Simultaneous multiple view high resolution surface geometry acquisition using structured light and mirrors

Hector R.A. Basevi,1,2,* James A. Guggenheim,1,2 Hamid Dehghani,1,2 and Iain B. Styles2

1 PSIBS, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom
2 School of Computer Science, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom
*hrb886@bham.ac.uk

Abstract: Knowledge of the surface geometry of an imaging subject is important in many applications. This information can be obtained via a number of different techniques, including time of flight imaging, photogrammetry, and fringe projection profilometry. Existing systems may have restrictions on instrument geometry, require expensive optics, or require moving parts in order to image the full surface of the subject. An inexpensive generalised fringe projection profilometry system is proposed that can account for arbitrarily placed components and use mirrors to expand the field of view. It simultaneously acquires multiple views of an imaging subject, producing a cloud of points that lie on its surface, which can then be processed to form a three dimensional model. A prototype of this system was integrated into an existing Diffuse Optical Tomography and Bioluminescence Tomography small animal imaging system and used to image objects including a mouse-shaped plastic phantom, a mouse cadaver, and a coin. A surface mesh generated from surface capture data of the mouse-shaped plastic phantom was compared with ideal surface points provided by the phantom manufacturer, and 50% of points were found to lie within 0.1mm of the surface mesh, 82% of points were found to lie within 0.2mm of the surface mesh, and 96% of points were found to lie within 0.4mm of the surface mesh.

© 2013 Optical Society of America

OCIS codes: (120.2830) Height measurements; (120.6650) Surface measurements, figure; (170.0110) Imaging systems.

References and links


#182075 - $15.00 USD   Received 19 Dec 2012; revised 5 Mar 2013; accepted 7 Mar 2013; published 14 Mar 2013
(C) 2013 OSA 25 March 2013 / Vol. 21, No. 6 / OPTICS EXPRESS  7222

28


1. Introduction

The field of non-contact surface capture allows the measurement of the geometrical shape of objects and scenes. It has many applications in a number of areas ranging from archaeology to industrial quality control, and consequently is both broad and mature. Non-contact surface imaging is a high throughput type of non-contact surface measurement, and typically makes use of cameras and sources of structured illumination. Sinusoidal fringe patterns are one of the most common classes of structured illumination used for non-contact surface imaging, and allow simultaneous height measurement of all areas within the fields of view of both the cameras and the structured illumination sources, using a number of different sinusoidal patterns. Standard sinusoidal fringe pattern systems make use of a crossed-axis assumption (where the camera and projector pupils lie in a plane parallel to the platform on which sits the object to be imaged, and their axes intersect at the platform), which simplifies data processing but also restricts the placement of the system components. In addition, imaging of discontinuous surfaces is problematic and the area that can be imaged is limited to the intersection of the direct fields of view of the cameras and structured light sources.

In this paper we present a sinusoidal fringe projection surface capture system which can acquire the absolute coordinates of points on a subject surface that may be discontinuous from the perspective of the system camera, for regions visible directly to the camera or visible only through reflective surfaces, and further can relax the crossed-axis constraint allowing arbitrary instrument geometry. This is accomplished through the use of mirrors, a general optical model, and phase measurement using patterns at multiple phases and frequencies, and allows simultaneous acquisition of direct and multiple mirror fields of view using a single camera without the use of any moving parts.

This work was developed to be used as part of a preclinical multi-modality optical imaging system [1, 2]. Optical imaging modalities such as Bioluminescence Tomography (BLT) [3] and Diffuse Optical Tomography (DOT) [4] use visible or near infrared light to probe the properties of living tissue, allowing the imaging of structural and functional features such as tissue composition and blood oxygenation. In order to model the physics of light propagation in a specific subject it is necessary to know the surface geometry of the subject. This can be accomplished via the use of secondary medical imaging modalities which can provide surface information, such as Computed Tomography [5] or Magnetic Resonance Imaging [6–8], or dedicated surface capture imaging modalities [9–11]. In addition to this, it is also necessary to be able to map measurements to the locations on the surface of the subject from which they originated. If measurements are acquired using a contact measurement device such as a Photomultiplier
Tube coupled to an optical fibre, one end of which is in physical contact with the subject, then measurement locations can be measured through the process of placing the detectors. However, if a non-contact measurement device such as a camera is used in the measurement process then the task of assigning measurements to physical locations is non-trivial. The target optical medical imaging system [1, 2] acquires measurements using a high sensitivity CCD-based camera, and uses mirrors to extend the field of view of the camera to regions not directly visible. However in order to use these measurements it is necessary to know the geometry of the regions visible only through the mirrors. The surface capture system presented here addresses these challenges.

A survey of surface capture techniques can be found in Section 2. The design of the surface capture system is detailed in Section 3, and results including measurement accuracy and imaging of realistically shaped objects are presented in Section 4.

2. Surface capture techniques

The set of techniques that measure the shape of surfaces can be broadly divided into two categories: contact, and non-contact techniques. Contact methods use physical probes to make contact with a surface, and measure the location of the probes as they are moved along the surface. This requires sometimes complex articulation of the probe, and is problematic when measuring deformable surfaces as deformation under forces imparted by the probe renders measurements unreliable.

Non-contact techniques use information that can be obtained from a subject without physical contact, in order to reconstruct height information. These methods include laser range finding, photogrammetry, and fringe projection. It should be noted that classes of systems that will be discussed are not mutually exclusive and so it is possible for a particular system to contain elements of multiple classes.

2.1. Light detection and ranging

Scanning Light Detection and Ranging (LIDAR) systems such as Laser Range Finding can be conceptually simple, and typically involve measuring the length of time that light takes to travel from the subject being measured to the detector, and converting that into a distance. These systems usually scan a pulsed laser beam over the imaging subject and measure the photon travel time using a high-speed photodetector. They possess certain disadvantages: they require moving parts (typically a scanning mirror), accuracy suffers when measuring near objects due to the difficulty of measuring the time of flight of light over small distances, and the presence of reflective or transparent objects with unknown refractive index invalidates the assumption that the light travels to the imaging subject from the laser in a straight line and at a constant speed.

Non-scanning LIDAR systems such as Time of Flight cameras [12, 13] use non-point illumination and a spatially resolved imaging sensor to calculate the distance of objects from the imaging instrument. These systems can acquire the distance from the camera to multiple objects simultaneously through the use of temporally modulated illumination and optical sensing. Temporally modulated illumination allows the measurement of the change in the phase of a modulated signal due to the signal travel time from the camera to the object and then back again. This measured phase is proportional to the distance travelled by the light and so can be converted to a point in space relative to the camera position and orientation in conjunction with knowledge of the optics of the camera. These systems require complex hardware such as a temporally modulated light source and a high-speed camera shuttering system [12, 13].
2.2. Photogrammetry

Photogrammetry is one of the oldest forms of non-contact geometry measurement techniques and in its simplest form uses multiple images taken from different points of view to perform triangulation. If the position and orientation of the camera for each image is known, then image analysis techniques can be used to isolate common points within multiple images. The spatial coordinates of these common points can be extracted by calculating the intersection of the camera rays corresponding to the common points for each image [14].

This method is limited by the need to move the camera between images (and measure its new position and orientation). Other means of taking images from multiple viewpoints include the use of multiple cameras, or mirrors, but each has disadvantages. In addition, only points that can be isolated in multiple images can be extracted using this technique.

Instead of using image analysis techniques to find points common to multiple images, it is possible to use characteristic markers placed within a scene. These markers can be physical objects, or light-based. For example, a laser can be shone into the scene and then imaged with a camera, or structured patterns can be projected onto the scene using a projector in a manner that enables segmentation of the resulting images into a number of regions [15]. Then spatial locations can be extracted using a single camera, through knowledge of the laser or projector’s position.

The laser-based version is limited by the time and equipment required to reorient the laser to a new position in the scene, and the need to acquire a separate image for each position measured. The projector-based version is limited by the camera’s ability to resolve changes in the projected intensity, requiring in practice either a large number of images, or a sacrifice in segmentation resolution.

2.3. Fringe projection profilometry

Fringe projection profilometry requires a camera and a source of structured light. This light typically has a sinusoidal spatial intensity distribution, but other types of structure can be used [16]. For simplicity, this discussion will be limited to sinusoidal fringe projection.

Sinusoidal fringes can be generated using a laser and diffraction grating to produce an interference pattern, or using a projector. These fringe patterns, when projected onto a surface and imaged from a different location, will appear to be deformed with respect to the patterns projected onto a plane orthogonal to the axis of the projector and imaged from the pupil of the projector. Analysis of the deformation of the pattern can be used to extract spatial coordinates at full camera resolution using fewer images than is typically necessary for other techniques [17–19]. The system presented in this paper is of this type.

In two dimensions, the intensity of a fringe pattern projected onto a plane can be described as:

\[ I(x) = r(x) \frac{1}{2}(1 + \cos(2\pi fx + \psi_b(x))) \]  

\( I(x) \) is the spatial coordinate in the plane, \( r(x) \) is a function representing the spatially dependent reflectance of the plane, \( f \) is the spatial frequency of the fringe pattern, and \( \psi_b(x) \) is the change in apparent phase of the pattern as a result of the plane not being orthogonal to the projector orientation. It is possible to calculate \( y \), the spatial coordinate perpendicular to the plane, through knowledge of \( \psi_b(x) \).

If an object is placed on the surface, the image acquired is now:

\[ I(x) = r(x) \frac{1}{2}(1 + \cos(2\pi fx + \psi_b(x) + \psi_o(x))) \]  

where \( \psi_o(x) \) is the apparent phase change as a result of the object, and \( r(x) \) now contains the spatially dependent reflectance of the surface and object. The value of \( y \) at various points on
The object can be calculated using $\psi_o(x)$, but first it is necessary to extract $\psi_o(x)$ from $I(x)$. This task is non-trivial as the $\psi_o(x)$ function is one of three additive terms operated on by a cosine function, which is itself coupled to a spatial reflectance term. A number of methods can be used to extract the argument of the cosine function, including Fourier Filtering [19, 20] and Phasor-based techniques [17]. Due to the periodic nature of the cosine function the argument cannot be extracted uniquely, and the quantity that can be extracted is $(2\pi f x + \psi_b(x) + \psi_o(x)) \mod 2\pi$. This must then be “unwrapped” in order to extract $2\pi f x + \psi_b(x) + \psi_o(x)$ which can then be used in a height calculation.

Phase unwrapping is a procedure that has applications in a number of fields, and so has been the subject of much research [21–24]. The largest class of techniques examine a wrapped phase image and attempt to determine locations at which a phase wrapping event has occurred (where phase in adjacent pixels changes from a value near $\pi$ to $-\pi$, or vice versa). Once these locations have been determined, and a reference pixel has been selected for which it is assumed that no phase wrapping has occurred, then offsets can be added to regions isolated by phase wrapping events to correct for the lost multiples of $2\pi$. The unwrapping process moves outward from the reference pixel in an iterative manner, so that the correction applied to a pixel is derived from an adjacent pixel which has already been phase unwrapped [24]. The first limitation of these techniques is the necessity of distinguishing phase wrapping events from real large changes in phase, which is further complicated by the presence of measurement noise. Secondly, a large height gradient may result in a change in phase between adjacent pixels of much greater than $2\pi$. It is not possible to uniquely unwrap phase changes of more than $2\pi$ between adjacent pixels resulting from large distances between the spatial points imaged by the pixels, and so these events necessarily create errors. Finally, due to the iterative nature of the unwrapping process, errors made during the unwrapping process propagate to affect the unwrapping of subsequent pixels.

If the structured illumination source is sufficiently flexible further images can be acquired to provide extra information for the phase unwrapping problem. For example, by projecting a series of binary images (such as Gray coded images [15]) where particular phase regions (of size $2\pi$) are illuminated it is possible to uniquely specify the degree of phase wrapping at each pixel using $\lceil \log_2 n \rceil$ extra images, where there are $n$ phase periods of size $2\pi$ [25]. Depending on the fidelity of the illumination source, however, pixels that lie on the edges of phase periods may remain ambiguous due to blurring of the encoding boundaries. It is also possible to acquire the wrapped phase at different frequencies, which provides additional information which can be used in phase unwrapping. In the extreme case, imaging at a frequency less than one phase period over the entire projected area produces a phase map that is always unwrapped, as the maximum phase change is less than $2\pi$. While this phase map may be too noisy to use for the intended application, it can be used in the unwrapping process of a higher frequency. The use of one frequency or several frequencies to aid the unwrapping process is called temporal phase unwrapping [23].

Once an unwrapped phase map has been extracted successfully the final step in the process is to convert the phase, and knowledge of the properties of the measurement system, into spatial coordinates. Standard methods use an approximation that places the projector at infinity in order to simplify the model of the system, and require the object phase $\psi_o(x)$ as an input to the process. The coordinates of camera pixels imaging the background plane are another prerequisite, and must be obtained prior to imaging. If the phase extracted by the previous processes results in $2\pi f x + \psi_b(x) + \psi_o(x)$, then it is necessary to first separate $\psi_o(x)$ from $2\pi f x + \psi_b(x)$. This can be accomplished by imaging the scene without the object of interest, and then subtracting the resulting phase map from the original one. Using the projector approximation, the height of the object from the background, $h(x)$, can be expressed as a function of the phase,
\[ \psi_{o}(x), \] the height of the camera and projector pupils from the background plane, \( l \), the distance between the camera and projector pupils, \( d \), and the frequency of the pattern being projected, \( f \) [26]:

\[ h(x) = \frac{l\psi_{o}(x)}{\psi_{o}(x) - 2\pi f d} \]  

(3)

Generalised systems where the camera and projector are placed in arbitrary positions have been investigated. Wang et al. [27, 28] rearranged the relevant equations to collect terms involving the imaging system characteristics such as projector and camera location and orientation into a number of constants, and fitted for these constants using calibration imaging, without explicitly determining or using the imaging system characteristics such as camera and projector locations and orientations. Mao et al. [29] generalised the geometry to the case where the projector and camera pupils are neither coplanar nor at the same height from a reference plane.

The system that is described here is generalised in that it allows arbitrary placement of optical components provided that their locations and orientations are known. It further enables the use of multiple mirrors to measure the entire field of view of the projectors, including regions occluded from the direct field of view of the camera, simultaneously in a common coordinate system. This allows the creation of a unified point cloud comprising the camera’s direct field of view and the fields of view visible through any mirrors without the use of point cloud registration.

2.4. Post-processing

Once a set of three dimensional points lying on the surface of the object has been collected, it may be necessary to further convert the set of points into a three dimensional surface or volume mesh to enable the use of the data for the desired application. For example, modelling of light transport in tissue can be performed using Finite Element Analysis [30], which requires a volume mesh.

2.5. Small animal imaging instruments

A diverse set of small animal optical imaging instruments are in use, a number of which include a type of surface capture system. These include photogrammetry [11, 31–34], scanning laser profilometry [10, 35], optical backprojection imaging [9, 36, 37], and structured light methods [35, 38–41]. A number of such systems are commercially available, such as those made by PerkinElmer [41], Berthold Technologies [42], and Biospace Labs [43].

A number of methods have been used to improve the field of view of the optical imaging instruments, including the use of single mirrors to shift cameras’ whole fields of view [31, 33, 38] or allow full angular view of a subject [10, 32]; the use of multiple mirrors to allow simultaneous collection of data [42, 44, 45]; and the use of a rotating stage for the imaging devices or the subject to acquire data from multiple views sequentially [35–37, 39, 40]. However, to the best of our knowledge mirrors have not been used for sinusoidal structured light surface capture.

3. Methods

The surface capture system presented here consists of a camera, several projectors that project structured light patterns onto an object, and mirrors that are used to extend the field of view of the camera to regions that are out of its direct line of sight. An example configuration can be seen in Fig. 1, but all components of the system can be placed in arbitrary positions and orientations.

The addition of mirrors to a surface capture system adds a number of complications to the acquisition process.
Fig. 1. Instrumental configuration. A camera is located above the object to be imaged. Two projectors project structured illumination patterns upon the object. Mirrors placed at either side of the object expand the field of view of the camera to include edges and concave regions of the object. Mirrors can be placed in arbitrary positions, orientations and number to expand the field of view. The black dotted lines represent light leaving the projectors and scattering off of the object’s surface. The blue lines indicate the regions of the object that are illuminated by at least one projector and visible directly to the camera. The red dashed lines indicate the regions of the object that are illuminated by at least one projector and visible to the camera through a mirror. It can be seen that there are regions of the object that are illuminated and visible through the mirror, but not directly.

1. The use of a mirror introduces a virtual camera position from which rays arriving at the mirror appear to be observed. The standard crossed-axis instrument geometry is less appropriate in this case, as its use imposes restrictions on the positions of any mirrors and requires the use of background planes placed in specific locations and orientations.

2. The imaging of two or more distinct regions on the camera adds complexity to the use of Fourier domain phase extraction techniques.

3. The addition of extra views increases the probability of imaging regions which appear to be discontinuous in height with respect to the camera’s perspective, and which are spatially separate from the other regions, which prevents the use of standard phase-unwrapping techniques.

The system addresses these issues through the acquisition of absolute phase, which is then unwrapped in a spatially independent manner, and then processed using a general geometric inversion formula to convert phase to absolute spatial coordinates given the positions and orientations of camera and projector, which can be arbitrary so long as they are known.

3.1. Instrument geometry and conversion of phase information to spatial coordinates

To simplify the treatment, assume that the camera and projector optics can be approximated as pinholes. Typical instruments use a crossed-axis configuration, within which the camera and projector pupils lie in a plane parallel to the platform on which sits the object to be imaged. The camera is oriented normal to this plane, and the projector is pointed towards the intersection of the camera axis and the platform [26]. This configuration allows geometrical simplifications to be made and under the assumption that the projector illumination is approximately planar when it reaches the object, results in a simple relationship between phase and height. In addition, imaging of the platform allows subtraction of background measurements, which simplifies the process and corrects for some artifacts resulting from the simplifications.
The system solves for absolute coordinates in three dimensions using phase and the direction of the rays entering the camera. As a result, the system is free of artifacts associated with geometrical simplifications (although still subject to lens aberration and other instrument-related artifacts). In addition, mirrors are used in order to increase the camera field of view. The use of a mirror results in an effective virtual camera placed behind the mirror. The position of this virtual camera is determined by the position of the camera and mirror, potentially invalidating the crossed-axis configuration assumption. By using absolute phase and solving for the relation between phase and position in a general optical configuration it is possible to convert phase to three dimensional spatial coordinates from a non-crossed-axis configuration. This allows mirror data to be treated in the same manner as direct data by replacing the real camera with a virtual camera (which is the reflection of the real camera about the plane of the mirror).

A generic optical configuration is shown in Fig. 2, and contains a camera with a pupil at \( c \), and a projector with a pupil at \( p \). The projector projects a sinusoidal pattern with a known spatial frequency \( f \) in a plane centred around the point \( o \) (at which location the pattern phase is zero), with \( \hat{v} \) the direction of increasing projector pattern phase. Then, the pattern projected (see Fig. 3) can be described:

\[
p(y, f) = \frac{1 + \cos(2\pi f(y - o) \cdot \hat{v})}{2}
\]

where \( y \) is a point within the projected plane.

![Fig. 2. Schematic of system.](image)

\( c \) and \( p \) are the locations of the camera and projector pupils respectively. \( o \) is a point on a plane orthogonal to the axis of the projector on which the spatial frequency, and direction as encoded by \( \hat{v} \), of the projected pattern is known. The direction of the pattern is the direction in the plane of maximal increasing pattern phase. \( \hat{r} \) is a ray originating at the camera pupil, corresponding to a camera pixel. The unknown point \( x \) lines somewhere along the line containing \( c \) and spanned by \( \hat{r} \). \( y \) is the intersection point of the plane containing \( o \) and the line between \( x \) and \( p \) (spanned by \( \hat{q} \)). For simplicity, this schematic is given in two dimensions, but this representation is also valid in three dimensions.

We desire to know the coordinates of a point \( x \). If \( \hat{r} \) is the direction of the ray from the camera to \( x \) (which can be deduced via knowledge of the camera), a ray from the projector may also intersect \( x \). Let the ray from the projector be \( \hat{q} \), and the intersection of that ray with the projector...
Fig. 3. An XPM-2 Phantom Mouse (Caliper Life Sciences, A PerkinElmer Company, Hopkinton, Massachusetts, United States of America) placed in the imaging system, utilising two mirrors, under structured illumination generated by one projector.

plane be $y$. Then:

$$y = p - \frac{|p - o|^2}{q \cdot (p - o)} \hat{q}$$  \hspace{1cm} (5)

where:

$$\hat{q} = \frac{x - p}{|x - p|}$$  \hspace{1cm} (6)

The phase, $\psi$, associated with $y$ is:

$$\psi = 2 \pi f (y - o) \cdot \hat{v}$$  \hspace{1cm} (7)

Using Eqs. (5) to (7), and the knowledge that $x$ lies on the line defined by $c$ and $\hat{r}$, it is possible to derive an expression for $x$:

$$x = c + \frac{\alpha \cdot (p - c)}{\alpha \cdot \hat{r}}$$  \hspace{1cm} (8)

where:

$$\alpha = \frac{\psi}{2 \pi f |p - o|} + |p - o| \hat{v}$$  \hspace{1cm} (9)

Once the phase values measured by the instrument have been converted to spatial coordinates, a mesh of the object can then be generated.

### 3.2. Extraction of wrapped phase

Fourier domain techniques are unsuited for extracting phase from disjoint regions containing projected patterns with different fundamental frequencies. Consequently it becomes desirable to extract the phase of spatial patterns at each spatial point independently of any other spatial points, especially in the case where it is expected that there exist spatial discontinuities. This is possible through the projection of multiple sinusoidal patterns offset at a number of phases for each spatial frequency [26].

The projected sinusoidal spatial patterns, $p_n$, at spatial frequency $f$ and phase offset $\phi_n$ can be expressed:

$$p_n(y, f) = \frac{1 + \cos(2 \pi f (y - o) \cdot \hat{v} + \phi_n)}{2}$$  \hspace{1cm} (10)
The imaged patterns, \( g_n \), take the form:

\[
g_n(x, f) = b(x) + a(x) \frac{1 + \cos(\psi(x, f) + \phi_n)}{2}
\]  

(11)

where:

\[
\psi(x, f) = 2\pi f(y - a) \cdot \hat{v}
\]  

(12)

and \( y \) is a function of \( x \). It is not possible to extract \( \psi(x, f) \) from \( g_n(x, f) \) directly, but it is possible to extract \( \psi(x, f) \) \( \mod \ 2\pi \) using the following equation:

\[
\psi(x, f) \mod 2\pi = -\arctan\left(\frac{\sum_{n=1}^{N} g_n(x, f) \sin(\frac{2\pi n}{N})}{\sum_{n=1}^{N} g_n(x, f) \cos(\frac{2\pi n}{N})}\right)
\]  

(13)

The phase maps extracted are termed “wrapped” because they are unique up to a constant, which is an unknown multiple of \( 2\pi \). Examples of wrapped phase maps can be found in Fig. 4.

3.3. Phase unwrapping

Standard phase unwrapping techniques use the spatial distribution of wrapped phase to determine locations of phase wrapping, and then correct accordingly. Each line of phase wrapping signifies that the phase for a spatial region must be increased or decreased by \( 2\pi \). Accordingly, errors in the determination of phase wrapping events propagate through spatial regions. It is also necessary to know a priori the phase of one pixel in the image so that it can be used as a reference. Without this, only relative phase can be determined. In addition, it is not possible to unwrap measurements suffering from phase wrapping that results from phase discontinuities of \( 2\pi n \) where \( n > 1 \), as these discontinuities are degenerate in the wrapped phase representation. An advantage of these techniques however, is that they require only a single wrapped phase map to unwrap (which itself can be calculated from a single image).

Given the likelihood of locally poor phase data and large discontinuities, it is preferable to unwrap the phase of each spatial point independently, which requires more phase data. Our system acquires a number of wrapped phase maps at different spatial frequencies and applies temporal phase unwrapping, which has previously been used in conjunction with phase shifting measurement techniques (for example, see Li et al. [46]). The spatial frequencies can be defined as:

\[
f_n = \frac{n}{d}
\]  

(14)

where \( n \) is the number of sinusoidal periods across the field of projection, and \( d \) is the size of the field of projection in a plane of interest. The maximum phase difference for the spatial frequency \( f_1 \) is \( 2\pi \), and so phase wrap events cannot occur when using this frequency. The wrapped phase measurement error is constant in average absolute magnitude independent of frequency, and so at low frequencies is relatively large in comparison to the range of phase values. Consequently, it is undesirable to reconstruct surface geometry at \( f_1 \). However, it is possible to use this as prior information to aid in unwrapping phase maps for higher spatial frequencies because the phase at a spatial point \( x \) is linearly dependent on the spatial frequency.

For example, by acquiring at frequencies \( f_1 \) and \( f_2 \) we can obtain \( \psi(x, f_1) \) and \( \psi(x, f_2) \) \( \mod \ 2\pi \). Due to the linear relationship, we know that \( \psi(x, f_2) = 2\psi(x, f_1) \) in the absence of measurement noise. In the presence of measurement noise, we can estimate \( \psi(x, f_2) \) as:

\[
\tilde{\psi}(x, f_2) = 2\psi(x, f_1)
\]  

(15)

where \( \tilde{\psi}(x, f_2) \) is the estimate of \( \psi(x, f_2) \). By then simulating the wrapping process on \( \tilde{\psi}(x, f_2) \), it is possible to compare \( \tilde{\psi}(x, f_2) \) \( \mod \ 2\pi \) and \( \psi(x, f_2) \) \( \mod \ 2\pi \) and correct \( \tilde{\psi}(x, f_2) \) accordingly for the difference and so calculate \( \psi(x, f_2) \). It is possible that error in \( \tilde{\psi}(x, f_2) \) or
ψ(x, f_2) (mod 2π) may be sufficient to remove a phase wrapping event or add one extra. In these cases it is necessary that the spatial frequencies chosen are close enough that it is possible to identify and correct for this.

By imaging at multiple frequencies and unwrapping between adjacent frequencies in order of increasing frequency, it is possible to produce an unwrapped high frequency (and so possessing small relative error) phase map in a spatially independent manner. An example of this can be found in Fig. 4.

3.4. Point cloud processing

After unwrapping the high frequency phase map, and using the geometric relations between projectors and real or virtual cameras to convert phase to spatial coordinates, the resulting point cloud may be processed further to facilitate visualisation and physical modelling. Conversion of the point cloud into a mesh format allows easier visualisation and may be necessary for physical
modelling. Processing and visualisation of point clouds in this paper were performed using Matlab (The MathWorks, Natick, Massachusetts, United States of America) and Meshlab [47].

3.5. Experimental instrument and test cases

The experimental instrument used in this paper has been described in Guggenheim et al. [2]. The components utilised for surface capture are:

- one C9100-14 ImageEM-1K camera (Hamamatsu Photonics K.K., Hamamatsu City, Japan), which is used in both the surface capture and bioluminescence imaging, but was chosen due to the high sensitivity requirement of bioluminescence imaging
- two Pocket Projector MPro120 (3M United Kingdom, Berkshire, United Kingdom) units to project the patterns used in the surface capture process
- one L490MZ Motorised Lab Jack (Thorlabs, Ely, United Kingdom), which is utilised in the surface capture system calibration procedure, and is also used in bioluminescence imaging
- one NT59-871 25mm Compact Fixed Focal Length Lens (Edmund Optics, York, United Kingdom), which is also used in bioluminescence imaging
- two N-BK7 75mm Enhanced Aluminum Coated Right Angle Mirrors (Edmund Optics, York, United Kingdom), which can be placed in any suitable position within the surface capture system, and are also used in bioluminescence imaging
- one FB580-10 Bandpass Filter (Thorlabs, Ely, United Kingdom), which is used to attenuate the projector signal to prevent saturation of the camera

Control and operation of the system is performed using LabVIEW (National Instruments, Austin, Texas, United States of America) and Matlab (The MathWorks, Natick, Massachusetts, United States of America).

The imaging process on the above instrument requires approximately 240 seconds, and approximately 60 seconds are required for the image processing and surface reconstruction, which are executed on a standard desktop computer. Imaging time could be greatly reduced through the use of a video-rate camera, and efficient synchronisation of pattern projection and imaging. As imaging time was not a significant concern with this instrument and application, the system was not optimised for imaging time. Calibrating the system may require a significant amount of time, but system calibration only needs to be performed once.

The camera and projectors are pre-calibrated using a geometric approach. Surface capture data is acquired for each of the two projectors sequentially. A total of 14 pattern frequencies are used, consisting of 0.78, 1.1, 1.6, 2.2, 3.1, 4.4, 6.3, 8.8, 13, 18, 25, 35, 50, and 70 waves per projected pattern. This frequency set has not yet been optimised as imaging time has not been a significant concern in the use of the prototype instrument. In order to measure the phase at each frequency, 6 phase shifts are used. The only exception to this is the highest frequency in the second test case (that of the two pence coin) for which 24 phase shifts were used.

To illustrate the performance of the system, several test cases are shown. The first of these involves an XPM-2 Phantom Mouse (Caliper Life Sciences, A PerkinElmer Company, Hopkinton, Massachusetts, United States of America) which is a plastic mouse-shaped phantom designed for use in BLT imaging. This demonstrates the performance of the system when imaging an optically homogeneous object. Quantitative analysis was performed by calculating the distances between points on an ideal mesh of the XPM-2 phantom, provided by the manufacturer, and the closest points to them forming part of a mesh created from the surface capture point
cloud. The full set of ideal points was first manually processed to remove those points judged not to lie within the region of the XPM-2 phantom imaged by the surface capture system. This reduced the number of ideal points from 1503 to 878 points. The ideal point set was then registered to the surface capture data using rigid transformations (translations and rotations), and then the measurement error in the surface capture system was calculated by taking the smallest distances between the ideal points and a mesh created from the surface capture point sets using Meshlab.

The second test case involves a United Kingdom bronze two pence coin. The two pence coin used has a diameter of 25.9mm and a maximum thickness of 1.85mm, with an embossed surface. This demonstrates the ability of the system to resolve small features on the order of hundreds of micrometres. For this test a greater number of phase shifts (24) was used to measure the final frequency of the phase map in order to maximise accuracy.

The third of these involves a sacrificed mouse and demonstrates the performance of the system when imaging a realistic subject.

4. Results and discussion

Figures 5 and 6 show a surface capture of an XPM-2 Phantom Mouse (Caliper Life Sciences, A PerkinElmer Company, Hopkinton, Massachusetts, United States of America), which is of similar size to the intended imaging subjects of the dual modality instrument. This surface cap-

![Image](a) Photograph of XPM-2  (b) Mesh of XPM-2  (c) Ideal points used for error calculation  (d) Error between ideal points and surface capture point cloud

Fig. 5. Surface capture of an XPM-2 Phantom Mouse. The phantom (see Fig. 5(a)) was imaged using the surface capture system, generating a point cloud which was then converted into a surface mesh using MeshLab [47] (see Fig. 5(b)). A collection of points lying on the XPM-2 Phantom was kindly provided by Caliper Life Sciences (see Fig. 5(c)) and used to estimate the measurement error in the surface capture acquisition (see Fig. 5(d)).

The current task involved data acquisition using two projectors and two mirrors. The mirrors provide a significant increase to the field of view and allowed almost all of the surface of the subject to be recovered, except for the lower and underside of the jaw (see Fig. 6). The addition of the mirrors improves the surface area coverage by 15.2%. The degree of improvement will in general depend on the geometry of the imaging subject and the placement of the mirrors. The
Fig. 6. Surface capture coverage of an XPM-2 Phantom Mouse. The addition of mirrors increases the surface area imaged by 15% over the surface area imaged directly.

Inability to image the lower side of the jaw was due to the insufficient length of the mirrors and so could be remedied with the use of longer mirrors. The underside of the jaw was obscured by the imaging platform and so could not be imaged without modifications to the experimental configuration (such as the addition of more mirrors to redirect projector light to the underside of the jaw).

The surface capture process resulted in tens of thousands of surface points acquired from six views (a direct view and two mirror views for each projector). These points were then converted into a high quality surface mesh using MeshLab [47].

The surface mesh was compared with 878 ideal surface points provided by the manufacturer, and ideal surface points were found to lie on average 0.14mm from the surface capture mesh, with a standard deviation of 0.14mm (see Fig. 5(d)). A small number of outliers were found to lie up to 1.3mm away from the surface capture data. 50% of points were found to lie within 0.1mm of the surface mesh, 82% of points were found to lie within 0.2mm of the surface mesh, and 96% of points were found to lie within 0.4mm of the surface mesh. The extreme points may result from errors in the surface capture meshing (the ears in particular are complex and concave and so the surface capture system provides incomplete coverage of this region, which may result in meshing errors).

Figure 7 shows a surface capture of a two pence coin and demonstrates that the system is sensitive to small changes in height. The relief features on the coins are on the order of tens to hundreds of micrometres.

The imaging performance was sufficient to capture the primary features of the design of the coin including the presence of text, despite the small size of the features. The system was not modified in order to image the coin and no extra lensing/magnification was applied, resulting in the image of the coin covering a small portion of the camera CCD. The addition of appropriate lensing would allow higher resolution measurement to the extent that the text may be rendered readable. Vertical and to a lesser extent horizontal line artifacts can be observed in the reconstructions, and may be reduced by improving the projector calibration process, or choosing a projector with a larger depth of field.

The data shown in Fig. 7 is from a single view (one camera, one projector, and no mirrors). Multiple views were not integrated because calibration errors between multiple views were significant enough to de-align the data and obscure the fine detail, which is of a scale on the
Fig. 7. Surface capture of a two pence coin. The surface capture data was acquired using one projector and no mirrors. The colour maps in Figs. 7(b) and 7(d) indicate height in millimetres. Note that no texture was applied to the renderings in Figs. 7(b) and 7(d).

order of hundreds of micrometres. The inclusion of a model of the aberrations present in the optics of the camera and projectors, or a more accurate system calibration, may reduce this, along with the line artifacts.

Figures 8 and 9 show a surface capture of a mouse cadaver. The projector coverage of the mouse can be seen in Fig. 8. The two projectors provide complementary information as they illuminate different regions of the subject. The two mirrors also provide complementary infor-
Fig. 9. Point cloud of a mouse cadaver (see Fig. 8). All views using projector 1 were acquired simultaneously, as were all views using projector 2. The mesh in Fig. 9(b) was created from a sub-sampled version of the point cloud in Fig. 9(a).

mation, although for each projector there is one mirror that provides little new information. This is a result of the projectors being placed at a significant angle to the vertical axis, resulting in shadowing which adversely affects the mirror opposite to it. However, the projectors are positioned so that each projector illuminates part of the shadowed region of the other projector, and thus increasing the total field of view. The resulting phase can be converted into absolute spatial coordinates, the sets of which form a point cloud for each view. These point clouds can be seen in Fig. 9(a). Here it is more obvious that the different views are providing complementary information. The point clouds were combined without the need for registration (as each provides absolute coordinates in a common coordinate system) and used to produce a mesh, which can be seen in Fig. 9(b). Areas of almost solid colour indicate regions of very dense points. It can be seen that the mesh generation algorithm has fused parts of the back legs to the rear and tail, and artifacts are also present at the ears, where there are fewer points. This is undesirable but indicate a challenge regarding the post-surface capture processing, rather than the imaging process itself.

5. Conclusions

The imaging system described in this paper uses a general geometric model and mirrors to allow simultaneous acquisition of multiple views of an imaging subject using a single camera, limited only by the fraction of the imaging subject illuminated by structured light patterns. It produces high density and accurate measurement of structures on the order of centimetres down to hundreds of micrometres. The general imaging technique is applicable to objects on
even larger scales through the use of appropriate lensing and sufficiently bright projectors.

The generalised model used for phase-to-spatial coordinate conversion allows flexible instrument design, and the addition of new optical components or modification to existing components without recalibration of the entire imaging system, provided that the new orientation of these components is known. The instrument used in this paper contains a number of expensive components as these are used in BLT measurements. However, it is possible to construct a standalone system with comparable performance using inexpensive components, which at a minimum may be a single consumer webcam and projector with a combined cost of less than £200.

The system is capable of imaging surfaces with large discontinuities and holes due to the method of phase acquisition, enabling measurement of complex surfaces and thus providing applicability to a variety of applications.

The current system is challenged by objects possessing surfaces which are specularly reflective, and by objects where secondary reflections are significant, such as translucent objects and objects possessing particularly concave regions. Addressing the issue of separating primary from secondary reflections is a direction for future work.

Additionally, the acquisition of all images necessary for phase extraction requires tens of seconds, and this precludes the imaging of moving objects. Optimisation of the phase measurement to allow the measurement of dynamic objects is another direction for future work.

Acknowledgments

Funding for this work was provided by EPSRC grant EP/F50053X/1 through studentships to Hector Basevi and James Guggenheim at the PSIBS Doctoral Training Centre at the University of Birmingham, and the University of Birmingham Capital Investment Fund. The authors would like to thank Laura Morton, David Millar and Mark Cobbold (University of Birmingham School of Immunity and Infection) for their assistance in the mouse cadaver imaging study. This work has a pending UK patent application (application number 1222361.6, filed on the 12th of December 2012).