Advances in Computational & Experimental Engineering & Science

New images from 3D Scanning

João Ramos[1], Nuno Martins[2], Fernando Silva[3]

Summary

The reconstruction of 3D world scenes problem based on images is being studied in so diverse scientific areas such as engineering, medicine and modern industries in general. In this article we present a shape and colour reconstruction, using a volumetric technique known as space carving. Based on that information, it is shown how to create new images of the reconstructed scene, from any world position, in real time. This allows overcoming the problem of blind or lack of images in a computer vision process.

Introduction

The use of computers in the industrial production of physical objects from digital models is greater every day, specially in the geometric design and creation areas. However, to the inverse problem, meaning the inference of the digital description from the same physical objects is being paid less attention. This problem is generally referred as reverse engineering or, more specifically, 3D scanning. There are several 3D characteristics in an object that it would be interesting to recover. Those characteristics include the shape, the colours and the properties of the object material. This article focuses the problem of 3D shape recover, as well as its colour.

There are many ways to recover shape and colours from a 3D world scene. In computer vision area, in which we include this article content, that reconstruction is accomplished through the processing of light sensors data (cameras or lasers). So, it is necessary to acquire images of the 3D scene. It is also necessary to know the geometric relationships between the camera and the surfaces that must be recovered (the intrinsic parameters of the camera and the extrinsic parameters - position of the camera, referred to the origin of the 3D world). With this knowledge a 3D world point could be related with an image point, more specifically, with its projection point in the image. This is the key of any reconstruction process. In order to reconstruct the object in a global way, there should be some kind of movement between the sensor and the object.

Images are the fundament for any task in computer vision. If there is a way to give images to blind or image fault processes, those tasks have a better performance. For example, some times there is the need of acquiring 3D scene images from places that are

[1]Systems and Informatics Engineer Department, Superior Institute of Engineering Institute of Coimbra, Polytechnic Institute of Coimbra, Portugal. jramos@alunos.deis.isec.pt http://sigma.gamedev-pt.net

[2] Systems and Informatics Engineer Department, Superior Institute of Engineering Institute of Coimbra, Polytechnic Institute of Coimbra, Portugal. ncmartin@isec.pt

[3] Systems and Informatics Engineer Department, Superior Institute of Engineering Institute of Coimbra, Polytechnic Institute of Coimbra, Portugal. fsilva@isec.pt

Advances in Computational & Experimental Engineering & Science

physical impossible to take with real cameras, or to give sight to blind devices that could not support cameras. Thus, to overcome this problem there must be a way to create new images from other real images for those tasks. This paper proposes a solution to that problem.

This paper results from a proposed task in the computer vision class in the System and Informatics Engineer Department in the Superior Institute of Engineering of Coimbra, and is organized as follows. First it is described the image acquisition and camera calibration process. Then it is presented what is the space carving theory based on silhouettes, and shown its application to shape and colour reconstruction. Finally, it is described how to obtain new images from the recovered scene.

Image acquisition and camera calibration

The images are the base of all reconstruction process in computer vision. In this work it is used a unique camera to capture all the images from the 3D scene. Those images are acquired from several different positions. As the number of images of the scene used in the reconstruction is greater, better is the result. To have a full form recover it is advised to acquire images around the scene.

Another important issue in the recovery process is the definition of a linear relation between a 3D point and an image point. To use an image in the reconstruction process, this projective relation between the world scene and that image must be known. This relation is given in homogeneous coordinates by

$$\mathrm{kp_h} = \mathrm{TP_h} = \mathrm{C}\,[\mathrm{R}\ \mathrm{t}]\mathrm{P_h} = \begin{bmatrix} \mathrm{fk_x} & \gamma & \mathrm{c_x} \\ 0 & \mathrm{fk_y} & \mathrm{c_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_t \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \mathrm{P_h} \tag{1}$$

where P_h is a 3D homogeneous point, p_h is its projection in the image, T is the projection matrix and k is a multiplicative factor. As one can see in equation (1), the projection matrix is composed by matrix C, known as calibration matrix, and by a rotation matrix R and a translation vector t, both referred to the origin of the world referential. Thus, if it is known a 3D point one can calculate the position of its projection in the image.

The camera calibration process can be described as the estimation of the elements of the projection matrix T, and use all the images that take part of the reconstruction process. The matrix C represents the intrinsic parameters of the camera. If it is used only one camera in the acquisition of the images, the matrix C for each image is the same. So, instead of having to estimate all the elements of the matrix C multiplied by the number

Advances in Computational & Experimental Engineering & Science

of images, one have only the elements of that matrix to calculate, optimizing the task. That is the reason for the use of only one camera in our work.

In this work, the camera calibration process uses the routines from Intel computer vision library, known as OpenCV routines [1]. This toolkit provides a mechanism that enables the user to create a projection matrix associated with each image used. It use information provided by the user for each image and calculates the camera's intrinsic and extrinsic parameters. To accomplish this phase, for each image, the user is asked to provide four coplanar points in each image (the same points must always be used). All parameters are then calculated through the equation (1) and the knowledge of the correspondence between chosen points in all used images.

Reconstruction by space carving based in silhouettes

Although there are several techniques to solve the problem of the 3D reconstruction of a surface from a set of images, these solutions are numerically steady for restricted and well-defined problems [2] [3]. Recently, some alternatives to these techniques have been developed with volumetric modelling algorithms based on the three-dimensional space carving principle. It was demonstrated that very complex structures with algorithms could be recovered, based on this principle [4] [5].

The space carving theory is a technique that can get a solution for the problem of the surface reconstruction in a general way and from a set of images [4]. This technique defines, at the beginning of the excavation process, a volumetric space where the surface to be reconstructed fits. This volumetric space is carved until the resulting surface is consistent with the acquired images, that is, until every point within the volumetric space is consistent with every point projected on those images. A 3D point, belonging to the volumetric space, is consistent with a point in an image if the colour of this 2D point is the result from the calculation of the three-dimensional point radiation. This is valid for reflection models, which are estimated locally, being generated a 3D volume consistent with the captured radiation in the image acquired from that point of view.

The goal of the global consistence test is to decide if there is a calculated radiation value that can be attributed to a point in space, so it can be considered consistent with all the used images. It is estimated the consistence criterion for a voxel and, if it is consistent, a radiation value is attributed and the voxel is accepted. Otherwise, it will be eliminated from the surface we are reconstructing. This process is renewed till there is no more voxels to eliminate or accept. The remnant of the volumetric space must be the shape of the surface we would wish to recover. Kutulakos and Seitz proved that [4], if this surface is found, then it is consistent with all the images on the surface collected from different points of view.

We have chosen this volume intersection approach due to its capacity for describing objects with a more complex topology (objects with holes, for instance). Thus, we follow

Advances in Computational & Experimental Engineering & Science

the work of [5] and [6], a shape reconstruction algorithm which creates an octree from several images. In this work, like in [6], we use a variation to the consistence test presented in [5]. Instead of using the function that models the scene radiation, we use the silhouettes (or outlines) and the projections of the voxels in the images. It is obtained, with this criterion, the occupation rate of the surface in each voxel. The silhouettes often are a dominant characteristic in the images and can be extracted easily. They give important information concerning the shape of the objects and it is, in reality, the sole information available regarding the smooth surfaces almost without a texture [5].

Beyond the recovered surfaces, we introduced in the described process the recovery of the surface colour. The idea associated to this recovery is that the colour of a surface is just a set of colours that can be seen in the images of the surface. So, this information follows the colour of the projection of the voxels in the images.

The algorithm used in the space carving process of the presented work, is:

1. Establish a cube that contains the scene to recover and set it as the grey root node of the octree;
2. While there are grey nodes in the current level of the octree do,
    a. For each grey node of the current level do,
        i. Subdivide it into eight sub nodes;
        ii. For each sub node do,
            1. Set its colour to black;
            2. For each image in the sequence do,
                a. Project the cube onto the image;
                b. If the projection lies completely outside the silhouette, set the colour of the correspondent node to white and ignore the rest of the images;
                c. Else
                    i. If the projection lies partially inside the silhouette set the colour of the correspondent node to grey;
                    ii. Else keep the colour of the node;

New image creation

Since the information regarding the object's topology in now known, one can view that topology from any wanted position, and in real time. To obtain new images one must use the information that results from the camera calibration process. Choosing the matrix

Advances in Computational & Experimental Engineering & Science
Copyright 2004 Tech Science Press

associated to one of the images and applying to that matrix the recovered points multiplied by a matrix M, the projection point of the new image is calculated. Mathematically, we multiply the 3D homogeneous point $\mathrm{P_h}$ in equation (1) by M, obtaining

$$\mathrm{kp_h = C[R \quad t]MP_h = C[R \quad t][R_{wanted} \quad t_{wanted}]P_h} \tag{2}$$

where matrix M aggregate a rotation and a translation, which locate the camera in the wanted position. The new images are consistent with the recovered model.

Experimental results

Having all the theory laid out, we can now present an example of its implementation. In such, figure 1 shows the images that were captured with the same camera, so that the surfaces on them could be reconstructed.

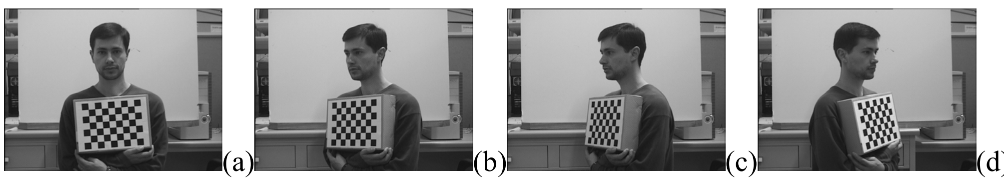

Figure 1 - The images used in the reconstruction process.

Using OpenCV routines to calculate the images of figure 1, we obtain the following intrinsic parameters of the camera

$$C = \begin{bmatrix} 1539.9 & 0 & 167.1 \\ 0 & 1433.4 & 394.1 \\ 0 & 0 & 1 \end{bmatrix}$$

The obtained rotation matrices and translation vectors, related to each images, are

$$R_a = \begin{bmatrix} -0.07 & -0.99 & -0.11 \\ -0.99 & 0.07 & -0.01 \\ 0.02 & 0.11 & -0.99 \end{bmatrix} R_b = \begin{bmatrix} 0.02 & -0.67 & -0.73 \\ -0.99 & 0.03 & -0.06 \\ 0.06 & 0.73 & -0.67 \end{bmatrix} R_c = \begin{bmatrix} 0.09 & -0.39 & -0.91 \\ -0.99 & 0.00 & -0.10 \\ 0.03 & 0.91 & -0.39 \end{bmatrix} R_d = \begin{bmatrix} -0.20 & -0.67 & 0.70 \\ -0.97 & 0.15 & -0.12 \\ -0.02 & -0.71 & -0.69 \end{bmatrix}$$

$$t_a = \begin{bmatrix} 34.1 \\ 7.8 \\ 158.6 \end{bmatrix} \qquad t_b = \begin{bmatrix} 26.6 \\ 7.8 \\ 155.5 \end{bmatrix} \qquad t_c = \begin{bmatrix} 23.0 \\ 7.9 \\ 156.2 \end{bmatrix} \qquad t_d = \begin{bmatrix} 37.5 \\ 7.9 \\ 169.9 \end{bmatrix}$$

Applying the implemented space carving algorithm to the images of figure 1 and using the previous matrix information, we obtain the results shown in figure 2. Analysing this figure we verified that the form of the bust is not smooth. The reason is the number of levels of the octree, 9 in this example, is less than the necessary. More levels in the octree improves the reconstruction precision (the bigger is the level, the little is the

Advances in Computational & Experimental Engineering & Science

voxel). The perfect smooth is accomplished when the voxel is a 3D point. Comparing figure 1 with figure 2, it is possible to verified that the colour it's also recovered. In spite of all, the process shows great potential.

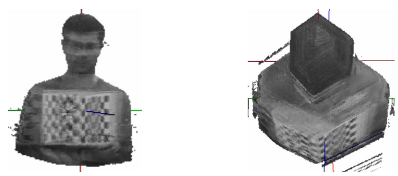

Figure 2 - Views of the reconstructed surfaces with colour

What one sees in figure 3(a) is the new image produced if the camera were rotated a 45º angle around the Z axis from the position defined by R_a and t_a, when the camera tool the image 3(b). Figure 3(c) is the new image created if the camera were translated in the Z axis -40 units from the position defined by R_b and t_b, when the camera tool the image 3(d). Please note that the model, showed in figure 2, isn't centred at the origin.

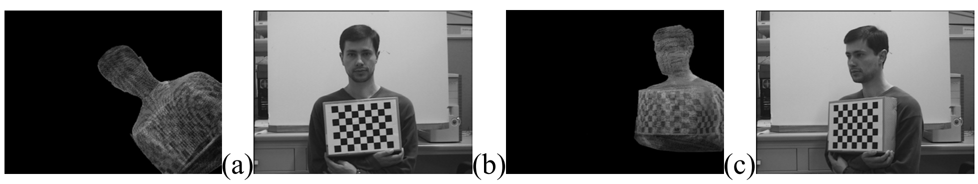

Figure 3 - Images (a) and (c) are created from the recovered structure and uses the position where the camera took images (b) and (c).

Reference

1 OpenCV url: http://www.intel.com/research/mrl/research/opencv/

2 D. Marr and T. Poggio, "Cooperative computation of stereo disparity", Science, Vol. 194, pp. 283-287, 1976.

3 R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours", Int. Journal of Computer Vision, Vol. 9(2), pp. 83-112,1992.

4 K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving", In Proc. 7th Int. Conf. on Computer Vision, pp. 307-314, Kerkyra, Greece, September 1999.

5 K.-Y. K. Wong and R. Cipolla, "Structure and motion from silhouettes", In proc. 8th Int. Conf. on computer vision, Vol. 2, pp. 217-222, Vancouver, B C, Canada, July, 2001.

6 R. Szeliski, "Rapid *octree* construction from image sequences", Computer Vision, Graphics and Image Processing, Vol. 58(1), pp. 23-32, July 1993.