

Industrial Phase-Shifting Profilometry in Motion

P. Schroeder, R. Roux, J.-M. Favreau, M. Perriollat, and A. Bartoli

ISIT – Image Science for Interventional Techniques, UMR 6284 UdA – CNRS
28, place Henri Dunant, 63001 Clermont-Ferrand, France
pierre.schroeder@etu.udamail.fr
<http://isit.u-clermont1.fr/>

Abstract. Phase-Shift Profilometry (PSP) provides a means for dense high-quality surface scanning. However it imposes a staticity constraint: The scene is required to remain still during the acquisition of multiple images. PSP is also not applicable to dynamic scenes. On the other hand, there exist active stereo techniques which overcome these constraints but impose other limitations, for instance on the surface’s continuity or texture, or by significantly reducing the reconstruction’s resolution.

We present a novel approach to recover reconstructions as dense and almost as accurate as PSP but which allows for a translational object/scene motion during the acquisition of multiple input frames, study its performance in simulations, and present real data results.

1 Introduction

Active stereo reconstruction techniques provide a means to acquire 3D reconstructions of textureless surfaces with few or no approximations in regard to the observed objects and surfaces. Other approaches are passive stereo (camera to camera) techniques [9], which require amply textured surfaces, or time-of-flight cameras [5] and infrared active stereo systems [8], which currently provide less precise results at low resolution.

There are significant differences amongst active stereo techniques: patient respiratory motion surveillance for instance requires real-time but not necessarily texture-insensitive high-density reconstructions and surface smoothness constraints are valid. Multimedia applications however (e.g. gaming industry) might use free-hand motion capable, texture-aware approaches yielding sparse low-resolution approximations of physical models onto which the texture is mapped in the application, which suffices for visualisation purposes. Digital historical heritage archiving (e.g. busts or reliefs) in turn requires dense reconstructions but can fulfil staticity constraints.

Industrial metrology and inspection systems commonly rely on Phase-Shift Profilometry (PSP) using sinusoidal fringe patterns, yielding dense and accurate reconstructions regardless of presence or absence of texture. However, it requires the scene to remain still in regard to the camera-projector-setup while multiple sequential images are acquired as the projected pattern changes. Solutions which overcome the staticity constraint either drastically reduce the reconstruction resolution in regard to the projector’s native resolution and raise issues in regard to merging multiple single-shot point clouds [3], render the system sensitive to texture [4] or even colour [14], or require expensive high speed hardware [12].

Many industrial inspection systems require the reconstruction density of PSP but cannot afford staticity on production lanes due to other manufacturing steps or time-consuming mechanical limitations (e.g. acceleration, deceleration, and stabilisation intervals of the production lane or part). In return though, the inspection system is only required to handle mainly omni-directional and controlled/monitored motion, rather than coping with free-hand motion. We found literature would lack a suitable method, which led to developing the one presented hereafter. It acquires 3D reconstructions as dense and almost as accurate as PSP but allows the scene to move along a main translation axis while acquiring multiple frames. Besides a powerful stroboscope, instead of the projector’s standard light source (to avoid motion blur in images), our method doesn’t require particular hardware like telecentric optics or high frame-rate projectors and cameras. Furthermore it remains robust to the scene’s texture and colour.

Paper. Section 2 reviews related work. Section 3 introduces our approach, the performance of which we then study on synthetic data in Sect. 4 and which we demonstrate with a real example in Sect. 5. We discuss our results and provide an outlook for future work in Sect. 6 and elucidate our conclusions in Sect. 7.

2 Related Work

Single-frame binary projection patterns either aim at reducing decoding ambiguities (camera-projector setups) as do Hall-Holt and Rusinkiewicz in [3] or maximise discrimination of disparity (projector supported camera-camera setup) of the projection pattern as in Konolige’s work [6]. For low-textured scenes with neutral colour, setups providing a colour projector and camera can multiplex multiple patterns onto the RGB-channels as do Zhang, Curless, and Seitz in [14].

Weise, Leibe, and Van Gool [12] put into relation reconstructed objects’ velocity and acceleration with the phase measuring error of 3-step high-speed PSP and provide an analytically derived correction LUT. Spatio-temporal approaches [13,1] aim at temporally validating spatial observations.

A more extensive overview of active stereo techniques, covering inter alia approaches capable of handling motion, is provided by Salvi et al. [7].

Although the aforementioned techniques provide good results, we have not been able to identify a suitable state-of-the-art approach for our particular industrial application (schematically illustrated in Fig. 1). Techniques relying on discrete patterns underly projectors’ physical limitations (native resolution) and require decoding-/matching-windows in camera-images, reducing attainable level of detail (resolution), while single-frame phase measuring methods have difficulties in handling surface discontinuities and texture. Our approach though targets on an industrial metrology application which provides

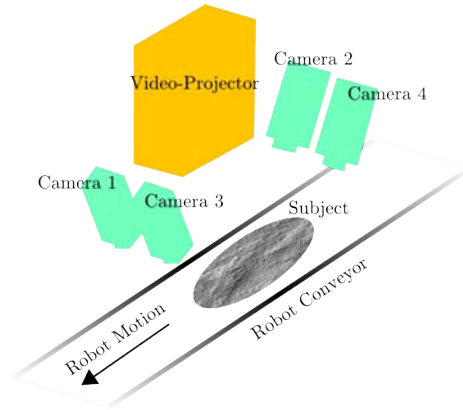


Fig. 1: Schematic representation of our hardware setup.

a well known environment (e.g. calibrated projector and camera, and actively controlled motion) in which staticity is expensive to afford. Our industrial partner requires us to conceal the real application though. We therefore reconstruct a credit-card with added colour/texture (conserving a similar scale and textural topology) as real data example instead.

3 Our Approach

3.1 Hardware Setup

Figure 1 schematically illustrates our hardware setup. It is composed of an optical and a mechanical part and can be considered isolated against external light sources and ambient illumination. The optical head contains one SVGA or XGA video-projector facing the scene perpendicularly, together with a set of four—or two—synchronised cameras equipped with common (entocentric) lenses and 4 mega-pixel grayscale sensors, accurately calibrated and pointed to the region illuminated by the projector, such that the Field Of Views (FOV) of the optics best overlap in the target Field Of Depth (FOD).

The mechanical part is a highly precise robot conveyor required by preceding manufacturing steps which provides accurate estimates of reconstruction targets' motion.

3.2 Choice of Projection Pattern

As projectors are in general of lower resolution in contrast to comparable cameras, using binary projection patterns results in a reconstruction resolution considerably lower than the cameras' resolution. That is, in the scene, several camera pixels image the same projector pixel—hence only duplicate information—and an entire neighbourhood of pixels is required to decode/match a patch.

Continuously encoded projection patterns in contrast intrinsically permit precise sub-pixel matching (in regard to the projector's pixels). Requiring a continuous and periodical pattern providing maximum information along mutual epipolar lines, a sinusoidal fringe pattern appears to be a conclusive choice.

Static PSP shifts n times the phase of the pattern with period λ by $\frac{\lambda}{n}$. This maximises the information over n images, which lead us to synchronise the acquisitions on translations corresponding to $\frac{1}{n} + k$ times the sine-wave's period at the target depth.

3.3 Reference Coordinate Frame in Motion

The target, subject to reconstruction, roughly represents a planar surface, hence we define the reference coordinate frame to be a planar point-grid in the world coordinate frame of the first shot and aligned to the a priori depth of the target. This results in m points $\mathbf{Q}_i^1(h_i) = [x_i \ y_i \ h_i]^T$ in space defined in terms of their variable height h_i . The corresponding points \mathbf{Q}_i^t at timestep t in regard to the reference grid at timestep $t = 1$ suffice the relation

$$\mathbf{Q}_i^t(h_i) = \mathbf{R}_t \mathbf{Q}_i^1(h_i) + \mathbf{t}_t, \quad (1)$$

with the rigid three-dimensional rotation \mathbf{R}_t and translation \mathbf{t}_t realigning the target at timestep $t = 1$ to itself at timestep t , such that, for a camera C , the pixel location

$${}^C \mathbf{q}_i^t(h_i) \sim \mathbf{P}_C \mathbf{Q}_i^t(h_i). \quad (2)$$

of a point over time can be computed (with the projection matrix \mathbf{P}_C of camera C). The dot under a letter refers to a point's homogeneous coordinates and \sim stands for the equality up to scale. The projector-camera setup is calibrated and radial distortion in images is corrected, enabling us to express the setup's geometry accurately using the pin-hole model.

Considering $\mathcal{I}_i^C(\mathbf{q})$ being the (linear interpolated) gray-level of the t -th image out of n , shot by camera C , at the pixel-coordinates \mathbf{q} , we construct the vector

$$\mathcal{I}_i^C(h_i) = [\mathcal{I}_1^C({}^C \mathbf{q}_i^1(h_i)) \dots \mathcal{I}_n^C({}^C \mathbf{q}_i^n(h_i))]^T. \quad (3)$$

Assuming we knew the height of a point on the surface, this vector would contain the gray-level intensities of the physical point on the target as captured by a camera at different timesteps; otherwise it suffers from parallax error indeed. Additionally, the illumination a physical point encounters varies from step to step, because the projection remains still in regard to the camera but not the scene. Therefore the same point is imaged with different intensity values throughout the image sequence.

These vector valued functions $\mathcal{I}_i^C(h_i)$ will be used to define cost functions (developed in Sect. 3.4 and 3.5), subject to minimisation in terms of the heights of the points on the reference grid. Thus, for r cameras we formulate the 3D-reconstruction problem for each point on the reference grid independently as

$$\arg \min_{\hat{h}_i} \mathcal{C}(\hat{h}_i) = \arg \min_{\hat{h}_i} \mathcal{C}'(\mathcal{I}_i^1(\hat{h}_i), \dots, \mathcal{I}_i^r(\hat{h}_i)). \quad (4)$$

3.4 Intensity Matching Cost

The Intensity Matching cost (IMx) does not exploit any knowledge about the projected pattern and directly matches the intensities between multiple cameras over all timesteps altogether.

$$\mathcal{C}_{\text{IM2}}(h_i) = \|\mathcal{I}_i^1(h_i) - \mathcal{I}_i^2(h_i)\|_2^2 \quad (5)$$

for two cameras, or for four cameras we write

$$\mathcal{C}_{\text{IM4}}(h_i) = \sum_{k=1}^4 \left\| \mathcal{I}_i^k(h_i) - \frac{1}{4} \sum_{l=1}^4 \mathcal{I}_i^l(h_i) \right\|_2^2 \quad (6)$$

3.5 Pseudo Phase Matching Cost

Our second cost, Pseudo Phase Matching (PPMx), is very analogue to the principle of PSP. Actually we are tempted to state it would be an altered PSP approach, such that the phase-shift is exerted through the motion of the scene instead of shifting the signal

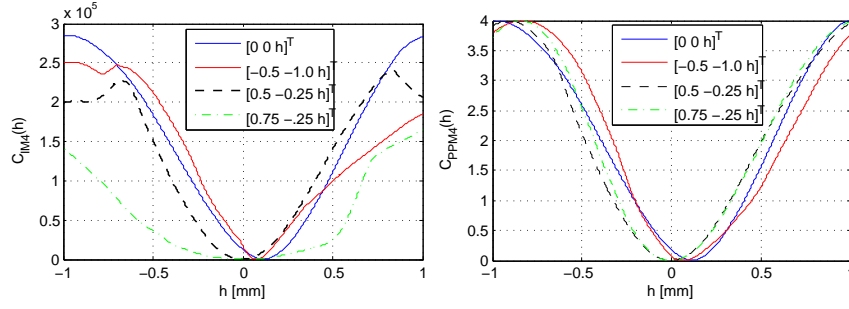


Fig. 2: Plot of cost functions IM4 (left) and PPM4 (right) at four different points in the simulated scene and using ideal synthetic images (no camera or motion noise).

in the projection device's image plane. This certainly causes a number of difficulties which require an adapted solution too.

The projection pattern is defined in terms of the pixel location \mathbf{p} by

$$L(\mathbf{p}) \sim \frac{1}{2} + \frac{1}{2} \cos(\boldsymbol{\omega}^T \mathbf{p}) \quad (7)$$

with the two-vector $\boldsymbol{\omega}$ defining the pattern's frequency along the main dimensions of the projector image. We assume that for a known height of a physical point on the surface, the relation between projected and acquired signal as it is modelled in static PSP holds even in the presence of the very limited viewpoint changes in our context and state that

$$\mathcal{I}_t^C(C \mathbf{q}_i^t(h_i)) = A + B \cos(\boldsymbol{\omega}^T \mathbf{p}_i^t(h_i)) \quad (8)$$

with the corresponding projector pixel $\mathbf{p}_i^t(h_i) \sim P_P \mathbf{Q}_i^t(h_i)$. This may be rewritten

$$\mathcal{I}_t^C(C \mathbf{q}_i^t(h_i)) = A + B \cos(\underbrace{\boldsymbol{\omega}^T \mathbf{p}_i^1(h_i)}_{\varphi_i(h_i)} + \underbrace{\boldsymbol{\omega}^T (\mathbf{p}_i^t - \mathbf{p}_i^1)}_{\delta\varphi_i^t(h_i)}). \quad (9)$$

Equations (8) and (9) introduce the notion of phase-shifting by viewpoint change—in contrast to phase-shifting by projection [10]—but conserve the notion of a point's passive properties A , texture/colour under ambient light, and B , a reflection coefficient; and the active and altering illumination referred to by the cosine. After trigonometric expansion the preceding equation can be written under matrix notation as

$$\mathcal{I}_i^C(h_i) = \underbrace{\begin{bmatrix} 1 & \cos(\delta\varphi_i^1(h_i)) & -\sin(\delta\varphi_i^1(h_i)) \\ \vdots & \vdots & \vdots \\ 1 & \cos(\delta\varphi_i^m(h_i)) & -\sin(\delta\varphi_i^m(h_i)) \end{bmatrix}}_{\Delta_i(h_i)} \underbrace{\begin{bmatrix} A \\ B \cos(\varphi_i(h_i)) \\ B \sin(\varphi_i(h_i)) \end{bmatrix}}_{\mathbf{x}_i^C(h_i)}. \quad (10)$$

Thus we can compute an estimator $\hat{\mathbf{x}}_i^C(h_i) = [\hat{\gamma}_i^C(h_i) \hat{\alpha}_i^C(h_i) \hat{\beta}_i^C(h_i)]^T$ for $\mathbf{x}_i^C(h_i)$:

$$\hat{\mathbf{x}}_i^C(h_i) = \Delta_i^\dagger(h_i) \mathcal{I}_i^C(h_i), \quad (11)$$

the components of which are then used to derive

$$\begin{bmatrix} \hat{c}_i^C(h_i) \\ \hat{s}_i^C(h_i) \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_i^C(h_i) \\ \hat{\beta}_i^C(h_i) \end{bmatrix} / \left\| \begin{bmatrix} \hat{\alpha}_i^C(h_i) \\ \hat{\beta}_i^C(h_i) \end{bmatrix} \right\| \quad (12)$$

which does represent the pseudo-phase at a given height. The fact that at a point's real height $\hat{c}_i^C(h_i)$ and $\hat{s}_i^C(h_i)$ represent estimators for $\cos(\varphi_i(h_i))$ and $\sin(\varphi_i(h_i))$ respectively, explains our choice of referring to them as *pseudo-phase*.

We state that at the real height in the ideal case both terms from one camera should equal their counterparts from another camera and that they diverge (for non-adjacent cameras) with increasing parallax error.

$$\mathcal{C}_{\text{PPM2}}(h_i) = \left\| \begin{bmatrix} \hat{c}_i^1(h_i) \\ \hat{s}_i^1(h_i) \end{bmatrix} - \begin{bmatrix} \hat{c}_i^2(h_i) \\ \hat{s}_i^2(h_i) \end{bmatrix} \right\|_2^2 \quad (13)$$

or, for four cameras, we rewrite the cost as

$$\mathcal{C}_{\text{PPM4}}(h_i) = \sum_{k=1}^4 \left\| \begin{bmatrix} \hat{c}_i^k(h_i) \\ \hat{s}_i^k(h_i) \end{bmatrix} - \frac{1}{4} \sum_{l=1}^4 \begin{bmatrix} \hat{c}_i^l(h_i) \\ \hat{s}_i^l(h_i) \end{bmatrix} \right\|_2^2. \quad (14)$$

For maximum information, the acquisition is synchronised to an inter-frame translation which corresponds to an n -th of the pattern's sine-wave's period modulo 2π in the reference plane. Figure 2 shows $\mathcal{C}_{\text{PPM4}}$ and \mathcal{C}_{IM4} of several points in regard to their supposed height on ideal, noise-free synthetic images and confirm our statement. In our implementation we used the Levenberg-Marquardt algorithm to optimize the costs, but other optimization algorithms might be suitable as well.

4 Simulation

4.1 Data

In order to test our approach, we used our in-house simulation software aiming at reflecting the available hardware in order to synthesise images as they would be acquired by the cameras while the scene is illuminated by a particular projection pattern. It uses ray-casting to determine a pixel's imaged intensity, based on the Lambertian reflectance model, according to the scene's colour and illumination at that point.

Additionally the simulation allows us to control the strength of several sources of uncertainty one by one and investigate the validity of our approach in the ideal case and its robustness to camera noise and motion estimation uncertainties.

The simulated scene pictured in Fig. 3 is composed of a planar patch with an out-bevelled "3" as profile and a texture-mapped "D". Thus, texture is present but not sufficiently for being exploited on its own by the reconstruction algorithm.

For each camera we rendered batches of sixteen scene-shifted images with a static sine-wave projection pattern, as well as eight images with a static scene but phase-shifted projection patterns, and sixteen scene-shifted images with a static binary pattern. All of them with different levels of camera noise ($\sigma_{GL} \in \{0, 0.1275, \mathbf{0.255}, 0.51, 1.02\}$)

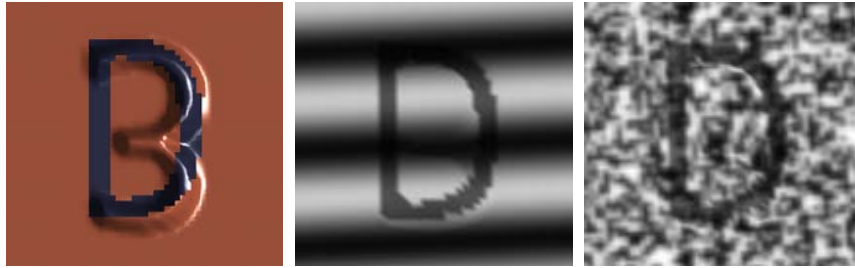


Fig. 3: Visualisation of the synthetic target: a plane with an out-bevelled "B" and a "D" mapped onto the surface in a different colour (left); region of interest of synthetic image with sinusoidal pattern projection (centre); region of interest of synthetic image with random pattern projection (right).

on a graylevel scale ranging from 0 to 255 and motion estimation noise ($\sigma_{ME} \in \{0, 0.0037, \mathbf{0.0075}, 0.015, 0.03\}$ millimeters)—the reference noise is typeset in bold. Two other methods allowed us to compare ours to and situate latter in regard to state-of-the-art techniques. These are our Static PSP (SPS) implementation of [10], yielding very good high-density reconstructions in the static case, and the implementation of Furukawa and Ponce's passive stereo [2] with random binary pattern projection followed by merging the sequential one-shot-reconstructions, as a candidate for reconstructing scenes in motion—which we had to fall back to, because [6], the only published implementation of that kind of approaches, appeared to be no more accessible.

4.2 Results

Comparison to binary projection techniques. Figure 4 shows a reconstruction based on the images with random-texture projections, reconstructed the scene at each timestep independently by using Furukawa's code and merged the sixteen one-shot-reconstructions afterwards.

Although this reconstruction looks very good, it confirms that the method is not suitable in our context: the input images were noise-free and the transformations between timesteps were exactly those used to synthesise the images and, nonetheless, the Root Mean Squared Error (RMSE)—between the reconstructed z-map and the ground truth—is 0.0467mm by tuning the parameters manually to find a suitable trade-off between smoothing and level of detail. Comparing this result under ideal conditions to the averaged RMSE of our PPMx and IM4, below 0.004mm including reference camera noise and motion uncertainties, we did not include this approach for further evaluation.

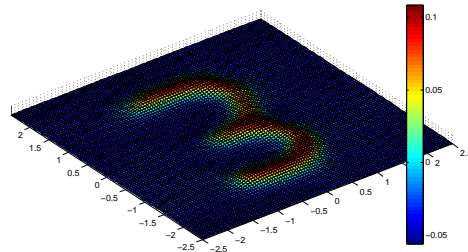


Fig. 4: Synthetic target reconstructed with random projection pattern (see Fig. 3) using Furukawa's code.

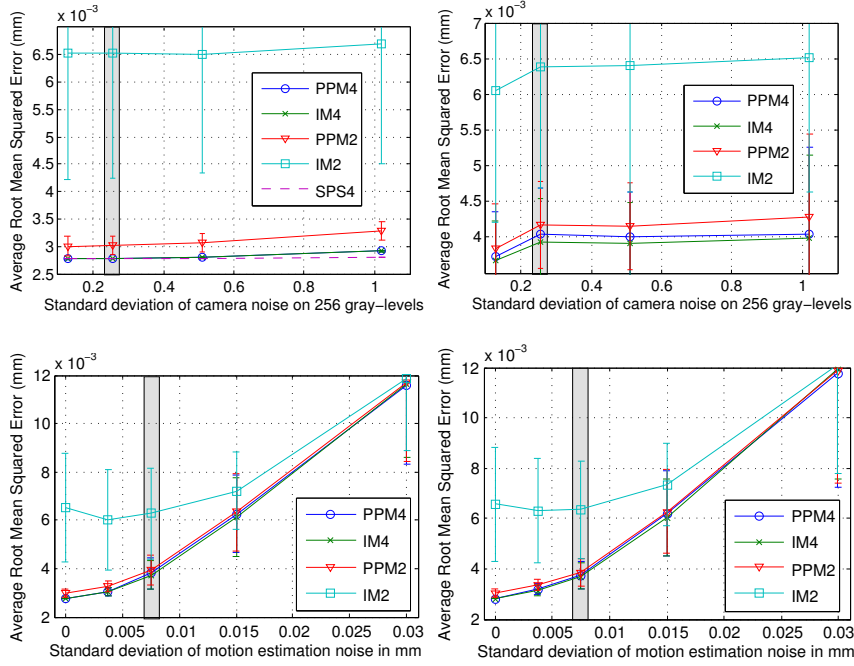


Fig. 5: RMSE in terms of camera noise with ideal motion estimation (top left) and with reference motion estimation noise (top right); RMSE in terms of motion estimation noise without (bottom left) and with reference camera noise (bottom right); measures averaged over 100 samples.

Camera noise. Figure 5 (top) shows two graphs picturing the average RMSE from 100 samples of reconstructions with our two cost functions, both in their two- and four-camera versions. The graph to the left shows the RMSE in regard to increasing camera noise while the ideal motion was used and, in order to provide a reference to the reader, includes results obtained with 4-camera SPS (SPS4) on a static scene. It shows that PPM4 and IM4 are slightly more sensitive to camera noise than SPS4 and perform both similar. PPM2 is generally more sensitive to noise and IM2 performs considerably worse than the other methods.

Enabling the reference motion estimation noise in addition to increasing camera noise, all the methods perform worse and less constant, considering the standard deviation of the RMSE over the samples as can be seen in the lower graph. Meanwhile IM2 remains comparable with or without motion noise, hence seems mostly sensitive to camera noise. It turns out that the bad performances of IM2 right from the beginning is probably caused by an elevated number of outliers, which can not be identified as such, because the reconstructed values remain inside the target FOD.

For the second case, SPS4 is no more included in the experiments, as there is no meaningful notion of motion estimation uncertainty.

Motion estimation noise. The following experiment reveals the methods' sensitivity to motion estimation noise. Figure 5 (bottom) pictures the RMSR in terms of increasing motion estimation noise while leaving camera noise disabled (left) and using the reference camera noise (right).

In both cases PPM4 and IM4 perform similarly, leaving PPM2 slightly behind. IM2 remains much worse than the other methods in this experiment too although, with heavy motion uncertainties, the performances of all four methods converge.

The vanishing differences between both graphs indicate less robustness against motion uncertainties than camera noise, in regard to their respective reference noises.

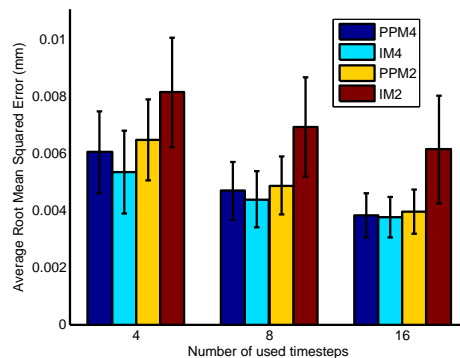


Fig. 6: RMSE averaged over 100 samples with reference camera noise and reference motion estimation noise.

Number of timesteps. Finally we investigate to what extent the methods' performances depend on the number of scene-shifts (resp. timesteps) under presence of camera noise and motion uncertainties. This is reflected by Fig. 6 revealing that with increasing number of steps PPM4, IM4, and PPM2 tend to comparable performances outperforming IM2 in any way.

5 Real Data Example

5.1 Data

The credit card onto which we added sparse texture with different white-board markers has been placed onto the obfuscated real target. It then passed an acquisition cycle of sixteen scene-shifts yielding the sixteen images for each of the four cameras as well as the motion estimation delivered by the robot conveyor.

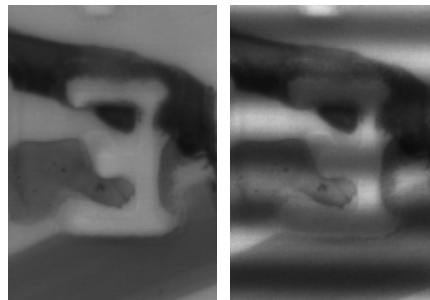


Fig. 7: Region of interest on real scene, subject to reconstruction, with uniform illumination (left) and fringe-projection (right).

5.2 Results

Figure 8 shows the 3D-reconstructions obtained by SPS4 with no motion (Fig. 8a and 8b) as best available approximation to the ground-truth and the reconstructions obtained by minimising our cost-functions (Fig. 8c to 8f). All reconstructions are colour-encoded in regard to their height maps except for Fig. 8b, the colour-encoding of which corresponds to the reconstructed texture.

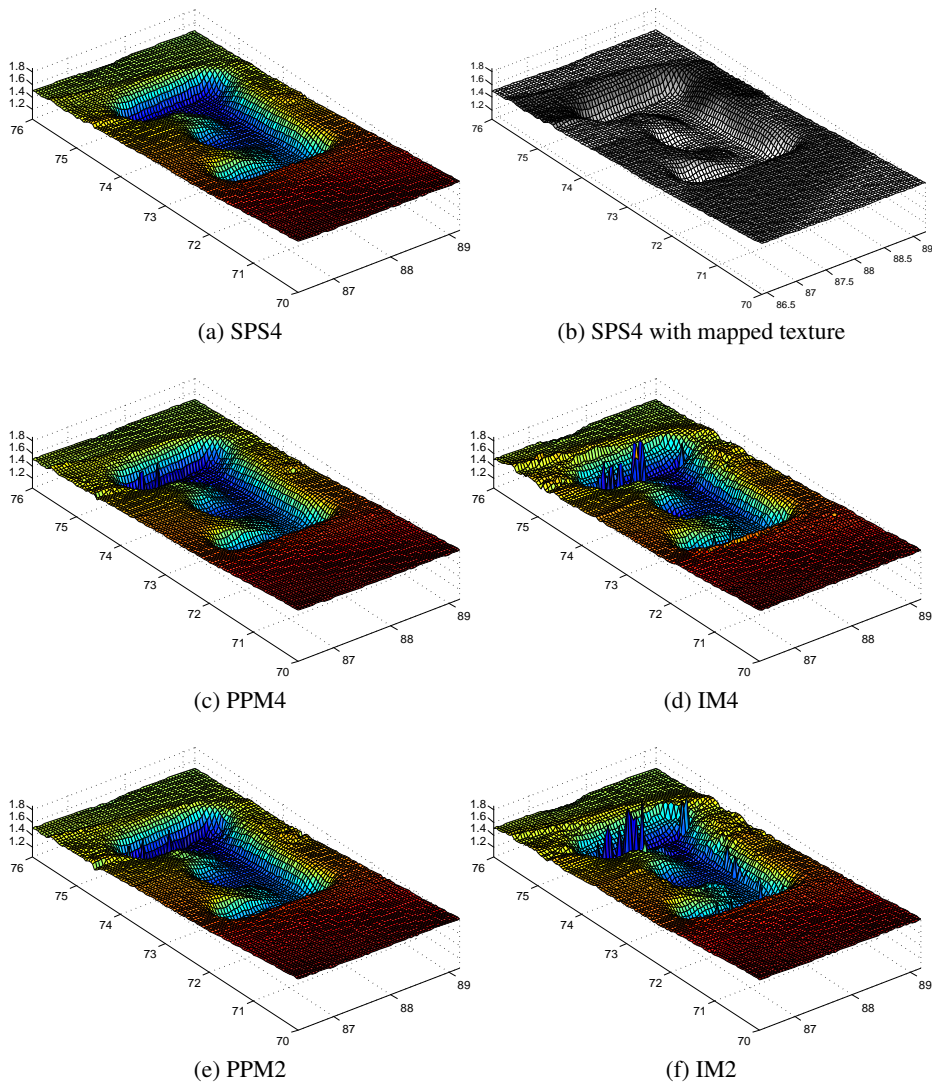


Fig. 8: (a)-(d): Reconstructions on real data images using all four cameras. (e)-(f): Reconstructions on real data images using camera 1 and 3.

All four reconstructions obtained from scene-shift images are in general consistent with the result from SPS4, although the reconstructions obtained with PPM4 and PPM2 do more resemble SPS than do the ones obtained with IM4 and IM2.

On one hand, IMx is more sensible to aggressive texture edges (e.g. upper left corner) than PPMx (also visible in SPS4 in this example) resulting in stronger artefacts along the texture edges. Looking at the texture edge located at the middle left, IMx

suffers from artefacts along that edge while PPMx hardly shows any artefacts along that edge and SPS4 shows none at all; or at the lower right of the reconstruction, which results in artefacts only for IM4.

On the other hand, on the dark coloured stripe in the upper left corner, providing low signal strength, IMx suffers from elevated reconstruction noise while PPMx is considerably less impacted. Furthermore, PPMx performs almost as good as SPS4 on the planar surface on the lower right in terms of reconstruction noise than does IMx.

The valley on the upper left shows some artefacts in PPMx and numerous in IMx, which we suspect to be caused by interreflections, specularities, or a combination of both, as the bottom of the valley acts as—in a manner of speaking—a convex mirror.

Figure 9 shows the distributions of the reconstruction differences between our methods and SPS. All four methods yield reconstructions, the modes of which are off by approximately $-5\mu\text{m}$ to $-10\mu\text{m}$. The distributions for PPM4 and PPM2 are more peaked than those for IM4 and IM2, which are both stronger biased towards lower height-values.

6 Discussion

The introduced approaches PPMx and IMx are distinct in whether or not they exploit a priori knowledge about the projected pattern. IMx can be considered to follow an unstructured illumination spatio-temporal paradigm like [13] (although the pattern is structured by design) and serves more as the realisation of a state of the art approach for our application. PPMx in contrast represents a novel realisation of known multi-camera PSP techniques by exploiting knowledge about the projected signal, notably the phase-shift of the pattern solely borne in regard to the scene.

Our method suits well the controlled/monitored industrial environment which gave rise to the method’s development in the first place. It yields maximum dense reconstructions of low-textured surfaces in motion without forfeit of level of detail, nor the expense of high-speed cameras or video-projectors. It even rather reduces the projection-device’s complexity as the pattern remains static, which could remedy further disturbance factors like projector gamma, sine-pattern asymmetry due to asymmetric alignment and relaxation times of Liquid Crystals in LCD projectors, mechanical calibration of piezoelectric actuated projectors, or video-projectors’ inter-pixel spacing.

Furthermore our system does not require expensive telecentric lenses which would either drastically reduce the field-of-view or increase the size of the optical setup.

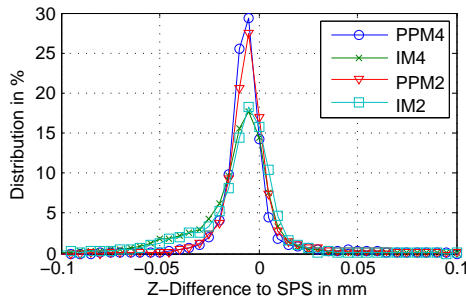


Fig. 9: Distributions of point-wise differences between the z-maps obtained with each one of our methods and the z-map obtained by SPS.

Table 1: Mean and standard deviation of differences of z-maps regarding SPS.

	Mean (mm)	σ (mm)
PPM4	-0.0052	0.0168
IM4	-0.0144	0.0352
PPM2	-0.0040	0.0185
IM2	-0.0127	0.0342

7 Conclusion

We introduced a novel method which closes a gap between state-of-the-art PSP and one-shot techniques for scenes in motion, which we previously identified in the domain of industrial metrology and surface inspection. We demonstrated its validity in simulations and on real data, and quantified its accuracy, the order of which is close to that one of SPS and out-performs current one-shot techniques.

Open issues, subject to future studies and development, may include handling of specularities, more analytical insight to the cost functions, exploitation of reconstructed texture [11], spatial regularisation of neighbourhood including cost-scores, extending the costs by a reflectance model, or investigating the processing speeds of (non-prototype) optimised implementations.

References

1. Davis, J., Nehab, D., Ramamoorthi, R., Rusinkiewicz, S.: Spacetime stereo: A unifying framework for depth from triangulation. *PAMI* 27(2), p. 296–302 (2005) [2](#)
2. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multi-view stereopsis. *PAMI* 32(8), p. 1362–1376 (2010) [7](#)
3. Hall-Holt, O., Rusinkiewicz, S.: Stripe boundary codes for real-time structured-light range scanning of moving objects. *ICCV* (2001) [1](#), [2](#)
4. Huang, L., Kemaq, Q., Pan, B., Asundi, A.K.: Comparison of Fourier transform, windowed Fourier transform, and wavelet transform methods for phase extraction from a single fringe pattern in fringe projection profilometry. *Optics and Lasers in Eng.* 48, p. 141–148 (2010) [1](#)
5. Jongenelen, A., Bailey, D., Payne, A., Dorrington, A., Carnegie, D.: Analysis of errors in tof range imaging with dual-frequency modulation. *Instrumentation and Measurement* 60(5), p. 1861–1868 (2011) [1](#)
6. Konolige, K.: Projected texture stereo. *ICRA*, p. 148–155 (2010) [2](#), [7](#)
7. Salvi, J., Fernandez, S., Pribanic, T., Llado, X.: A state of the art in structured light patterns for surface profilometry. *Pattern Recognition* 43(8), p. 2666–2680 (2010) [2](#)
8. Smisek, J., Jancosek, M., Pajdla, T.: 3d with kinect. *ICCV Works.*, p. 1154–1160 (2011) [1](#)
9. Taguchi, Y., Wilburn, B., Zitnick, C.: Stereo reconstruction with mixed pixels using adaptive over-segmentation. *CVPR* (2008) [1](#)
10. Wang, Y., Liu, K., Hao, Q., Lau, D., Hassebrook, L.G.: Multicamera phase measuring profilometry for accurate depth measurement. *SPIE*, vol. 6555 (2007) [5](#), [7](#)
11. Wang, Y., Liu, K., Hao, Q., Wang, X., Lau, D., Hassebrook, L.: Robust active stereo vision using kullback-leibler divergence. *PAMI* 34(3), p. 548–563 (2012) [12](#)
12. Weise, T., Leibe, B., Gool, L.V.: Fast 3d scanning with automatic motion compensation. *CVPR* (2007) [1](#), [2](#)
13. Zhang, L., Curless, B., Seitz, S.: Spacetime stereo: shape recovery for dynamic scenes. *CVPR* 2, p. 367–74 (2003) [2](#), [11](#)
14. Zhang, L., Curless, B., Seitz, S.M.: Rapid shape acquisition using color structured light and multi-pass dynamic programming. *3D Data Proc. Visual. and Trans.* (2002) [1](#), [2](#)