projection and (b) laser speckle projection.

3. EXPERIMENTS

Experiments were realized to evaluate the performance of the Inverse Triangulation method with both types of projection for temporal correlation. The influence of number of *N* images in the temporal correlation and in the error of a measured plane with the both types of projection patterns are shown. Earlier evaluation of the method using sinusoidal fringe projection is used for comparison as described by Pinto T. *et al.*, 2012.

3.1 Number of images for temporal correlation

The number *N* of images used for temporal correlation, or the length of vectors **g** and **g'**, influence the sharpness and the evidence of the correlation peak. An experiment to evaluate different *N* numbers of images in the correlation calculation was done for both types of projections. Figure 4 shows a graph of temporal correlation curves, using band-limited random pattern projection, for N = 10 (red) and N = 40 images (blue). The tested grid point is *X=0, Y=0* with *Z* ranging from *Zmin* = -100 mm to *Zmax* = 100 mm with scanning steps of ΔZ = 0,01 mm. The coordinate system is defined during camera calibration and its origin is at the center of measurement volume about 600 mm of the cameras with Z axis pointing toward measurement system. Each tested point is related to a different pixel coordinate at each image.

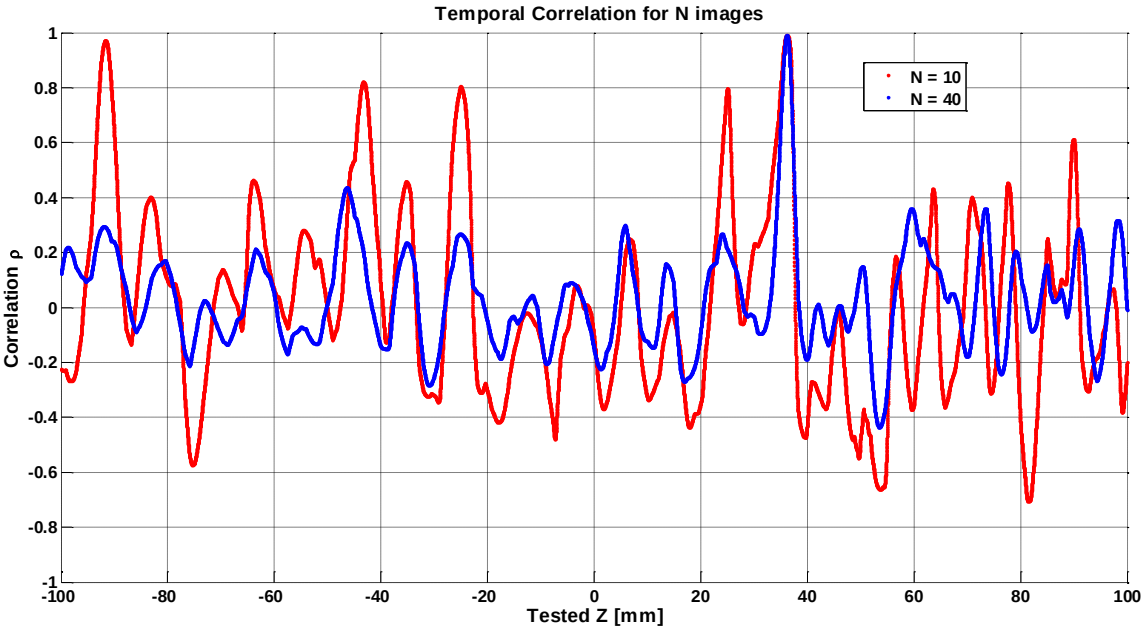

Figure 4: Temporal correlation curves for N = 10 (red) and N = 40 images (blue) for band-limited random projection.

Higher number of *N* images gives a more distinct correlation peak, but both peaks are detected at the approximately same location (Z ~ 36,20 mm).

A more detailed view of the peak region for a varied number of *N* images for temporal correlation can be seen in the Figure 5 below. In this graph the greater the number of *N* images, the narrower is the peak and above *N* = 20 images the difference in the shape of the peak is insignificant.

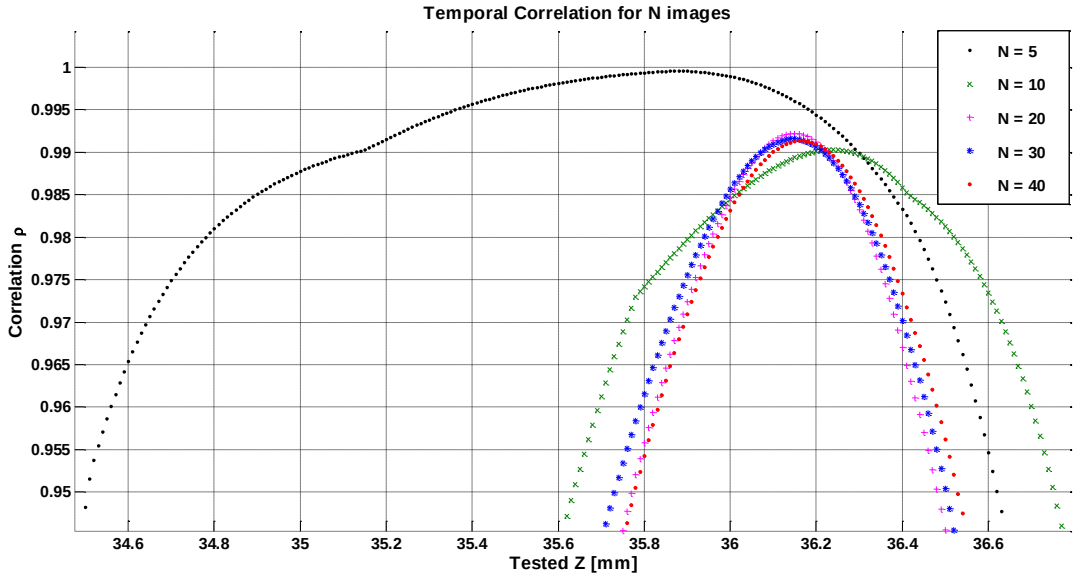

Figure 5: Peaks region of temporal correlation curves for band-limited random projection.

For laser speckle projection the Figure 6 shows the correlation curves for N=10 and N=40 images. In this case the peak for N = 10 is in different Z position of the peak with N = 40 images. This is caused by poor correlation quality in this kind of projection. Figure 7 show a more detailed view of the peak region for a varied number of *N* images for temporal correlation.

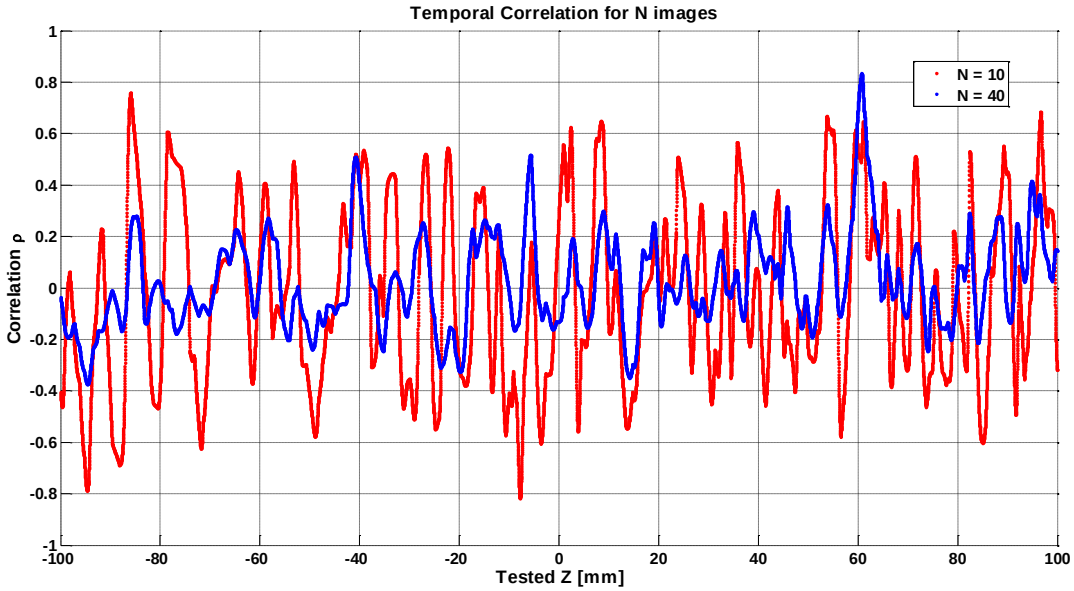

Figure 6: Peaks region of temporal correlation curves for laser speckle projection.

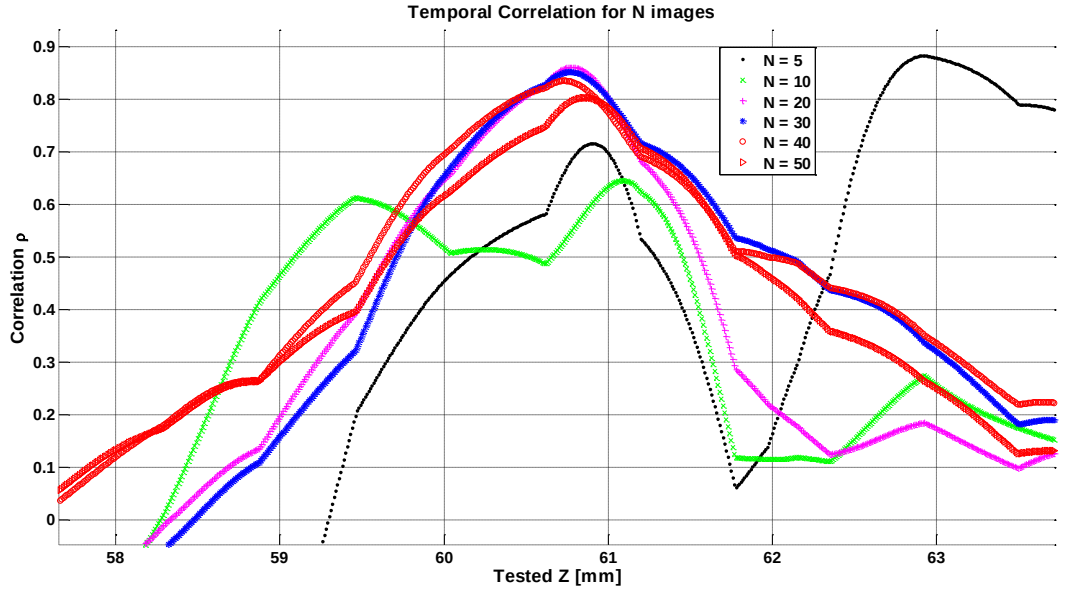

Figure 7: Peaks region of temporal correlation curves for laser speckle projection.

For laser speckle projection, as for band-limited random pattern projection, $N \leq 10$ images results in poor and ambiguous peak detection. It's also evident, for this system configuration, that laser speckle projection has lower performance. This is caused by the subjective speckle also present, causing different intensities depending on the point of view for both cameras.

Even with larger number of N images, the position of the peak has small lateral displacements in both cases, resulting in different Z's for each number of N images used. To analyze the influence of N number in the 3D measurement itself a reference plane is measured.

3.2 Plane measurement

Measurement of a reference plane was done to evaluate the influence of the number of images in the error of the points measured in reference of a best-fit plane. The standard deviation of the error and the RE parameter, as described in the VDI-VDE 2634 guidelines was evaluated for one position of the reference plane. RE is the amplitude between the

most negative error and the most positive error of the measured points. Figure 8 and Figure 9 shows the influence of *N* images in the standard deviation and RE parameter for BLP and LS projection respectively.

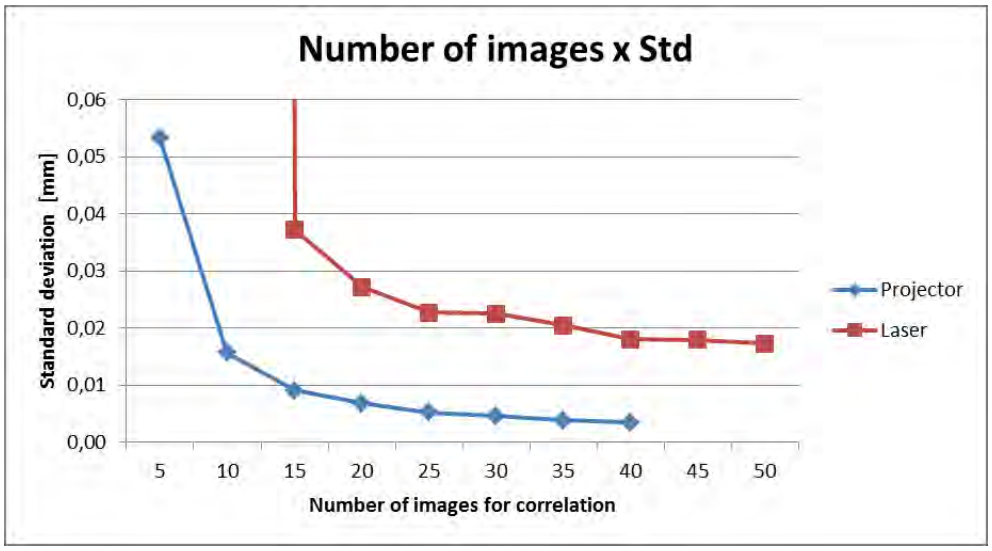

Figure 8. Number of images influence in standard deviation for a fitted plane.

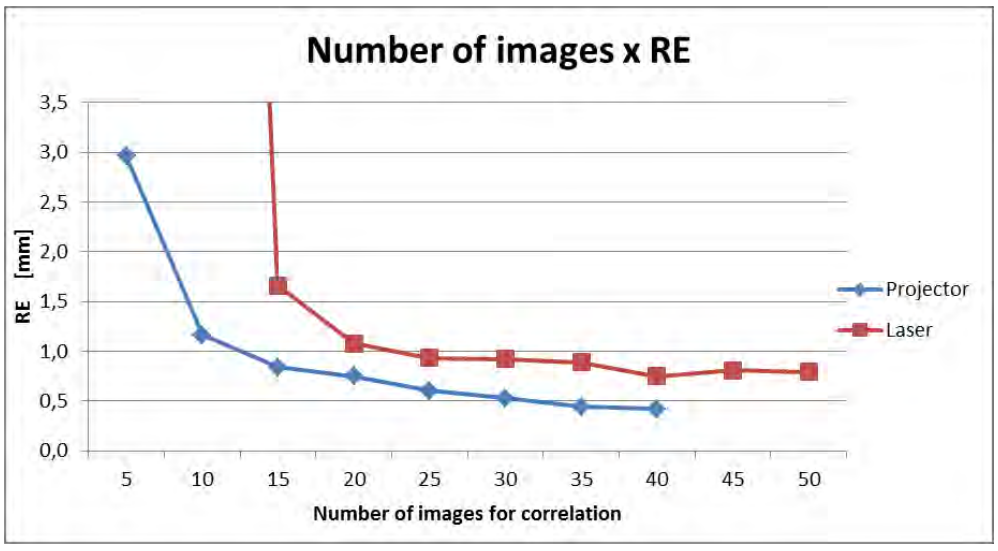

Figure 9. Number of images influence in RE parameter for a fitted plane.

The higher the number of images, the smaller standard deviation and RE parameter. For BLP projection the standard deviation is below 0,01 mm and RE below 1,0 mm for N = 15 images. For LS projection the standard deviation is below 0,025 mm and RE below 1,0 mm for N = 25 images. The worst results for LS projection are mainly caused by subjective speckle and non-uniform intensity distribution in the image. Previous measurements with sinusoidal fringe projection, as described by (Pinto T. *et al*, 2012), achieve RE = 0,10 mm.

3.3 Measurement exemples

Figure 10 show a 3D point cloud of a measured plane and Figure 11 show a 3D point cloud of a dummy face. Acquisitions where made with BLP projection and *N* = 40 images. Measurement

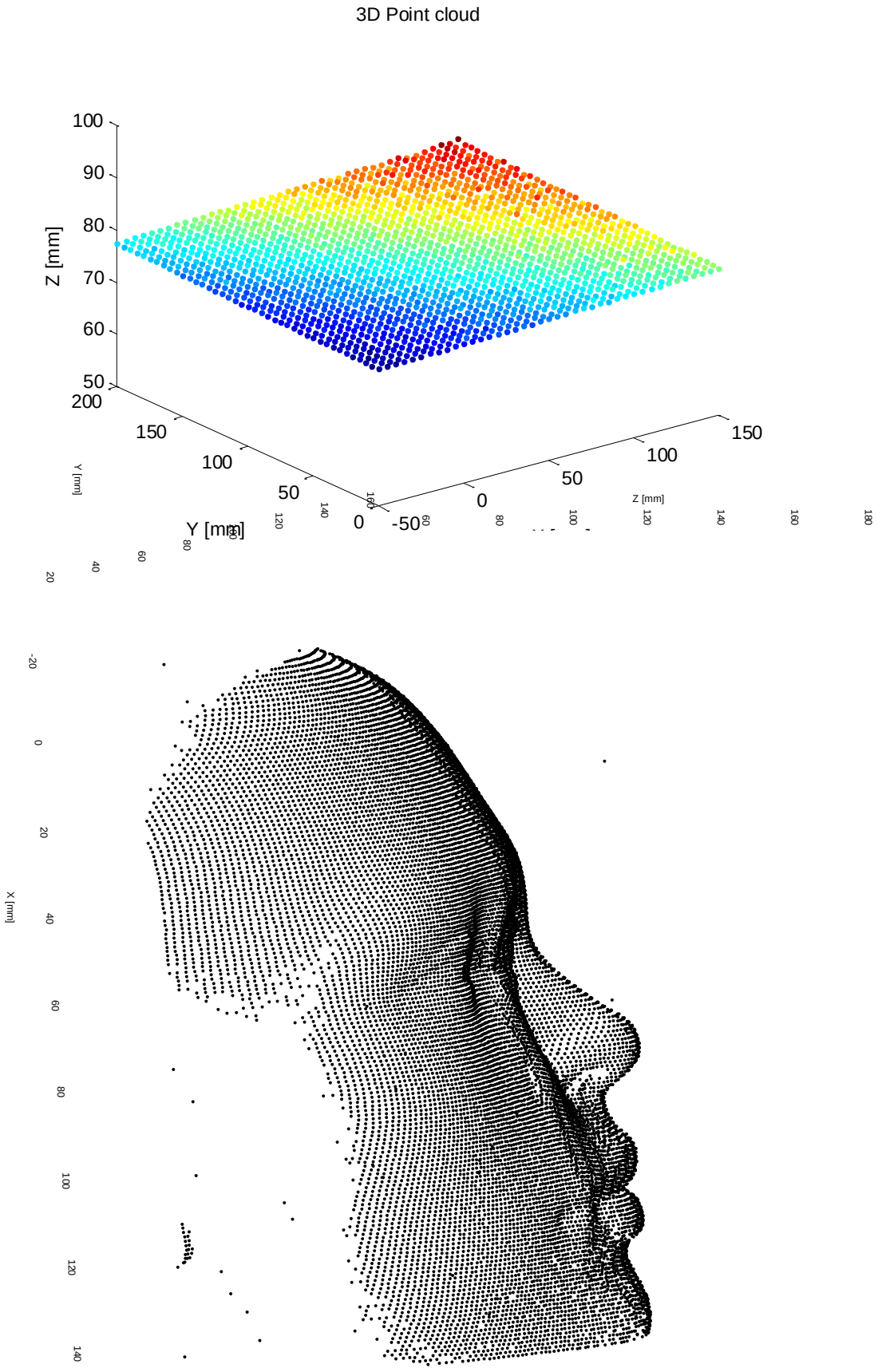

Figure 11. 3D Point cloud of a dummy face.

4. CONCLUSION

The Inverse Triangulation method can be used with temporal correlation technique. The band limited random pattern projection have better performance over laser speckle projection. This is mainly caused by subjective speckle and non-uniform intensity distribution in projected patterns.

Future work includes optimization of the laser speckle projector, tests with spatial-temporal correlation and underwater measurements.

5. REFERENCES

Bouguet, J. 2010 Bouguet, J. 2010 “Camera Calibration Toolbox for Matlab” 1 June 2013, <http://www.vision.caltech.edu/bouguetj/calib_doc/>

Bräuer-Burchardt, C., Breitbarth, A., Kühmstedt, P., Schmidt, I., Heinze, M., Notni, G., "Fringe projection based high-speed 3D sensor for real-time measurements," Proc. SPIE 8082, 808212 (2011).

Fantin, A., et al “Efficient mesh oriented algorithm for 3D measurement in multiple camera fringe projection,” Proc. SPIE, 6616, p. 6616 1B (2007).

Geng J., "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon. 3, 128-160 (2011).

Gorthi S., Rastogi, P., "Fringe Projection Techniques: Whither we are?," Opt. L. Eng., 48(2):133-140, (2010).

Hartley R. and Zisserman A., Multiple View Geometry, (Cambrige University Press, 2003).

Heikkilä J. and Silvén O. “A Four-step Camera Calibration Procedure with Implicit Image Correction,” In Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition, 1997.

Kanatani, K., Sugaya Y., and Niitsuma H., “Triangulation from two views revisited: Hartley-Sturm vs. optimal correction” Proceedings of the 19th British Machine Vision Conference, 2008.

Lei, S., Zhang, S., "Flexible 3-D shape measurement using projector defocusing," Opt. Lett. 34, 3080-3082 (2009).

Lindstrom, P. "Triangulation Made Easy," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (San Francisco, CA, 2010), pp. 1554- 1561, 2010.

Liu, Y., Su, X., Zhang, Q., "A novel encoded-phase technique for phase measuring profilometry," Opt. Express 19, 14137-14144 (2011).

Pinto T., Kohler. C. and Albertazzi A. “Regular mesh measurement of large free form surfaces using stereo vision and fringe projection” Optics and Lasers in Engineering, Vol. 50, Issue 7, 2012, http://dx.doi.org/10.1016/j.optlaseng.2012.03.003

Pinto, T., Kohler, C., Albertazzi, A. "Optical measurement and comparison of large free form surfaces through a regular mesh", Proc. SPIE 8082, 80821A (2011).

Schaffer, M., Große, M., Harendt B. and Kowarschik, R. “Fast 3D shape measurements using laser speckle projection”, Proc. of SPIE Vol. 8082, 808219, 2011 http://dx.doi.org/10.1117/12.889471

Schaffer, M., Grosse, M., Kowarschik R., "High-speed pattern projection for three-dimensional shape measurement using laser speckles," Appl. Opt. 49, 3622-3629 (2010).

VDI/VDE 2634 Part 1 “Optical 3D measuring systems – Imaging systems with point-by-point probing”, Düsseldorf, 2002.

VDI/VDE 2634 Part 2 “Optical 3D measuring systems – Optical systems based on area scanning”, Düsseldorf, 2002.

Wiegmann, A., Wagner, H. and Kowarschik, R., “Human face measurement by projecting bandlimited random patterns” Optics Express, 7692 Vol. 14, No. 17, 2006.

Zhang, S and Huang, P. “Novel method for structured light system calibration” Opt. Eng. 45(8), 083601 (2006).

6. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.