Short Papers

Bayesian Helmholtz Stereopsis with Integrability Prior

Nadejda Roubtsova and Jean-Yves Guillemaut

Abstract—Helmholtz Stereopsis is a 3D reconstruction method uniquely independent of surface reflectance. Yet, its sub-optimal maximum likelihood formulation with drift-prone normal integration limits performance. Via three contributions this paper presents a complete novel pipeline for Helmholtz Stereopsis. First, we propose a Bayesian formulation replacing the maximum likelihood problem by a maximum a posteriori one. Second, a tailored prior enforcing consistency between depth and normal estimates via a novel metric related to optimal surface integrability is proposed. Third, explicit surface integration is eliminated by taking advantage of the accuracy of prior and high resolution of the coarse-to-fine approach. The pipeline is validated quantitatively and qualitatively against alternative formulations, reaching sub-millimetre accuracy and coping with complex geometry and reflectance.

Index Terms—Helmholtz Stereopsis, 3D, complex reflectance, MAP

1 Introduction

3D geometry reconstruction is both challenging and much desirable for practical applications. State-of-the-art permits sub-millimetre accuracy given tailored capture conditions and surface properties. Multiple images are used to resolve depth ambiguity but methods differ in image acquisition, formulated constraints, whether to characterise vertices by depth or normal (or both) and the degree of neighbourhood support utilised in the estimation. Conventional [1], [2], [3], [4], [5], [6] and photometric [7], [8], [9], [10], [11], [12] stereo are mature techniques with exceptional results within a wide application scope. Their universal applicability is prohibited by underlying assumptions. Conventional stereo relies on uniformly Lambertian reflectance for invariance of feature point appearance, which is restrictive as purely Lambertian photometric behaviour is uncommon in reality. Photometric stereo allows any parametric invertible reflectance model as long as it is *a priori* known. As estimation/parametrisation of complex spatially-varying reflectance is challenging, the overly simplistic Lambertian assumption is often made by photometric techniques. Shape-from-Silhouette [13] is the only classical method with the rare property of true reflectance independence, but this geometric technique is limited by its fundamental inability to reconstruct concavities.

Helmholtz Stereopsis (HS) [14] tackles photometric complexity by exclusively exploiting the generic BRDF symmetry of reciprocity instead of a specific reflectance model through tailored acquisition. Its normal constraint is uniquely independent of the reflectance model which in addition provides a likelihood of the sampled depth. Hence the method's surface characterisation is more complete, with an unexplored potential for further reconstruction improvements, than the one-sided depth maps of conventional stereo or the normal fields of photometric stereo. The original formulation of HS [14] has shown the ability to cope with specularities. Yet, the sub-optimal per-point reconstruction approach and the failure to make full combined use of depth and normal information leads to global and local artefacts. A noisy depth map obtained by independent per-point depth search indexes normals. Normal integration reveals a surface of a much higher resolution than the original depth map and camouflages the errors by enforcing integrability *a posteriori*. The result is however incorrect as the integrated normal field is comprised of inaccurate spatially inconsistent normals reflecting the noise of the indexing depth maps, whose continuity is never enforced. Normal integration itself is prone to drift.

We propose a novel HS pipeline through several contributions addressing the shortcomings of standard HS. First, the per-point *maximum likelihood* (ML) reconstruction of standard HS is replaced by a Bayesian *maximum a posteriori* (MAP) formulation using neighbourhood support to tackle noisy depth maps. Second, a novel prior tailored to HS is formulated unlike standard HS exploiting both depth and normal per-point estimates and enforcing their consistency. The proposed prior ensures accuracy and integrability already at reconstruction de-emphasising explicit surface integration. In our pipeline the drift-prone surface integration stage is removed. We argue for the integration-free approach by highlighting the artefacts introduced by integration methods in standard pipelines (e.g., the Frankot-Chellappa algorithm in standard HS).

The current paper builds on our previous publications [15], [16] on Bayesian HS but also significantly improves and extends it. One such improvement is the re-formulation of the depth-normal consistency prior in a more theoretically principled way relating it directly to the physically meaningful and universally desirable surface integrability from [17]. The merit of the new formulation over the previously proposed is verified experimentally, both quantitatively and qualitatively, on synthetic and real data. Further Bayesian HS is embedded into a coarse-to-fine framework to achieve resolution sufficient for our integration-free pipeline enabling to showcase the true capabilities of the proposed prior. The proposed integration-free approach is validated against established explicit surface integration methods.

2 Related Work

We give an overview of related work focussing on the BRDF dependence aspect to justify the choice of HS for dense 3D reconstruction with arbitrary unknown reflectance.

Geometric methods involve multiview geometry of camera arrangement. These are both Shape-from-Silhouette (SfS) [13] deriving the visual hull [18] without intensity sampling as well as the intensity-based methods. Although SfS is reflectance independent and recently much improved [19], it suffers from the inherent inability to reconstruct concavities and severe dependence on number of views. Conventional stereo (CS) [1], [2] is a classical intensity-based geometric method that computes disparity by feature-point matching between calibrated views with sufficient texture and the restrictive Lambertian BRDF assumption to guarantee photo-consistency. Among CS methods with relaxed BRDF restrictions are the interest-point-based ones (e.g., photo-tourism [20]),

- *N. Roubtsova was with the CVSSP, University of Surrey, Guildford GU2 7XH, United Kingdom. She is now with the CAMERA, University of Bath, Bath BA2 7AY, United Kingdom. E-mail: n.s.roubtsova@bath.ac.uk.*
- *J.-Y. Guillemaut is with the CVSSP, University of Surrey, Guildford GU2 7XH, United Kingdom. E-mail: j.guillemaut@surrey.ac.uk.*

Manuscript received 24 Nov. 2016; revised 13 June 2017; accepted 29 Aug. 2017. Date of publication 21 Sept. 2017; date of current version 13 Aug. 2018.
(Corresponding author: Nadejda Roubtsova.)
Recommended for acceptance by Y. Matsushita.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2017.2749373

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/

those hard-coding more complex models in the CS constraint (e.g., Ward's model in [21]) and the implementations of joint shape/reflectance estimation [22]. Interest-point-based methods facilitate only a sparse reconstruction. Dense non-Lambertian CS, although conceptually interesting, fails to deliver accurate high resolution results. CS describes the surface as a depth map only and, for a limited number of views and more general scenes, lacks in structural resolution of the globally accurate shape.

Structured light methods tackle texture dependence of CS by projecting a known pattern. Fusion of these techniques with diffuse/specular component separation enables pore-level resolution geometry recovery on top of the globally accurate shape facilitated by structured light [23], [24]. These methods typically require the specialist Light Stage for dense reflectance sampling and are limited to non-metallic surfaces relying on light polarisation on reflection.

In contrast, KinectFusion is a structured light implementation using commodity hardware [25], based on projecting infra-red structured light patterns for instantaneous depth map acquisition. While the real-time feedback is impressive, the obtained geometries lack resolution and require fusion of several depth maps to become complete making KinectFusion unable to reconstruct dynamic scenes without supporting algorithms. Building on efficient depth-normal fusion for surface integration in [26], to improve resolution there has been successful research on combining Kinect with photometric stereo [27], [28] and shape-from-shading [29], [30], often under uncalibrated illumination but with the assumption of Lambertian reflectance.

Photometric stereo (PS) constraints [7] are derived from a point's response to varying illumination at a *constant* viewpoint avoiding the feature point matching problem. PS seeks to reconcile intensity measurements and predictions determined by the BRDF choice. The technique is mature with state-of-the-art multiview variants [8], [9] and extensions to dynamic non-rigid full-3D scenes [10] including surfaces with both uniform albedo and texture [11], [12].

In PS the reflectance can be an arbitrary parametric model, yet it must be a priori known to formulate predictions. The reflectance dependence of CS is hence not solved by photometric stereo but rather the burden is shifted to reflectance estimation which is often impossible to the desired precision for real surfaces. Hence the Lambertian assumption is wide-spread in PS. PS for non-Lambertian reflectance is an active research direction [31]. PS by example [32] recovers unknown geometry by relating its intensity responses under different illumination to a reference geometry object of the same material. The method pioneers surface reconstruction with anisotropic reflectance given the limitations of needing the *a priori* knowledge of the material or the assumption of it being a linear hybrid of the chosen diffuse and specular references. Other PS methods exploit fundamental BRDF symmetries to generalise to arbitrary reflectance. Alldrin and Kriegman [33] recover gradients assuming reflectance isotropy only, without restricting to a specific model, and advance no further than the gradient-defined iso-contour structure. The disambiguation of the Euclidean structure by ordering the iso-depth contours requires additional isotropic symmetries [34]. Tan et al. [34] study radiometric constraints from the symmetries observed under directional illumination of a curved surface given uniformly isotropic BRDF. PS ambiguities (e.g., the generalised bas-relief ambiguity) are resolved as the linear transform mapping between photometrically equivalent normal fields breaks the symmetry constraint. Chandraker et al. [35] factorise differential image formation equations into geometry and BRDF-dependent terms deriving a determinant condition valid regardless of the functional form of the BRDF. The PS problem is made well-constrained by assuming isotropic BRDF. BRDF isotropy is also the only reflectance constraint in [36] which combines reconstruction by structure-from-motion with PS by propagating sparse stereo estimates along iso-depth contours. Shadows as another BRDF invariant are explored by Okabe et al. [37] who encode reference points by shadow response to varying illumination and assign normals based on code similarity under the assumption of uniformly distributed lights and a convex surface. As another invariant, Shi et al. [38] enforce monotonicity of reflectance fall-off from the axis of specular lobe optimising for normal elevation given its azimuthal angle: the geodesic distance between per-pixel intensity profiles under uniform directional illumination is linearly related to the normal angular difference via a material dependent coefficient derivable from the profile intensity distribution. The methods provide robustness to non-Lambertian behaviour, but not universally as all except [37] rely on common, yet not generic, isotropy.

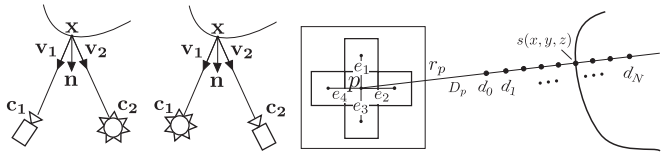

Fig. 1. Left: Reciprocal pair capture; Right: Depth sampling for virtual camera pixel p in a 4-connected neighbourhood.

With known BRDF, PS outputs descriptive fine detail normals. Direct normal field integration is however prone to drift with the resultant global shape distortion. Hybrid systems fusing multiview CS for global shape and photometric stereo for resolution have proven effective [10], [39]. The fusion can be a weakness as failure of either negatively impacts the whole and complex non-Lambertian reflectance is an unsolved challenge for both CS and PS.

Helmholtz Stereopsis [14] is photo-geometric [40] utilising both changing viewpoint and illumination. BRDF modelling is bypassed by enforcing consistency of *reflectance-model-independent* Helmholtz reciprocity: a light ray and its reverse will undergo identical optical processes [41]. Let $\mathbf{v}_1$ be the unit vector directed from the surface point to the camera and $\mathbf{v}_2$ the corresponding vector from the surface point to the light source (Fig. 1, left). Its implication, first observed by Zickler et al. [14] in the context of multiview reconstruction, is that interchanging the light source and camera in the set-up has no effect on the point's reflective behaviour. The BRDF f_r is reciprocal: $f_r(\mathbf{v_2}, \mathbf{v_1}) = f_r(\mathbf{v_1}, \mathbf{v_2})$. The image formation equation for image I_1 in the reciprocal pair $i_1 = f_r(\mathbf{v_2}, \mathbf{v_1}) \frac{\mathbf{n} \cdot \mathbf{v}_2}{\|\mathbf{c_2} - \mathbf{x}\|^2}$ expresses intensity i_1 of surface point $\mathbf{x}$ as a function of BRDF, surface normal $\mathbf{n}$, the two reciprocal unit vectors and the camera/light source positions $\mathbf{c}_1/\mathbf{c}_2$ (analogously for I_2 by interchanging the indices). Reciprocity of BRDF in conjunction with the image formation equations result in constraint $\mathbf{w}$ *without* any dependence on the BRDF:

$$\left(i_1 \frac{\mathbf{v_1}}{\|\mathbf{c_1} - \mathbf{x}\|^2} - i_2 \frac{\mathbf{v_2}}{\|\mathbf{c_2} - \mathbf{x}\|^2} \right) \cdot \mathbf{n} = \mathbf{w} \cdot \mathbf{n} = 0. \tag{1}$$

With one $\mathbf{w}$ per reciprocal pair, 3 or more reciprocal pairs result in constraint matrix W suitable for singular value decomposition (SVD): $SVD(W) = U\Sigma V^\top$ where U, V are orthogonal and Σ is a diagonal matrix. The last column of V gives the normal at the sampled point. The last diagonal value of Σ, the SVD residual σ_3, tends to 0 when there is mutual constraint consistency. For outlier elimination, Zickler et al. formulate the consistency measure as the quotient $\frac{\sigma 2}{\sigma 3}$, which tends to infinity for true surface points. HS has recently been extended to dynamic scenes by making use of wavelength multiplexing for simultaneous acquisition of the reciprocal pair triplet [42].

Standard HS is promising to tackle photometric complexity but with shortcomings. First, depth label estimation at each point is *maximum likelihood*, computed independently of its neighbours. The approach results in lack of smoothness and detail due to the inherent depth ambiguities, measurement noise, sensor saturation and calibration/discretisation errors. Further, the unique ability of HS to characterise surface points by both depths and normals is

exploited at neither the estimation nor integration stage of the standard pipeline: the consistency of depth and normal estimates is not verified at estimation or during direct normal field integration despite its proneness to drift.

In this paper the shortcomings are addressed by proposing a novel framework. First, neighbourhood support is introduced to enforce depth map continuity: the ML formulation is replaced by a *maximum-a posteriori* one, solvable for the first time in the context of HS by numerous mature MRF optimisation techniques. The MAP formulation is the core of our Bayesian HS. Addressing the unexploited correlation of depths and normals, we formulate a tailored prior for Bayesian HS to enforce consistency and comparable accuracy levels of depths and normals. This revises our original depth-normal consistency prior formulation in [15], [16], which was *correlation-based*, seeking similarity between the depth transition normal and photometric normal projections. Here we propose a more principled *distance-based* depth-normal consistency prior formulation, enforcing consistency between the tangential plane to the photometric normal and the depth transition plane. Directly related to surface integrability [17], the new distance-based formulation is shown theoretically superior to the correlation-based one. The prior's superiority, against former formulations and the classical priors, is thoroughly verified experimentally. To address drift-prone integration of standard HS, we embed Bayesian HS into a coarse-to-fine framework to achieve resolutions rendering surface integration redundant. In the integration-free pipeline the full potential of the proposed prior for detail resolution is showcased in contrast to pipelines with direct integration back-ends [26], [43], [44].

3 Methodology

3.1 Problem Statement and Notation

Our work addresses dense surface recovery in 3D. Let us introduce a virtual camera projecting 3D points $s(x, y, z)$ onto pixel p via projection ray r_p: $p = P(s(x, y, z))$. With the point's depth d_p along r_p, pixel p can be back-projected onto $s(x, y, z)$ using $\mathbf{P_{back}}$: $s(x, y, z) = \mathbf{P_{back}}(p, d_p)$. The virtual camera, perspective or orthographic, defines the viewpoint for the 2.5D reconstruction.

Each pixel p represents a random variable D_p - a measure of depth for the visible surface points projecting to p. The surface reconstruction problem is formulated as a labelling problem where depth label d_p is assigned to each random variable D_p of pixel p. For each p there is a set of N depth hypotheses $\{d_1, \ldots, d_N\}$, possible values for D_p. The set of hypotheses is obtained by sampling the reconstruction volume V along r_p (Fig. 1, right). If the sampling is orthographic as in our work, the framework is analogous to the voxel representation where a surface is embedded in a volume V of $N_X \times N_Y \times N_Z$ discrete voxels $v(x, y, z)$ sampled at resolution $\delta x \times \delta y \times \delta z$. Let us define set $\mathcal{F}$ to be all virtual camera pixels. The solution to the defined labelling problem is the label configuration $d = \{d_p \mid \forall p \in \mathcal{F}\}$ where $d \in S$ with S being the set of all possible label configurations.

The problem lends itself to representation as a Markov Random Field (MRF). In the MRF graph $G = (\mathcal{F}, \mathcal{E})$, each virtual camera pixel $p \in \mathcal{F}$ is a node. The nodes are connected by edges $e \in \mathcal{E}$ to neighbouring nodes modelling spatial dependencies (Fig. 1, right). These dependencies define the prior probability distribution of the framework's state variables D_p. Each D_p individually is also characterised by a plausibility distribution over its set of depth hypotheses $\{d_1, \ldots, d_N\}$, based on local observation. In the MRF, the local observation gives rise to the data term whereas the prior distribution defines the smoothness term.

3.2 Helmholtz Stereopsis

Let us now look at HS in the context of the adopted notation. To assign label d_p to each virtual camera pixel p, 3D points $s(x, y, z)$ along r_p are sampled. Only the projection rays intersecting with the visual hull are considered. Each sampled point is defined by its position along r_p i.e., the depth hypothesis value d. The depth hypothesis set $\{d_1, \ldots, d_N\}$ for random variable D_p, confined to within the visual hull, is thus accumulated with an arbitrary sampling resolution.

In HS, each $s(x, y, z)$ along r_p is sampled by projection onto reciprocal images to acquire a set of N_w intensity 2-tuples $\{(i_1, i_2)_1, \ldots (i_1, i_2)_{N_w}\}$ and formulate N_w constraints $\mathbf{w}$ as in (1). Only the $s(x, y, z)$ at the optimal depth d_p^* will have N_w mutually consistent constraints. The consistency measure, i.e., the local observation of D_p, is distributed over $\{d_1, \ldots, d_N\}$ as defined by the SVD-residual-based coefficient $\frac{\sigma_2}{\sigma_3}$. Hypothesis d_p has the likelihood $E_{data}(p, d_p)$ defined by the coefficient $\frac{\sigma_2(\mathbf{P_{back}}(p,d_p))}{\sigma_3(\mathbf{P_{back}}(p,d_p))}$ of $s(x, y, z)$ that p back-projects to at d_p along r_p. The coefficient tends to infinity as d_p approaches the correct depth d_p^*. For compatibility with MRF minimisation, $E_{data}(p, d_p)$ is formulated as a decaying function of $\frac{\sigma_2(\mathbf{P_{back}}(p,d_p))}{\sigma_3(\mathbf{P_{back}}(p,d_p))}$ with the factor $\mu = 0.2\ln(2)$ and bounded in the range $[0, 1]$:

$$E_{data}(p, d_p) = e^{-\mu \times \frac{\sigma_2(\mathbf{P_{back}}(p,d_p))}{\sigma_3(\mathbf{P_{back}}(p,d_p))}}. \tag{2}$$

Standard HS solves a non-Markovian *maximum likelihood* problem optimising each D_p independently, based solely on the data term without involving the prior distribution from spatial dependencies:

$$d_{ML}^* = \arg\min_{d \in S} \sum_{p \in \mathcal{F}} E_{data}(p, d_p). \tag{3}$$

The resultant solution is sub-optimal which leads to noisy depth maps as well as lacking surface smoothness, structural detail and global shape accuracy.

3.3 Bayesian Helmholtz Stereopsis

Instead of the sub-optimal ML, we propose to formulate the labelling problem as a *maximum a posteriori* optimisation. In this formulation, for pairs of neighbouring pixels p and q, in addition to the respective data costs $E_{data}(p, d_p)$ and $E_{data}(q, d_q)$ (Section 3.2), a smoothness cost $E_s(p, d_p, q, d_q)$ is defined. The smoothness cost is computed for each $(p, q) \in \mathcal{E}$, given the 4-connected neighbourhood abstraction. With α balancing the relative data and prior contributions, the MAP solution is

$$\begin{aligned} d_{MAP}^* = \arg\min_{d \in S} & \sum_{p \in \mathcal{F}} ((1 - \alpha) E_{data}(p, d_p)) \\ & + \sum_{(p,q) \in \mathcal{E}} \alpha E_s(p, d_p, q, d_q)). \end{aligned} \tag{4}$$

A Bayesian framework is more suitable because of the strong statistical dependency between neighbouring depth estimates. Compared to standard HS, Bayesian formulation in (4) produces cleaner depth maps improving accuracy by correct normal indexing. As $E_{data}(p, d_p)$ has been defined in Section 3.2, we now focus on the formulation and comparison of several priors as candidates for $E_s(p, d_p, q, d_q)$.

3.3.1 Depth-Based Prior (Dprior)

The prior is known from conventional stereo. We define the depth-based smoothness cost $E_{s\_d}(p, d_p, q, d_q)$ for neighbouring pixels p and q (Fig. 2a) as the discontinuity-preserving truncated squared difference of their respective depth labels d_p and d_q

$$E_{s\_d}(p, d_p, q, d_q) = \min(E_{s\_d}^{max}, (d_p - d_q)^2), \tag{5}$$

where the truncation value $E_{s\_d}^{max}$ of half the reconstruction volume squared weakly moderates depth penalties. Through penalties for large label jumps while disregarding the available normals, the

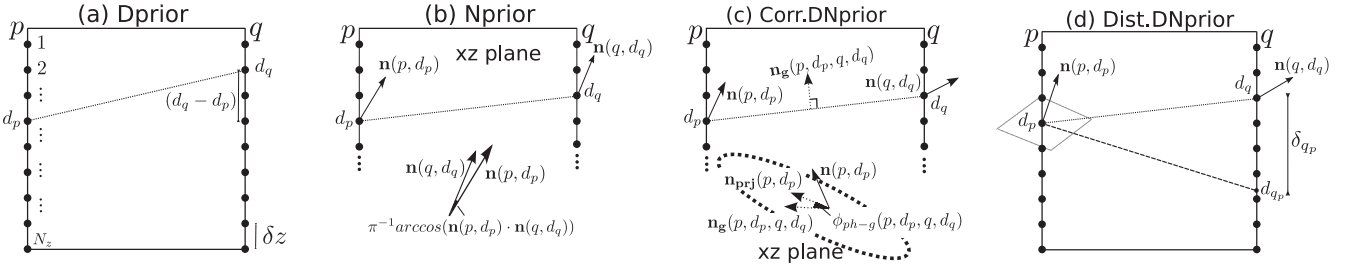

Fig. 2. Schematic representation of the priors for two laterally neighbouring pixels p and q in the xz plane.

prior encourages piece-wise constant depth biasing towards a fronto-parallel representation, particularly if unchecked by a low α.

3.3.2 Normal-Based Prior (Nprior)

Surface characterisation by normals is typical of photometric stereo. A suitable normal-based prior would enforce locally constant normals encouraging locally flat, though not necessarily fronto-parallel, surfaces and making Nprior in theory less restrictive of reconstructed surfaces than Dprior.

We define a normal-based prior where similarity of corresponding normals is used to assess neighbouring label compatibility with discrete depth hypotheses as the labels. Let $\mathbf{n}(p, d_p)$ be the normal estimate associated with depth hypothesis d_p, i.e., the normal vector estimated by HS at 3D point $\mathbf{P_{back}}(p, d_p)$ which is the back-projection of pixel p at d_p along r_p. Given photometric normals $\mathbf{n}(p, d_p)$ and $\mathbf{n}(q, d_q)$ corresponding to the labels of neighbouring pixels p and q (Fig. 2b), Nprior can be formulated as follows:

$$E_{s\_n}(p, d_p, q, d_q) = \big(\pi^{-1} \arccos\big(\, \mathbf{n}(p, d_p) \cdot \mathbf{n}(q, d_q) \,\big)\big)^2, \tag{6}$$

which is the squared normalised correlation angle between normals, notably not making use of depth information. Complications in using normal-based priors arise because normals are susceptible to noise and their correlations are irregular expressions not optimisable by graph cuts [45]. Instead, sequential tree re-weighted message passing (TRW-S) [46], [47] is used for MRF optimisation consistently in the paper because it does not require regularity of prior.

3.3.3 Depth-Normal Consistency Prior (DNprior)

Where Dprior seeks to de-noise *depth maps* by enforcing their smoothness, Nprior promotes gradual spatial *normal field* evolution. Both approaches are *one-sided*: the depth is optimised indexing the normals or vice versa. Depth and normal estimation processes are however not independent and must be consistent with each other. We postulate that a superior prior in the context of HS explicitly enforces consistency between depths and normals, performing *joint* optimisation of the depth map and normal field. Hence, we formulate the depth-normal consistency prior (DNprior).

DNprior has had two formulations: a correlation-based and a distance-based one. The superiority of the latter over the former will be argued both theoretically and experimentally.

Correlation-based DNprior (corr.DNprior) based on our previous work [15], [16] is formulated from the normalised correlation angle between the geometric normal $\mathbf{n_g}(p, d_p, q, d_q)$ to the $d_p - d_q$ depth transition from p to q and the normalised projections of the estimated photometric normals $\mathbf{n}(p, d_p)$ and $\mathbf{n}(q, d_q)$ onto the plane embedding the $d_p - d_q$ transition. These projections are denoted $\mathbf{n_{prj}}(p, d_p)$ and $\mathbf{n_{prj}}(q, d_q)$ respectively (Fig. 2c). We work with an orthographic virtual camera and 4-connected pixel neighbourhoods which means that p and q are lateral neighbours. If the depth transition is correct, $\mathbf{n_g}(p, d_p, q, d_q)$ correlates well with $\mathbf{n_{prj}}(p, d_p)$ and $\mathbf{n_{prj}}(q, d_q)$. The smoothness cost of assigning d_p to pixel p given q at d_q is expressed via the correlation angle ($\phi_{ph-g}(q, d_q, p, d_p)$ is analogous):

$$\begin{aligned}&\phi_{ph-g}(p, d_p, q, d_q)\\ &\quad = \pi^{-1} \arccos(\mathbf{n_{prj}}(p, d_p) \cdot \mathbf{n_g}(p, d_p, q, d_q)).\end{aligned} \tag{7}$$

The orientation of $\mathbf{n_g}$ must be out of the surface. The prior's end smoothness cost $E^{corr}_{s\_dn}(p, d_p, q, d_q)$ with l_2-norm is

$$\begin{aligned}&E^{corr}_{s\_dn}(p, d_p, q, d_q)\\ &\quad = \frac{1}{2}\Big(\big(\phi_{ph-g}(p, d_p, q, d_q)\big)^2 + \big(\phi_{ph-g}(q, d_q, p, d_p)\big)^2\Big).\end{aligned} \tag{8}$$

Distance-based DNprior (dist.DNprior) is an alternative formulation of depth-normal consistency derived from fundamental perpendicularity of the normal to the surface. The estimated photometric normal $\mathbf{n}(p, d_p)$ of pixel p at depth d_p suggests a surface transition from the back-projection $\mathbf{P_{back}}(p, d_p)$. Considering one direction for simplicity, we observe that according to $\mathbf{n}(p, d_p)$ the surface will continue from $\mathbf{P_{back}}(p, d_p)$ to some $\mathbf{P_{back}}(q, d_{q_p})$. Note that d_{q_p} does not have to be part of the discrete label set of q.

By definition, $\mathbf{n}(p, d_p)$ is perpendicular to the transition plane from $\mathbf{P_{back}}(p, d_p)$ to $\mathbf{P_{back}}(q, d_{q_p})$:

$$(\mathbf{P_{back}}(q, d_{q_p}) - \mathbf{P_{back}}(p, d_p)) \cdot \mathbf{n}(p, d_p) = 0. \tag{9}$$

The corresponding back-projections $\mathbf{P_{back}}(q, d_{q_p})$ and $\mathbf{P_{back}}(q, d_q)$ deviate only in depth:

$$\mathbf{P_{back}}(q, d_{q_p}) = \mathbf{P_{back}}(q, d_q) + [0, 0, \delta_{q_p}]^\top, \tag{10}$$

where δ_{q_p} is the depth discrepancy between d_q and d_{q_p}. This δ_{q_p} is also the error metric decreasing as d_{p_q} suggested by $\mathbf{n}(p, d_p)$ approaches the assigned d_q, with the growing confidence of d_p and d_q being compatible labels for p and q respectively. If (10) is substituted into (9) we can re-arrange given that δ_{q_p} is along the depth axis (i.e., ray r_p):

$$(\mathbf{P_{back}}(q, d_q) - \mathbf{P_{back}}(p, d_p)) \cdot \mathbf{n}(p, d_p) + \delta_{q_p} n_{r_p}(p, d_p) = 0, \tag{11}$$

where $n_{r_p}(p, d_p)$ is the depth component of the photometric normal along r_p. The following expression defines depth discrepancy δ_{q_p} between d_q and d_{q_p} (Fig. 2d):

$$\delta_{q_p} = \frac{(\mathbf{P_{back}}(q, d_q) - \mathbf{P_{back}}(p, d_p)) \cdot \mathbf{n}(p, d_p)}{n_{r_p}(p, d_p)}. \tag{12}$$

Our new distance-based formulation of DNprior minimises the depth discrepancy of (12) in the energy function

$$E^{dist}_{s\_dn}(p, d_p, q, d_q) = \frac{1}{2}\big((\delta_{q_p})^2 + (\delta_{p_q})^2\big). \tag{13}$$

Every edge $(p, q) \in \mathcal{E}$ is characterised by the symmetrical pair of equally contributing discrepancy scores, δ_{q_p} and δ_{p_q}. Though not an integral part of formulation, truncation of $\delta_{q_p}/\delta_{p_q}$ may be necessary for some reconstruction scenarios e.g., fine initial spatial resolution (see Section 4).

The physical meaning of dist.DNprior is intuitively easier to grasp than the correlation of corr.DNprior. Its formulation is

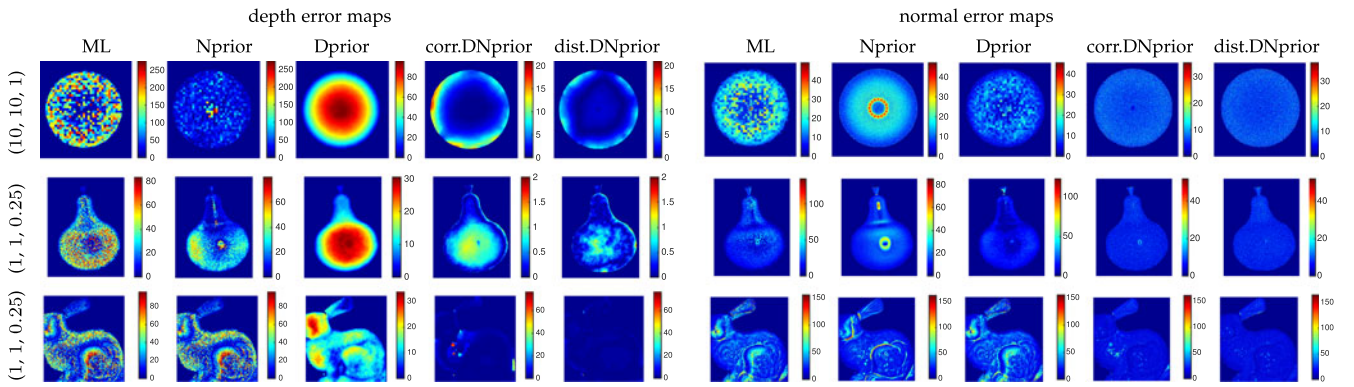

Fig. 3. Depth and normal error for reconstructions from noisy synthetic data. Initial sampling [mm] (δx, δy, δz) as indicated.

plausible being derived from the surface normal axiom and is consistent with the well-established depth priors. In the following section, we relate the distance-based formulation to integrability as theoretical grounds to support intuition in showing it to be more principled.

3.3.4 Relation to Integrability

We argue superiority of dist.DNprior on the grounds of its smoothness function $E^{dist}_{s\_dn}(p, d_p, q, d_q)$ being equivalent to the surface integrability constraint. Integrability is the least restrictive prior effective unless the problem is under-constrained [17],[48]. Unlike priors biasing towards fronto-parallel (Dprior) or locally flat (Nprior) surfaces, the integrability constraint is essentially just a check of mathematical plausibility and will only definitively bias against rare extremes (e.g., Dirac peaks, step-like transitions).

The integrability constraint forms the *height-from-gradient* recovery framework from the classical paper by Horn [17]. In [17], surface gradient (g_x, g_y) is the derivative of surface $z(x, y)$: $g_x = \frac{\delta z}{\delta x}$ and $g_y = \frac{\delta z}{\delta y}$ The surface normal is related to the gradient as $\mathbf{n} = (1 + g_x^2 + g_y^2)^{-\frac{1}{2}}(-g_x, -g_y, 1)^\top$. The surface gradient can be related to the unit surface normal via: $g_x = -\frac{n_x}{n_z}$ and $g_y = -\frac{n_y}{n_z}$. The following cost function formulated by Horn is a joint optimisation of surface height $z(x, y)$ and normal over the entire surface:

$$\int\int \left(\left(\frac{\delta z}{\delta x} - g_x \right)^2 + \left(\frac{\delta z}{\delta y} - g_y \right)^2 \right) \mathrm{d}x \; \mathrm{d}y. \tag{14}$$

Clearly, the cost function encapsulates the idea of depth-normal consistency of our DNprior. For clarity, let us consider surface evolution along the x dimension. Numerical approximation of gradient discretises the derivative to: $\frac{\delta z}{\delta x} = \frac{z_2 - z_1}{\delta x}$ where $z_1 = z(x, y)$ and $z_2 = z(x + \delta x, y)$. In the terminology of our priors z_1 and z_2 are depth labels d_p and d_q of neighbouring p and q. Using gradient discretisation, let us formulate a symmetrical, with respect to neighbouring points, integrability prior function E^{Horn}_{dn} based on (14)

$$E^{Horn}_{dn} = \frac{1}{2} \left(\left(\frac{z_2 - z_1}{\delta x} - g_{x,1} \right)^2 + \left(\frac{z_2 - z_1}{\delta x} - g_{x,2} \right)^2 \right), \tag{15}$$

where $g_{x,1}$ and $g_{x,2}$ are the x gradients of neighbouring $s_1(x, y)$ and $s_2(x + \delta x, y)$ expressed via normal components.

Let us now relate dist.DNprior to the integrability prior E^{Horn}_{dn}. Depths d_{q_p} and d_{p_q} become z_{21} and z_{12} i.e., the neighbouring depth labels suggested by the normal-imposed gradients: $z_{21} = z_1 + \delta x g_{x,1}$ and $z_{12} = z_2 - \delta x g_{x,2}$. Substituting the definitions, the corresponding depth discrepancies $\delta_{q_p} = \delta_{21}$ and $\delta_{p_q} = \delta_{12}$ are: $\delta_{21} = (z_{21} - z_2) = (z_1 + \delta x g_{x,1} - z_2)$ and $\delta_{12} = (z_{12} - z_1) = (z_2 - \delta x g_{x,2} - z_1)$. Hence, the distance-based DNprior cost from (13) becomes

$$E^{dist}_{dn} = \frac{1}{2}((z_1 + \delta x g_{x,1} - z_2)^2 + (z_2 - \delta x g_{x,2} - z_1)^2). \tag{16}$$

The equation can be manipulated into the following:

$$E^{dist}_{dn} = E^{Horn}_{dn} \delta x^2. \tag{17}$$

Dist.DNprior in (17) boils down to Horn's surface integrability in (15) at a given resolution. This direct equivalence to the least restrictive existing prior supports the claim of the distance-based DNprior being a principled formulation. No such relation to integrability can be derived for the correlation-based formulation (see supplementary material, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2017.2749373).

3.4 Coarse-to-Fine Integration-Free Pipeline

Bayesian HS with integrability prior facilitates previously unattainable accuracy. For matching resolution given finite memory, the system is embedded into a coarse-to-fine framework with gradual sub-sampling both spatially (up to pixel resolution, e.g., 0.25 mm in Section 4) and in depth. Combined high accuracy and resolution of the point cloud de-emphasise explicit integration: instead vertices are arranged into facets based on known geometric relationships within the voxel volume.

4 Results

Synthetic data with specular reflectance is used to tune parameter α from (4) per prior/dataset (see the supplementary materials, available online) and to quantitatively assess the depth/normal accuracy of each reconstruction method at the best α setting on noiseless and noisy data (Gaussian noise, normalised variance of 0.001 or ± 2072 intensity levels). Based on the Middlebury accuracy [2] for depth and normal in Table 1 and the corresponding results in Fig. 3, both DNprior formulations outperform one-sided priors and standard ("ML") HS by at least an order of magnitude. The proposed integrability promoting dist.DNprior is more robust to noise than the earlier corr.DNprior resulting in depth accuracy improvements from just under 0.5 mm (pear) to 2-4 mm (bunny/sphere), though a similar normal error. The two formulations perform comparably on noiseless data. Dist.DNprior costs were truncated on bunny/pear at just under 300 percent of initial spatial sampling barring transitions with excessive elevation.

Real Data: Reconstruction Method. We qualitatively evaluate the performance of the five reconstruction methods on eight real datasets from [49]: the three specular (teapots /vase) and one Lambertian with intricate geometry (doll) here as well as four others of varying complexity in the supplementary material, available online. Parameter α is set per prior within the average optimal range tuned on synthetic data. In Fig. 4, Nprior corrupts shape while keeping ML HS noise. Dprior also distorts the global shape whereas corr.

TABLE 1
Middlebury Accuracy [mm, °] and Completeness [%] at 90 Percent Threshold in Terms of Depth (**D**) and Normal (**N**) Error on Synthetic Datasets for ML HS and Bayesian HS with α (Sphere, Pear, Bunny) Set as Follows per Prior: Nprior (1, 1, 0.1), Dprior (0.3, 0.4, 0.03), corr.DNprior (1, 1, 1), dist.DNprior (1, 0.9, 1)

		no noise					noisy data				
		ML	Nprior	Dprior	corr.DNprior	dist.DNprior	ML	Nprior	Dprior	corr.DNprior	dist.DNprior
sphere	D	28.29/**64.5**	20.82/64.1	86.17/**64.5**	0.55/64.2	**0.46**/64.1	200.52/61.0	62.25/63.7	86.82 /**64.4**	9.42/64.2	**5.71**/64.3
	N	6.55/**65.1**	5.75/64.7	15.13/**65.1**	**0.35**/64.8	0.37/64.7	25.22/61.6	13.67/64.3	20.94/**65.0**	**11.69**/64.8	11.87/64.9
pear	D	15.20/82.2	23.16/77.2	24.97/**83.1**	**0.29**/**83.1**	0.39/**83.1**	62.63/76.3	38.36/80.7	27.69/**83.1**	0.97/**83.1**	**0.70**/**83.1**
	N	12.62/82.9	41.85/77.8	17.96/**83.8**	**0.49**/**83.8**	0.57/**83.8**	31.78/76.9	19.32/81.3	23.09/**83.8**	**11.64**/**83.8**	12.09/**83.8**
bunny	D	6.50/84.5	5.78/84.5	4.18/84.7	**0.59**/ 84.7	0.74/**84.8**	63.34/81.0	63.29/81.1	19.88/84.5	4.50/**84.7**	**1.99**/**84.7**
	N	15.10/85.0	13.65/85.0	11.72/85.2	**4.98**/85.2	5.11/**85.3**	58.91/81.5	57.66/81.6	48.90/85.0	27.60/**85.2**	**27.00**/**85.2**

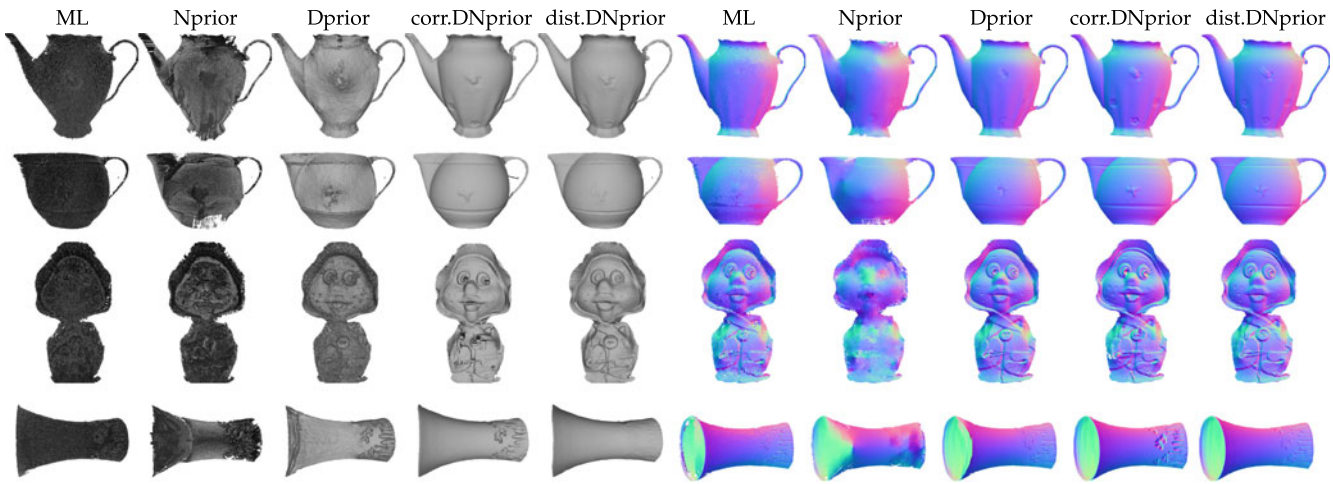

Fig. 4. Reconstructions and normals from real data; initial sampling [mm]: (1, 1, 0.25); α (left to right): (1, 0.05, 1, 0.8).

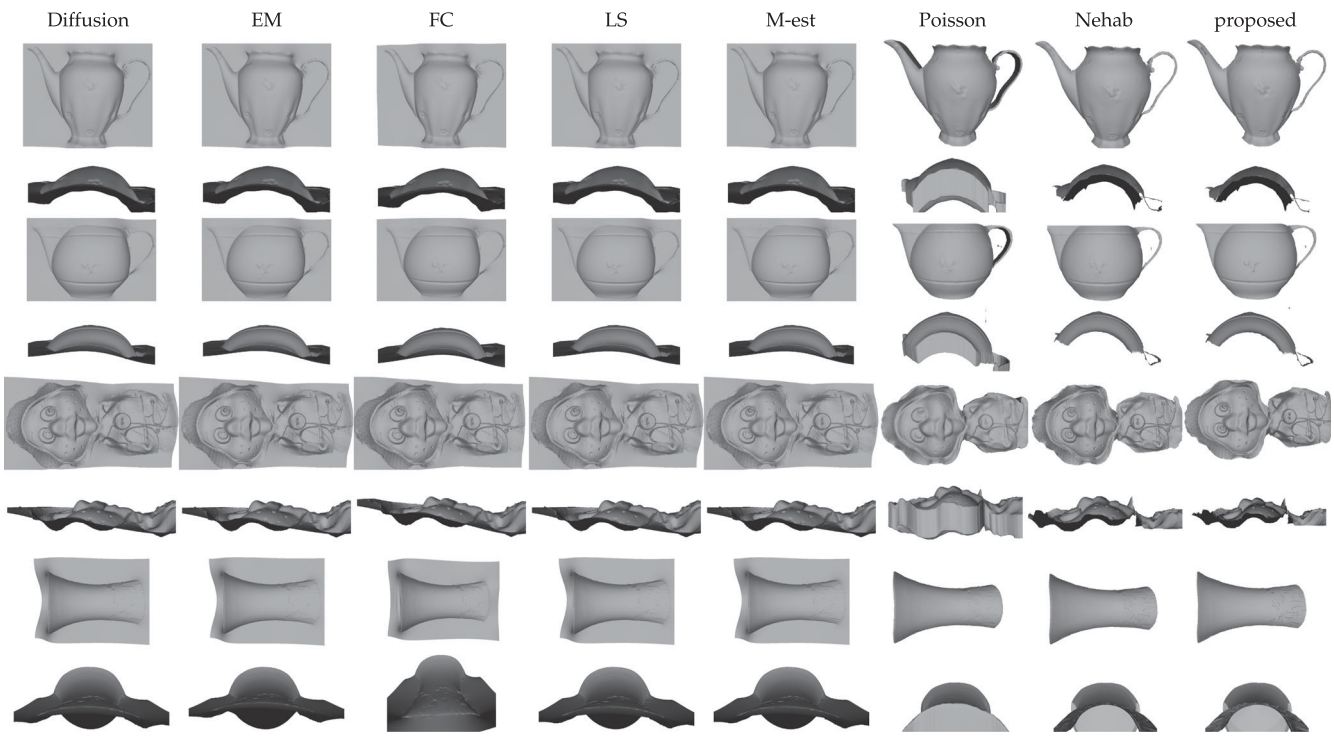

Fig. 5. Integration using diffusion tensor, energy minimisation, Frankot-Chellappa, least squares, M-estimator[43], Poisson [44], Nehab [26] and no explicit integration (proposed) of the point clouds obtained by Bayesian HS with dist.DNprior.

DNprior shows a tendency for depth jumps (doll, vase). Only the proposed dist.DNprior reconstructs globally accurate high resolution meshes. The pseudo-geometry artefact in textured regions can be lessened by intensity averaging over patches [16].

Real Data: Integration Method. The point cloud generated by Bayesian HS with dist.DNprior can be converted into a continuous surface without explicit integration. The accuracy of the integration-free approach is compared in Fig. 5 against five normal integration methods from [43] and two based on depth-normal fusion (Poisson[1] [44], Nehab[2] [26]). Normal integration over-flattens and distorts global shape unlike the fusion methods while Poisson over-smooths (doll). The subtle detail improvement without integration over Nehab is shown in the supplementary material, available online.

5 Conclusion and Future Work

We have presented Bayesian HS with integrability prior without explicit surface integration for high quality reconstruction of geometrically and photometrically complex objects. Three contributions define the method. First, standard ML formulation of HS is replaced by a novel MAP (Bayesian) one combining local observation with neighbourhood support. Second, to utilise the neighbourhood, a suitable prior enforcing consistency between depth and normal information is proposed. The prior is tailored to the unique ability of HS to generate both depth and normals. The distance-based formulation of the consistency prior related to integrability has been shown superior to the correlation-based formulation, one-sided depth and normal priors as well as standard HS. With the least-restrictive regularisation basis of integrability defining the prior, the results show accuracy, high resolution and robustness against intensity noise. Third, facilitated by the accuracy of prior and the coarse-to-fine framework for previously unattainable point cloud densities, the final surface is assembled without integration to avoid shape distortion and oversmoothing.

The integrability prior seeks mathematical surface continuity and cannot cope with infinite transitions. The difficulty with such discontinuities can be addressed by locally increasing the sampling resolution. Yet, as this pertains to outlier cases only, the coarse-to-fine integration-free Bayesian HS with integrability prior is already widely applicable for dense reconstruction with complex surface reflectance properties and geometries.

Acknowledgments

We thank our industrial partners *The Foundry* and *Double Negative* for useful discussions. This work was supported by the *EPSRC* (grant EP/M021793/1). Our synthetic data is available at https://doi.org/10.15126/surreydata.00841369. Source of 3D models used (pear/bunny): *Suggestive Contour model* database, http://gfx.cs.princeton.edu/proj/sugcon/models/ and *Stanford 3D scanning* repository, https://graphics.stanford.edu/data/3Dscanrep/ respectively.

References

[1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1–3, pp. 7–42, 2002.
[2] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 519–528.
[3] J. Liu, C. Li, F. Mei, and Z. Wang, "3D entity-based stereo matching with ground control points and joint second order smoothness prior," *Vis. Comput.*, vol. 31, pp. 1–17, 2014.
[4] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.
[5] J.-Y. Guillemaut and A. Hilton, "Joint multi-layer segmentation and reconstruction for free-viewpoint video applications," *Int. J. Comput. Vis.*, vol. 93, no. 1, pp. 73–100, 2011.
[6] J. Starck and A. Hilton, "Surface capture for performance-based animation," *IEEE Comput. Graph. Appl.*, vol. 27, no. 3, pp. 21–31, May/Jun. 2007.
[7] R. J. Woodham, *Shape from Shading*. Cambridge, MA, USA: MIT Press, 1989, pp. 513–531.
[8] G. Vogiatzis, C. Hernández, and R. Cipolla, "Reconstruction in the round using photometric normals and silhouettes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 1847–1854.
[9] C. Hernández, G. Vogiatzis, and R. Cipolla, "Multiview photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 3, pp. 548–554, Mar. 2008.
[10] D. Vlasic, et al., "Dynamic shape capture using multi-view photometric stereo," in *Proc. SIGGRAPH Asia*, 2009, Art. no. 174.
[11] G. Brostow, C. Hernández, G. Vogiatzis, B. Stenger, and R. Cipolla, "Video normals from colored lights," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 2104–2114, Oct. 2011.
[12] Z. Jankó, A. Delaunoy, and E. Prados, "Colour dynamic photometric stereo for textured surfaces," in *Proc. Asian Conf. Comput. Vis.*, 2010, pp. 55–66.
[13] B. Baumgart, "Geometric modeling for computer vision," Ph.D. dissertation, Dept. Comput. Sci., Univ. Stanford, Stanford, CA, USA, 1974.
[14] T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *Int. J. Comput. Vis.*, vol. 49, no. 2/3, pp. 215–227, 2002.
[15] N. Roubtsova and J.-Y. Guillemaut, "A Bayesian framework for enhanced geometric reconstruction of complex objects by Helmholtz stereopsis," in *Proc. 9th Int. Conf. Comput. Vis. Theory Appl.*, 2014, pp. 335–342.
[16] N. Roubtsova and J.-Y. Guillemaut, "Extended Bayesian Helmholtz stereopsis for enhanced geometric reconstruction of complex objects," (Eds.) S. Battiato, S. Coquillart, J. Pettré, R. Laramee, A. Kerren, J. Braz, Computer Vision, Imaging and Computer Graphics - Theory and Applications, Communications in Computer and Information Science, Springer, Cham, vol. 550, 2015, pp. 223–238.
[17] B. K. Horn, "Height and gradient from shading," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 37–75, 1990.
[18] A. Laurentini, "The visual hull concept for silhouette-based image understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 2, pp. 150–162, Feb. 1994.
[19] C. Liang and K.-Y. Wong, "3D reconstruction using silhouettes from unordered viewpoints," *Image Vis. Comput.*, vol. 28, no. 4, pp. 579–589, 2010.
[20] S. Agarwal, et al., "Building Rome in a day," *Comm. ACM*, vol. 54, no. 10, pp. 105–112, 2011.
[21] H. Jin, S. Soatto, and A. Yezzi, "Multi-view stereo reconstruction of dense shape and complex appearance," *Int. J. Comput. Vis.*, vol. 63, no. 3, pp. 175–189, 2005.
[22] G. Oxholm and K. Nishino, "Multiview shape and reflectance from natural illumination," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2163–2170.
[23] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec, "Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination," in *Proc. 18th Eurographics Conf. Rendering Techn.*, 2007, pp. 183–194.
[24] A. Ghosh, G. Fyffe, B. Tunwattanapong, J. Busch, X. Yu, and P. Debevec, "Multiview face capture using polarised spherical gradient illumination," *ACM Trans. Graph.*, vol. 30, no. 6, 2011, Art. no. 129.
[25] S. Izadi, et al., "KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera," in *Proc. Annu. ACM Symp. User Interface Softw. Technol.*, 2011, pp. 559–568.
[26] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi, "Efficiently combining positions and normals for precise 3D geometry," in *Proc. ACM SIGGRAPH*, 2005, pp. 536–543.
[27] Q. Zhang, M. Ye, R. Yang, Y. Matsushita, B. Wilburn, and H. Yu, "Edge-preserving photometric stereo via depth fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 2472–2479.
[28] S. M. Haque, A. Chatterjee, and V. M. Govindu, "High quality photometric reconstruction using a depth camera," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 2283–2290.
[29] Y. Han, J. Y. Lee, and I. S. Kweon, "High quality shape from a single RGB-D image under uncalibrated natural illumination," in *Proc. Int. Conf. Comput. Vis.*, 2013, pp. 1617–1624.
[30] L. F. Yu, S. K. Yeung, Y. W. Tai, and S. Lin, "Shading-based shape refinement of RGB-D images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 1415–1422.
[31] B. Shi, Z. Wu, Z. Mo, D. Duan, S. K. Yeung, and P. Tan, "A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3707–3716.
[32] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape reconstruction with general, varying BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1254–1264, Aug. 2005.
[33] N. G. Alldrin and D. J. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.
[34] P. Tan, L. Quan, and T. Zickler, "The geometry of reflectance symmetries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2506–2520, Dec. 2011.
[35] M. Chandraker, J. Bai, and R. Ramamoorthi, "On differential photometric reconstruction for unknown, isotropic BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2941–2955, Dec. 2013.

1. unscreened, with tree depth of 10 (teapots, vase) and 12 (doll)
2. with normal correction, 3 iterations, kernel size 2×median edge

[36] Z. Zhou, Z. Wu, and P. Tan, "Multi-view photometric stereo with spatially varying isotropic materials," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 1482–1489.
[37] T. Okabe, I. Sato, and Y. Sato, "Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 1693–1700.
[38] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 455–468.
[39] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I.-S. Kweon, "Multiview photometric stereo using planar mesh parameterization," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 1161–1168.
[40] T. Zickler, P. Belhumeur, and D. Kriegman, "Toward a stratification of Helmholtz stereopsis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2003, pp. I-548–I-555.
[41] H. Helmholtz, *Treatise on Physiological Optics*. New York, NY, USA: Dover, 1925.
[42] N. Roubtsova and J.-Y. Guillemaut, "Colour Helmholtz stereopsis for reconstruction of dynamic scenes with arbitrary reflectance," *Int. J. Comput. Vis.*, vol. 124, no. 1, pp. 18–48, 2017.
[43] A. Agrawal, R. Raskar, and R. Chellappa, "Surface reconstructions from a gradient field?" in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 578–591.
[44] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proc. Eurographics Symp. Geometry Process.*, 2006, pp. 61–70.
[45] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.
[46] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "MAP estimation via agreement on trees: Message-passing and linear-programming approaches," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3697–3717, Nov. 2005.
[47] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1568–1583, Oct. 2006.
[48] C. Hernández, G. Vogiatzis, and R. Cipolla, "Overcoming shadows in 3-source photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 419–426, Feb. 2011.
[49] J.-Y. Guillemaut, O. Drbohlav, R. Šára, and J. Illingworth, "Helmholtz stereopsis on rough and strongly textured surfaces," in *Proc. 2nd Int. Symp. 3D Data Process. Vis. Transmiss.*, 2004, pp. 10–17.