The 11th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2010)
A. Artusi, M. Joly-Parvex, G. Lucet, A. Ribes, and D. Pitzalis (Editors)

High Resolution Acquisition of Detailed Surfaces with Lens-Shifted Structured Light

M. Ritz[1], M. Scholz[2], M. Goesele[2] and A. Stork[1,2]

[1]Fraunhofer IGD, Germany [2]TU Darmstadt, Germany

Abstract

We present a novel 3D geometry acquisition technique at high resolution based on structured light reconstruction with a low-cost projector-camera system. Using a 1D mechanical lens-shifter extension in the projector light path, the projected pattern is shifted in fine steps at sub-pixel scale with a granularity of down to 2048 steps per projected pixel, which opens up novel possibilities in depth accuracy and smoothness for the acquired geometry. Combining the mechanical lens-shifter extension with a multiple phase shifting technique yields a measuring range of 120x80 mm while at the same time providing a high depth resolution of better than 100 micron. Reaching far beyond depth resolutions achieved by conventional structured light scanning approaches with projector-camera systems, depth layering effects inherent to conventional techniques are fully avoided. Relying on low-cost consumer products only, we reach an area resolution of down to 55 micron (limited by the camera). We see two fields of benefit. Firstly, our acquisition setup can reconstruct finest details of small Cultural Heritage objects such as antique coins and thus digitally preserve them in appropriate precision. Secondly, our accurate height fields can be viable input to physically based rendering in combination with measured material BRDFs to reproduce compelling spatially varying, material-specific effects.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Shape, Range data

1. Introduction

Digitizing objects in 3 dimensions opens up a rich set of possibilities to their eternal and appropriate preservation, especially looking at cultural heritage objects, as opposed to 2D-only digitization as images. Due to the human vision system relying on visual depth cues, 3D models can reveal a far more immersive and tangible impression of the object's features and significance due to the infinite possibilities of choosing viewpoint, lighting and rendering settings, e.g., at any time after acquisition, as opposed to the rather fixed nature of 2D images. 3D scanning with structured light is limited, however, and does not resolve fine details at sufficient quality. We present and analyze a 3D scanning approach that has a high depth resolution, allowing for reconstruction of finest surface details. Be it the fine marble structure of sculptures or structures found on very small and movable objects like coins, capturing of finest details can be crucial to appropriate long-time preservation in this object domain. Besides digitization, analysis of aging effects like micro cracks in objects is in the interest of museums, as well as the reconstruction of fine engravings, which current techniques have not mastered yet. Applications reach even into the domain of art where determining the depth of paint can reveal specific information on the artist. Material acquisition at fine-granular scale is also a vital necessity in another field, the one of realistic 3D rendering, where the focus changes due to more powerful hardware from using approximate close-to-reality shading models towards rendering physically-acquired reflexes of materials and thus delivering compelling realism. A required core information is the height field of material samples, providing the basis necessary for realistic rendering of Spatially Varying Bidirectional Reflectance Distribution Functions (SVBRDFs), distributed BRDFs, or providing rendering systems with geometry information for physically correct light throughput regarding traversal of fine material structures.

We propose a fully automatic geometry acquisition system that brings about several contributions:

- Novel lens-shifting reconstruction technique based on a precise hardware shifting extension [SDS09], combined with multiple phase shifting [LM07], lifting common structured light limitations.
- Accuracy analysis.
- High resolution at comparatively large measuring range and low cost due to off-the-shelf components.
- High precision geometry reconstruction for cultural heritage object preservation and aging effect analysis.
- Accurate height field acquisition for material-realistic rendering.

2. Related Work

Various methods address 3D digitization [AAA*05]. A very prominent one is structured light, others include shape from specularity, a representative of photometric stereo methods, where the magnitude of specular reflections is used to derive a normal field by moving a light source over the object [CGS06]. Francken et al. [FCM*08] project several binary patterns on the object using an LCD-display to obtain a unique identification of the incoming light direction at maximum intensity for surface normal computation. A disadvantage of the above techniques is their relying on specular reflectance. An advantage over other methods is the possibility to measure the surface of translucent objects.

The field our approach is based upon is structured light which has been widely investigated and has some significance in the industry. Techniques in this field run on projector-camera systems, and several cameras can be used to increase robustness and accuracy. Acquisition is done by projecting a series of patterns onto the object to obtain a unique correspondence between projector-planes and camera rays to determine 3D points by intersection. The basic approach is using gray code patterns to shut out 1-bit measuring errors due to special encoding. The sequence in which a pixel is overlaid by dark and bright stripes reveals the plane ID. A method extending mainly the patterns projected by using color coding is applied by [RCM*01]. The downside of these methods is a layering effect in depth, since all surface positions lying on a pattern stripe are clustered together, leading to a discretization in space during intersection. This can be improved by multiple phase shifting [LM07], where three different stripe patterns formed by different wavelengths of a sinusoidal luminance function are used instead of binary patterns. Each pattern is then shifted by fractions of its wavelength over the object, revealing a certain phase for each surface point, and all patterns together allow for globally unique determination of the correlation between surface points and projector plane IDs. Shifting the patterns removes the layering effect somewhat as the discretization is more fine granular, by allowing the determination of stripe IDs on a sub-pixel basis. Problems arise in general when objects show a high percentage of indirect light transport, as well as when there are dark materials and surfaces with a high amount of specularities.

Global illumination effects are tackled by Chen et al. [CSL08] using modulated high- and low-frequency pattern projections. The observation that the integration of a normal field mostly yields low frequency errors, while structured light scanners are mainly susceptible to high frequency errors, leads to a combination of photometric methods with multiple phase shifting methods, as done by Nehab et al. [NRDR] who fuse both domains of surface descriptions using an optimization technique that increases the quality of 3D scans.

Finally, going into the domain of line shifting scanners, a digital projector is used in [GÖ1] to project thin stripes in regular distances, which then are moved over the object, and for each camera pixel, the maximum luminance over scanning time and the stack of all patterns is determined, enabling a precise correlation between projector and camera pixels. In comparison to Phase Shifting approaches, this technique is more robust against indirect light transport due to larger distances between projected stripes, but shows problems with specular reflections.

Even though the field of phase-shifting has been widely explored, there are currently no publications closely related to our acquisition setup. We build on the multiple phase shifting technique but drastically increase its depth resolution by adding an additional pattern at the projector's Nyquist limit. Since phase shifting is in this case no longer possible using different projected patterns, we use a hardware lens-shifter to shift the pattern with sub-pixel accuracy. As a result, we fully avoid depth layering effects inherent to structured light techniques and in addition break limits in depth accuracy, going far beyond 100 microns.

3. Hardware Setup and Analysis

The acquisition setup consists of a DLP LED projector, a consumer DSLR (digital single-lens reflex) camera, the lens-shifter and the sample mount, these constituents being mounted fixedly into an alloy framework which is covered with a matte black anti-reflective foil to minimize stray light. The framework itself rests on rubber pads to decouple it from external vibrations and to damp vibrations caused by the camera. This setup makes the system insensitive against minor external vibrations.

The angle between camera and projector is approx. 20°, the baseline measures about 170 mm, and the average measurement distance to the sample mount is 250 mm. Together with the projector and camera intrinsic parameters, this leads to a measurement range of about 120x80 mm covered by 2176x1434 sample positions.

Figure 1: *Acquisition setup, consisting of alloy framework, camera and projector with lens-shifter, and material mount.*

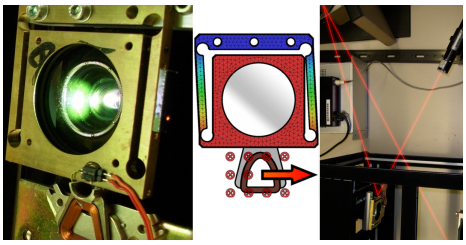

Figure 2: Left: *Lens-shifter.* Middle: *Scheme with color coded displacement of the frame under current-induced Lorenz force due to magnetic field.* Right: *Laser light path (sketched) from source (upper right) over lens-shifter mirror and external mirror to projection surface .*

3.1. Projector and Camera

The projector used in the system is an LG-HS101 Pocket Projector with a resolution of 800x600 pixels using high performance LEDs in the three color channels. It is directly fed with image data by the reconstruction algorithm. As capturing device, a Nikon D300S was chosen with a resolution at 4352x2868 pixels (12 MPix). Its image sensor consists of periodical groups of 2x2 color sensors, 2 for green color, and one for red and blue, respectively. Rather than using the data post processed by the camera, which interpolates the color values for each of the four positions in the groups, we directly read out raw data from the camera and form a per-color weighted average manually from the exact readings per sensor unit to obtain a grayscale measurement for the pixel group. Since every physical pixel of the sensor chip can detect only one of the three primary colors, this procedure is necessary and leads to an effective resolution of 2176x1434. We use a Zeiss Distagon macro objective with low distortion and small color aberration. The camera is integrated into the system over a USB port, used both for fully automatic camera control and image data gathering. The camera as the only measuring component in the system is strongly influenced by image noise, which has a direct influence on the accuracy of reconstruction. Thus, the representative measure of signal to noise ratio (SNR) was optimized over the space of exposure settings in order to maximize reconstruction quality.

3.2. Lens-Shifter

The major contribution of our method is the use of a lens-shifting reconstruction technique to significantly increase depth resolution, by controlling a mechanical lens-shifter extension [SDS09]. It is a brass frame with an embedded lens and 4 flexure joints at the corners, realized by very thin regions of the material, that allow for nearly friction-less 1 DOF movement of the frame and thus accounting for smooth and accurate shifting of the lens, ensuring precise repeatability. The controller driving the shifter is integrated into the system over the serial port which is used to transmit steering commands. The lens-shifter mechanism serves two goals. Firstly, the lens itself improves the projector's depth of focus and allows for a smaller distance between projector and object to increase reconstruction area resolution. Secondly, the projected image can be moved along one axis at sub pixel accuracy over a range of ± 1 projection pixel in a granularity of 2048 steps per pixel. With the precise shifting of the projection, high frequency patterns can be used for the scan process, which amplifies luminance changes measured for small surface profile differences. These two features make a more dense sampling of the projected signal possible, which results in a higher depth accuracy of the scanner. In addition, smaller pattern wavelengths increase the system's robustness against indirect light transport as observed by [CSL08].

The lens-shifter is responsible for precise shifting of the projection. Positioning errors have direct influence on the luminance measured by the camera per pixel and thus on the reconstruction, making repeatability a highly significant measure for consistent and accurate scans. To assess repeatability, a small mirror was mounted at the side of the lens-shifter frame to redirect a laser beam onto the projection surface captured by the camera, and extend the light path length to register even minimal elongations of the shifting device, measurable as the movements of the laser dot seen by the camera (cf. Figure 2). The shifter was repeatedly driven from the neutral position to predefined positions, and overall repeatability was found to be within an error of 0.001 projector pixels. The laser method also confirmed that the linearity between the desired position, conveyed by the current driven by the controller, and the actual position of the lens-shifter, is granted and sufficient.

4. Meso-Structure Acquisition

The process of capturing images with patterns overlaid over the object is done in two phases, one for capturing phase-shifted images for three different wavelengths according to the method of Lilienblum et al. [LM07], and one for capturing lens-shifted images. The first phase assigns surface points to a certain period and resolves the uniqueness problem of the second phase, which is responsible for exact determination and line-plane-intersection based on the Lens-Shifting information. The capturing phase is followed by

normalizing all captured images (see 5.1), and the actual reconstruction of all surface points seen by the camera. Finally, a surface mesh is constructed between the samples of the reconstructed point cloud. As each 3D point correlates with the 2D image point it was reconstructed from, and all 2D image points lie in a regular grid, the connectivity information applies to the 3D point cloud as well, and triangulation can be done in a straightforward way by inserting two triangles between any rectangular neighborhood of four points.

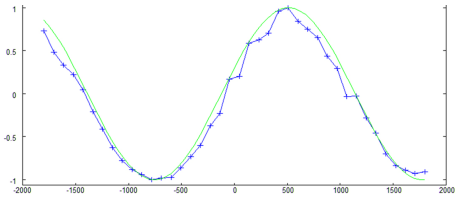

Figure 3: *Luminance of pixel over a complete Lens-Shifting sequence and fitted model function.*

4.1. Lens-Shifting

The second acquisition step is also based upon the Phase Shifting approach similar to [LM07], but with one phase and finer steps. A sine wave pattern with fixed wavelength is projected and shifted by a fixed distance after each capturing. Instead of shifting the pattern digitally, the Lens-Shifter is used for this task, while the projector output remains unbiased. Since it can position the pattern independent of the projector resolution at sub-pixel accuracy at a theoretical step size of $\frac{1}{2048}$ of a projector pixel, the use of high-frequent patterns is now possible. Thus, the sine wave pattern with the smallest displayable wavelength of two was used for reconstruction, consisting of one pixel wide black and white bars. Acquisition results in a luminance profile per pixel which is interpreted as a time-dependent signal. There is a significant difference between digital and analogous shifting of the signal. Regular Phase Shifting produces the same luminance profile for each pixel due to the time dependent signal not being coupled to its spatial dimension. However, in the case of the optical Lens-Shifting where a coupling between time and space is given, the time dependent signal is impacted by lens distortions and other optical influences, and both temporal and spatial differences in the wavelength for different image positions are the consequence. The conventional Phase Shifting approach of using Fourier analysis for phase determination is thus not applicable because it requires the exact wavelength. Instead, an optimization problem is solved to fit the luminance profile into a model function per pixel. With the resulting accurate phase, projector and camera pixel can be correlated uniquely within the domain of a wavelength.

4.2. Fitting the model function

Despite steering the projector with a rectangular signal, the luminance pattern of each pixel over a lens shifter sequence resembles a sine wave. The reason is probably the optics and projector image generator that lead to a rather washed-out reproduction of the image, with the consequence of a sine wave shaped luminance flow per pixel. Determining the phase of the signal requires the model function to be periodic and represent the captured data well. The analysis of the normalized temporal signal of several distinct image positions is shown in Figure 5 (upper left). Normalization is necessary to account for inhomogeneous distribution of reflectance of most objects, especially textured surfaces, where the signal amplitude varies strongly over space. The following model function is thus used, parameterized in frequency f, amplitude a and phase ϕ.

$$b(x) = a * \sin(\phi + 2\pi f x) \tag{1}$$

These three parameters are then found by optimization using the Levenberg-Marquardt-Algorithm for nonlinear least squares problems [Lou05]. Figure 3 shows an example of a fitting result. Appropriate starting values are necessary for a robust optimization (see 5). For the spatially varying phase, the value is left variable. The starting value for the amplitude is set to $a = 1$ due to the normalization, the frequency is set empirically.

4.3. MPS and LS combined: High Depth Accuracy at Large Range

The relative and more accurate phase ϕ determined by Lens-Shifting is now combined with the absolute and coarser phase ω resulting from classical phase shifting. The globally unique phase ω is transformed to the range $[0..1]$, pointing to a pixel row between the bounds of the projected image. ϕ points to a local position within the wavelength λ and has the same range as ω, but in this case pointing between the bounds of a wavelength.

$$\alpha = \lambda \left(\left\lfloor \omega \frac{w}{\lambda} - \phi + \frac{1}{2} \right\rfloor + \phi \right), \quad \omega, \phi \in [0 \ldots 1] \tag{2}$$

To combine the two phases and benefit both from global uniqueness and high accuracy, ω is first transformed to point to a multiple of λ by relating it to the projection width w and dividing by λ. The non-integral remainder is then removed by subtracting ϕ, now that the two phases are expressed in the same unit, and the result is rounded to the next integral multiple of the wavelength λ. Finally, The accurate phase ϕ is added, and the projector coordinate α is obtained by multiplying with the Lens-Shifting wavelength. Subtraction of the phases in a common unit rather than more intuitively just rounding the course phase is necessary since especially at boundaries of wavelengths, image noise can lead to erroneous classification of projector plane IDs. Now that for each camera point the absolute projector coordinate is known, 3D plane/ray intersection between projector planes identified by the projector coordinate and rays defined by picture posi-

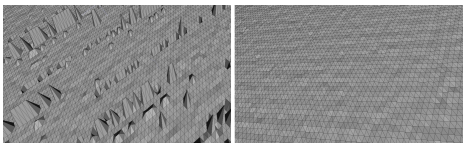

Figure 4: *Reconstruction of a plain surface before* (left) *and after* (right) *2-step fitting.*

tions results in 3D positions of the surface points according to the basic idea of Structured Light reconstruction.

5. Optimization

Reconstruction is subject to several influences, for instance camera noise or variable average luminance over captured images. Several optimizations were introduced to make reconstruction robust and more accurate.

Reconstructing a plain surface revealed instabilities of the numerical optimization used for fitting of the model function, which lead to outliers visible in Figure 4 (left). These problems were addressed by dividing the optimization process into two steps that differ by the set of parameters defined to be fixed while letting the others be subject to optimization. In the first stage, the fitting process is done by optimization with only the spatially varying phase being variable, while the frequency remains fixed and the amplitude is set to $a = 1$ due to normalization. The resulting phase is then used as more accurate starting value for the second step with the remaining parameters, including the phase itself, being subject to optimization (except from the amplitude) and thus variable, achieving a robust optimization as can be seen in Figure 4 (right).

5.1. Image Normalization and Luminance Profile Smoothing

Even though the luminance flow over a full Lens-Shifter sequence shows nearly sinusoidal behavior, there are still deviations recognizable, as depicted in Figure 5 (upper left). The vertical highlight bar shows that deviations occur simultaneously for all pixels at the same Lens-Shifter position, captured by the same image. Since the phase is spatially dependent, this cannot be an effect of possible deviation from linearity of the Lens-Shifter, which would also contradict the linearity test conducted (3.2). Neither can it be a consequence of image noise that is spatially independent. The reason is found within a captured image compared to neighboring frames in time. The change of luminance over the course of frames deviating from the expected sinusoidal behavior is apparently caused by a per-frame bias in overall luminance, which is an effect of the consumer camera used not claiming to be apt for precise repeatability, and the missing synchronization with the oscillating projector signal. As

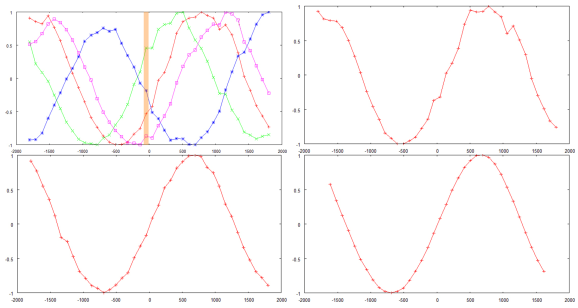

Figure 5: Upper left: *Luminance for 4 distinct pixels over common full Lens-Shifter sequence. Luminance flow over full sequence before* (upper right)*, after image normalization* (lower left) *and after 5px box filtering* (lower right)*.*

a solution, the capturing sequence is normalized to the average luminance of the first image in the sequence, which shows significant improvement, visible when comparing the upper right and lower left charts in Figure 5. The final optimization that leads to a behavior sufficiently close to the target model function is filtering the luminance profiles of each pixel with a box filter of size 5px (cf. the lower right chart in Figure 5). After this step, the fitting process of captured data to the model function is robust, and reliable parameters are determined automatically.

6. Results

As our acquisition system targets the two domains of high accuracy digital preservation of cultural heritage objects and acquisition of height fields for material rendering, we show results of the two domains and provide an accuracy comparison between our approach and the pure phase-shifting technique. An analysis on depth accuracy supports that the acquisition accuracy is indeed higher than 100 micron, and tries to give an upper limit on depth accuracy beyond.

6.1. Meso-Scale Reconstructions

In Figure 6 (left) we show the rendering of a reconstructed leaf under green illumination. The leaf sets itself apart from the planar background geometry, jumping out into the third dimension. Note that we explicitly did not render true color, but used virtual illumination to highlight the fine geometry. The close-up shows a wireframe subregion of the rendering, outlining the accurate reconstruction of the fine leaf branch structure. The reconstruction of a different leaf is pictured in gray-scale shaded rendering under white illumination in Figure 6 (middle). The deep meso-structure 'valleys' formed by the fine leaf cell structure are plastically visible. White paint was used to exclude specular highlights during reconstruction. The reconstruction of two somewhat more historic objects, a 1955 Sixpence coin (silver) and a 1899 One Penny coin (bronze), are shown as renderings in Figure 7.

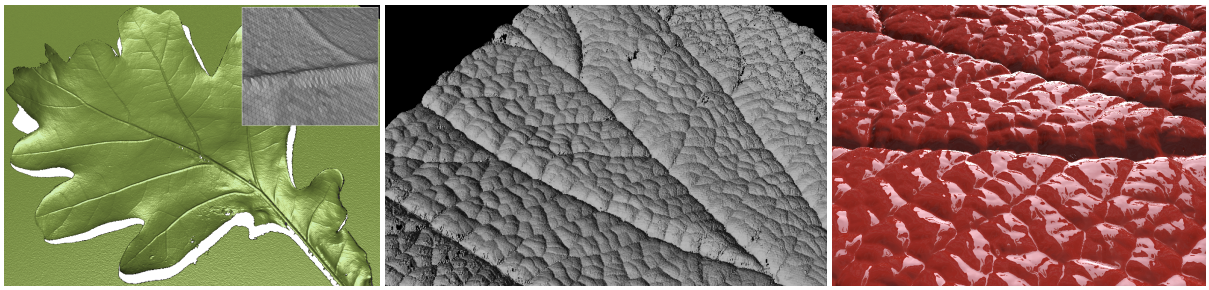

Figure 6: Left: *Rendering of a leaf scan under green illumination with close-up of triangle mesh.* Middle: *Gray-scale shaded rendering for different leaf (2x2 cm section).* Right: *Ray-traced geometry of the same leaf (1x1 cm section), combined with BRDF material probe of red car paint and Uffizi environment light probe. The orientation of a given BRDF material sample is controlled by the acquired fine-structure geometry over the entire object surface. Ray-Tracer: [HNG*10], Probes: Spheron VR.*

Figure 7: *Reconstruction of shiny objects without treatment (blending between textured and gray-scale shaded rendering).* Left: *Sixpence coin from 1955, front and back.* Right: *One Penny from 1899.*

Figure 8: *Reconstruction of a 5 cent euro coin with no treatment* (left) *and with prior application of white paint* (middle)*; reconstruction of 2 euro coin using modeling clay* (right).

Despite the strong reflectance due to the shiny metallic material, no prior treatment was applied before reconstruction; yet fine details can be seen in the results. See Figure 8 for a comparison of some common object preparation techniques. Small letters and complex details of the coined pattern are visible in the reconstruction of the 19 mm Six Pence coin. The One Penny comes at an estimated maximum coining elevation of 100 micron, with the major part of the coin having been ground down to zero over time. Only the right arm and torso of Victoria are left at maximum elevation, with a totally smooth transition into the zero-level coin background. These features and their smooth transition are captured well in the rendered reconstruction. Even though the year letters and the 'ONE' inscription are left at estimated 50 and 20 micron elevations, respectively, both coining details are still measured and reproduced by our scanning system.

6.2. Comparison to Pure Phase-Shifting

Since our acquisition system applies two independent techniques, a comparison of our approach to pure phase-shifting is easily done by disabling the lens-shifting functionality, enabling a direct comparison by leaving the hardware setup invariant. To objectively compare the accuracy of the two approaches in relation to the number of images captured, we introduce the following metric:

$$\Delta = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} \|P_1(x,y) - P_2(x,y)\|^2 \qquad (3)$$

For two reconstructions P_1, P_2 of the same object (a white planar surface) acquired immediately in sequence, the mean quadratic deviation is computed for each pair of 3D samples $(P_1(x,y), P_2(x,y))$ and accumulated over the reconstructed range, which for the evaluation was 500x500 pixels. Since our acquisition system is calibrated such that one unit in the reconstructed model corresponds to the length of 1 m in physical object space, we express the results of the metric in micron. The metric expresses the positional deviation between two subsequent scans which is an important upper bound on accuracy due to the heavy influence of noise, impacting reconstruction indeterministically. The pure Phase-Shifting method was first evaluated using 3 wavelengths at 7, 11, and 13 pixels, and capturing 3 images per wavelength, coming to a total of 9 images. In the next run, 7 images per wavelength were acquired, which corresponds to the maximum number of shifts possible for the smallest wavelength, leading to 21 input images. For the comparison between the

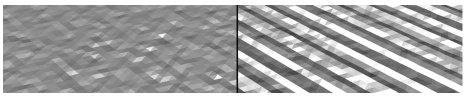

Figure 9: *Close-up of rendered reconstruction of a planar surface, divided into the extended Lens-Shifting* (left) *and the pure Phase-Shifting result* (right).

approaches, our lens-shifting technique was evaluated, once for 20 and 40 captured images at a wavelength of 2 pixels, which means that the lens-shifter was driven to as many positions distributed evenly within the theoretic lens-shifting range of 2048 steps. The additional 9 images required for phase-shifting do not influence the accuracy measure as they serve for global phase determination only. Table 1 shows the evaluation results. Already 20 images are sufficient for

Table 1: *Accuracy comparison of pure Phase-Shifting and the extended approach with a Lens-Shifter.*

Scan mode	Image count	Δ [micron]
Phase-Shifting only	9	9.596
Phase-Shifting only	21	5.057
Lens-Shifting (+PS)	20 (+9)	2.760
Lens-Shifting (+PS)	40 (+9)	2.251

our approach to nearly double the accuracy according to the measure in comparison to the pure phase-shifting technique that uses one additional image, while increasing the number of images further leads to slight improvements in the accuracy metric.

The qualitative results of the evaluation are visually compared in Figure 9 which is a rendered image of the reconstruction divided into two parts. The left part shows the result as achieved by our extended Lens-Shifting technique, continuously going over to the right part that is the result of the pure Phase-Shifting reconstruction. In the pure phase-shifting result, stripe pattern artifacts are apparent that follow the pixel structure of the projected image, leading to deviation from the sine wave luminance flow. Our extended lens-shifting method does not show these artifacts and is significantly more error resilient. This comes at the cost of longer acquisition time needed in comparison to Phase-Shifting alone, which is due to the additional Phase-Shifting image acquisition and the higher computational cost for fitting. Reconstruction time for the full acquisition range of 120x80 mm is about 18 min (4 core 64 bit system, 3.2 GHz). Parallelization of the algorithm on a GPU or several CPUs is straightforward, however, and can largely reduce the time needed.

6.3. Breaking the 100 Micron Mark

Measuring depth accuracy is a complex task in such small dimensions. To quantitatively support the impression of high

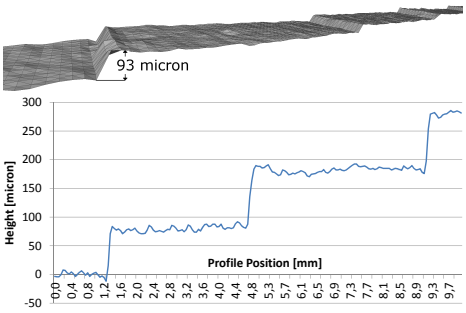

Figure 10: *Rendering of paper ramp reconstruction and corresponding height profile alongside ramp. The profiled length is about 1 cm, height of steps is approx. 93 micron.*

accuracy resolution given by the reconstructed coin profiles in Figures 7 and 8, we constructed a ramp of paper sheets *(5staroffice Re-Move notes)*. The measured height of a 100 sheet block allows the conclusion that the height of one sheet is approx. 93 micron thick. Shifting the sheets apart and placing the resulting ramp on the acquisition surface allows reconstruction of steps going into depth at a known height of 93 micron each. The rendering of the ramp reconstruction (Figure 10, top) reveals that our scanning system is able to reconstruct the fine steps. Figure 10 (bottom) is a height profile plot over a length of about 1 cm along the ramp. The steps are clearly distinguishable and at significant distance from the extent of noise over the sheet area, which is caused by the uneven micro-structure of the paper and impacted by camera noise.

This observation lets us assume that the depth accuracy is even beyond 100 micron. To determine an upper limit on depth accuracy, we constructed a continuous ramp going into depth. We used a square bar steel over the length of the measuring range and placed it on two pivots, resting on the scanning surface. The flat side of the steel facing the camera was covered with a white plastic layer from a camera color calibration chart. In this setting, we reconstructed the white surface between the two pivots. Then we elevated the steel bridge at one pivot by placing one of the sheets with known height between steel and pivot, and again reconstructed the white surface. When comparing the two reconstructions, there are two surfaces intersecting at the first pivot's position and deviating in depth at the second pivot's position (by 93 micron). In between, we thus achieve continuous depth differences between 0 and 93 micron. The position where the difference of the two height profiles sets itself apart from the noise leads to a minimum depth difference distinguishable by our 3D scanner, due to the distance between pivots and height difference at the second pivot known, and thus to an upper bound on its depth accuracy. Figure 11 visualizes the height difference between the profiles through the two reconstructions, for the region between

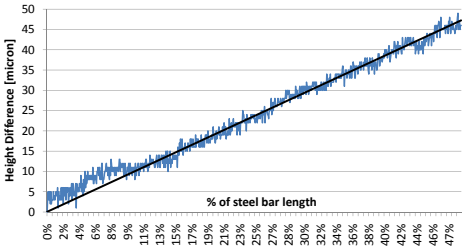

Figure 11: *Plot of height differences between profiles of initial reconstruction and reconstruction with second pivot elevated (by 93 micron), for first half of ramp. The fitted line reveals that deviances are small against height differences already after 30% of ramp length.*

the fixed pivot and the middle of the bar. A fitting line shows the linear profile trend and makes the magnitude of deviations clear. Already at 50% of the bar, deviations are small compared against the height difference of the two profiles. As this position corresponds to an object height difference of about $\frac{93}{2}$ micron (half of the height difference achieved by elevation), it is obvious that depth accuracy is at least as high as 50 micron, which is a new lower bound on accuracy with still a high safety tolerance. Setting deviances in relation to distance, the two reconstructions can still be clearly distinguished at 30%, which is about 28 micron. However, deviances shown here apply to reconstruction of the white material used and vary in magnitude for different materials, meaning that depth accuracy is dependent upon materials scanned.

7. Conclusion and Future Work

We developed a meso-scale acquisition system targeting two domains, one being the high accuracy, fine scale and comparatively large range digital preservation of cultural heritage objects, the other one being the support of realistic rendering with measured meso-structure height fields of the materials to be reproduced. Using a 1 DOF mechanical lens-shifter extension allows entering novel domains of depth accuracies while avoiding limitations known from common structured light approaches or artifacts introduced by projection devices. A very low positional error is achieved, as the evaluation of deviations in the geometry resulting from two consequent scans showed, which is significantly below the one of the comparison technique. The area resolution of 55 micron is bounded by the camera used, while the depth resolution is influenced by camera noise and the projection technique. Our acquisition setup uses off-the-shelf-components and delivers high accuracy at low cost. Being bounded by the precision of hardware, the accuracy of our system could further be increased by gray scale cameras without SLR, avoiding vibration due to mirror movement and coming at higher luminance sensitivity and lower noise, or by adding a second camera which would increase precision and help in regions of high specularities not visible to the first camera. While we currently use a simple calibration model with manually measured parameters, not including lens distortion or similar effects and abstracting from lens-shifter motion, an automatic calibration procedure using a test pattern would be a valuable step further to avoid geometry distortions. Using a laser-based projector like the ultra-miniature, scanning projection Microvision PicoP projector would raise the accuracy due to sharp projection and avoiding dark regions between pixels. With miniature size and no need for focus adjustment, minimum distance to the object could be decreased, increasing the projected resolution and thus the area resolution of the reconstruction.

Acknowledgment

This work was partially supported by the European project 3D-COFORM (FP7-ICT-2007.4.3-231809) and the DFG Emmy Noether fellowship GO 1752/3-1.

References

[AAA*05] ADAMSON A., ALEXA M., ANDERSEN O., ATTENE M., BAREQUET G.: *State-of-the-Art Report Survey Acquisition and Reconstruction*. Tech. rep., AIM@SHAPE, 2005. 2

[CGS06] CHEN T., GOESELE M., SEIDEL H.-P.: Mesostructure from specularity. In *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Washington, DC, USA, 2006), pp. 1825–1832. 2

[CSL08] CHEN T., SEIDEL H.-P., LENSCH H. P. A.: Modulated phase-shifting for 3d scanning. In *CVPR* (2008), IEEE Computer Society. 2, 3

[FCM*08] FRANCKEN Y., CUYPERS T., MERTENS T., GIELIS J., BEKAERT P.: High quality mesostructure acquisition using specularities. *CVPR 2008* (June 2008), 1–7. 2

[GÖ1] GÜHRING J.: Dense 3-d surface acquisition by structured light using off-the-shelf components. In *Proc. Videometrics and Optical Methods for 3D Shape Measurement* (2001), pp. 220–231. 2

[HNG*10] HUFF R., NEVES T., GIERLINGER T., KUIJPER A., STORK A., FELLNER D.: A general two-level acceleration structure for interactive ray tracing on the gpu (to appear in cgi 2010), 2010. 6

[LM07] LILIENBLUM E., MICHAELIS B.: Optical 3d surface reconstruction by a multiperiod phase shift method. *Journal of Computers (JCP 2* (2007), 73–83. 2, 3, 4

[Lou05] LOURAKIS M. I. A.: A brief description of the levenberg-marquardt algorithm implemented by levmar. foundation for research and technology, 2005. 4

[NRDR] NEHAB D., RUSINKIEWICZ S., DAVIS J., RAMAMOORTHI R.: Efficiently combining positions and normals for precise 3D geometry. *Proc. of ACM SIGGRAPH 2005*. 2

[RCM*01] ROCCHINI C., CIGNONI P., MONTANI C., PINGI P., SCOPIGNO R.: A low cost 3d scanner based on structured light. *Computer Graphics Forum 20*, 3 (2001), 513–522. 2

[SDS09] SCHOLZ M., DANCH D., STORK A.: An analog lens-shifting device for digital fringe projection. In *PROCAMS 2009. IEEE International Workshop on Projector-Camera Systems in conjunction with CVPR 2009* (2009). 2, 3