

Generating 3D Models of Paintings Through the Combination of 2D, 3D and RTI Data

Xavier Aure
Centre for Fine Print Research
University of the West of England
School of Art and Design
Bristol UK
xavier2.aurecalvet@live.uwe.ac.uk

Paul O'Dowd
Centre for Fine Print Research
University of the West of England
School of Art and Design
Bristol UK
paul3.o'dowd@uwe.ac.uk

Joseph Padfield
The National Gallery London
Trafalgar Square
London
UK
joseph.padfield@ng-london.org.uk

The National Gallery in London has recently been testing the potential of 3D scanning technology to record and measure the surface of paintings. To view and interact with the high-resolution scans requires expensive computational hardware. The proposed workflow borrows some of the techniques used in the gaming industry to provide a computationally efficient interactive interface, even suitable for online viewing. The workflow synergises multiple imaging techniques, and therefore provides better texture representation than 3D scanning alone. The process provides a new way of visualising possible relations across paint layers by combining normal maps with existing image based techniques.

3D laser scanning. 3D models. 3D visualisation. Normal maps. Reflectance transformation imaging. Paintings.

1. INTRODUCTION

The National Gallery in London has recently been testing the potential of 3D scanning technology to record and measure the surface of paintings, using a Lucida laser scanner custom built by Factum Arte. The Lucida software outputs 3D information as 2D grayscale depth-map files. These files can be used to generate highly detailed 3D triangular meshes of the textured surface. However, to view these meshes in an interactive real-time format requires costly computational hardware, which is likely to be cost-prohibitive even within important heritage institutions.

The proposed workflow borrows some of the techniques used in the gaming industry to provide a computationally efficient interactive interface to highly detailed 3D models of the surface of paintings. A number of normal maps, generated with the Reflectance Transformation Imaging (RTI) software, are stitched together and incorporated into the 3D models. The normal maps create the illusion of a highly detailed texture without the strain of a multimillion polygonal mesh.

The workflow (Figure 1) exploits the intrinsic potentialities of each individual imaging technique to facilitate the efficient visualisation and interaction of

captured surface texture without compromising on the quality of the 3D model. Furthermore, the synergistic combination of 2D colour images, RTI and 3D data generates high-resolution coloured 3D models of paintings that remain computationally efficient enough for online viewing on ordinary devices.

We also describe briefly a simple method for combining the mosaicked normal maps with existing Infrared, UV and X-ray images of a painting to produce new sets of interactive RTI images. This final consideration allows for further deeper understanding of painting composition useful within heritage.

2. SET-UP

The basic requirements include a 3D laser scanner (Figure 2) and a custom built RTI dome with a digital camera (Figure 3). Although the presented method makes use of some specific hardware, the techniques and ideas behind it are transferrable to other types of equipment and set-up. This means other forms of 3D scanners might be used to capture 3D data and the more traditional RTI method with a camera, a single light source and a reflective sphere might be used to produce the RTI images.

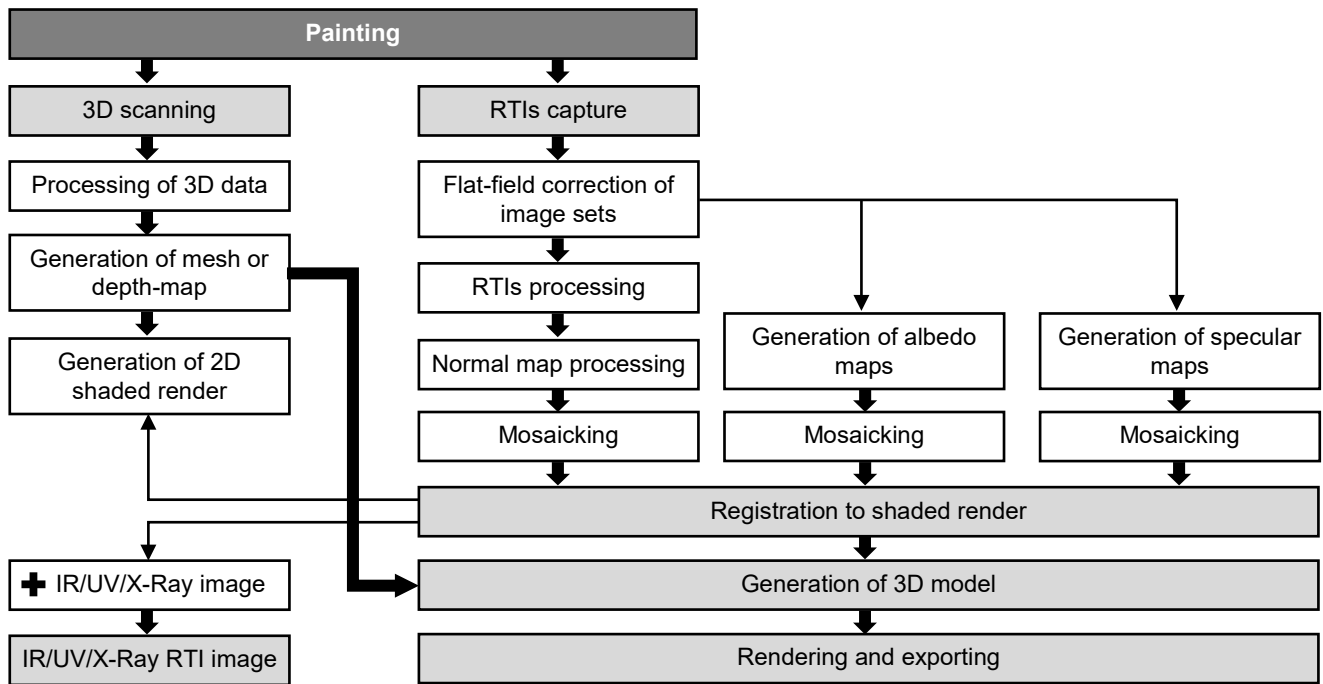


Figure 1: Workflow steps followed to generate the 3D models of paintings

2.1. Hardware

2.1.1. 3D scanner

The paintings are recorded with the Lucida 3D scanner, a high-resolution 3D laser scanner specifically design to record the surface of paintings and low relief objects. The Lucida projects a beam of red light (650nm) onto the surface. The deformation of the beam is recorded by two monochrome video cameras as it rakes across the surface. The files are saved as raw black and white video data and processed as a tone depth-map. The scanner records individual tiles which are then edited, stitched and merged to create a whole record of the surface at a resolution of 254 DPI. After scanning and processing the 3D data, a final 32-bit grayscale depth-map file is obtained. The different grayscale tones in the file define the depth of each pixel recorded of the surface of the painting.

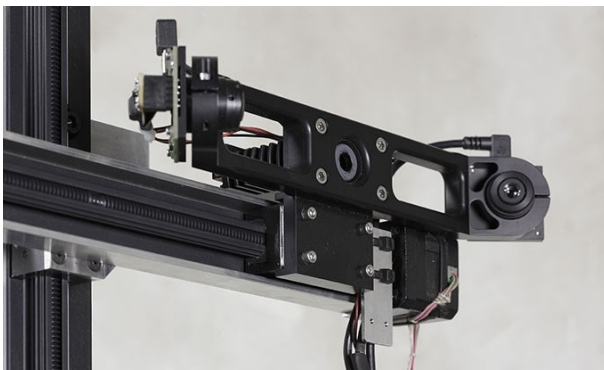


Figure 2: The Lucida 3D scanner dual camera system
(Source: Roglab, 2014)

2.1.2. RTI system

The dome system used to produce the RTI images comprises a 700 mm acrylic hemisphere with 90 Cree® XLamp LEDs (Table 1). The LEDs are geodesically distributed around the dome in 6 horizontal tiers.

Table 1: LEDs specifications

Colour Temperature*	3852K
Luminous Flux/Radiant Flux	308 lm
Colour Quality Scale (CQS)*	94.2
D_{uv} *	0.0013
If - Forward Current	400mA
Viewing Angle	115deg
Vf - Forward Voltage	9V
*Measured at the National Gallery with a GL Optic Spectis Touch 5.0 spectroradiometer.	

A low cost electronic solution automates the sequence of 90 illuminated photographs of the painted surface. An Arduino waits for a button press from the user. When activated, the Arduino drives a chain of twelve shift registers. The shift register provides a means to send a signal to 90 different LEDs. The shift register, a TPIC6B595, is of interest because it operates as open drain, and is therefore capable of handling the relatively high power LEDs. The LEDs draw power in turn from a constant current source supply to ensure consistent illumination. As each LED is illuminated, the Arduino also powers a relay to activate the remote trigger input of a digital camera and take a photograph. The circuit, bill of materials and Arduino software is freely available online (O'Dowd, 2017).

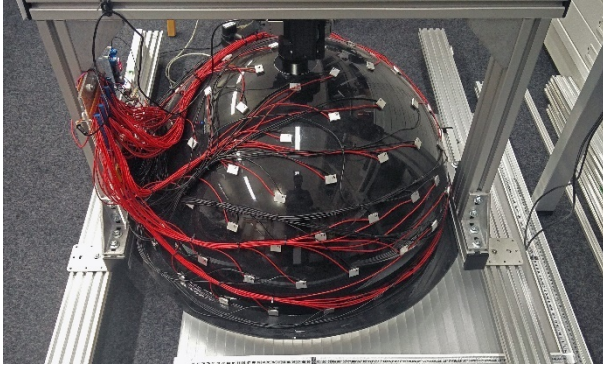


Figure 3: The RTI dome

The camera used is a high resolution digital camera (16.3 MP sensor) fitted with a 50mm lens which produces images of 4896 x 3263 pixels with a resolution of about 660 DPI. The camera is mounted on a slider rail fixed to the aluminium frame that holds the dome. The dome is mobile and can be used horizontally with the painting positioned underneath or it can be fixed to a camera stand with the painting placed on an easel. The system can be configured to shoot with any standard digital camera. Upgrading the camera increases the spatial resolution of surface details.

2.2. Software

To process and render the data a variety of different applications are used. Because the intention is to allow other practitioners to replicate the workflow, all software is off-the-shelf. The software is grouped into four parts:

- 3D scanning and processing: Lucida custom software is used during capture and processing of 3D data. PTGui and Photoshop are used for stitching the various tiles.
- RTI capture and processing: CHI RTIBuilder is used to process the images captured with the dome, and the CHI RTIViewer is used to open the RTI files and save the normal maps generated.
- Image processing: Knald and Photoshop are used to process the normal maps, Fiji to process the albedo and the specular maps. All these image textures are stitched with PTGui. The various composite outputs are registered with Nip2 and Photoshop.
- 3D modelling and render: Blender is used to convert the depth-maps into 3D meshes, map their corresponding image textures and produce the final 3D model ready for visualisation and exporting.

3. WORKFLOW

The proposed workflow breaks the problem of efficient visualisation into four components: a 3D model, fine surface details, an albedo map and a specular map.

The 3D models are produced by using the 3D data only to generate a relatively low computationally intensive polygonal mesh. This mesh provides the model with the overall 3D shape sufficient to show the possible deformations of the canvas or wooden support and enough detail to display the thickest impasto on the surface of the painting.

Finer surface details are provided by high resolution normal maps generated from Reflectance Transformation Imaging data. Normal maps in computer graphics are 2D images used to compute the angle at which light reflects off the surface at per pixel level (Figure 4b). Each pixel in the image stores the X, Y, Z coordinates of that angle in relation to the point of origin at the surface of that pixel. The coordinates are represented by RGB values. With this information, a 3D rendering engine can create the illusion of highly detailed textures when the image illumination is computed, and provides an effective visual compensation for a relatively low-polygon mesh.

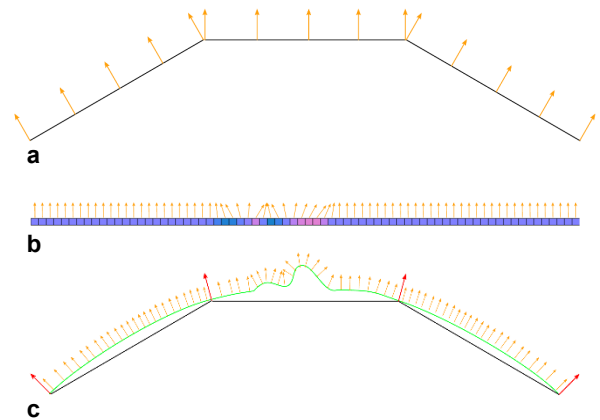


Figure 4: a) Normal vectors in a polygon mesh. b) Normal vectors at each pixel in a normal map. c) Combination of vectors from a mesh and normal map
(Source: Unity 3D, 2017)

The painting (Figure 5) is captured in several RTI images and the resulting normal maps from each section are then processed and stitched together to form a composite normal map of the complete surface of the painting.

A mosaicked albedo map, a 2D texture containing the colour information without specular highlights or shadows, is also generated from each group of 90 images for each RTI section. The same images are used to produce a greyscale specular map of the

painted surface, a 2D texture used to simulate how shiny a surface is. The resulting normal, albedo and specular maps are registered and mapped to the 3D data to produce the final 3D/2D model hybrid.

The final model can be both interacted with instantly via the 3D software environment or it can be rendered using current physically based rendering standards to produce 2D images or videos of the painting. The models can also be exported for online visualisation and interaction.



Figure 5: Test painting used in this workflow

3.1. 3D scanning and processing

The Lucida is mounted onto x, y and z axis motorised linear guides and can either be fixed to a camera stand or a movable lightweight aluminium structure. The painting is placed parallel to the scanner on an easel at about 10 cm. The scanner records tiles of 48 cm x 46.8 cm with a 5 cm overlap. Either the painting or the scanner are moved from one section to the next until all areas are recorded. The depth of field is limited to 2.5 cm, but the software allows re-scanning of specific regions if distorted supports or high relief make it necessary. The scanning process is controlled from a portable computer allowing laser intensity and exposure to be increased or decreased depending on the characteristics of the surface being recorded.



Figure 6: Depth-map of the test painting

Each tile is processed individually to remove spikes, noise, rain and other possible artefacts and re-scanned areas are merged together. The software outputs a 32-bit depth-map image for each tile. These are then stitched (x y overlap) and blended together (adjustment of z difference) to produce a final complete file (Figure 6). This depth-map can be imported into a 3D software to generate both a 3D mesh or 2D renders.

3.2. RTI capture

The paintings captured so far have been placed flat under the dome. We used a simple lightweight board moved along x and y linear guides. This system is simple and affordable and avoids unwanted rotational movements of the painting while moving from one section to another. However, in the future paintings will be captured vertically as this avoids possible sagging of the canvas resulting in a fake deformation of the support. The ideal set-up will include the National Gallery computer-controlled motorised easel which allows the paintings to move with repeatable accuracy while capturing the various RTI tiles.

The light position (LP) file necessary to process the images with the RTIBuilder is produced only once with the reflective black sphere method described in Cultural Heritage Imaging (2017). The sphere is placed perpendicular to the camera in the centre of the field of view. The camera captures the specular highlights reflected on the sphere, these are used by the software to locate the light sources in space. The same LP file is used to process all sections of the painting.

The camera is set to manual mode throughout each capture sequence with constant focus, aperture and shutter speed. Each painting will require slightly different settings depending on the type of surface being captured. Once the optimal settings are selected the Arduino is triggered and the sequence starts. The LEDs *on time* is configured to allow sufficient time for the camera to save the file before it triggers the next. Images are saved in RAW format and postprocessed to output JPG images as this is the only format RTIBuilder recognises.

Once the first set of 90 images is captured the painting is moved to the next position with a 10 per cent overlap to allow stitching. The sequence is repeated for each tile or section until the whole painting is photographed.

3.2.1. RTI processing

To process the various RTI images we follow the steps in Cultural Heritage Imaging (2017) while implementing a flat-field correction method before processing. If required by the camera and lens other image correction techniques may need to be

applied. Colour correction can be incorporated during RTI capture and processing.

3.2.2. Flat-field correction

In view of the work of Huang et al. (2015) the images captured with the RTI dome are corrected for illumination variance to generate higher quality normal maps. This step corrects for the RTI software far-light assumption which assumes the object in a dome is illuminated from the same angle and with the same intensity across the entire field of view.

To compensate for this error in illumination, a correction image set consisting of 90 images of a flat white surface is initially captured with the dome. The images are applied a median filter with Fiji to remove any imperfections of the surface while still retaining light intensity changes across the surface. These images are used in Nip2 to apply a flat-field correction to each individual RTI group of 90 images. Once all images are corrected they are processed with the Dome LP file (HSH Fitter) option within the RTIBuilder.

3.3. Normal maps processing

Each individual normal map generated in the RTIViewer is saved in PNG format (Figure 7). Before stitching, the normal maps need to be processed to remove low frequency components. This step is necessary to correct for normal discontinuities across the tiles and avoid waviness on the final normal map composite.

This process is done with Knald, a 3D software application used to generate textures for computer graphics visualisations. The software applies digital frequency filters to the image data for a specifiable number of iterations to produce a variety of high quality texture maps for 3D rendering engines.

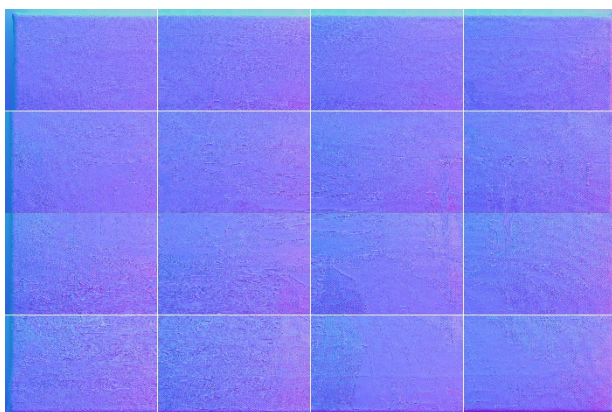


Figure 7: All normal maps before processing

The maps are processed individually using the same settings across all images. A low number of iterations will only retain high frequency components which are the surface texture details we need to

keep (Figure 8). For paintings with medium to high impasto 20 to 50 iterations are usually sufficient. Paintings with low relief and gilded areas require lower iterations from 2 to 5.

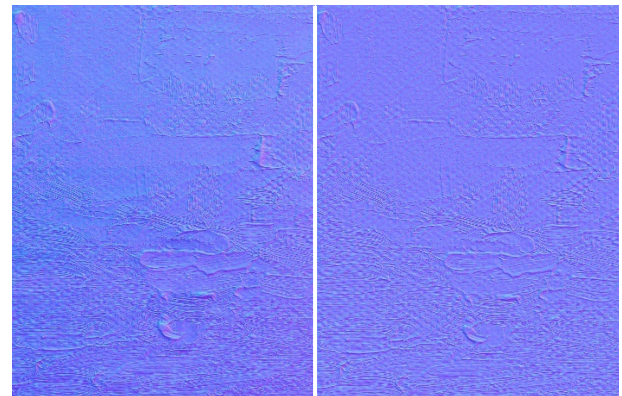


Figure 8: A normal map detail before (left) and after processing (right)

3.4. Normal maps mosaicking

The processed normal maps are uploaded to PTGui. We ignore the EXIF data, and instead select the rectilinear lens option with a focal length set to infinite (around a value of 20,000) and a flat rectilinear projection in the panorama settings. This allows mosaicking the images with no perspective corrections. The software will attempt to align the images automatically and correct for small rotations if necessary. If unsuccessful, additional matching point pairs can be assigned manually until all tiles are properly aligned.

The final mosaic is exported in PSB format which allows files of large size. In Photoshop the mosaic is cropped to the edges and save as a PNG file.

3.5. Generation of the albedo maps

The same groups of RTI images are used to produce the albedo maps (Figure 9). This is an important step before producing the 3D models because we need those highlights and shadows to be rendered from our normal map and mesh.

To generate the albedo maps we import the first sequence of 90 images into Fiji, and under image stacks, we run a median projection filter. The histogram of the resulting image might need some adjustment. The same process is then repeated for all the remaining groups of images.

To generate the albedo mosaic we replace the normal map images in PTGui with the albedo ones, the software will align the images with the same settings and control points. We export as PSB and crop to the exact same size as the normal map.



Figure 9: Detail of the albedo and specular map.

3.6. Generation of specular maps

The specular maps generated approximate the brightness of the specular highlights of the painting (Figure 9). At this point the glossiness value is ignored, but will be later estimated with Blender's own roughness input.

To produce the specular map, the same groups of RTI sequences are imported into Fiji and converted to 8 bit greyscale. The brightness and contrast range are then adjusted for each group by reducing the minimum range, so only the brightest highlights are displayed. Finally, under image stacks, a maximum intensity projection filter is applied. The resulting images are then stitched with PTGui as described earlier.

3.7. Registration of texture maps

Images captured with different techniques need to be registered to correct for geometrical distortions and spatial resolution (Fontana et al., 2003). Here, we want to match our high-resolution mosaicked maps to the 3D geometry to produce accurate 3D models. We use measurements from the 3D data as ground truth, but rendered as projected 2D shaded images. Because the Lucida scanner outputs a depth-map we only need to import it into Blender to render it as a shaded image from an orthographic lens camera view. If we had a triangular mesh or point cloud from a different source of 3D data, the same shaded image could be rendered out. Due to the difference in spatial resolution we need to render the shaded image to a similar resolution to the mosaicked maps so we do not down-sample the images when registering.

To register the images we use a similar method described by Saunders and Cupitt (1995). We use Nip2 to carry out a perspective match transformation to correct size, translation, rotation and perspective distortion differences between the first pair of images. It is useful to do the first registration with the normal map in monochrome colour space, then replace it with the other texture maps to apply the

same transformation. If the registration needs tweaking, Photoshop allows for multi-layer free transformations with smart objects, and then exporting individual layers as files.

Once the images are registered to the shaded render, the depth-map, which has a smaller spatial resolution, needs to be cropped to proportionally match the other texture maps. We do not match the depth-map to the size of the other images as this would modify the 3D information.

3.8. Generation of 3D models

To generate the 3D model, we first need to extract the 3D information and then integrate our 2D texture maps into the 3D model.

To produce a triangle mesh, the depth-map is imported into Blender as a 2D image plane. The plane is subdivided into a number of polygons and then displaced on the z axis according to the greyscale information (Figure 10b). To keep the model manageable the plane is only subdivided about half a million polygons. Bigger paintings might require further subdivisions, in that case the mesh is subsequently decimated to reduce the polygon count. The final number of polygons should provide the model with enough overall detail to display the deformations of the painting support and the thickest brushwork on the surface. The finer surface details will be later provided by the normal map.

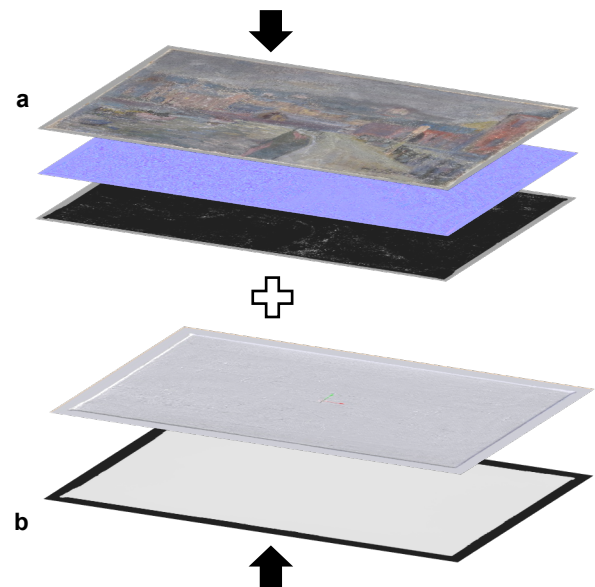


Figure 10: Generation of the 3D model: a) mapping of the image textures to the mesh and b) subdivision and displacement of the depth-map to produce a mesh

To integrate the 2D images onto the geometry we use a simple mapping orthographic projection method. The 3D mesh is first flattened onto a 2D (UV) coordinate space which has the exact size and

resolution as our texture maps and is proportional to the depth-map from which we extracted the mesh. The albedo, specular and normal map are then mapped onto the UV space. Because we have previously registered all the images to the shaded render the 3D data is perfectly aligned with the 2D texture maps. This method provides excellent results for quasi-planar surfaces like paintings even though some stretching and interpolation of the images may occur on surfaces with high relief.

3.9. Rendering and exporting the 3D models

To achieve a photorealistic rendering of the 3D model, Blender's Cycles engine takes into account all the texture maps, the polygonal mesh, the lights and the environment background in the scene. These assets are defined by a network of nodes and are assigned specific values and properties. Our texture maps are connected to a series of

bidirectional scattering distribution function (BSDF) shader nodes that describe how light is reflected at the surface of the painting and allow the 3D model to be rendered realistically.

The node set-up for each specific painting may vary depending on the type of surface we are trying to render. For instance, metallic surfaces such as gold leaf will need special input textures that describe the reflectance values of the surface.

Once the 3D model is ready it can be both interacted with in real-time within the 3D software or it can be rendered to produce 2D images or videos of the painting. Blender allows exporting the model in different 2D and 3D formats that can be used in other applications and 3D suits. There are also several add-ons that prepare the models for exporting to online platforms such as Sketchfab.

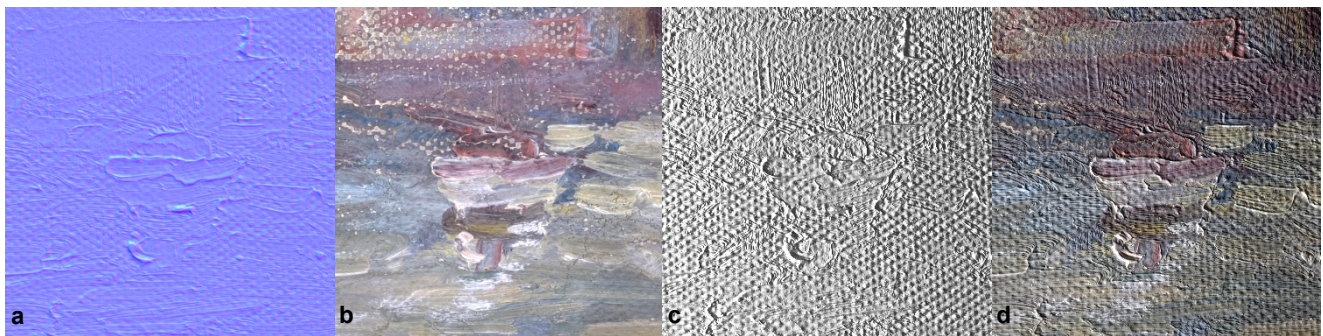


Figure 11: Same detail viewed as a) normal map, b) albedo map, c) shaded render and d) colour render in raking light



Figure 12: Final render of the surface of the painting

4. COMBINING NORMAL MAPS WITH EXISTING IMAGE BASED TECHNIQUES

The high-resolution normal maps generated with this method can be further explored by registering them to existing technical images (Infrared, UV and X-ray) of the painting to visually compare the relationship between the textured data and the other image based examination techniques.

To register the images, we follow the same method described earlier using Nip2. Once registered, the normal map is imported into Blender and assigned a diffuse colour containing either the IR, UV or X-ray information. The new image can be rendered in real-time while moving a virtual light to enhance the surface texture, or series of images with directional illumination can be generated and then used to produce a new interactive RTI image.

5. FUTURE WORK

In future, bespoke software could automate and control the low frequency flattening of the normal map based on the 3D data to improve the processing and final composite image.

A more rigorous approximation to mapping specular and glossy areas as described in MacDonald (2014) could be incorporated in the workflow. This would also improve the quality and accuracy of the normal maps generated.

Because this method is suitable for planar surfaces like paintings it would be possible to capture normal

map composites of similar objects such as oversized documents to reveal fine surface details.



Figure 13: Infrared reflectogram combined with a normal map rendered as a shaded image. Detail of NG1473 Doña Isabel de Porcel, Francisco de Goya, before 1805.

6. CONCLUSIONS

By merging 3D data, normal maps and colour images we have been able to generate photorealistic high-resolution 3D models of paintings. The better spatial resolution of digital cameras used in RTI capture outperforms the resolution of most 3D scanner systems. Hence, by using normal maps we are not only able to compensate for the low polygonal meshes necessary for optimal interaction and visualisation, but we are able to render texture with much more detail than we would with only 3D data.

We have also demonstrated a simple combination of normal maps with existing technical images to generate further RTI images that provide a new way of visualising possible relations across paint layers and whether hidden layers are revealed on the surface of a painting.

7. REFERENCES

- Cultural Heritage Imaging (2017) RTI: Guide to Highlight Image Capture and RTI: Guide to Highlight Image Processing.
http://culturalheritageimaging.org/What_We_Offer/Downloads/ (Retrieved 24 January 2017)
- Fontana, R. et al. (2003) Integrating 2D and 3D data for diagnostics of panel paintings. In R. Salimbeni, ed. *Optical Metrology for Arts and Multimedia*. SPIE, pp. 88-98.
- Huang, X. et al. (2015) Near light correction for image relighting and 3D shape recovery. In *Digital Heritage*, 2015 (Vol. 1, pp. 215-222). IEEE.
- MacDonald, L. (2014) Realistic Visualisation of Cultural Heritage Objects. PhD, University College London.
- O'Dowd, P. (2017) RTI Dome (Version 3.0) [Code] Available at:
https://github.com/paulodowd/RTI_Dome_v3 (Retrieved 27 March 2017)
- Roglab (2014) Lucida 3D Scanner. [Photograph] <http://roglab.si/en/fresh/2014/carlos-bayod> (Retrieved 24 March 2017)
- Saunders, D. and Cupitt, J. (1995) Elucidating Reflectograms by superimposing Infra-red and Colour images. *National Gallery Technical Bulletin*, 16, pp.61-65.
- Unity 3D (2017) Normal map (bump mapping). [Diagram] <https://docs.unity3d.com/Manual/StandardShaderMaterialParameterNormalMap.html> (Retrieved 20 March 2017)
- Software:**
- Adobe Photoshop
<http://www.adobe.com/uk/products/photoshop.html> (Retrieved 20 March 2017)
- Blender <https://www.blender.org/> (Retrieved 25 March 2017)
- ImageJ/Fiji <https://fiji.sc/> (Retrieved 23 March 2017).
- Knald <https://www.knaldtech.com/> (Retrieved 25 April 2017)
- Nip2
<http://www.vips.ecs.soton.ac.uk/index.php?title=VI> PS (Retrieved 25 March 2017).
- PTGui <https://www.ptgui.com/> (Retrieved 15 March 2017)
- RTI Viewer and RTI Builder
http://culturalheritageimaging.org/What_We_Offer/Downloads/ (retrieved 11 March 2017)