RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Simultaneous capture of the color and topography of paintings using fringe encoded stereo vision

Tim Zaman[1*], Pieter Jonker[2], Boris Lenseigne[2] and Joris Dik[3]

Abstract

Introduction:　Paintings are versatile near-planar objects with material characteristics that vary widely. The fact that paint has a material presence is often overlooked, mostly due to the fact that we encounter many of these artworks through two dimensional reproductions. The capture of paintings in the third dimension is not only interesting for study, restoration and conservation, but it also facilitates making three dimensional reproductions through novel 3-D printing methods. No single imaging method is ideally suited to capture the painting's color and topography and each of them have specific drawbacks. We have therefore designed an efficient hybrid imaging system dedicated to capturing paintings in both color and topography with a high resolution.

Results:　A hybrid solution between fringe projection and stereo imaging is proposed involving two cameras and a projector. Fringe projection is aided by sparse stereo matching to serve as an image encoder. These encoded images processed by the stereo cameras then help solve the correspondence problem in stereo matching, leading to a dense and accurate topographical map, while simultaneously capturing its color. Through high-end cameras, special lenses and filters we capture a surface area of 170 square centimeter with an in-plane effective resolution of 50 micron and a depth precision of 9 micron. Semi-automated positioning of the system and data stitching consequently allows for the capture of larger surfaces. The capture of the 2 square meter big Jewish Bride by Rembrandt yielded 1 billion 3-D points.

Conclusion:　The reproductive properties of the imaging system are conform the digitization guidelines for cultural heritage. The data has enabled us to make high resolution 3-D prints of the works by Rembrandt and Van Gogh we have captured, and confirms that the system performs well in capturing both the color and depth information.

Keywords:　3-D, Painting, Topography, Fringe encoding, Stereo vision

Introduction

The amount and variety of applied scientific research on paintings has intensified over the past decade. The impact of research on the material aspects of the painting often extends into the understanding of our cultural heritage. Advanced methods like X-ray fluorescence [1] or Terahertz imaging [2] give insights into the material below the surface. Such sub-surface features often reveal never before seen features or entire depictions. The canvas of a painting can also be subjected to study, and by counting its threads, separate paintings originating from the same roll of canvas can be matched [3]. All the data resulting from such applied methods will help build up a solid basis of information about the work of art. This information can help art historians draw more reliable conclusions about for example the meaning, provenance or even attribution of a painting. Since paintings are subject to the elements, they evolve over time. The research therefore fixes the work in time on the moment it is captured, allowing comparative studies about their ever developing condition. This data can then, for example, be used to extrapolate how the work will probably evolve in the future, and has evolved in the past.

*Correspondence: t.zaman@tudelft.nl
[1]Quantitative Imaging Group, Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, Delft, Netherlands
Full list of author information is available at the end of the article

Sculpting with paint

Paint is not only used by artists as a direct colorant, but, as with sculptures, its material presence can also be used to create texture or apply shadows or highlights. When enjoying the view of a painting, we are not always conscious of the impact that the painting's topography has on its depiction. Van Gogh sometimes painted flower petals with a single thick stroke, and the ambient lighting would do the rest by casting shadows and reflecting highlights. In other works, such as those of Rembrandt, the painting shows its age through craquelure, a three dimensional process where the paint has cracked apart, leaving dark canyons in the topography.

The interaction of light between illuminant, painting and observer is a dynamic process that changes constantly through relative movement. Through this process, a painting can appear to come to life, since it will look slightly different when observed from different angles. Taking a two dimensional photograph freezes this interaction and fixes the painting in a flat depiction. Such two dimensional images are then distributed through media like books, posters or the computer screen. The lack of a proper medium to depict the three dimensional data is one of the reasons why the 3-D capture of paintings seems not to have matured yet. Even though much work has been done in the 3-D digitization of cultural heritage [4-8], sculptures are often the main focus, rather than paintings. These imaging methods are usually not applied on full scale paintings, since they are inefficient (slow) for realistic production, or do not acquire the color data simultaneously.

Another reason for the immaturity of the 3-D scanning of paintings is the technical challenge. In order to capture the same depth information as our eyes can see at a normal viewing distance, a high-resolution capture is required. The high resolution, combined with an often large in-plane size of the painting means the amount of data to be captured is very high. Moreover, the scale of depth deviation of the paint versus this in-plane size is very small, as the depth deviation is rarely more than 1 cm. The varnish on paintings can be highly reflective, interfering with the capture. Furthermore, due to the fragility and value of paintings, their transportation should be avoided, requiring a portable and non-invasive imaging system. From these insights, our design requirements could be summarized as follows.

- Non invasive, portable and low cost.
- Size (XY): 2 × 2 m, depth (Z): 2 cm.
- Resolution: 50 μm/pixel; the resolving power of the human eye around 75 cm distance.
- Simultaneous capture of depth and color data.
- The color of the captured surface should not include reflections.

- Color accuracy conform the Technical Guidelines for Digitizing Cultural Heritage. Materials [9] (FADGI).
- Minimize the need for image stitching.

Results

Because we want to achieve a high resolution in both 3-D and full color, we wish to capture the topographical and color data simultaneously, avoiding image registration and misalignment. As we aspire to recreate the depth and color that we as humans observe in a painting, we could look at the way our eyes and brain retrieve this three dimensional data. We can mimic this process by employing a stereo vision approach consisting out of two cameras. If the scene is now observed by two different viewpoints, the correspondence problem needs to be solved in order to triangulate a point on the scene; what point that the left camera sees corresponds to the point that the right camera sees in the scene? We can solve this by extracting salient keypoints in both images, and then matching these keypoints. When two points are correlated, they can be triangulated in 3-D given the geometry and dimensions of the setup. This process is then repeated for all keypoints observed by both views. However, this approach will give us a sparse set of data as it only triangulates keypoints that are distinctive. Since we wish to represent each spatial point on the painting with depth data with a high certainty, an approach like passive (entirely non-invasive) stereo vision is not optimal.

Laser scanning A successful active method to capture topographical and color data simultaneously is by the use of a non-monochromatic (white-light) laser scanner [10-12]. The first 3-D scanning prototype that started our current research involved (red) line laser scanning. While the spatial accuracy proved to be high enough for scanning small features like paint craquelure, the monochromatic laser hindered our desire of sampling the full color on the exact position as the spatial measurement was taken. The beam width of the laser scanner introduces accuracy limitations and artifacts. Artifacts due to shifting of the laser beam because of shape and reflectivity discontinuities can be accounted for in post-processing, but these in turn can cause new artifacts [13]. Furthermore, the laser scanner projects either a point or a line on the surface that is then captured by a two dimensional sensor (the camera). This means that most of the sensor area is unused, and that the system would be faster if the entire sensor would have been used. In order to achieve a speed increase, one could employ multiple lasers. Instead, it makes more sense to use the power of modern projectors and project a two dimensional structured pattern on the surface.

Structured light projection Another technique that can simultaneously capture color and topography is the

structured light projection technique [10,12,14]. A common setup consists of a projector and a camera offset by a certain distance and aimed at the scene of interest. A structured light of known structure is then projected on the scene and captured by a camera. The camera can then compute the path of the light coming out of the projector, hitting the surface and then entering the camera. This triangulation allows for computing the topography while also capturing the color, if the projector's illuminant is neutral. Their main drawbacks are that they are limited to the resolution of the projector and problems due to specular reflections. The constraints set by the projector's resolution can be circumvented by using the fringe projection technique. In order to make the quantization or pixel pattern invisible, the projector's projection can be blurred by defocusing the lens, resulting in a smooth fringe pattern (convolution theorem). The resolution of the capture is now constrained by that of the camera, and so we can employ cameras with a large pixel count and capture a large quantity of data per capture. Such large captures both increase speed and reduce the need for image registration (data stitching). Multiple cameras can be used to observe the projected pattern from different angles. Such arrangements often employ the same algorithm for each individual camera without much synergy, and so are only used to reduce occlusions and increase accuracy. One common problem with fringe projection is the fact that the its intensity pattern has to be exactly sinusoidal. This is difficult due to the non-linearity in the illumination of the projector and the sensor of the camera. Not only do these have to be accurately calibrated and accounted for; both the projector and the camera need to be geometrically and optically calibrated. The projector calibration can be achieved by projecting patterns, however, this projection is then again limited to the resolution of the projector, making precise calibration difficult. Instead of using structured light to do projector-to-camera light-ray calculations, we can also use the projection to encode the surface, uniquely labeling each point on the surface [15]. This unique label then has to be observed by multiple cameras, immediately solving the correspondence problem we encounter in stereo vision. Having solved this, camera-to-camera triangulation can be performed as is done in a stereo vision setup. Fringe projection can be used for this labeling as our system is then not limited by the projector's resolution. Each point is then labeled with phase values. However, since these fringes were repetitive, the cameras do not yet know which fringe in the left image belongs to the fringe observed by the right image. In other words, there exists an offset between the phase value observed in the left and right phase images. This offset can be calculated if we compare phase values of at least one part of both images that we know that correspond. This can be done with the common stereo-vision approach, where we first search for keypoints in both color images, and then find corresponding matches. The offset of phase value observed from both cameras can now be nullified as we know the phase value should be the same for each point in the scene. This fringe projection aided stereo vision approach is therefore our selected method, since it can be both highly efficient and accurate.

Lighting and optics The problem with local illuminants, such as projectors, is that they can cast shadows and specular reflections when the surface is glossy. This can be avoided by illuminating the painting perpendicular to the surface, and we can account for light reduction on slanted surfaces through Lambert's cosine law.

Since paintings are often varnished, specular reflections are generally abundant and pose a problem to every imaging setup. Many 3-D imaging approaches assume an entirely diffuse (Lambertian) surface, and therefore do not cope well with reflections. Reflections are generally easy to detect, but information will be lost at reflective locations. Therefore, prevention of reflections in our captures images is important. It is relevant to note that the (amount of) reflections is not of much importance to the depiction of the artwork itself, as it often applied by museum professionals instead of the original artists. Because we work with the projector as a local illuminant, there is a straightforward way to suppress specular reflections. By using the fact that light reflecting off a surface is polarized with a direction perpendicular to the plane of incidence. We can filter out most of this particular polarized light by using a polarization filter mounted on the camera. In order to maximize this effect, we can mount another polarization filter in a crossed direction with the one in front of the camera. This will effectively cancel out all specular reflections, resulting in an image of the entirely diffuse surface; specular reflections induced by a film of varnish will not be visible in our data. However, such a film can influence the observed depth due to the refraction of the projector's light. This contribution is considered negligible due to the small thickness of the film. A non-transparent varnish will introduce errors similarly to those of a defocused camera, and will decrease our triangulation accuracy on high frequency features like sharp edges.

To avoid shadow formation, we place the projector perpendicular with the surface of the painting. This means that the cameras will be observing the painting at an angle, leading to the depth of field of the camera not being parallel with the surface. The part of the scene that is not within the cameras depth of field will be out of focus, and it should therefore be extended or rotated. This extension can be enforced by changing the aperture, and the rotation by using Scheimpflug (tilt-shift) lenses. The tilt property of such lenses allow us to rotate the depth of field

in parallel with the surface of the painting, while putting the entire scene safely within focus.

A common drawback of the previously mentioned laser scanning and structured light projection is that there can only be a measurement if the point of interest is observed or illuminated by all parts of the system. This means that by using pure stereo vision, both camera needs to be able to observe a point in order to triangulate it. A pronounced protruding feature and its surroundings might therefore not be measurable. We found little of such occurrences in the particular paintings we have studied here, as the painting's surface is continuous, and the entire surface was observable by both the cameras' positions.

System design

The final design was entirely made from off-the shelf parts that are available to any consumer, and are usually already owned by institutions that display paintings. The main components are two cameras and a projector. The cameras we had used were borrowed, while for the software we made use of multiple open-source libraries. Therefore, our costs for this project were around € 800 because we already owned two camera sets and a projector. We have tested multiple NIKON and CANON models and lens combinations with our system and all of them worked with our software.

Hardware

The configuration of our system is depicted in Figures 1 and 2. Our camera of choice was the NIKON D800E, as it features one of the highest commercially available sensors with 36 megapixels. The best suiting corresponding lens is the 80 mm PC-E macro (tilt-shift) model and was fitted with a polarization filter. The OPTOMA PK301 was used as the projector because it has a very short throw distance, since it can already focus on a plane a few decimeters in distance. This is needed because the size of the projection on the surface will be small since we aspire a high resolution. A polarization filter was put in front of the projector's lens and the camera's lenses. This configuration resulted in a large decrease of light intensity, which was countered with a larger exposure time of the camera, typical between 3 and 10 seconds and a corresponding ISO setting of 100 and an aperture of around f/11. The maximum light intensity on its surface does not exceed 1000 lux, and the illumination time for one capture is approximately half a minute. A meter long linear axis was used to automatically translate the scanner system horizontally. The translation in the vertical direction was done manually, as the whole system can be mounted on any studio tripod. The framework of the system is in its current state prone to vibrations, and therefore requires a stable floor to stand on.

Software

The first implementation of the software that drives our 3-D scanner was written in MATLAB. Our latest implementation runs in C++ because of speed, stability and versatility improvements. Both implementations used the same implementation, as schematically depicted in Figure 3. After setting up the system, it has to be calibrated once and this semi automatic process takes around 15 minutes. We used a Colorchecker Passport color reference
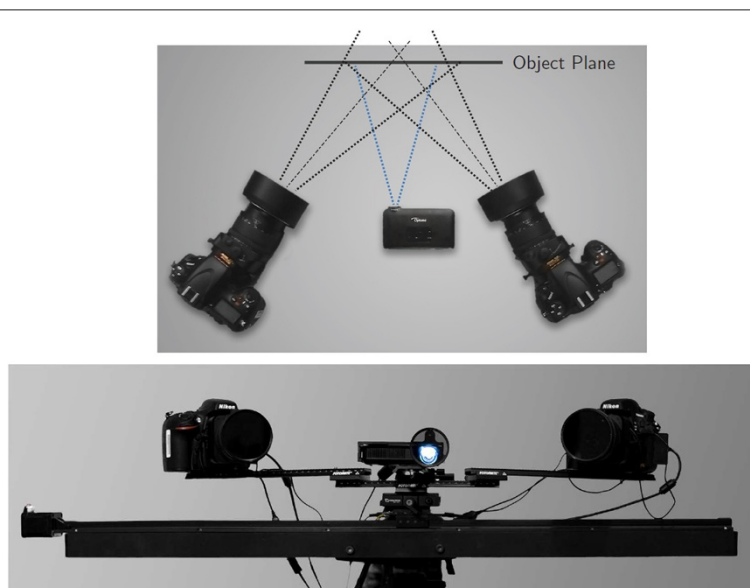
Figure 1 Top and front view of the setup. The imaging setup consists our of two cameras and a projector, with the projector oriented perpendicular to the surface, amounted on a horizontal translation axis.

Figure 2 Scanning of the self-portrait by Rembrandt. The transportation, setting up and scanning of this particular painting took just over one morning.

chart to automatically generate an ICC profile for the cameras, with the projector as the illuminant. Although we acquire a satisfactory color reproduction (in relation to the FADGI guidelines) if we benchmark our color values on the color checker, the small amount of reference patches in the chart, and the non-neutrality of the projector's illumination decrease our performance. Problems with metamerism easily arise and we will therefore incorporate an extra capture with a neutral illuminant in the capture routine for our next design. Furthermore, we take one image of a white balance chart to measure the projector's

illumination non-uniformity over the projection, which can later be modeled and accounted for and balanced. The vignetting caused by the cameras lenses is accounted for by the camera itself.

Then the spatial calibration is performed with a flat checkerboard with blue checkers of which the cameras take multiple images. The projector also projects a red checkerboard which is used to calibrate the projector. Since the checkerboard is planar, the checkers lie in a fixed grid and the size of the checkers is known, the corners of each checker can be automatically located when they are
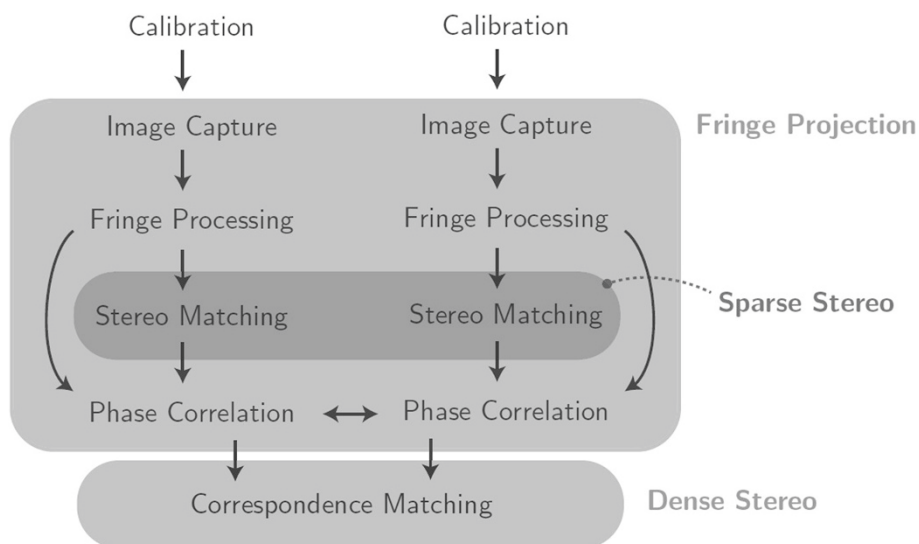
Figure 3 Software workflow. This schematic reveals the individual parts in the pipeline of the data in our software.

observed by both cameras. We can use these corners as features to locate the camera with respect to the checkerboard. Around 20 captures with the checkerboard held in a different orientation allow us to solve the system and give us all the relevant parameters [16]. This calibration data will later be converted to an optical ray map per camera sensor that contain the vector of light coming into each pixel of the camera. Now, we can also calibrate the projector in the inverse way as we have calibrated the camera (although the projector calibration is not strictly necessary for our system as we only use the projected pattern as an encoder). The red projected checker pattern that was projected onto the (blue) checkerboard should be extracted in all images. Through the camera calibration, we know the exact orientation of the plane of the physical (blue) checkerboard. Therefore, we also know the exact 3-D position of the corners of each of the red checkers that have been projected onto this board. If we repeat this extraction on multiple images with the physical checkerboard at different orientations, we can trace back the rays of light into the optical center of the projector and repeat the camera calibration procedure in an inverse manner.

The calibration results arising from this procedure are their exact relative spatial position and orientation, and the distortion of the lens systems in the cameras and projector. After the calibration we can perform the structured light projection. Because we wish to use the projection as an encoder for the camera, we have to encode the horizontal and vertical dimension. The two cameras are horizontally spaced apart, and through epipolar geometry we know which horizontal lines in both images belong to each other. The horizontal dimension therefore does not need to be encoded, which means only the vertical dimension has to be captured and processed. For fringe projection this means we only need to project vertical fringes. The fringe spacing (period) was set around 4 mm. The three-step fringe projection method [17] was used as this is fast and sufficient for the use of encoding. Because paintings often have a large intensity range, problems arise from the modulating intensity of the fringes that are projected. In terms of using a fringe as an illuminant, it should be seen as the projection of taking an image with a low and a high intensity - at the same time. This is problematic for the cameras are they already have a limited dynamic range. Our solution was to shoot two images of each fringe projection with different exposure values, resulting in a high dynamic range. Therefore we employed the double three-step fringe projection method [18], where the second projection sequence had a different exposure setting. The capture of one scene currently takes around a minute, depending on the settings. After the fringes have been captured, they are processed using the corresponding algorithmic for our chosen fringe projection method. A series of fringes then results in a wrapped phase map. The

phase unwrapping process then produces a continuous phase image with unique labels at each spatial position. However, since the fringes are repetitive, these phase values are not yet correlated between the two cameras. These are then set to exactly match through sparse stereo matching. We use a method called SIFT to extract keypoints [19] from images of both cameras. We require to find at least 1000 keypoints (potential matches) in both images, and decrease the keypoint strength threshold until at least this amount has been found. This assures enough matches for captures of both light and dark areas. The same keypoints observed by both cameras are then matched, resulting in pixel locations in the images that we know are the same spatial positions. We can then use this information to correlate the relative phase maps deriving from our fringe projection. Each unique value in the left image, is looking at the exact same point as a point in the right image with the same value. As we have now solved the correspondence problem, we can compute the 3-D position of each point in the image. We need to do this for all 36 megapixels that are present in our images. The 3-D position of each point can be computed through ray tracing by taking in account the information from our camera calibration. We managed to speed up this process by a few orders of magnitude by making a look-up table with these phase values and pre-computing the optical ray maps. We can then construct ray vectors for each pixel in each camera, after which the ray intersection approximation is trivial. The processing of a single image currently takes around one minute or up to 10 minutes in our MATLAB implementation.

A graphical user interface allows for intuitive operation for both the capture as the processing. A large painting like the Jewish Bride by Rembrandt consisted out of 240 captures, which all need to be stitched to each other in order to have one coherent capture. Since the registration between our depth and color data perfectly overlaps, we can use the color information to stitch multiple captures in 3-D, since this is directly linked to the color information. This was again done through SIFT keypoint extraction with sub-pixel accuracy and matching. Since our captures are highly accurate, only basic rotation and translation between consecutive captures were necessary for the stitching, without any need of scaling or warping. Around 20% overlap was taken between each neighboring capture. After stitching, investigation of the overlap of neighboring captures revealed there was no spatial discrepancy visible between separate captures. On a global scale, we also visually verified the absence of shape deformation caused by our stitching through single images acquired by the institutions themselves. Due to the enormous amount of information captured of this specific painting, data stitching took almost two weeks on a high end PC.

Table 1 Investigated paintings

Title	Year	Artist	Size (w x h, cm)	Size (points)
Flowers in a Blue Vase	1887	Van Gogh	39 × 62	8.000 × 9.000
Self-portrait	c. 1668	Rembrandt	65 × 60	12.000 × 13.000
Jewish Bride	c. 1667	Rembrandt	160 × 120	33.000 × 24.000

Information about the paintings that were chosen for our pilot.

Because of small errors and accuracy limitations that arise in each part of the system, the effective resolution is different from the sampled resolution. The ISO guideline 12233:2000 was used to compute the in-plane sampling efficiency. The effective resolution was found to be 65 μm and 46 μm for respectively the horizontal and vertical axis. This method of computing the sampling efficiency is usually applied for sampling camera sensors, and it involves the capture of a planar slanted edge target. The slanted edges in this target is then measured for its 'sharpness' and results in its sampling frequency response, from which the sampling efficiency is derived. At each coordinate, our system also captures the color information at the exact same location as the depth is sampled. The spatial accuracy of the color information is therefore directly related to the accuracy of our three dimensional data. We can directly relate these metrics if we project our three dimensional data to a two dimensional plane. From this two dimensional plane, we then take the color map, resulting in a common 2-D image. We can then proceed with extracting the sampling frequency response following the ISO guideline.

No similar standard exists yet to repeat this procedure for the depth axis, but through testing for planarity [20] precision, we measured the effective depth resolution to be 9.2 μm. The planarity target was made on a flat plane with a few hundred flat-topped cylinders sticking out of it with a discrete and known height. Our effective depth resolution then resulted from the observed planarity of the tops of each of these cylinders. The accuracy was

measured through the capture of several checkerboards. The 3-D locations of all corners in the checkerboard were then measured, and it was then checked how well they fit into the real flat checkerboard (with the checkers in a fixed grid). The difference in distance of the real position of each corner and our sampled position then produced the accuracy. This was repeated for multiple orientations. The depth accuracy was found to be 38 μm.

The final data used for visualization and 3-D printing was scaled down to this effective resolution in order to keep the files manageable.

Experiment

After scanning and printing multiple mock-up paintings with good results, we started a pilot where we scanned three works each from a different museum, with varying three dimensional properties. The selected paintings and their approximate dimensions are listed in Table 1 in chronological order of scanning.

The Self-portrait by Rembrandt proved to be highly reflective due to its varnish. The cross-polarization setup was employed and effectively suppressed all specular reflections as expected. Transportation, deployment, calibration and the actual scanning was done in around 6 hours. Apparently, the intensity range of the image was very high, featuring deep black and bright white features. This caused artifacts in dark regions because at this time we had not employed high dynamic range imaging yet. These artifacts reveal themselves as harmonic patterns in the depth dimension. The capture of high dynamic range solved this issue. A detail of the turban in this painting is depicted in Figure 4. The only artifact and source of error since found were subtle harmonic patterns caused by external lighting. These were encountered in the scans of the Jewish Bride, where external lighting was accidentally cast on the painting at multiple occasions. Dim external lighting is not necessarily a problem, but when the amount of external illumination changes within one specific capture, the fringe algorithms
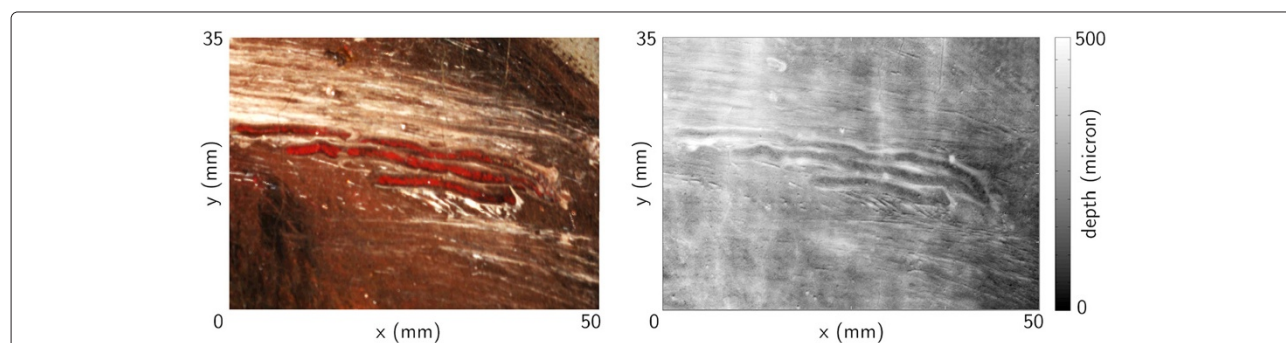

Figure 4 Detail of the color image (left) and depth map (right) of the self-portrait by Rembrandt. Note the depth features like the three scratches made in the paint that reveal the red ground layer and elevate the paint around the scratches. In the top-right corner the canvas weave pattern can be distinguished and distributed over this sample we see large individual pigment particles protruding.

Figure 5 Software 3-D render of the flowers in a blue vase by Van Gogh with artificial lighting. Different techniques of painting the flower petals can be identified, and in the top left corner the individual hairs of Van Gogh's brush can be distinguished.

fail and produce these patterns. Scanning with our system should therefore preferably take place in a static or closed environment.

The most straightforward way to test the validity of our data was comparing it to its reference, the original. This was done through high resolution, full color 3-D reproductions made by OCÉ using an ink-jet based process. This system can only deposit from above and no overhangs can be created (sometimes denoted as 2.5D printing). The previous limitation did not influence the printing of our artworks as our system is unable to capture such overhangs, nor were there overhangs present in the particular paintings. Software rendering and virtual visualizations of the large amount of data in each painting proved to be difficult, and although efficient massive point-cloud visualization solutions exist, looking at this data in its entirety on a screen with a limited resolution seemed sub-optimal compared to the 3-D print. A detail of a software render is depicted in Figure 5. A
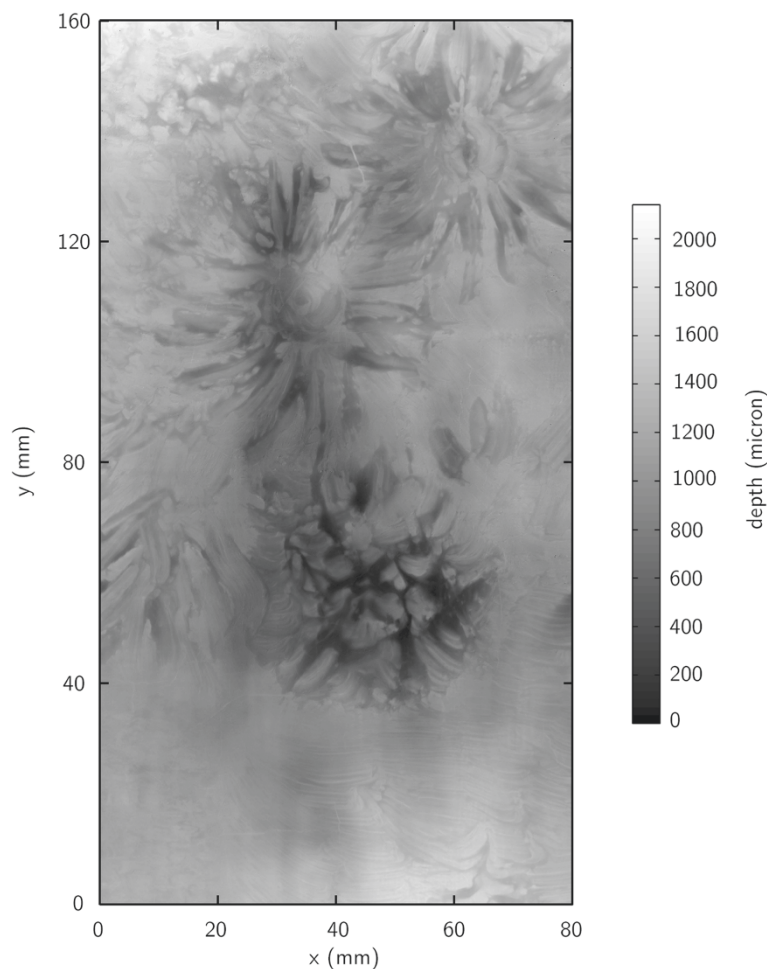
Figure 6 Depth map of the bottom flowers in Van Gogh's flowers in a blue vase. Even in a depth map the flower petals can be identified, which are all protruding as they are painted with thick strokes of paint. We can also see more global depth features that are caused by an underpainting.

more objective way of analyzing the 3-D data is through a depth map, like depicted in Figure 6. In this depth map, we can clearly distinguish Van Gogh's protruding flower petals. We can also see smooth vertical features that have nothing to do with the paint or the canvas of the depiction of the surface. These features originate because of an underpainting; an image of the resulting 3-D print of the Van Gogh is depicted in Figure 7.

Each reproduction was then taken to the museum for a side-by-side evaluation. The three dimensional data of each of the paintings is valid, but on very close inspections the very finest of cracks are not preserved and on even closer inspection the quantization of the ink drops from the printing can be seen. The color reproduction was fair, but clearly needs improvement due to the non-linearity of the projector's illuminant and the small amount of color reference patches for ICC profiling.

Conclusions

Paintings are versatile near-planar objects that have material characteristics that vary widely. The fact that paint has a material presence is often overlooked, mostly due of the fact that we encounter many of these artworks through two dimensional reproductions. The capture of paintings in the third dimension is not only interesting for

Figure 7 The 3-D printed van Gogh painting. The reflection of the raking light enhances our perception of the brush strokes. This particular print was made on foam board, which was later framed.

study, restoration and conservation, but it also facilitates making three dimensional reproductions through novel 3-D printing methods. In order to print in full color as well, the color of the painting has to be captured. To sample the color and depth simultaneously, we have designed an imaging system using a hybrid approach. Fringe projection and stereo imaging are combined to yield an accurate, fast and reliable method. The usage of high-end cameras, special lenses and filters accommodate the capture of a large amount of 3-D data per capture, while simultaneously capturing color information.

We have demonstrated that fringe projection is a very effective method of encoding images independent of the projector's resolution. The projection facilitates that each pixel observed in one camera of a stereo setup can be easily matched to a pixel in the other camera. Through our combined camera and projector calibration, we can use these pixel locations to triangulate an absolute 3-D point in space. A prerequisite of using fringe projection as image encoder is that both observed fringe images are correlated to each other. This was done through sparse stereo matching. The consequent triangulation of around 36 million data points per capture was done through a special look-up table construction and by using pre-computed optical ray maps for each camera sensor. The processing and capture both take around a minute to finish. The reproductive performance for both color and spatial properties were satisfactory related to digitization guidelines, although the color performance in practice needs to be improved. We obtained an in-plane precision of around 50 μm. An effective depth resolution of 9.2 μm then allows the typical observer for resolving very small three dimensional features in a painting. If an even higher spatial resolution is desired, the configuration of the system could be adapted by adding magnification. In our system, this can be done by adding close-up filters on the lenses and moving the cameras closer to the canvas.

The preliminary results from the capture and 3-D printing of works by Rembrandt and Van Gogh using our design indicate that the system works well in practice. The depth map and color image resulting from this will be used for the construction of a reproduction in full color and full dimension.

Improvements and future work

We haven focused on the selection, refinement and fusion of the best 3-D imaging technique for paintings. The fact that we have designed such a device that is ideally suited to do so, does not exclude the viability of other methods. Requirements and standards regarding the 3-D scanning of paintings still remain to be set. Studies have to be carried out into which performance is needed for the purpose of study, conservation and restoration, but also for reproduction. Such solid requirements are important in the

design of an ideal system. For example, detailed inspection also indicated the very finest of cracks, those smaller than our effective resolution, are of course not preserved in our data. We have already established that our color reproduction performance can be improved by using a neutral illuminant, as we are currently using the three-color illuminant in the projector. In order to further increase that performance, color calibration should be done with a reference chart containing more color patches. Another interesting self-assessment would be the scanning of the print of a scanned object, especially for investigating the quality of the print. The feasibility of using this scanner for more purposes than paintings should be investigated. So far, successful 3-D captures were taken of fingerprints, textiles and wax seals. The data itself could be a source of further quantitative research as well, like for retrieving the canvas weave pattern. The global deformation shape of the canvas can also be used for investigating the stretch and strain of the canvas. The side-by-side comparisons of our 3-D reproductions and the original paintings have shown that although a 3-D reproduction is a lot more lively and impressing than a common poster, but the original painting still has features that our reproduction lacks. Such features seem to be differences in reflectivity and transmission of light in the material. These differences are being quantified and applied in our current research.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
TZ designed, built and deployed the imaging device. JB conceived of the study, shaped design and design requirements, arranged partnerships and scanning pilot. PPJ and BL both supervised and participated the design and used methods. All authors read and approved the final manuscript.

Author details
[1]Quantitative Imaging Group, Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, Delft, Netherlands. [2]Faculty of Mechanical Engineering, Department of BioMechanical Engineering, Delft University of Technology, Mekelweg 2, Delft, Netherlands. [3]Faculty of Mechanical Engineering, Department of Materials Science, Delft University of Technology, Mekelweg 2, Delft, Netherlands.

References
1. Dik J, Janssens K, Van Der Snickt G, van der Loeff L, Rickers K, Cotte M: **Visualization of a lost painting by Vincent van Gogh using synchrotron radiation based X-ray fluorescence elemental mapping.** *Anal Chem* 2008, **80**(16):6436–6442.
2. Adam AJ, Planken PC, Meloni S, Dik J: **TeraHertz imaging of hidden paint layers on canvas.** In *Infrared, Millimeter, and Terahertz Waves, 2009. IRMMW-THz 2009. 34th International Conference on*. Busan: IEEE; 2009:1–2.
3. Johnson DH, Johnson C, Klein AG, Sethares WA, Lee H, Hendriks E: **A thread counting algorithm for art forensics.** In *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*. Marco Island, FL: IEEE; 2009:679–684.
4. Fontana R, Gambino M, Pampaloni E, Pezzati L, Seccaroni C: **Panel painting surface investigation by conoscopic holography.** In *8th International Conference on Non Destructive Investigations and Microanalysis for the Diagnostics and the Conservation of the Cultural and Environmental Heritage*. Lecce, Italy; 2005:15–19.
5. Sansoni G, Trebeschi M, Docchio F: **State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation.** *Sensors* 2009, **9**:568–601.
6. Pavlidis G, Koutsoudis A, Arnaoutoglou F, Tsioukas V, Chamzas C: **Methods for 3D digitization of cultural heritage.** *J Cultural Herit* 2007, **8**:93–98.
7. Pieraccini M, Guidi G, Atzeni C: **3D digitizing of cultural heritage.** *J Cultural Herit* 2001, **2**:63–70.
8. Boehler W, Marbs A: **3D scanning and photogrammetry for heritage recording: a comparison.** In *Proceedings of the 12th International Conference on Geoinformatics*. Gävle, Sweden; 2004:291–298.
9. FADGI-Still Image Working Group: **Technical guidelines for digitizing cultural heritage materials.** Tech. rep. U.S National Archives, Washington, 2009.
10. Lahanier C, Aitken G, Pillay R, Beraldin A, Blais F, Borgeat L, Cournyer L, Picard M, Rioux M, Taylor J, Breuckmann B, Colantoni P, de Deyne C: **Two-dimensional multi-spectral digitization and three-dimensional modelling of easel paintings.** Article published at the ICOM-CC Preprints of the 15th Triennial Meeting, New Delhi, 22–26 September 2008, Vol. I.
11. Blais F, Taylor J, Cournoyer L, Picard M, Borgeat L, Godin G, Beraldin JA, Rioux M, Lahanier C: **Ultra high-resolution 3D laser color imaging of paintings: the Mona Lisa by Leonardo da Vinci.** In *Lasers in the Conservation of Artworks; Proceedings of the International Conference Lacona VII, Madrid, Spain, 17–21 September 2007*. Edited by Ruiz J, Radvan R, Oujja M, Castillejo M, Moreno P: CRC Press; 2008:435–440.
12. Bunsch E, Sitnik R, Michonski J: **Art documentation quality in function of 3D scanning resolution and precision.** In *IS&T/SPIE Electronic Imaging*. San Francisco: International Society for Optics and Photonics; 2011:78690D–78690D.
13. de Jong F: **Range imaging and visual servoing for industrial applications.** *PhD thesis*. Delft University of Technology, Delft, The Netherlands, 2008.
14. Karaszewski M, Adamczyk M, Sitnik R, Michoński J, Załuski W, Bunsch E, Bolewicki P: **Automated full-3D digitization system for documentation of paintings.** In *SPIE Optical Metrology 2013*. Munich: International Society for Optics and Photonics; 2013:87900X–87900X.
15. Scharstein D, Szeliski R: **High-accuracy stereo depth maps using structured light.** In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference, Volume 1*: IEEE; 2003:I–195.
16. Zhang Z: **A flexible new technique for camera calibration.** *Pattern Anal Mach Intell IEEE Trans* 2000, **22**(11):1330–1334.
17. Wyant J: **Interferometric optical metrology: basic principles and new systems.** *Laser Focus* 1982, **18**(5):65–71.
18. Huang P, Hu Q, Chiang F: **Double three-step phase-shifting algorithm.** *Appl Opt* 2002, **41**(22):4503–4509.
19. Lowe D: **Distinctive image features from scale-invariant keypoints.** *Int J Comput Vis* 2004, **60**(2):91–110.
20. Zaman T: **Development of a topographic imaging device for near-planar surfaces.** *Master's thesis*. Delft University of Technology, Delft, The Netherlands, 2013.