Creating Realistic 3D Models From Scanners by Decoupling Geometry and Texture

1 Antonio Adán, 2 Pilar Merchán, 2 Santiago Salamanca
1 Escuela Superior de Informática. Universidad de Castilla La Mancha. Spain.
2 Escuela de Ingenierías Industriales. Universidad de Extremadura. Spain.
1 Antonio.Adan@uclm.es, 2 {PMerchan,SSalamanca}@unex.es

Abstract

To date the methods to create accuracy dense realistic 3D models of outdoors by using laser scanners are highly dependent on the on-site conditions in the very moment of the 3D data collection. Thus, researchers put in a lot of effort on eliminating colour incoherencies (sunny/shady, bright/dark, non sensed areas, etc.) or modelling the light of the scene to obtain free-illumination models. This paper proposes a new strategy that aims to separate the moments in which geometry and colour are taken, making the modelling process independent of the time and external circumstances in the data collection stage, and on-site calibration-free. Our approach is based on matching anytime suitable external images of the scene into the complete 3D geometric model and carrying out an iterative weighted colour mixing algorithm which gradually builds the textured model. The method has been mainly tested in archaeological sites; particularly in rest of ancient Roman buildings, yielding excellent results. Comparison with other commercial solutions is also shown in the paper.

1. Introduction.

Building accuracy and complete (geometry+texture) 3D models of large outdoor scenarios from scanners is still a challenging and unclosed matter in which many researchers work nowadays in many problems [1, 2]. Great part of these problems comes from accepting the conditions that exist in the very moment of the 3D data acquisition, which makes the fusion of geometry and colour very prone to errors (see Figure 1 left).

Restricted technical and environmental conditions can make the creation of realistic models quite impossible. For example, scanners with on-board cameras are restricted to use a specific image quality. In case of using external cameras, they have to be placed on a specific position with respect to the scanner and must be previously calibrated. On the other hand, the colour is severely affected by the weather variability (sun, rain, humidity, etc.) as well as the moment of the daytime in which the photos are taken. Here we propose a non-dependent 3D modelling method. Figure 1 shows a comparison of the 3D outdoor model obtained throughout several days of work with a scanner on-board camera and the model obtained without technical and time restrictions.

Figure 1. Low-quality textured model from usual 3D scanning techniques (left) and model created following our approach (right).

The problems in texture fusion are mainly caused by the colour disparity between images corresponding to the same sensed surface. In the case of non-controlled scenarios, which is really the most challenging case, researchers have devised only partial solutions. Many researchers have addressed solutions based on pairwise colour correction [3, 4], colour mixing [5] algorithms and image illumination change detection [6]. Most of these solutions are based on global approaches, in which the complete colour image from one position of the scanner is processed (and therefore, modified) [7], or on local strategies, in which only the colours at certain points of the image are corrected [4].

The other research line is focused on obtaining a surface reflectance map of the scene or modelling the light. Thus, the 3D model is illumination-free and can be rendered with any illumination conditions. Some
interesting approaches to obtain a reflectance function have been presented under laboratory conditions [8] and outdoors scenarios [9].

2. Building textured models: An overview.

Our strategy is based on getting the processes of geometrical data acquisition and colour acquisition apart, so they become independent. Thus, the acquisition of colour information with external high-resolution cameras could be executed at chosen moments, separated from the data acquisition for the geometrical modelling process, that is not sensitive to time or weather variations.

Let us assume the geometry of a complete 3D model \( \mathcal{M}(V,F,N) \) (\( V \)=vertices, \( F \)=patches, \( N \)=normals). Our procedure consists of four main stages: image matching, generation of the coloured ortho-image, colour assignment in the geometrical model and seam elimination.

In the first stage, two images are chosen: an ortho-photo of \( \mathcal{M} \), which will be denoted as \( I_1 \), and an external image \( I_2 \). \( I_1 \) must be taken from a similar viewpoint to the one of the \( I_2 \). The transformation between pixels of \( I_1 \) and \( I_2 \) can be modelled as

\[
P_e = AP_m \quad (1)
\]

\( P_m \) and \( P_e \) being the homogeneous coordinates of the pixels in \( I_1 \) and \( I_2 \). Matrix \( A \) is a transformation that includes rotation, translation and deformation of image \( I_1 \). The components of matrix \( A \) can be obtained from

\[
A = (M^T M)^{-1} M^T X_p \quad (2)
\]

\[
M = \begin{bmatrix}
x_{m1} & y_{m1} & 1 & 0 & 0 & -x_{m1}x_{e1} & -y_{m1}x_{e1} \\
0 & 0 & x_{m1} & y_{m1} & 1 & -x_{m1}y_{e1} & -y_{m1}y_{e1} \\
x_{m2} & y_{m2} & 1 & 0 & 0 & -x_{m2}x_{e2} & -y_{m2}x_{e2} \\
0 & 0 & x_{m2} & y_{m2} & 1 & -x_{m2}y_{e2} & -y_{m2}y_{e2} \\
... & ... & ... & ... & ... & ... & ...
\end{bmatrix}
\]

\[
X_p = (x_{e1}, y_{e1}, x_{e2}, y_{e2}, ..., x_{en}, y_{en})^T
\]

In equation (2), the pairs \((x_{m1}, y_{m1})\) and \((x_{e1}, y_{e1})\) signify pixel coordinates of \( I_1 \) and \( I_2 \) which have been put in correspondence manually.

Once \( A \) is calculated, the final corrected image, \( I'_1 \), will be the one obtained after modifying the RGB components with the RGB values of the corresponding pixels in the image \( I_2 \). Let \( C \) be the vector which contains the three colour components, \( C = [R \ G \ B]^T \). For each pixel of \( I_1 \) the colour is modify as follows.

\[
C(P'_k) = C(P_k), \quad P_k = \frac{A P_k^e}{x_k}, \quad \forall k \quad (3)
\]

\( P_k \) and \( P_k^e \) being the coordinates of the \( k \)-th pixel in \( I_2 \) and \( I'_1 \), respectively. \( C(P_k) \) is the colour stored in the \( k \)-th pixel of \( I_2 \) and \( C(P'_k) \) the colour assigned to the \( k \)-th pixel in \( I'_1 \).

The following stage consists of mapping \( I'_1 \) onto the model \( \mathcal{M} \). To do this, we use a function which maps 2D coordinates of a graphical window to the corresponding 3D coordinates of the model. At the end of this process, a \( nx4 \) matrix \( T \) contains the vertex’s indexes of the 3D model and their colour in \( I'_1 \).

When only the first external image is matched to \( \mathcal{M} \), we have a coloured model \( \mathcal{M}(V,F,N,C) \) in which only a part of its vertices has got colour information. To match and mix multiple external images into \( \mathcal{M} \) we execute an iterative procedure that adds colour information to some of the nodes as new images are matched onto the geometrical model.

Let us assume a particular step \( t \) in the iterative process. At this point, we denote \( \tilde{T}_t \) as the matrix generated throughout the earlier \( t-1 \) iterations which stores the colour information extracted from \( t-1 \) external images (\( \tilde{T}_t \) contains the RGB values of each vertex of \( \mathcal{M} \)). A new external image could provide texture for some of the vertices that have not been assigned yet, leading to a matrix \( T_{t+1} \) that must be combined with \( \tilde{T}_t \) to produce a new matrix \( \tilde{T}_{t+1} \). This process is denoted by the symbol \( \otimes \) in equation (4).

\[
\tilde{T}_{t+1} = \tilde{T}_t \otimes T_{t+1} \quad (4)
\]

Somehow, matrix \( \tilde{T}_t \) is an accumulated matrix after \( t \) iterations. The colour fusion process is defined depending on the overlapping between the new view and the current status of the model.

If \( \tilde{T}_t \) y \( T_{t+1} \) have no nodes in common, equation (4) is reduced to

\[
\tilde{T}_{t+1} = \tilde{T}_t U T_{t+1} \quad (5)
\]

Nevertheless, if \( \tilde{T}_t \) y \( T_{t+1} \) have got common nodes, the colour merging then becomes a more complex process.
The colour mixing process entails making a colour assignation by using a weighted expression. The colour updating for a node $N$ that is common in $T_t$ and $T_{t+1}$ is given by the expression:

$$\bar{C}_{t+1}(N) = \frac{C_{t+1}(N) \cdot w_{t+1}(N) + \bar{C}_t(N) \cdot w_t(N)}{w_{t+1}(N) + w_t(N)}$$

(6)

Where $C_{t+1}(N)$ is the colour of $N$ in $T_{t+1}$ and $\bar{C}_t(N)$ is the colour of $N$ in $T_t$, the weights $w_{t+1}(N)$ and $\bar{w}_t(N)$ are calculated with equations (8) and (9) from the observation angles $\theta_{t+1}(N)$ and $\bar{\theta}_t(N)$, respectively. $\bar{\theta}_t(N)$ is calculated in equation (7). A brief explanation about these weights and angles follows.

$\theta_{t+1}(N)$ is the observation angle of $N$ in position $t+1$, so that $\theta_{t+1}(N) = (\bar{n}, \bar{l})$. $\bar{l} = \bar{N}\bar{R}_{t+1}$, $\bar{n}$ being the normal at $N$ and $\bar{R}_{t+1}$ being the camera position in iteration $t+1$. The weighted-observation-angle of $N$, $\bar{\theta}_t(N)$, is computed as follows:

$$\bar{\theta}_t(N) = \frac{\theta_{t+1}(N) \cdot w_{t+1}(N) + \bar{\theta}_t(N) \cdot \bar{w}_t(N)}{w_{t+1}(N) + \bar{w}_t(N)}$$

(7)

in which weights $w_{t+1}$ are calculated by means on angles $\theta_{t+1}$ as follows

$$w_{t+1}(N) = \begin{cases} 1 & \theta_{t+1}(N) \leq a \\ \frac{\theta_{t+1} - b}{a-b} & a < \theta_{t+1}(N) < b \\ 0 & \theta_{t+1} \geq b \end{cases}$$

(8)

and $\bar{w}_t$ is based on previous weights in the following manner:

$$\bar{w}_t(N) = \sum_{k=1}^{t-1} w_k$$

(9)

In equation (8) $a$ and $b$ are two thresholds imposed to the observation angles (in our case, $a=10^\circ$ and $b=60^\circ$).

With this algorithm, matrix $\bar{T}_{t+1}$ is obtained taking into account the colour and the observation angle for each common node in $T_t$ and $T_{t+1}$.

3. Results

The approach presented in this paper is mainly being testing in archaeological sites; particularly in rest of ancient Roman building from to the first century A.D. As an example, we present here the results yielded in the reconstruction of the Forum Portico of Emerita Augusta from Merida (Spain).

For the acquisition of the geometry, we used a Faro Photon 80 laser scanner which takes 7 minutes to scan a panoramic sample of the scene. The density of data provide by the sensor is 10000x4307 points per scan. This site needed 25 scans to be completely reconstructed; having an average of 6 million points each. These data were taken during four sessions, eight hours per session. The final point cloud of the definitive model was made up of 48560845 points for the highest resolution, with a memory size of 1148 Gb.

The model building stage was run over two PC Intel® PentiumTM 4 3.2 GHz, 4 GB RAM with a NVIDIA Quadro FX 3700.

To generate our complete geometry-texture model we took, in a different day and during one hour, 126 shots of the site and applied our method. We chose a clouded but light day so that illumination conditions were the best. Table I presents some figures regarding the coloured model generation process. We include data concerning further refinement processes like colour hole detection and filling. Figure 3 a) illustrates some moments of the sequence of steps carried out to assign colour from external images to the geometrical model.

**Table I. Computational times in the multi-view colour-fusion algorithm.**

<table>
<thead>
<tr>
<th>Task</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-processing of each view (normal calculation, format conversion, etc.)</td>
<td>5 min</td>
</tr>
<tr>
<td>Definition of the ortho-image from the 3D model</td>
<td>2 min</td>
</tr>
<tr>
<td>Definition of corresponding pairs between the model and the external image</td>
<td>3 min</td>
</tr>
<tr>
<td>New colour modification/assignation</td>
<td>1 min</td>
</tr>
<tr>
<td>Colour-holes filling</td>
<td>10 min</td>
</tr>
<tr>
<td>Colour merging with the whole coloured model</td>
<td>3 min</td>
</tr>
</tbody>
</table>

We also present the results using the on-site texture for comparison. In this case, the colour information was obtained with a Nikon D200 on-board camera with Nikon AF DX Fisheye lens. This system, synchronized with the scanner, composes a single panoramic image from 10 photos, 10.2 Mpxixels each. The system takes 8 minutes to take the sequence of photos. The complete textured model was generated with the help of scanner commercial software. The procedure is here totally different since the information provided by the scanner (points with their associate colour) is registered for each scan. The total coloured point cloud is then meshed assigning to each patch the mean colour of the vertices of the patch. Therefore there is not a proper mixing colour algorithm since the points keep the original colour. Usually the result is poor.

Figure 3 b) shows the results obtained with the on-board/on-site camera (on the left) compared with the model built with external images following our modelling method. Some details of colour errors in the on-site model can be noticed in such figures. For example, note the lack of colour in the places where the scanner was positioned (noted as A in the picture).
the contrast between sunny and shady areas (noted as B) and the erroneous mix of the colours from different positions of the scanner (noted as C). As can be seen our approach eliminates all these errors and yields a much more photorealistic 3D model.

![Image of 3D model comparisons](image)

**Figure 3.** a) Selection of points in the external image and colour mapped onto the 3D model. b) Left) On-site model generated from images provided by the scanner on-board camera. Deficiencies are highlighted. Right) Our 3D model, created with off-site cameras.

4. Conclusions

Strategies based on decoupling colour and geometry can provide better photorealistic outdoor 3D models than those based on on-site data. They avoid the tedious on-site scanner-camera calibration, the limitations and restrictions imposed by the weather conditions and makes the colour integration algorithms more efficient.

We here present an approach based on matching sets of external off-site images on a previously built 3D geometrical model. The results presented show that our method provides higher-quality coloured 3D models than those built with commercial software and that they can be compared with intricate solutions based on illumination models and colour mixing algorithms.

**Acknowledgments**

This work has been supported under National and Regional Governments projects DPI2009-14024-C02-01, PCJ100402 and PRI09C088.

**References**