EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2021)
A. Chalmers and V. Hulusic (Editors)

*Short Paper*

# Fully Automatic Mechanical Scan Range Extension and Signal to Noise Optimization of a Lens-Shifted Structured Light System

H.Kutlu[1], M.Ritz[1], P.Santos[1], and D. W. Fellner[1,2,3]

[1]Fraunhofer IGD, Germany
[2]Graz University of Technology, Institute of Computer Graphics and Knowledge Visualization, Austria
[3]TU Darmstadt, Germany

**Abstract**

*Digitization of cultural heritage is of growing importance, both for its preservation for coming generations in the face of looming dangers of natural decay or intentional destruction, and current generations, that increasingly have access to virtual cultural heritage for interactive exploring or scientific analysis. These goals can only be achieved by 3D replicas at reasonable quality and resolution, to come as close as possible to the original. This brings about several challenges to overcome. The challenge of digitizing huge numbers of artefacts is addressed by CultLab3D, the first fully automatic 3D digitization system. Another challenge is the size of objects, as each digitization system is designed for a certain optimum measurement range, leaving which results in loss of quality. Due to optical and mechanical constraints, most systems are not able to faithfully reconstruct objects under a certain size limit in their full geometric detail. Historic coins are one good example, where the deterioration of the surface structure in most cases has progressed to a degree that it not even is perceptible through the fingernail. This challenge is addressed by a modular extension of CultLab3D, the MesoScanner, which is a structured light system that breaks limits in depth resolution through a mechanical lens-shifting extension, allowing physically shifting of fringe patterns on top of the well-known multi-period phase shift method. This is where this work adds two major improvements: First, the signal to noise ratio and thus reconstruction quality has been improved significantly through several algorithmic processing steps. Second, the physical limitation of the measurement range was removed using a 2D actuator steering the object mount, thus allowing for a measurement range at theoretically arbitrary size. This opens up the fully automatic handling of two scenarios: Complete digitization of objects exceeding the measurement range, and unsupervised digitization of large collections of small objects in one run.*

**CCS Concepts**

• ***Hardware*** → ***Scanners;*** • ***Computing methodologies*** → ***Reconstruction; 3D imaging;*** *Visual inspection;*

## 1 Introduction and related work

Digitization of cultural heritage for the sake of its preservation for coming generations is of growing importance, looking at the looming dangers of natural decay or even intentional destruction. For current generations, digital replicas of artefacts play an important role too, especially thanks to current technology that opens up ways of interactive exploring and scientific analysis for everyone, even in the web. When it comes to digitization, 2D is not sufficient anymore – only 3D replicas at reasonable quality and resolution allow for virtual true-to-reality interaction with the original. This brings about several challenges to overcome. The challenge of digitizing huge numbers of artefacts is addressed by CultLab3D [SRT*14], the first fully automatic 3D digitization system developed by the Fraunhofer Institute for Computer Graphics Research. Another challenge is the size of objects, as each digitization system is designed for a certain optimum measurement range. If an object does not meet the size requirement, the consequence is loss of quality. Due to optical and mechanical constraints, most systems are not able to faithfully reconstruct objects under a certain size limit in their full geometric detail. One example are coins: While their form and shape can be an interesting aspect, the focus is mostly on inscriptions and engravings on their surface, which are not reconstructed accurately by scanners designed for typically larger scanning volumes. The deterioration of the relief found on historic coins in most cases has progressed to a degree where it not even is perceptible by the fingernail anymore. This challenge is addressed by a modular extension of CultLab3D, the MesoScanner [RLS*12], which is a structured light system that breaks limits in depth resolution through a mechanical lens-shifting extension, allowing physically shifting of fringe patterns on top of the well-known multi-period phase shift method [LM07]. On the other hand, digitizing high-detail geometry typically is bound to smaller scanning volumes, which in turn leads to the problem that larger objects cannot be completely digitized, and especially for the sce-





DOI: 10.2312/gch.20211407

nario of entire coin collections, digitizing each coin in turn requires a huge time effort. This is where this work, based on the thesis [Kut20], adds major improvements by introducing the MesoScannerV2, with the following contributions:

- Removal of physical limitation of measurement range through automatic mechanical translation of object, leading to:
  - Complete digitization of objects exceeding the measurement range, as well as unsupervised digitization of large collections of small objects in one run
- Significant improvement of signal to noise ratio, and thus reconstruction quality, through several algorithmic optimizations:
  - Detection of optimum start values for phase fitting leading to cleaner, faster, and more accurate depth information
  - Statistical point cloud filters for cleaning while preserving surface density and fine detail
- Revised system design featuring higher resolution components

## 2 Improved system design

Two ISVI C25 color cameras with square 25 *megapixel* resolution look down onto the object mount, which can be mechanically translated arbitrarily within a horizontal plane (two degrees of freedom), controlled by the software and fully embedded into the 3D scan workflow (Fig. 1). A digital light processing (DLP) projector at full HD resolution is used for projecting fringe patterns vertically onto the surface. In front of the projector output lens, an additional lens is mounted on an industrial-grade linear axis which can be moved over about 162 *mm* in up to 35,000 *steps*, leading to a step size of 4.629 $\mu m$. Thanks to this mechanism, lens-shifted structured light encoding of the surface at sub-pixel level is possible, breaking the depth accuracy limit defined by the resolution of projection pixels. The second camera allows for acquisition of an alternative

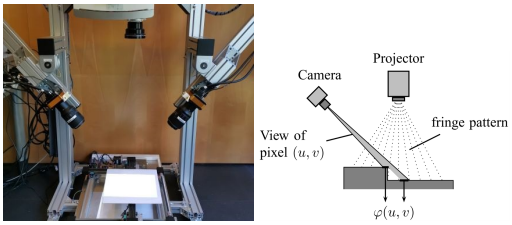


**Figure 1:** *MesoScannerV2 setup with two 25 megapixel color cameras and a DLP projector, looking down onto an object mount that can be mechanically translated along two dimensions (left). Problems can occur due to occlusions of the object (right) [LM07].*

depth reconstruction sample, identically computed as for the first camera. Based on a calibration between the two cameras and the projector, a correspondence between measured samples from either camera corresponding to the same point on the object surface point can be established and used for two purposes. First, in order to choose the better reconstruction based on the residual error of the respective sample reconstruction, and second, to resolve occlusions and self-shadowing of the object, given that at least one camera can see the encoded point on the surface.

## 3 From phase-coded image acquisition to 3D reconstruction

The complete 3D scanning process, visualized in Fig. 2, begins with image capture of fringe patterns projected on the object surface. This step is a combination of the multi-period phase shift method [LM07], typically using three different wavelengths to derive a globally unique phase coding over the entire measurement range, and the lens-shifted phase shift method [RLS*12] which uses the mechanical lens-shifter. Both methods generate fringe images with projection columns at constant intensity and sinusoidal intensity distribution across the second dimension of the projection image. The second method uses a fixed wavelength of two, i.e., consecutive black and white columns, that are perceived by the camera as a sinusoidal signal. In combination with the high resolution step size (see section 2), sub-pixel accuracy is reached, leading to a significantly higher depth resolution as with multi-period phase shifting, but comes at the cost of ambiguities across all periods, that are resolved through combination with a basic version of the first method as detailed in [RLS*12]. The resulting phase value corresponds to one specific mathematical plane emanating from the projector that is intersected with camera rays to obtain 3D points on the object surface, which in turn leads to a 3D point cloud for all valid camera pixels. The implicit neighborhood of the sample points, provided by the correlating image pixels, is used to generate a triangle topology as well as per-point color information, which is the final output 3D model.

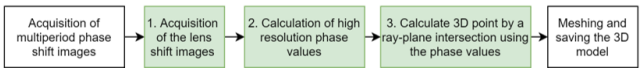


**Figure 2:** *3D scanning process flow: From phase-coded image acquisition, a combination of multi-period and lens-shifted phase encoding, over decoding, to 3D reconstruction of a colored point cloud with triangle topology. The optimizations proposed affect the process steps colored green described in the sections 3.1 – 3.3.*

### 3.1 Robust lens-shifted phase encoding

The phase values for each camera ray, leading to a reconstructed 3D point on the object surface observed by the camera, are decoded using Levenberg-Marquardt fitting to determine the parameters of a sinusoidal curve. The convergence of the iterative process towards a correct solution depends on two factors: Noise in the lens-shifted phase-encoded images, and how suitable the initial start values are chosen for the fitting scheme to avoid divergence to wrong model parameters. The first aspect could be improved by applying a temporal noise filter (TNF) over a set of $20 - 50$ captured images and using the aggregation function 'mean'. Fig. 3 shows a comparison of different aggregation functions applied for the TNF and split into black and white by thresholding. Outliers are visible in the form of single pixels sticking out from their neighborhood. The effect of

**Figure 3:** *From left to right: Noisy Base, Mean 5 Images, Mean 20 Images, Mean 50 Images, Median 5 Images, Median 20.*

the optimization can be observed in the final 3D reconstruction as shown in Fig. 4 (qualitative).

### 3.2 Stable decoding of lens-shifted phase values

After the optimization described in 3.1, the Levenberg-Marquardt fitting algorithm still does not consistently converge against correct

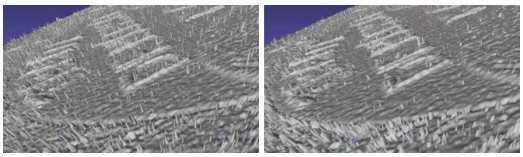

**Figure 4:** *Rendered 3D scanning of a 2 Euro coin without (left) and with (right) TNF.*

parameters for all camera rays, which is due to suboptimal choice of starting values for the iterative fitting scheme that are guessed using the lens-shifted phase-encoded images. The following metric is used to assess the outcome of the fitting process [Gav19]:

$$\chi = \frac{\sum_{k=1}^{N} (\sin(\omega \cdot k + \varphi) - im_k(i,j))^2}{N - m + 1}, \qquad \chi \in \mathbb{R}^+. \tag{1}$$

Systematic tests revealed a threshold of $\chi > 0.2$ as indicator for fitting divergence. Fig. 5 shows two cases (left, middle) where the red fitted curve fails to approximate the measured intensity profile for a camera ray, while the third case (right) shows a successful fitting result. Once a failed fitting result is detected for a specific

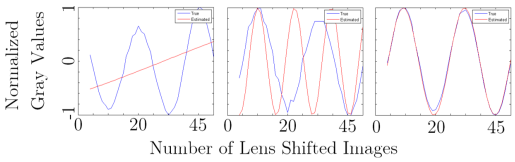


**Figure 5:** *Two divergent fits (left, middle) and one successful fit (right). Blue curve: measured intensity for camera ray over sequence of images captured, red curve: fitted curve.*

camera ray based on Eq. 1, the parameter values $(\varphi, \omega)$ for fitting are applied as starting values from a previous camera ray for which fitting succeeded. Fig. 6 visualizes an example face decoding process, where for one pixel ray (indicated in red) fitting fails and is detected, such that fitting can be reiterated using the parameter values of the previously successful fitting process. If left uncorrected, the red pixel would have been detected as white, and thus represent an outlier (compare with Fig. 3). Even when using a previously

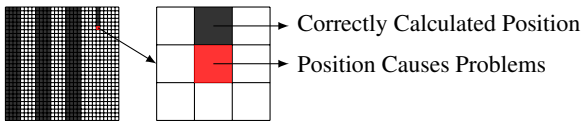


**Figure 6:** *A detected divergent fitting (red pixel) is handled by using parameter values as starting values from a previously successful fit.*

correctly calculated phase value $\varphi$ and period $\omega$ in case of a divergent fit (Eq. 1) fitting can fail. This situation is handled by restarting fitting repeatedly within an iteratively increasing region around the initial estimated starting values. The effect of the optimization can be observed in the final 3D reconstruction as shown in and Fig. 7 (quantitative and qualitative).

### 3.3 Clean 3D model through methods of statistical analysis

While the previous optimizations targeted phase encoding and decoding in order to derive robust phase values for 3D reconstruction, this section addresses filtering the set of reconstructed 3D points

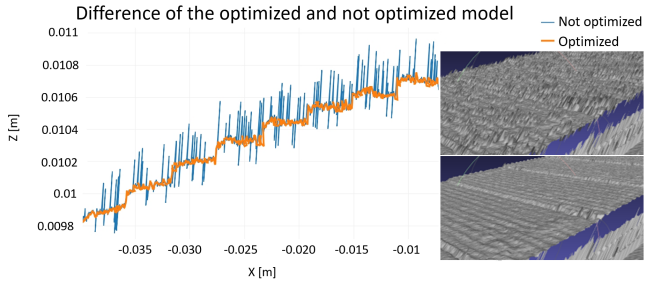


**Figure 7:** *Rendered 3D scanning of a plastic piece without (top) and with (bottom) optimization of fitting starting values. The blue line visualizes a stripe of the surface of the plastic model without any optimizations, while the orange line shows the same line but optimized. It is clear, that the optimized model is much smoother and cleaner than the model without optimization.*

directly, each of which is a result of intersection of a camera ray and a plane defined by the phase value derived from the intensity profile seen in that pixel. The entire set of 3D points reconstructed still contains a significant amount of noise, i.e., outliers that are not part of the object surface, but are caused by an incorrect derivation of phase values. This effect is mostly caused by cavities on the object surface, but also material-specific effects such as reflectivity. The goal of the filtering methods proposed are to provide

**Figure 8:** *Outliers in the unfiltered point cloud (left), after applying the Interquartile range filter (middle) and after alignment of the point cloud using a principal component analysis, followed by application of Interquartile range filter (right).*

a performant means to exclusively target outliers caused by incorrect phase derivation, while preserving the digitized surface using adaptive thresholds such that no parameter tuning is required. The filters described assume that the number of outliers is significantly less than the number of correct 3D points on the object surface.

**Interquartile range (IQR) filter** One possibility to gather a data-driven threshold is to use information of the box plot, in detail the whiskers $(W_{\text{down}}, W_{\text{up}})$ in combination with the interquartile range ($IQR$). Given a set of points $(x_i, y_i, z_i) \in D \subset \mathbb{R}^3$ that represents the reconstructed 3D point cloud, a box plot of each coordinate axis is created to analyze the distribution and density of the point cloud [Mit12]:

$$IQR := x_{(0.75)} - x_{(0.25)}, \tag{2}$$

where $x_{(q)}$ is the q-th quantile of the data $x$. Using Eq. (2) results in the thresholds $W_{\text{down}}$ and $W_{\text{up}}$

$$W_{\text{down}} := x_{(0.25)} - 1.5 \cdot IQR, \qquad W_{\text{up}} := x_{(0.75)} + 1.5 \cdot IQR. \tag{3}$$

Coordinate values outside $[W_{\text{down}}; W_{\text{up}}]$ qualify a 3D point as outlier. Subsequently, the point is removed from the set, as visualized for three examples in Fig. 8. As this filter is not rotation invariant,

an axis alignment is carried out beforehand, using principal component analysis (*PCA*).

**Gaussian mixture models** Another method based on Gaussian mixture models is used to cluster a pre-defined number of Gaussians on the data by optimizing the parameters of all Gaussian functions using an Expectation-Maximization algorithm [Bis06]. The idea is to classify the jumping edges outlier as Gaussian functions, whereby the real surface points are classified by one single Gaussian function. Using Gaussian mixture models with three Gaussians proved best on the 3D point cloud (Figure 9). All points assigned

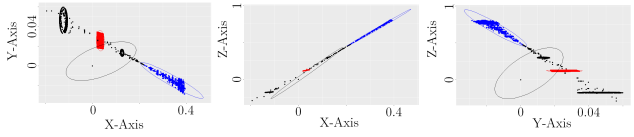


**Figure 9:** *Example of reconstructed object with outliers, classified using Gaussian Mixture Models. The ellipses represent the Gaussians for the corresponding classified object points, whereby the actual model is classified by a single Gaussian (red).*

to one of the two Gaussian functions corresponding to the lower number of points are removed based on above assumption that the number of outliers is significantly lower than the number of 3D surface points.

## 4 Automation and extension to arbitrary scanning range

To handle two very common problems of high-resolution digitization systems with one solution, caused by limited measurement range, a mechanically movable object mount was integrated into the system, allowing to translate the object horizontally along two dimensions. This extension was fully embedded in the process flow such that it is automatically controlled by the 3D scanner software. The two scenarios covered are objects that are larger than the measurement range, and collections of objects that are to be scanned in one run. The object mount is actuated by two orthogonally aligned linear axes with a step size of 18.182 $\mu m$. For both scenarios, the object mount is driven into all positions in turn that correspond to the sub-areas, as indicated in Fig. 10 (bottom right). For each of the cells, a quick analysis is done requiring capturing of one image. The image, filtered by a median blur kernel, is analyzed for any deviation from a constant color distribution, which indicates that an object is present in the current cell, in which case a 3D scan is initiated, otherwise the cell is skipped. For the scenario of objects larger than the scanning range, the final 3D scan is then stitched together from the single cell scans after prior alignment using Iterative Closest Points (ICP). For the scenario of collections of objects as in Fig. 10 (bottom left), a set of disjunct scans is generated. Hybrid cases of both scenarios as well as situations of the second scenario, in which objects cover several cells partially, are subject to future work. Current development includes detecting contours during the phase of identifying presence of objects in cells to move away from the approach described of fixed cells to adaptive clustering of partial scans at arbitrary locations on the object mount surface.

## 5 Conclusion and future work

We propose a redesign of an existing lens-shifted structured light system which features higher resolution components and an additional camera, plus an automation component to actuate the target object laterally in two dimensions. Furthermore, we propose several algorithmic optimizations and statistic filters along the process flow that together significantly improve 3D scan quality, which we support both quantitatively and qualitatively. The contributions are higher signal to noise ratio than achieved the previous system. In fact, the algorithm in its proposed form has reached a robustness that removes the need for post-processing (cleaning) of the scan and does not require parameter tuning. The automation component introduced resolves the problem of limited scanning range of comparable high-resolution systems by allowing theoretically arbitrary scanning ranges, and at the same time allows for fully automatic batch-process scans of entire collections. While hybrid and special cases still need development, the method proposed is a large step towards high-quality, fully automatic digitization of cultural heritage, especially for small objects, such has coin collections.

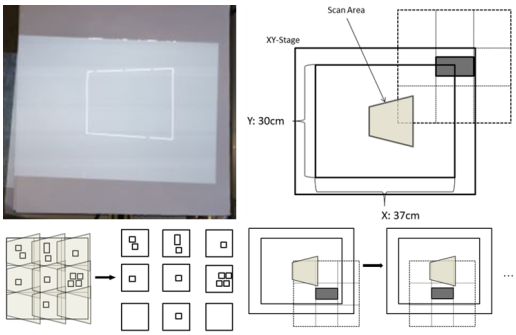


**Figure 10:** *Movable object mount seen from above with measurement range visualized (white projected frame, top left), segmented into sub-areas coverable by measurement range (top right), scenario of collection of objects (bottom left), and example of iterative coverage of entire object mount surface by iterative translation into the measurement range (bottom right).*

## References

[Bis06] BISHOP C. M.: *Pattern recognition and machine learning*. Springer, 2006. 4

[Gav19] GAVIN H. P.: The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering, Duke University* (2019), 1–19. 3

[Kut20] KUTLU H.: Fully automatic mechanical scan range extension of a lens-shifted structured light system, 2020. 2

[LM07] LILIENBLUM E., MICHAELIS B.: Optical 3D surface reconstruction by a multi-period phase shift method. *JCP 2*, 2 (2007), 73–83. 1, 2

[Mit12] MITTAG H.-J.: *Statistik: eine interaktive Einführung*. Springer-Verlag, 2012. 3

[RLS*12] RITZ M., LANGGUTH F., SCHOLZ M., GOESELE M., STORK A.: High resolution acquisition of detailed surfaces with lens-shifted structured light. *Computers & Graphics 36*, 1 (2012), 16–27. Cultural Heritage. doi:10.1016/j.cag.2011.10.004. 1, 2

[SRT*14] SANTOS P., RITZ M., TAUSCH R., SCHMEDT H., MONROY R., STEFANO A. D., POSNIAK O., FUHRMANN C., FELLNER D. W.: CultLab3D - On the Verge of 3D Mass Digitization. In *Eurographics Workshop on Graphics and Cultural Heritage* (2014), Klein R., Santos P., (Eds.), The Eurographics Association, pp. 65–73. doi:10.2312/gch.20141305. 1