

Colour Helmholtz Stereopsis for Reconstruction of Complex Dynamic Scenes

Nadejda Roubtsova
Centre for Vision, Speech and Signal Processing
University of Surrey
Guildford, United Kingdom
n.roubtsova@surrey.ac.uk

Jean-Yves Guillemaut
Centre for Vision, Speech and Signal Processing
University of Surrey
Guildford, United Kingdom
j.guillemaut@surrey.ac.uk

Abstract—Helmholtz Stereopsis (HS) is a powerful technique for reconstruction of scenes with arbitrary reflectance properties. However, previous formulations have been limited to static objects due to the requirement to sequentially capture reciprocal image pairs (i.e. two images with the camera and light source positions mutually interchanged). In this paper, we propose colour HS - a novel variant of the technique based on wavelength multiplexing. To address the new set of challenges introduced by multispectral data acquisition, the proposed novel pipeline for colour HS uniquely combines a tailored photometric calibration for multiple camera/light source pairs, a novel procedure for surface chromaticity calibration and the state-of-the-art Bayesian HS suitable for reconstruction from a minimal number of reciprocal pairs. Experimental results including quantitative and qualitative evaluation demonstrate that the method is suitable for flexible (single-shot) reconstruction of static scenes and reconstruction of dynamic scenes with complex surface reflectance properties.

Keywords—Helmholtz Stereopsis; wavelength multiplexing; 3D geometry; dynamic scenes; complex reflectance;

I. INTRODUCTION

3D reconstruction has been an active research area in computer vision in the past years due to the high demand in numerous industrial applications. For example, modern heritage preservation projects set high standards for geometric accuracy on challenging data striving for sub-millimetre resolution accuracy and impeccable global shape. Reconstructed surfaces often have complex unknown surface reflectance. There is also much interest in capturing dynamic often non-rigid deformation. This paper tackles the combined challenge of dynamic scene reconstruction with complex arbitrary reflectance properties.

Shape-from-Silhouette [1] is a classical geometric technique that is independent of surface reflectance. However, the resolution of structural concavities in the visual hulls [2] is poor compared to intensity-based methods. The well-established intensity-based methods for 3D geometry reconstruction such as single [3] and multi-view [4] conventional stereo as well as photometric stereo [5] have demonstrated remarkable sub-millimetre geometric accuracies on tailored data. The known limitation of both methods however is the inherent inability to deal with unknown surface reflectance. Conventional stereo requires Lambertian (purely diffuse)

Bi-directional Reflectance Distribution Function (BRDF) being unable to establish feature matches where surface specularities occur. Photometric stereo on the other hand requires the a priori knowledge of the BRDF that must be acquired as pre-processing by a cumbersome and often insufficiently accurate method. To our knowledge, Helmholtz Stereopsis (HS) is the only technique in existence capable of accurately modelling surfaces with an arbitrary BRDF. The technique's novel acquisition set-up, proposed in the seed paper by Zickler *et al.* [6], features reciprocal image pairs characterised by the mutually interchanged camera and light source. The reciprocity at acquisition allows to formulate a depth constraint with the dependence on the BRDF factored out and hence an expanded range of applicability.

Standard HS has been shown to achieve excellent results for rigid scenes with complex a priori unknown reflectance. However, standard HS is not scalable to dynamic scenes since it does not permit simultaneous acquisition of the minimum of 3 reciprocal pairs due the performed swap of the camera and light source. In this paper, we propose colour HS where wavelength multiplexing is used to enable simultaneous capture of reciprocal pairs. Signal separation is achieved by using 3 cameras and 3 coloured light sources and treating each camera channel as a separate image (Figure 1). The novel approach permits instantaneous capture of 3 reciprocal pairs but it also introduces a new set of challenges. The challenges are the acute need for photometric calibration, signal dependence on surface chromaticity and the ambiguity introduced by the drastically reduced number of reciprocal pairs per point. We address all these challenges by developing a complete practical pipeline for colour HS. The pipeline includes a generalisation of the white light photometric calibration procedure from Janko *et al.* [7] to accommodate for chromatic characteristics of the cameras and multispectral light sources. As surface colour will affect inter-channel compatibility we propose a novel method for surface chromaticity calibration. Further, to cope with the reduced number of reciprocal pairs, we incorporate the state-of-the-art Bayesian HS formulation from [8] into the pipeline. To the best of our knowledge, our colour HS pipeline is the first approach capable of reconstructing dynamic scenes with arbitrary unknown reflectance properties.

II. RELATED WORK

Amongst intensity-based methods Helmholtz Stereopsis (HS) is the only completely BRDF independent technique. The subsequent development of HS, after its introduction by Zickler *et al.* [6], included work on extensions for wider applicability and increased geometric accuracy. Guillemaut *et al.* [9] propose modifications for accurate geometric reconstruction of highly textured surfaces by HS. In [10], a more physically meaningful HS constraint resulting in a Maximum Likelihood (ML) surface is formulated. In the recent work by Roubtsova and Guillemaut [8] Bayesian formulation of HS with a tailored prior jointly optimising depth and normal information is shown to produce superior results to the original ML formulation in [6]. There has also been some work [11], [12] on full 3D reconstruction by HS.

A major limitation of HS has always been its controlled set-up and the slow acquisition speed. Some of the impracticalities preventing the reconstruction technique from gaining wider popularity were addressed by the HS inventors in the follow-up papers. Firstly, Zickler *et al.* [13] propose a binocular variant of HS where geometry is reconstructed from a single reciprocal pair by a differential approach. A partial differential equation of depth as a function of surface coordinates with prior initialisation produces a family of solutions, the ambiguity of which is resolved through cumbersome multi-pass optimisation. Although an excellent example of applied optimisation, binocular HS simplifies HS acquisition at the cost of vastly increased computational complexity and introduced reconstruction ambiguities. Another paper by Zickler [14] addresses automatic online geometric calibration of a HS set-up using stable regions of interest: the texture-based and the inherent to HS specularly-based features. The paper proposes a method to avoid pre-calibration of the set-up but does not deal with the bottleneck issue of tedious sequential image acquisition limiting the scope of the technique to static scenes.

The same paper also touches upon automatic radiometric calibration of HS set-up using the inherent specularly-based features. In [14], the definition of radiometric calibration is limited to measuring relative intensities of isotropic light sources. Provided the assumptions of equal camera responses and no spatial source intensity variation hold, the resultant calibration is good considering that it is done from HS data without calibration targets. A much more general radiometric set-up calibration for HS was proposed by Jankó *et al.* [7]. Using a sequence of localised calibration planes Jankó *et al.* calibrate for a spatially varying joint parameter describing sensitivity and radiance of a collocated camera and light source pair.

Unlike Zickler *et al.* [13] who for HS simplification modify the reconstruction algorithm only, we propose a complete novel pipeline tailored for processing HS input for the first time acquired using wavelength multiplexing.

Simultaneous multi-channel acquisition for dynamic scene reconstruction is known from the well-established colour photometric stereo. In [15] and its later extension [16], colour photometric stereo is shown to produce impressive reconstructions of dynamic scenes with untextured, uniform albedo objects, specifically cloth deformation and facial expression sequences. By enforcing spatio-temporal smoothness, Jankó *et al.* in [17] extend the technique to textured surfaces, hence allowing spatial albedo variations via its chromaticity component. Neither work however deals with albedo varying spatially as a result of a non-uniform reflectance model. Moreover, the reflectance model in both works is assumed to be known (Lambertian) and is hardcoded into the reconstruction pipeline. In contrast, our proposed colour HS is valid for arbitrary BRDF, thus naturally permitting albedo variability via the reflectance model. Our method also allows local surface colour variation which we estimate by integrating novel chromatic calibration procedures into the pipeline. We also generalise the previous work in photometric calibration of Jankó *et al.* [7] to multiple multi-chromatic cameras and light sources. Our colour HS is the first variant of HS to successfully reconstruct real dynamic scenes as demonstrated in the evaluation.

III. METHODOLOGY

We propose a novel approach that generalises Helmholtz Stereopsis (HS) to dynamic scenes. Traditionally, the HS reconstruction pipeline was based on white light intensity sampling. Hence, firstly, we introduce the theory of traditional white light HS together with its calibration procedure from [7]. Subsequently, we present our novel pipeline for coloured light HS (colour HS) complete with tailored calibration procedures and data processing algorithms.

A. White light Helmholtz Stereopsis

To introduce Helmholtz Stereopsis (HS), let us define a perspective camera \mathcal{C} and a light source \mathcal{S} centred at \mathbf{c}_1 and \mathbf{c}_2 respectively. In standard HS reciprocal image pairs are acquired with any \mathcal{C} and \mathcal{S} respectively first at location \mathbf{c}_1 and \mathbf{c}_2 and then at \mathbf{c}_2 and \mathbf{c}_1 i.e. with the camera and light source mutually interchanged. As in Jankó *et al.* [7] we define the concept of Helmholtz camera \mathcal{R} as a collocated camera and light source at some position \mathbf{c} . Traditionally, the collocation is simulated by either the camera/light source swap or by using a turntable to move the scene relative to the set-up. Let us define two physically collocated $(\mathcal{C}, \mathcal{S})$ pairs, $\mathcal{R}_1 = (\mathcal{C}_1, \mathcal{S}_1)$ and $\mathcal{R}_2 = (\mathcal{C}_2, \mathcal{S}_2)$ located at \mathbf{c}_1 and \mathbf{c}_2 respectively (Figure 2). The arrangement facilitates Helmholtz reciprocity as one image of the reciprocal pair is obtained with \mathcal{C}_1 and \mathcal{S}_2 at \mathbf{c}_1 and \mathbf{c}_2 and the other with \mathcal{C}_2 and \mathcal{S}_1 at \mathbf{c}_2 and \mathbf{c}_1 respectively.

Helmholtz camera \mathcal{R} is photometrically characterised by its radiance and sensitivity functions, ρ and σ respectively. Both ρ and σ vary as a function of ray \mathbf{v} from \mathcal{R} to the

surface point \mathbf{x} (Figure 2). Hence, intensity i_1 at surface point \mathbf{x} in the reciprocal pair image I_1 acquired with \mathcal{R}_2 as the light source and \mathcal{R}_1 as the camera can be expressed [7] as:

$$i_1 = \rho_2(\mathbf{v}_2)\sigma_1(\mathbf{v}_1)f_r(\mathbf{v}_2, \mathbf{v}_1)\frac{\mathbf{v}_2 \cdot \mathbf{n}}{\|\mathbf{c}_2 - \mathbf{x}\|^2} \quad (1)$$

where $\rho_2(\mathbf{v}_2)$ is the radiance of \mathcal{R}_2 along \mathbf{v}_2 and $\sigma_1(\mathbf{v}_1)$ is the sensor sensitivity of \mathcal{R}_1 along \mathbf{v}_1 . Intensity i_2 , which is the projection of \mathbf{x} in the other reciprocal pair image, is obtained by interchanging the vector indices 1 and 2 in (1). Jankó *et al.* [7] propose a method for photometric calibration of \mathcal{R} in traditional white light HS. Specifically, for \mathcal{R}_1 and \mathcal{R}_2 as in Figure 2 Jankó *et al.* calibrate for the radiance to sensitivity ratios:

$$\mu_k(\mathbf{v}_k) = \frac{\rho_k(\mathbf{v}_k)}{\sigma_k(\mathbf{v}_k)}, \quad k = 1, 2 \quad (2)$$

In HS, reciprocal intensity measurements i_1 and i_2 can be combined into a single surface normal constraint eliminating the dependence on the BRDF $f_r(\mathbf{v}_2, \mathbf{v}_1)$. The elimination is based on Helmholtz reciprocity [18] - the invariance of behaviour in medium of a light ray and its reverse. For BRDF the implication first observed by Zickler *et al.* is that: $f_r(\mathbf{v}_1, \mathbf{v}_2) = f_r(\mathbf{v}_2, \mathbf{v}_1)$. Via this equality reciprocal intensities i_1 and i_2 expressed as in (1) are linked, incorporating photometric calibration μ , to give the normal constraint:

$$\left(\frac{\mu_1(\mathbf{v}_1)i_1}{\|\mathbf{c}_1 - \mathbf{x}\|^2}\mathbf{v}_1 - \frac{\mu_2(\mathbf{v}_2)i_2}{\|\mathbf{c}_2 - \mathbf{x}\|^2}\mathbf{v}_2 \right) \cdot \mathbf{n} = 0 \quad (3)$$

Reconstruction is the process by which for every surface point depth values are selected from a set of hypotheses. As originally described in [6], in standard HS at least 3 constraints in the form $\mathbf{w} \cdot \mathbf{n} = 0$ (3) are required in order to solve $W\mathbf{n} = 0$ where W is the constraint matrix with \mathbf{w} as rows. Singular value decomposition of W gives a normal estimate \mathbf{n} and the confidence value for the estimate (i.e. the likelihood of the hypothesis being a surface point). Standard HS formulation is a maximum likelihood (ML) method where depth for each surface location is optimised independently by selecting the hypothesis with the highest confidence value.

B. Colour Helmholtz Stereopsis

In this section, we expand the theory of white light HS to formalise colour HS - the novel variant of HS we propose. In colour HS, the light sources \mathcal{S}_1 and \mathcal{S}_2 are characterised by different frequency spectra. For consistent frequency-independent response, chromaticity of the reconstructed surface must be factored into the intensity equation. The BRDF function at \mathbf{x} is also a function of the illumination frequency ω : $f_r(\mathbf{v}_2, \mathbf{v}_1, \omega)$. We propose to split f_r into the geometry-related part $f_r(\mathbf{v}_2, \mathbf{v}_1)$, dependent on \mathbf{v}_1 and \mathbf{v}_2 only, and the part related to the surface point chromaticity $p(\omega)$. We define the local chromatic constant $p_{1,2}$ as the

reflectance coefficient due to the inherent colour of a point when seen by camera \mathcal{C}_1 and lit by light source \mathcal{S}_2 . The camera is of importance due to possible differences in spectral sensor characteristics. For the coefficient to be 0, the illumination spectrum must exactly match the point's chromatic absorption spectrum. This is unlikely to happen exactly, although the signal quality will degrade for points of purer (R,G,B) colours. Incorporating chromaticity $p_{1,2}$, we can re-write intensity equation (1) for colour HS as:

$$i_1 = \rho_2(\mathbf{v}_2)\sigma_1(\mathbf{v}_1)p_{1,2}f_r(\mathbf{v}_2, \mathbf{v}_1)\frac{\mathbf{v}_2 \cdot \mathbf{n}}{\|\mathbf{c}_2 - \mathbf{x}\|^2} \quad (4)$$

For colour HS the normal constraint from (3) becomes:

$$\left(\frac{p_{2,1}\mu_1(\mathbf{v}_1)i_1}{\|\mathbf{c}_1 - \mathbf{x}\|^2}\mathbf{v}_1 - \frac{p_{1,2}\mu_2(\mathbf{v}_2)i_2}{\|\mathbf{c}_2 - \mathbf{x}\|^2}\mathbf{v}_2 \right) \cdot \mathbf{n} = 0 \quad (5)$$

Traditional HS set-ups feature just one camera-light source pair where either the equipment moves relative to the scene or the static scene is moved relative to the set-up for reciprocal pair acquisition. Colour HS we propose is a static configuration consisting of three pairs of collocated cameras and light sources. The cameras are equipped with an RGB sensor while the light sources all have different RGB characteristics. The three light sources must have the minimum frequency overlap to ensure signal separation. Signal separation allows simultaneous acquisition of the 3 required reciprocal pairs for normal estimation and enables generalisation to dynamic scenes.

Photometric calibration for colour HS with its physically different cameras and multispectral light sources is particularly important. Note that Helmholtz cameras in colour HS are essentially characterised as a sensor of one light frequency spectrum and a transmitter of another in different reciprocal pairs. Section III-C details how we generalise the algorithm from Jankó *et al.* for photometric calibration in colour HS and provide insights into its application in practice. Section III-D introduces the procedure we devised to calibrate surface chromaticity observed by each camera for the spectrum of each light source to eliminate intensity inconsistencies within the same reciprocal pair. Finally, we give a detailed description of the experimental set-up and the rest of the pipeline designed for reconstruction accuracy under the constraint of 3 reciprocal pairs per frame.

C. Photometric calibration of Helmholtz camera

Jankó *et al.* [7] photometrically calibrate each Helmholtz camera \mathcal{R}_1 using another Helmholtz camera \mathcal{R}_2 by linking HS constraints, obtained by gradual displacement of a calibration plane, via the ray of incident illumination \mathbf{v}_2 (Figure 2). In the original paper, the calibration was performed in a highly controlled environment with the plane translated in fixed vertical increments with a single camera and a light source suspended overhead and manually centred over the turntable with the plane. For photometric calibration in the colour HS pipeline we went for a more freehand approach

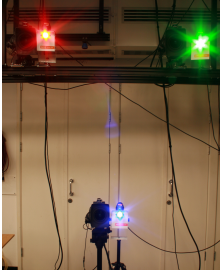


Figure 1. Experimental set-up

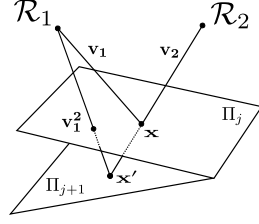


Figure 2. Geometry of the calibration procedure

featuring a hand-held calibration board which was randomly moved within the reconstruction volume and had markers (Figure 4) for 3D localisation. Our entire configuration consisting of 3 Helmholtz cameras was thus calibrated simultaneously.

As in Jankó *et al.* [7] for every position j of the calibration plane Π_j we establish a ratio of parameters μ_1 and μ_2 corresponding to the Helmholtz cameras \mathcal{R}_1 and \mathcal{R}_2 sampled at a surface point \mathbf{x} where rays \mathbf{v}_1 and \mathbf{v}_2 intersect (Figure 2). However, in colour HS, the resultant ratio κ is not the same as for white light HS as it is derived from (5) rather than (3) incorporating chromaticity:

$$\kappa(\mathbf{v}_1, \mathbf{v}_2 | \Pi_j) = \frac{p_{2,1} \mu_1(\mathbf{v}_1)}{p_{1,2} \mu_2(\mathbf{v}_2)} = \frac{\mathbf{n}^\top \mathbf{v}_2 \|\mathbf{c}_1 - \mathbf{x}\|^2 i_2}{\mathbf{n}^\top \mathbf{v}_1 \|\mathbf{c}_2 - \mathbf{x}\|^2 i_1} \quad (6)$$

For simplicity, our calibration object is uniformly coloured (except the masked out markers): $p_{2,1}$ and $p_{1,2}$ are constant for all surface points. Subsequently, point \mathbf{x} on plane Π_j is transferred onto the plane in the new position Π_{j+1} by finding the intersection \mathbf{x}' of ray \mathbf{v}_2 with Π_{j+1} . Hence for plane Π_{j+1} the ratio $\kappa(\mathbf{v}_1^2, \mathbf{v}_2 | \Pi_{j+1}) = \frac{p_{2,1} \mu_1(\mathbf{v}_1^2)}{p_{1,2} \mu_2(\mathbf{v}_2)}$ is established sharing the denominator with the corresponding relationship of plane Π_j . The shared denominator, together with chromaticity constancy, allows to obtain a relationship between $\mu_1(\mathbf{v}_1^2)$ and $\mu_1(\mathbf{v}_1)$ (two pixel locations (u_1, v_1) and (u_1^2, v_1^2) corresponding to rays \mathbf{v}_1 and \mathbf{v}_1^2 in the photometric map of \mathcal{C}_1 under \mathcal{S}_2):

$$r_1(\mathbf{v}_1, \mathbf{v}_1^2) = \frac{\kappa(\mathbf{v}_1, \mathbf{v}_2 | \Pi_j)}{\kappa(\mathbf{v}_1^2, \mathbf{v}_2 | \Pi_{j+1})} = \frac{p_{2,1} \mu_1(\mathbf{v}_1)}{p_{2,1} \mu_1(\mathbf{v}_1^2)} = \frac{\mu_1(\mathbf{v}_1)}{\mu_1(\mathbf{v}_1^2)} \quad (7)$$

Equation (7) provides constraints on a set of control points of the photometric map via bilinear interpolation. We chose to use a simpler regularisation kernel than Jankó *et al.* who perform bi-cubic interpolation between control points. A single constraint is of the form $(\mathbf{a}_i - \mathbf{b}_i)\boldsymbol{\lambda} = \delta_i$ where $\boldsymbol{\lambda} = [\ln(\mu_1(\mathbf{v}_1)), \ln(\mu_1(\mathbf{v}_2)), \dots, \ln(\mu_1(\mathbf{v}_N))]^\top$ is the vector of variables, \mathbf{a}_i and \mathbf{b}_i are the interpolation coefficients from control points to sample points and $\delta_i = \ln(r_1(\mathbf{v}_1, \mathbf{v}_1^2))$. The resultant linear system is:

$$(A - B)^\top (A - B)\boldsymbol{\lambda} = (A - B)^\top \Delta \quad (8)$$

where $A = [\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_M]^\top$, $B = [\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_M]^\top$ and $\Delta = [\delta_0, \delta_1, \dots, \delta_M]^\top$.

It has perhaps not been made explicit in [7] that Helmholtz cameras involved in a single reciprocal pair must be calibrated as a couple and not individually. Once the set of photometric values λ_1 of Helmholtz camera \mathcal{R}_1 has been obtained from (8), the set of *relative* values $\lambda_{1,2}$ for \mathcal{R}_2 is computed by transfer via (6). Values $\lambda_{1,2}$ are different from λ_2 computed by direct calibration of \mathcal{R}_2 .

In contrast to Jankó *et al.*, we have found that for its accurate calibration a Helmholtz camera must be observed in at least two Helmholtz camera pairs as in Figure 2. Jankó *et al.* mention ill-posedness of the calibration problem when a single Helmholtz camera pair is used due to constraints being sampled along the projection ray and the linked samples being along the same epipolar line. Jankó *et al.* do not deem multiple Helmholtz camera pairs essential using a strong bi-cubic regulariser to address the ill-posedness. In our case, we use the already available multiple pairs per Helmholtz camera to make the problem better posed allowing us to work with a weaker bi-linear regulariser and avoiding potential artefacts due to over-regularisation.

D. Surface chromaticity calibration

In this section, we propose a procedure for pointwise calibration of reconstructed surface chromaticity. The goal is to compute $(p_{1,R}, p_{1,G}, p_{1,B})$ per point \mathbf{x} : the reflectance coefficients viewed by camera \mathcal{C}_1 under the red, blue and green frequency spectra. The calibration method is based on sampling chromatic response of a reference white object and the arbitrarily coloured reconstruction object to be calibrated. The reference and calibrated objects are sequentially exposed to red, green and blue illumination from the same direction. The spectra are defined by the colour filters subsequently intended for colour HS data acquisition. The sum of the RGB spectra defines white illumination in this context.

From (4) we can write intensity equations for the RGB channels of \mathcal{C}_1 separately when under red (ρ_r), green (ρ_g) and blue (ρ_b) spectrum illumination respectively:

$$\begin{aligned} i_r &= \rho_r(\mathbf{v}_2, \mathbf{v}_1) \sigma_r(\mathbf{v}_2, \mathbf{v}_1) p_{1,R} f_r(\mathbf{v}_2, \mathbf{v}_1) \frac{\mathbf{v}_2 \cdot \mathbf{n}}{\|\mathbf{c}_2 - \mathbf{x}\|^2} \\ i_g &= \rho_g(\mathbf{v}_2, \mathbf{v}_1) \sigma_g(\mathbf{v}_2, \mathbf{v}_1) p_{1,G} f_r(\mathbf{v}_2, \mathbf{v}_1) \frac{\mathbf{v}_2 \cdot \mathbf{n}}{\|\mathbf{c}_2 - \mathbf{x}\|^2} \\ i_b &= \rho_b(\mathbf{v}_2, \mathbf{v}_1) \sigma_b(\mathbf{v}_2, \mathbf{v}_1) p_{1,B} f_r(\mathbf{v}_2, \mathbf{v}_1) \frac{\mathbf{v}_2 \cdot \mathbf{n}}{\|\mathbf{c}_2 - \mathbf{x}\|^2} \end{aligned} \quad (9)$$

By definition, the reference white object shows the same chromatic response for red, green and blue light: $p_{1,R}^w = p_{1,G}^w = p_{1,B}^w$. Hence, the RGB intensity equations (i_r^w, i_g^w, i_b^w) , formulated as in (9) but for a white surface, are chromaticity independent. As the camera-to-light source geometry is also identical between channel measurements, the ratio of two channels for a white point is determined by

the relative ρ - σ products. For example:

$$\frac{i_r^w}{i_g^w} = \frac{\rho_r(\mathbf{v}_2, \mathbf{v}_1)\sigma_r(\mathbf{v}_2, \mathbf{v}_1)}{\rho_g(\mathbf{v}_2, \mathbf{v}_1)\sigma_g(\mathbf{v}_2, \mathbf{v}_1)} \quad (10)$$

The equivalent ratio for the same point in space defined by rays \mathbf{v}_1 and \mathbf{v}_2 but occupied by the arbitrarily coloured object would be:

$$\frac{i_r}{i_g} = \frac{\rho_r(\mathbf{v}_2, \mathbf{v}_1)\sigma_r(\mathbf{v}_2, \mathbf{v}_1)p_{1,R}}{\rho_g(\mathbf{v}_2, \mathbf{v}_1)\sigma_g(\mathbf{v}_2, \mathbf{v}_1)p_{1,G}} \quad (11)$$

as only the geometry cancels out in this case. Observe that ρ and σ at the point in space defined by rays \mathbf{v}_1 and \mathbf{v}_2 are independent of the surface occupying it. Substituting (10) into (11) we obtain an expression for the ratio of two components of the chromaticity characterising triplet:

$$\frac{p_{1,R}}{p_{1,G}} = \frac{i_r i_g^w}{i_g i_r^w} \quad (12)$$

Having three such ratios $\frac{p_{1,R}}{p_{1,G}}$, $\frac{p_{1,R}}{p_{1,B}}$ and $\frac{p_{1,G}}{p_{1,B}}$ allows one to formulate a system of linear equations solved for the triplet $(p_{1,R}, p_{1,G}, p_{1,B})$. The set of corresponding RGB values for calibrated surface points visible to \mathcal{C}_1 constitutes a chromaticity map. Each RGB value is effectively the sampled surface point colour *under white light of our definition*.

E. Implementation of Colour HS

In this section, we discuss the details of colour HS implementation at various stages of the pipeline from acquisition to reconstruction. Figure 1 shows our acquisition set-up consisting of three collocated cameras \mathcal{C}_i and light sources \mathcal{S}_i where $i = \{1, 2, 3\}$. Each collocated pair partakes in 2 Helmholtz cameras resulting in a 6 Helmholtz camera set-up: $(\mathcal{R}_{1,r}, \mathcal{R}_{1,g}, \mathcal{R}_{2,b}, \mathcal{R}_{2,g}, \mathcal{R}_{3,r}, \mathcal{R}_{3,b})$. Sources \mathcal{S}_i are given different frequency characteristics by using colour filters - red, green and blue for maximal spectral separation. The filters were chosen to match RGB channel spectra of the cameras as much as possible and no ambient light is allowed. With the set-up we simultaneously acquire three reciprocal image pairs, each characterised by two Helmholtz cameras and two RGB signal channels.

Using these reciprocal pairs, constraints as in (5) are formulated. The constraints can be directly integrated into the original reconstruction pipeline proposed by Zickler *et al.* in the seed paper introducing HS [6]. However, standard ML HS is known to be prone to noise due to the lack of regional support in depth assignment. In our colour HS we are inherently limited to just 3 reciprocal pairs leaving room for reconstruction ambiguity. Additional intensity error may occur through channel cross-talk which we do not explicitly address in this work. Consequently, standard ML HS is inadequate in this case. In this work, we adopt the Bayesian HS formulation of Roubtsova and Guillemaut [8] who perform depth assignment by minimising the sum of data and prior costs of the entire configuration of reconstructed

surface points: $\alpha E_{data} + (1 - \alpha)E_{prior}$. We use their data term definition and the idea of the depth-normal consistency prior shown to produce the most accurate depth maps by enforcing consistency between sampled depth and normal characteristics. The optimisation process results in a point cloud of oriented vertices which is subsequently integrated into meshes by Poisson surface reconstruction [19].

IV. EVALUATION

The methodology of colour HS was validated using real datasets. Our evaluation comprises both static and dynamic scenes. The range of static objects was selected to demonstrate accuracy of the algorithm and its ability to cope with cases of different complexity. The objects are a white plane for quantitative evaluation (“Plane”) defined by the calibration board as in Figure 4, a white plaster statue of a monster head (“Monster”) in Figure 3 and a multi-coloured toy-dog (“Slinky”) in Figure 6 showing structural and radiometric complexity due to texture, transparency, specular reflectance and pure colours of its various materials. Having shown correctness of geometric reconstruction for the more controlled static scenery sets, we subsequently validate our claim of suitability of colour HS for dynamic scene reconstruction. We are particularly interested in reconstruction of such scenes with complex reflectance properties as these are inherently challenging for conventional and photometric stereo methods. In the dynamic scene evaluation we shall show temporal structural deformation of a white specular sheet and of a glossy white blouse cloth (Figure 8). For static scenery we are able to handle spatially varying chromaticity, while for dynamic scenes we are currently assuming spatially uniform (specifically white) chromaticity.

A. Static scenes

For static scene evaluation, each Helmholtz camera was radiometrically calibrated as described in section III-C using a calibration board with markers to define the position of the plane in each calibration shot. Position triangulation of any three of the four markers gives an anchor point and a normal to the calibration plane defined by the board. Using 7 plane positions we produced a photometric map of μ covering the region framed for static object reconstruction for each Helmholtz camera in $(\mathcal{R}_{1,r}, \mathcal{R}_{1,g}, \mathcal{R}_{2,b}, \mathcal{R}_{2,g}, \mathcal{R}_{3,r}, \mathcal{R}_{3,b})$.

Note that a uniformly white calibration surface is required for radiometric camera calibration in our case since we are simultaneously calibrating light sources of various colour spectra: light reflection may not be affected by the chromatic properties of the calibration surface. So, the black localisation markers are masked out together with the background outside the calibration boards in the images. Furthermore, data from more than just one pair of planes Π_j and Π_{j+1} is needed for continuous calibration coverage of the reconstruction frame. Figure 3 shows the resultant photometric maps in the region of interest for the monster

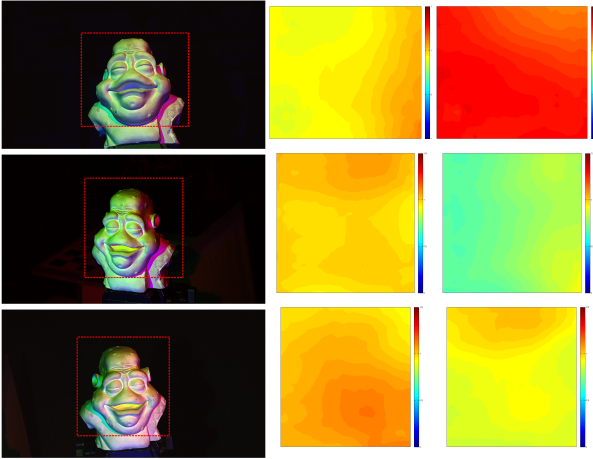


Figure 3. Photometric calibration for our six Helmholtz cameras in the image window of interest.

dataset. The maps are represented as heat maps to illustrate regional variation of μ within the scope of the reconstruction frame. The photometric maps characterise our 6 Helmholtz camera set-up and hence are invariant with the dataset.

Reconstruction of each shot is performed from 3 RGB images (3 reciprocal pairs) instantaneously acquired by cameras C_1 , C_2 and C_3 under concurrent multispectral (RGB) illumination of S_1 , S_2 and S_3 as described in Section III-E. Poisson surface reconstruction is used to generate surface meshes from our point clouds of oriented vertices (i.e. surface points with the corresponding normals). The method works on watertight point clouds. We convert our one-sided reconstructions into watertight point clouds by injecting visual hull vertices and computing geometric normals based on neighbouring vertices from shared facets.

1) *Quantitative evaluation:* The first static object considered is a plane defined by the calibration board. The plane example was selected from the radiometric calibration data, as the data contain markers to compute the ground truth plane orientation. The test plane was not one of the 7 planes used in the photometric calibration process (PhCalib). It was not necessary to perform chromatic calibration in this case as the surface is made of a similar material to the sheet used to define the reference white. Figure 7 shows the following 4 reconstructions of the plane segment: 1. the visual hull (VH) and Poisson meshes of 2. standard ML HS with photometric Helmholtz camera calibration (ML_HS_wPhCalib), 3. Bayesian HS without PhCalib (BayesianHS_w/oPhCalib) and 4. Bayesian HS with PhCalib (BayesianHS_wPhCalib). Despite full calibration, with just 3 reciprocal pairs standard HS performs poorly due to the absence of regularisation. Bayesian HS in contrast performs quite well. Quantita-



Figure 4. Calibration boards as viewed by C_1 , C_2 and C_3 .

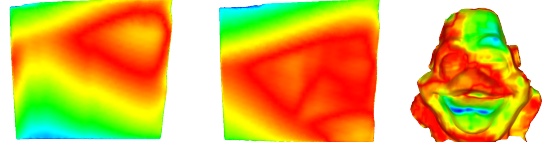


Figure 5. Heat maps of local Hausdorff distance variation for: Plane by BayesianHS_w/oPhCalib and GT (left), Plane by BayesianHS_wPhCalib and GT (middle), Monster by Bayesian HS with and without PhCalib and ChromCalib (right).



Figure 6. An intensity image of Slinky's appearance (left) and its chromaticity maps for cameras C_1 , C_2 and C_3 (remaining images).

tively, the accuracy of performance is expressed using the Hausdorff distance that measures the locational discrepancy between corresponding vertices of the given reconstruction and the overlaid ground truth (GT) mesh we obtain by marker localisation. We measured the RMS Hausdorff distance of 2.5 mm for BayesianHS_w/oPhCalib and 2 mm for BayesianHS_wPhCalib. The spatial error variation for BayesianHS_w/oPhCalib and BayesianHS_wPhCalib is visualised as heat maps in Figure 5 with the red indicating the closest correspondence and the blue the farthest. Both the RMS and the heat maps indicate the positive effect PhCalib has on reconstruction accuracy.

2) *Qualitative evaluation:* Let us first consider the Monster. The object is not exactly uniformly white exhibiting significant spatial chromaticity variations due to dirt warranting chromatic surface calibration (ChromCalib). In Figure 7 we consider the following reconstructions: 1. VH, 2. standard ML HS with both PhCalib and ChromCalib (ML_HS_wPh&ChromCalib), 3. Bayesian HS without PhCalib or ChromCalib (BayesianHS_w/oCalib), 4. Bayesian HS with PhCalib but without ChromCalib (BayesianHS_wPhCalib) and 5. Bayesian HS with both PhCalib and ChromCalib (BayesianHS_wPh&ChromCalib). As for Plane, Bayesian HS results are clearly superior to standard HS since the depth-normal consistency prior allows more accurate and robust surface localisation. The Bayesian result shows both global geometric accuracy and structural detail, which could be further improved on by increasing spatial sampling resolution.

At the first glance there seems to be little difference between the results of BayesianHS_w/oCalib, BayesianHS_wPhCalib and

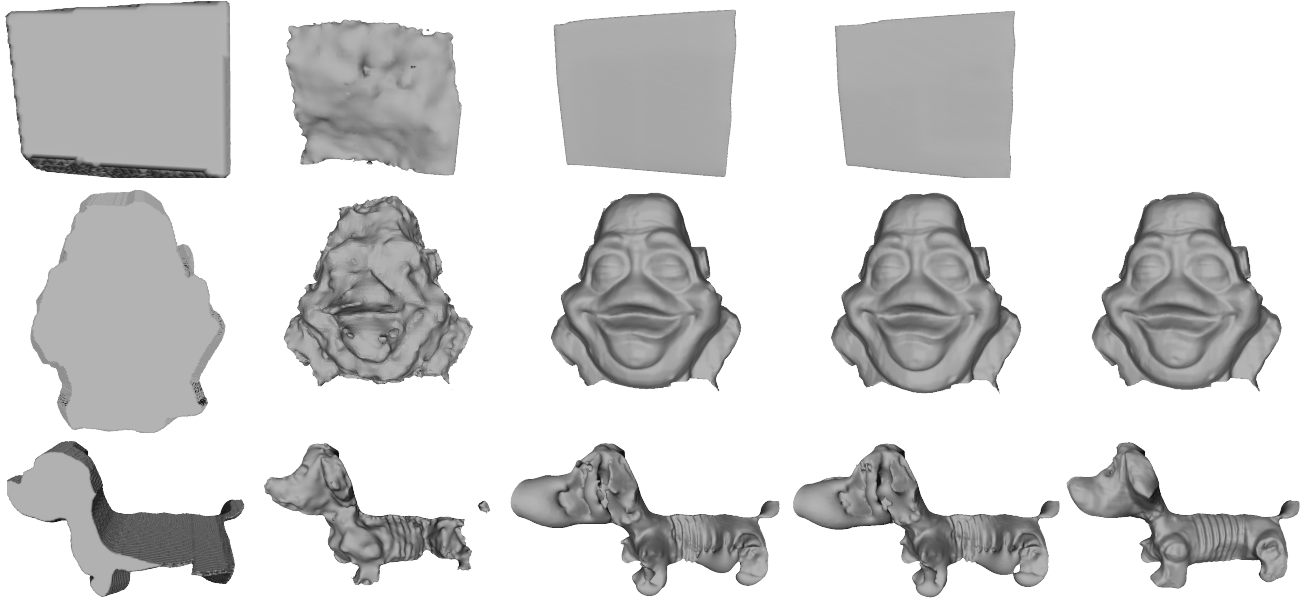


Figure 7. Reconstruction results for static object datasets in the order of increasing reconstruction algorithm complexity. For Monster and Slinky (left to right): VH, ML_HS_wPh&ChromCalib, BayesianHS_w/oCalib, BayesianHS_wPhCalib and BayesianHS_wPh&ChromCalib. For Plane (left to right): VH, ML_HS_wPhCalib, BayesianHS_w/oPhCalib and BayesianHS_wPhCalib

BayesianHS_wPh&ChromCalib. However, in reality, their mutual global shape variation is quite substantial as evidenced by the heat map of Hausdorff distance between the no calibration result (BayesianHS_w/oCalib) and the full calibration result (BayesianHS_wPh&ChromCalib). For example, note the significant difference in the lip. Full calibration clearly makes a difference to the result although it is not possible to quantify the improvement due to the absence of ground truth. The lack of more significant visual differences between the three results is attributed to the relatively mutually homogeneous characteristics of our colour-balanced cameras and light sources as well as the relatively uniform chromaticity of the test object overall. Although the pipeline is clearly tolerant to some degree of chromaticity variation, on close observation of the partial calibration result (BayesianHS_wPhCalib) a global shape distortion can be noticed indicating ChromCalib is not redundant even for this relatively uniformly white object.

The importance of calibration becomes more pronounced for the larger local chromatic variations of Slinky. The same 5 reconstruction methods as for the Monster are compared for this dataset. ML_HS_wPh&ChromCalib performs exceptionally poorly producing a fragmented surface. As expected, chromaticity calibration is crucial in this case. The insufficiency of PhCalib alone is clearly demonstrated by the fact that BayesianHS_wPhCalib performs as poorly as BayesianHS_w/oCalib. Figure 6 shows the surface chromaticity maps obtained for each camera individually. They appear to be largely in agreement about the chromatic properties of the corresponding regions apart from some

artefacts due to intensity sampling at grazing angles, sensor saturations and in the circular mirror region of the hind leg arguably having no inherent chromaticity. The chromaticity coefficients, although mutually consistent for reconstruction under the given illumination spectrum, may not exactly correspond to the absolute objects colours since our full frequency spectrum definition is the sum of RGB filter responses rather than the absolute white spectrum. The result of BayesianHS_wPh&ChromCalib obtained using these chromatic maps is clearly superior to the rest showing a plausible global shape and a substantial degree of structural resolution on this highly challenging object. The dealt with challenges of Slinky include its signal scattering texture of the face, high frequency geometry in the spring-like torso and the highly specular multi-coloured plastic of the rest of the body. BayesianHS_wPh&ChromCalib seems highly promising for multi-chromatic object reconstruction coping with substantial geometric and radiometric complexity.

B. Dynamic scenes

We have shown accurate geometric reconstruction of static scenes obtained from instantaneously acquired data. Clearly, colour HS demonstrates a great potential for dynamic scene reconstruction. We reconstructed temporal mesh deformations of a white specular laminated sheet and a white glossy blouse. The corresponding mesh deformation video sequences of 201 frames each are provided as supplementary material. Figure 8 shows a sample set of 5 reconstructions 25 frames apart from each video sequence. The algorithm being based on the reflectance model invariant Helmholtz Stereopsis easily copes with the specularities of the laminated sheet

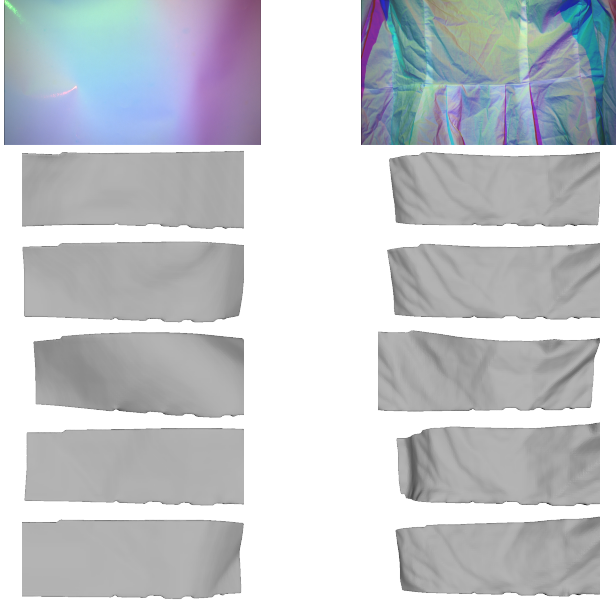


Figure 8. Example frames from the dynamic scene reconstruction sequences for the white laminated sheet and the glossy blouse cloth datasets.

that would cause immense problems for conventional stereo. The folds, creases and domes of the deformable objects plausibly materialise in the reconstructed meshes.

V. CONCLUSION AND FUTURE WORK

In this paper, we have proposed colour HS - a new variant of HS suitable for reconstruction of dynamic scenes with complex reflectance properties. We utilise wavelength multiplexing for simultaneous acquisition of three reciprocal image pairs via different RGB channels. To fully exploit this novel multispectral acquisition set-up, we have addressed problems with inter-channel signal consistency by multi-spectral Helmholtz camera photometric calibration extended to involve several reciprocal pairs and by proposing a novel surface chromaticity calibration procedure. We have also discussed Bayesian HS as the better alternative to standard ML HS for colour HS data with its minimal configuration of three reciprocal pairs per shot. The colour HS pipeline has been validated quantitatively and qualitatively on both static and dynamic real scenes of varying structural and reflectance complexity. Quantitatively, we demonstrated that the method is able to achieve the accuracy of 2 mm on the test plane object. Qualitatively, we have obtained high quality results on a range of static object with complex spatially varying reflectance properties and demonstrated the positive effect of the proposed calibration procedures in these cases. To validate colour HS in practice as a reconstruction method for dynamic scenes with complex reflectance, we have presented two video sequences of detailed dynamic white glossy material deformation reconstruction. Although able to cope with arbitrary reflectance properties, dynamic scene

reconstruction by colour HS is currently limited to objects with uniform surface chromaticity due to the static per-pixel nature of our chromaticity calibration procedure. As future work, we intend to remove this limitation by using dense surface point tracking to transfer statically calibrated chromaticity properties to points in their new positions across the entire video sequence.

REFERENCES

- [1] B. Baumgard, "Geometric modeling for computer vision," Ph.D. dissertation, University of Stanford, 1974.
- [2] A. Laurentini, "The visual hull concept for silhouette-based image understanding," *PAMI*, vol. 16, no. 2, pp. 150–162, 1994.
- [3] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV*, vol. 47, no. 1-3, pp. 7–42, 2002.
- [4] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *CVPR*, vol. 1, 2006, pp. 519–528.
- [5] R. J. Woodham, *Shape from shading*. MIT Press, 1989, ch. Photometric method for determining surface orientation from multiple images, pp. 513–531.
- [6] T. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *IJCV*, vol. 49, no. 2–3, pp. 215–227, 2002.
- [7] Z. Jankó, O. Drbohlav, and R. Šára, "Radiometric calibration of a Helmholtz stereo rig," in *CVPR*, vol. 1, 2004, pp. 166–171.
- [8] N. Roubtsova and J.-Y. Guillemaut, "A Bayesian framework for enhanced geometric reconstruction of complex objects by Helmholtz Stereopsis," in *VISAPP*, vol. 3, 2014, pp. 335–342.
- [9] J.-Y. Guillemaut, O. Drbohlav, R. Šára, and J. Illingworth, "Helmholtz stereopsis on rough and strongly textured surfaces," in *3DPVT*, 2004, pp. 10–17.
- [10] J.-Y. Guillemaut, O. Drbohlav, J. Illingworth, and R. Šára, "A maximum likelihood surface normal estimation algorithm for Helmholtz stereopsis," in *VISAPP*, vol. 2, 2008, pp. 352–359.
- [11] A. Delaunoy, E. Prados, and P. Belhumeur, "Towards full 3D Helmholtz stereovision algorithms," in *ACCV*, vol. 1, 2010, pp. 39–52.
- [12] M. Weinmann, R. Ruiters, A. Osep, C. Schwartz, and R. Klein, "Fusing structured light consistency and helmholtz normals for 3D reconstruction," in *BMVC*, 2012, pp. 1–12.
- [13] T. E. Zickler, J. Ho, D. J. Kriegman, J. Ponce, and P. N. Belhumeur, "Binocular Helmholtz stereopsis," in *ICCV*, vol. 2, 2003, pp. 1411–1417.
- [14] T. Zickler, "Reciprocal image features for uncalibrated Helmholtz stereopsis," in *CVPR*, vol. 2, 2006, pp. 1801–1808.
- [15] C. Hernández, G. Vogiatzis, G. Brostow, B. Stenger, and R. Cipolla, "Non-rigid photometric stereo with colored lights," in *ICCV*, 2007, pp. 1–8.
- [16] G. Brostow, C. Hernández, G. Vogiatzis, B. Stenger, and R. Cipolla, "Video normals from colored lights," *PAMI*, vol. 33, no. 10, pp. 2104–2114, 2011.
- [17] Z. Jankó, A. Delaunoy, and E. Prados, "Colour dynamic photometric stereo for textured surfaces," in *ACCV*, vol. 2, 2010, pp. 55–66.
- [18] H. Helmholtz, *Treatise on Physiological Optics*. Dover (New York), 1925, vol. 1.
- [19] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *SGP*, 2006, pp. 61–70.