A low-cost ChArUco-based 3D scanner for cultural heritage

Luca Puggelli\textsuperscript{1}, Rocco Furferi\textsuperscript{1}, Yary Volpe\textsuperscript{1}, and Lapo Governi\textsuperscript{1}

\textsuperscript{1} Department of Industrial Engineering, University of Florence, Firenze, Italy

Abstract. In the region of the cultural patrimony, operators use high-resolution orthophotos of paintings for the restoration, monitoring and electronic recording and exhibition purposes. Unfortunately, artworks that are to be restored and/or shown in digital museums are painted on canvas that are far from perfectly planar. Professional documents surrounding an artwork to be preserved in digital collections or museums can therefore be enhanced with information relating to the paintings' 3D structure. This paper proposes both the design of a portable low-cost device that enables the acquisition of 3D geometry of painting and the procedure for triangulation of 3D data. This process uses a set of fiducial markers to set and continuously control the mutual orientation of the laser source and the camera and works accordingly to the principle of laser-camera triangulation.

Keywords: Reverse Engineering, Cultural Heritage, Marker Detection, Pose Estimation, 3D Laser Scanner.

1. Introduction
In the field of cultural heritage, high-resolution images of paintings are often used for various purposes both for public benefit and for technical use \cite{1, 2}.

Although typical undistort digital photos are deemed more than adequate in the first example, specialized applications are more challenging both in terms of resolution and in terms of dimensional accuracy \cite{3}. These two aspects have a fundamental role to play, especially in the case of restoration of painting and monitoring purposes \cite{4}. The common workflow, based on the realization of a master orthophoto that will serve as primary reference in the successive collection of multiple images, is designed to achieve a picture that is at the same time measurable (i.e. a measurement taken in the image is equal to the corresponding measured value on the painting) and with a gigapixel resolution. As tiles of a mosaic, these images are stretched to match the orthophoto master's image precisely.

The resulting image is an exceptionally detailed photograph, on which the finer details of the art piece can be observed, which exactly fit the actual measurements of painting. The precision of the orthophoto is therefore crucial to the success of the entire procedure. Starting from an undistort image of the painting, the orthophoto is obtained by means of metric references (points or planes), typically obtained exploiting perspective properties \cite{5, 6}. One of the most popular methods for assessing metrics information from pictures is the so-called “cross-ratio” method \cite{7, 8}. This method aims at retrieving the plane on which the subject (in this case, the painting) lays by means of a series of proportions (the cross-ratio, see Equation 1) among known points, displaced along convergent lines (at least 3 points on each line).
\[
\frac{AC \cdot BD}{BC \cdot AD} = \frac{A'C' \cdot B'D'}{B'C' \cdot A'D'}
\] (1)

On the left side of the equation, distance between points in real-world; on the right side, distances between points on the image. Looking at Equation 1, let’s call A B C D four aligned points, of which D is on the painting surface. On the left side of the equation, coordinates are defined in real-world reference system (and distances are expressed in meters); on the right one, in image reference system (and distances are expressed in pixels). In order to make it easier to detect A’, B’, C’ and D’, a properly devised tool can be used, composed by a table – on which a series of convergent lines with corresponding points are drawn – and a laser line projector. This last is placed so that the laser line is tangent to the table. This tool is then placed between the camera and the painting, so that it is entirely framed into the picture. The laser line on the painting highlights the intersection between the painting surface and the “table”. Consequently, it is very easy to detect in the picture the intersection between each convergent line of the support and the painting surface, and so it is possible to apply the cross-ratio and proceed with the points and plane retrieval.

As simple as the method is, it has two major drawbacks. First, the operation of point detection is usually carried out manually and, consequently, it is slow and subject to error. In addition, since the works of art to be restored and/or reported in digital museums are generally painted on canvases or wood panels, the painting surface is far from being approximate with a planar surface.

The present paper proposes the design of a portable low-cost device to acquire a canvas’ 3D geometry reliably with an approach to overcoming the above-mentioned drawbacks. In addition, an appositely conceived semi-automated technique is defined to achieve an effective 3D acquisition. Indeed, the main idea is to realize a cheap device, capable of performing a 3D triangulation when combined with the pre-calibrated consumer single-lens reflex (SLR) camera, allowing high accuracy 3D measure of thousands of points on the surface. At the same time, the camera used together with the laser to triangulate the 3D scene may be used to acquire 2D data to create the orthophoto.

The rest of the paper is arranged as follows: in section 2 the device and the procedure are briefly described, accuracy analysis of pose estimation phase is treated in section 3, while final test and conclusions are discussed in section 4.

2. Device
A 3D scanning system based on 3D laser camera triangulation is composed by 2 main devices, camera and laser source. Focusing on this application, the most suitable layout consists on a camera maintained in the same position through all the acquisition process, while the laser swipes the painting surface, by rotating around the vertical axis.

In order to obtain a 3D reconstruction both intrinsic and extrinsic calibration of the acquisition system are required.

The intrinsic calibration (IC) concerns the image acquisition system (sensor of the camera and lens) and, in brief, it consists on obtaining the set of parameters that define the mathematical transformation between captured image and real-world, mostly due to sensor, perspective and lenticular distortions. Therefore, IC is fundamental to rectify the original image and eliminate, for example, barrel/pincushion distortion (Figure 1) [9].
On the other side, extrinsic calibration (EC) concerns the overall 3D acquisition system (camera and laser source) and it consists on obtaining the relative and global (i.e. w.r.t. the object to be scanned) positioning and orientation of the two devices. The output of this calibration procedure is the RT matrix, which describe both rotation (R) and translation (T) transformations.

While IC parameters are obtained by realizing a proper set of photos of a fiducial marker whose dimensions are known a-priori, generally a chessboard, EC depends on the architecture of the 3D acquisition system. Whatever the architecture, these data are mostly obtained by means of absolute encoders, which make it possible to evaluate both rotations and translations with an elevated level of accuracy. To make an example, the angular position can be easily measured with a rotary absolute encoder with an accuracy equal or superior to $10^{-2}$ deg. Such equipment is fundamental to obtain a high-accuracy 3D data. Unfortunately, each encoder is also highly pricey, with a cost that starts from hundreds of euros and that can reach some thousands, depending on the specs. With the aim of keeping costs low, it has been decided to avoid the adoption of such instruments and to obtain EC parameters in an alternative way.

An option to solve the lack of encoders is given by fiducial markers and relative computer vision procedures, whose development has seen a relevant growth starting from the arise of augmented reality (AR) applications [10, 11].

Observing Figure 2, is it possible to see how the AR application uses a fiducial marker as a reference system within the frame to re-project on it a virtual 3D figure. More specifically, the application detects the marker in the frame, estimates its pose and, since marker dimensions are known, its able to define the transformation from “world” to “camera” reference systems. Such transformation is then applied to the virtual model, so that it is possible to super-impose it on the image and obtain what is visible on the screen of the smartphone.

The idea is to exploit information about marker pose estimation to calculate EC parameters in a 3D acquisition system.
Differently from what have been stated about encoders, fiducial marker and relative AR function libraries are very affordable and, in some cases, completely open source.

OpenCV libraries have been considered for two main reasons:

1. OpenCV libraries are open-source and make use of open-source fiducial markers (ChArUco, see Figure 3);
2. ChArUco markers are considered among the most accurate and fast in marker detection phase [12, 13].

In addition, OpenCV libraries can be easily compiled to be implemented into MATLAB applications, which on its turn offer a large library of computer-vision and image-processing functions and its largely widespread in the scientific community.

Moving from these considerations, a laser line generator support has been properly designed, so that it can fulfil two principal functions: 1) mounting the laser line generator on a tripod/head and 2) allow the attachment and fine positioning of the fiducial marker (printed on adhesive paper). The specific design has been extensively treated in [14]. In brief, the device is composed by two main components – the marker-board and the laser support. This last is the interface that links the support to the tripod, which both laser line generator and marker-table are coupled with.

The marker-table has to be oriented w.r.t. the projected laser line, so the architecture of the support must allow the fine-tuning of its position has to be fine-tuned. This operation is made possible by linking marker-board to the support by means of 3 opportune devised “fasteners”, each one composed by: a bolt, a compression-spring and a couple of spherical-washers. The spring is included to provide pre-load to the bolt. Spherical washers admit a slight rotation (±5°) between marker-table and support (so that one can be slightly rotated w.r.t. the other). For a sake of clarity, the linkage scheme is reported in Figure 4.
The design of the device has been carried out paying attention to all the aspects that may influence the accuracy of the 3D measurements obtained with this scanning system. To this purpose, the influence of marker-board planarity has been assessed as the most influent. For this reason, it has been analysed in depth, as reported in the following section.

3. Marker-board planarity

In order to evaluate the accuracy of the pose estimation phase, a specific test-campaign has been performed. The scope of this analysis is to highlight the influence of the planarity error of the marker-board in the pose estimation procedure, included in Open-CV function libraries. The accuracy of this procedure is fundamental for the whole 3D acquisition process.

Tests have been carried out as follows:
- Two different supports have been used as marker-table: a plywood table (planarity error = 0.6 mm) and a calibrated grade-0 granite plate (planarity error < 0.012 mm, for other specifications see Table 1);
- Supports have been placed on a rotating base (yaw-angle);
- The angular displacement between two consecutive positions has been measured by means of a touch probe (3D Romer Absolute Arm 7520-SI) and by means of the Open-CV based procedures (i.e. by means of fiducial marker);
- Considering the probed measure as ground-truth, the error is considered as the difference with the measure obtained with the proposed procedure.
- A planarity error of 1.5 mrad has been considered acceptable [15].

Table 1. Supports for fiducial marker, used during tests.

|                  | Calibrated plate | Plywood table |
|------------------|------------------|---------------|
| Size [mm]        | 300x200 mm       | 300x200       |
| Planarity [mm]   | <0.012 mm (certified) | 0.6 mm (meas.) |
| Material         | granite          | Pine plywood |

Image acquisition has been performed using Fujifilm X-T1 camera, with a Fujifilm XF 18-55mm f/2.8-4R LM OIS lens mounted on, at a focal length of 18 mm, whose specifications are reported in Table 2.

Table 2. Cameras and lens specifications.

|                  | Fujifilm X-T1 |
|------------------|---------------|
| Sensor           | CMOS APS-C    |
| M. pixel         | 16.3          |
| Lens             | Fujifilm XF 18-55 f/2.8-4R LM OIS |
| Focal length [mm]| 18 mm         |
| Eq. focal length [mm]| 27 mm     |

In Table 3 and in Table 4, some of the obtained results are briefly reported.
Table 3. Test results, calibrated granite plate (Fujifilm X-T1).

| Detected angle | Ground (Romer) | Error (degrees) | Error (mrad) |
|----------------|----------------|-----------------|--------------|
| 14.713         | 14.7045        | 0.0085          | 0.1484       |
| 12.104         | 12.0235        | 0.0805          | 1.4050       |
| 15.499         | 15.4951        | 0.0039          | 0.0681       |
| 14.289         | 14.1766        | 0.1124          | 1.9618       |
| 16.719         | 16.7047        | 0.0143          | 0.2489       |

Table 4. Test results, plywood table (Fujifilm X-T1).

| Detected angle | Ground (Romer) | Error (degrees) | Error (mrad) |
|----------------|----------------|-----------------|--------------|
| 14.803         | 14.7641        | 0.0389          | 0.6789       |
| 13.978         | 13.8874        | 0.0906          | 1.5812       |
| 13.773         | 13.7495        | 0.0235          | 0.4101       |
| 18.033         | 17.9738        | 0.0592          | 1.0332       |
| 10.528         | 10.4573        | 0.0707          | 1.2339       |

Tables show only about 10% of all the obtained measurements. However, the results are significant, since the global behaviour of all measurements is similar to this one. Looking at data, two main observation arise:

1. The overall accuracy of measurements is globally satisfying, since only one measurement per-support exceeds the acceptance limit of 1.5 mrad (0.08°).
2. There is not an evident advantage in using supports with extremely low planarity error, as the calibrated granite plate. In fact, looking at data, the errors produced by using the two supports are approximately in the same range.

As shown in Table 3 and in Table 4, the results of these preliminary tests show a satisfactorily robustness of the procedure, which retrieve measurements with an acceptable accuracy, imposed at 1.5 mrad (0.08°).

Against expected results, tests demonstrated that support planarity error under 0.6 mm has only a very slight influence in the pose-estimation phase. Thus, no strict requirement on the flatness of the support is required. This allowed the adoption of cheaper components, as a plywood table, to build the marker-board.

On the other hand, test revealed a high sensitivity of the procedure to IC approximations. Therefore, calibration phase must be carried out carefully, by capturing a set comprising at least 32 pictures (see Figure 5). More in detail, calibration pictures have been shot by placing camera at a constant distance to the marker (to be maintained during scanning phase). The marker has been placed on the ground, upon a rotating support. In each picture, the marker must cover more than 80% of the frame. At least at three different angular position of the camera w.r.t. ground, the marker is rotated at list in 8 different positions. Focal length and focus distance of the lens have been maintained during scanning test.

During tests it also emerged that brightness and contrast of the marker in the pictures also influences the accuracy of the detection phase. High contrast particularly benefits ChArUco landmarks detection. In case of low contrast images, corners are not clearly detected, with an unusable result. From this observation, it derives the necessity of provide an adequate lightning for the marker, when the scanning phase happens in dark rooms.
4. Final test and conclusions
The prototype has been finally tested for the acquisition of the geometry of a painting canvas, in the Confraternita Museum, in Arezzo (Italy), as shown in Figure 6. The acquisition has been carried out by means of 112 pictures, one for each laser angular displacement (distance between adjacent laser lie projection around 10 mm).

Final test is meant to verify the accuracy emerged during preliminary tests. Both procedure and device presented in this paper revealed to be a smart alternative to the cross-ratio based method for the metric acquisition of 3D points. The physical prototype is cheap (general cost <200€) and it is easily portable, due to its reduced dimension (A3 format size) and its low weight (<1kg). 3D points are obtained directly from the set of pictures, without any user interaction, in a time of around 5-10 minutes for a set of 100 pictures by using a common notebook. Results have been appreciated by specialists.

References

[1] Vassilopoulou, S., Hurni, L., Dietrich, V., Baltzavias, E., Pateraki, M., Lagios, E., & Parcharidis, I. (2002). Orthophoto generation using IKONOS imagery and high-resolution DEM: a case study on volcanic hazard monitoring of Nisyros Island (Greece). ISPRS Journal of Photogrammetry and Remote Sensing, 57(1–2), 24–38. https://doi.org/10.1016/S0924-2716(02)00126-0

[2] Jacobsen, K., & Passini, R. (2003). Accuracy of digital orthophotos from high resolution space imagery | K Jacobsen. In Proceedings of the Workshop High Resolution Mapping from Space. Retrieved from https://www.researchgate.net/publication/229001679_Accuracy_of_digital_orthophotos_from_high_resolution_space_imagery

[3] Governi, L., Furferi, R., Volpe, Y., Puggelli, L., & Vanni, N. (2014). Tactile exploration of paintings: An interactive procedure for the reconstruction of 2.5D models. In 2014 22nd Mediterranean Conference on Control and Automation, MED 2014 (pp. 14–19).
[4] Markevicus, T., Olsson, N., Carfagni, M., Furferi, R., Governi, L., & Puggelli, L. (2012). IMAT project: From innovative nanotechnology to best practices in art conservation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 7616 LNCS). https://doi.org/10.1007/978-3-642-34234-9_83

[5] Barazzetti, L., Remondino, F., Scaioni, M., LO BRUTTO, M., Rizzi, A., & Brumana, R. (2010). GEOMETRIC AND RADIOMETRIC ANALYSIS OF PAINTINGS, 62–67. Retrieved from https://iris.unipa.it/handle/10447/51556#.XRsG8egzaUk

[6] Bimber, O., Coriand, F., Kleppe, A., Bruns, E., Zollmann, S., & Langlotz, T. (2006). Superimposing pictorial artwork with projected imagery. In ACM SIGGRAPH 2006 Courses on - SIGGRAPH '06 (p. 10). New York, New York, USA: ACM Press. https://doi.org/10.1145/1185657.1185805

[7] Gros, P. (1994). How to Use the Cross Ratio to Compute 3D Invariants from Two Images https://doi.org/10.1007/3-540-58240-1_6

[8] Mohr, R., & Gravir, B. T. (1996). Projective Geometry for Image Analysis. Retrieved from https://pdfs.semanticscholar.org/942c/eb81d8a39efa7cb108e6451c85e918786534.pdf

[9] Furferi, R., Governi, L., Volpe, Y., & Carfagni, M. (2013). Design and Assessment of a Machine Vision System for Automatic Vehicle Wheel Alignment. International Journal of Advanced Robotic Systems, 10(5), 242. https://doi.org/10.5772/55928

[10] Santachiara, M., Gherardini, F., & Leali, F. (2018). An Augmented Reality Application for the Visualization and the Pattern Analysis of a Roman Mosaic. IOP Conference Series: Materials Science and Engineering, 364(1), 012094. https://doi.org/10.1088/1757-899X/364/1/012094

[11] Köhler, J., Pagani, A., & Stricker, D. (2010). Detection and Identification Techniques for Markers Used in Computer Vision. In VLUDS 2010 - Visualization of Large and Unstructured Data Sets - Applications in Geospatial Planning, Modeling and Engineering (IRTG 1131 Workshop). https://doi.org/10.4230/OASIcs.VLUDS.2010.36

[12] An, G., Lee, S., Seo, M.-W., Yun, K., Cheong, W.-S., Kang, S.-J., ... Kang, S.-J. (2018). Charuco Board-Based Omnidirectional Camera Calibration Method. Electronics, 7(12), 421. https://doi.org/10.3390/electronics7120421

[13] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition, 47(6), 2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005

[14] Puggelli, L., Furferi, R., & Governi, L. (2020). Low Cost Device to Perform 3D Acquisitions Based on ChArUCo Markers. In Springer (Ed.), International Conference on Design, Simulation, Manufacturing: The Innovation Exchange ADM 2019: Design Tools and Methods in Industrial Engineering (pp. 189–200).

[15] Buonamici, F., Carfagni, M., Volpe, Y. Recent strategies for 3D reconstruction using reverse engineering: A bird’s eye view (2017) Lecture Notes in Mechanical Engineering, pp. 841-850.