Space carving with a hand-held camera

Anselmo Antunes Montenegro, Paulo C. P. Carvalho, Luiz Velho
IMPA–Instituto Nacional de Matemática Pura e Aplicada
Estrada Dona Castorina, 110, 22460-320 Rio de Janeiro, RJ, Brasil
{anselmo,pcezar,lvelho}@visgraf.impa.br

Marcelo Gattass
Puc-Rio - Departamento de Informática
Rua Marquês de São Vicente, 225, RDC, 22453-900, Gávea, Rio de Janeiro, RJ, Brasil
mgattass@tecgraf.puc-rio.br

Abstract

This paper presents a 3D scene reconstruction method, based on space carving, that works with a hand-held camera. In our system, the intrinsic and extrinsic parameters of the camera are determined at the moment of image capture, as opposed to other systems that rely on fixed pre-calibrated camera setups. In order to do this we place a special calibration pattern in the scene in such a way that it does not alter scene visibility. However, the calibration pattern may be partially occluded by the objects of interest in the scene. This has led us to adopt a calibration method based on model recognition. Scene reconstruction is obtained from the set of input images by an adaptive space-carving algorithm that uses not only photometric information but also segmentation information. The segmentation information of a given input image is determined by a robust statistical test based on an approximate model of the scene's background. Such model is computed from a set of images of the scene's background that are warped in such a way that they match the geometry of the desired camera.

1. Introduction

Space carving is a simple and efficient method for image-based scene reconstruction. Despite having some advantages, such as simplicity and conciseness, it has some drawbacks and fundamental requirements. One important requirement is knowledge of the parameters of the cameras associated to the input images. For this reason, most space-carving systems adopt fixed camera setups, which enable easy simultaneous camera calibration. Although the use of these setups constitutes a simple and adequate solution for many cases, there are situations in which this is not acceptable. This is the case, for instance, when one needs to reconstruct objects that cannot be moved from their site, or when it is not possible to carry the reconstruction apparatus to the place of interest. A common example of this is the reconstruction of archeological artifacts.

In this work we propose a solution to this problem based on a space-carving reconstruction system that uses a hand-held commodity camera and a calibration pattern. One of the main advantages of this system is that it can be used at any place and does not require moving the object of interest to the place where the reconstruction apparatus is located.

Our system uses a special calibration pattern that is inserted in the scene so that the acquisition camera can be calibrated at the moment of image capture. One of the main difficulties with this approach is that the calibration pattern may be partially occluded by the objects in the scene. To solve this problem we adopted a robust calibration method based on model recognition.

Object reconstruction is done by an adaptive space-carving algorithm that uses not only photometric information but also segmentation information. The segmentation information of a given input image is determined by a robust statistical test based on an approximate model of the background of the scene. Such model is computed from a set of background images taken from the scene space that are warped in such a way that they match the geometry of the desired camera.

This paper is organized as follows. In section 2, we describe the fundamental ideas related to space-carving methods. In section 3 we describe the adaptive space-carving approach used in our system and its advantages in relation to other space-carving approaches. In section 4 we describe in details the calibration process and the foreground segmen-

tation technique that was used. In section 5 we show and analyze some results. Finally, in section 6 we draw some conclusions and present some final comments about this work.

2. Space Carving

Space carving is a class of volumetric-based methods used to solve the problem of reconstructing the shape and color information of a scene from a set of input images.

Definition 1 (Image-based reconstruction problem) *Let S be a three-dimensional scene in a closed subset $U \subset \mathbb{R}^3$ and let I be a set of images of S taken from a set of cameras C. Determine the shape V, composed by a subset of properly colored points of U, that reproduces the input images when rendered from the original viewpoints of the cameras C.*

The problem of scene reconstruction from a set of images is an ill-posed problem. This occurs because, in general, there is more than one shape that is able to reproduce the set of input images of the scene when rendered from the original viewpoints [5].

In order to pose this problem appropriately, Seitz and Dyer [9], and, later, Kutulakos and Seitz [5] introduced the concept of point photo-consistency.

Definition 2 (Point photo-consistency) *An infinitesimally small point p on a Lambertian surface S is photo-consistent with a set of images I if the projections of p onto the images in which p is visible are of the same color.*

Based on a photo-consistency analysis of subsets of points in the 3D space, Seitz, Dyer and Kutulakos have identified the existence of a shape with unique properties. This shape consists of the maximal shape that subsumes all photo-consistent shapes, which was denominated *Photo Hull*. It was also proved that the *Photo Hull* can be obtained by the systematic removal of all non-photo-consistent points of the scene-reconstruction space. This strategy used to obtain the $PhotoHull$ is the key idea behind all space-carving methods.

In practice, some problems must be taken into consideration in the specification of a computational method based on this idea. First, we cannot work with an infinite subset of points. Second, we cannot rely on a photo-consistency criteria that is based on the color equality of infinitesimal points and that assumes that the scenes are perfectly Lambertian. Consequently, in space carving we represent the reconstruction space discretely as an array of voxels and use photo-consistency criteria based on statistic measures.

The statistic measures are used to verify if the difference observed among the colors in the projections of a voxel v are explained by a statistical model of error that takes into consideration the noise introduced by the camera sensors and the non-Lambertian components in the radiance of the scene surface.

Several statistic measures can be used. Seitz, for instance, proposed a likelihood ratio test, based on a χ^2 statistic. More recently, Broadhurst proposed the use of F-statistic measures in his multiple-threshold voxel coloring [2]. Other approaches, based on more complex probabilistic models, were proposed by Bhotika [1] and also by Broadhurst [3].

Another crucial problem in space carving is visibility determination. Seitz and Dyer's Voxel Coloring assumes that all points in the reconstruction space can be globally ordered in relation to the calibrated camera centers. This assumption considerably simplifies the determination of visibility of the partial scene, at the expense of greater generality. The existence of such ordering is also assumed in the present work.

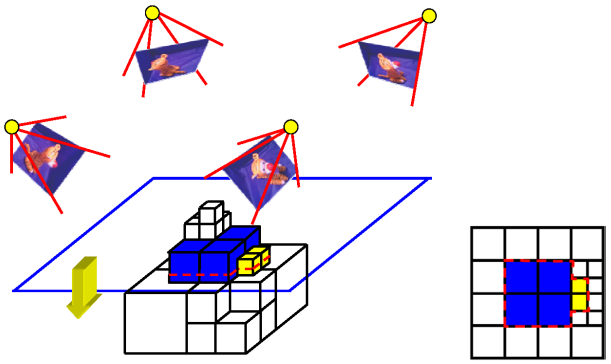
The simplest mechanism to deal with visibility issues is to associate visibility maps to each of the cameras. Initially all elements in the visibility maps are set to zero. Then, when a consistent voxel is found, all visible pixels in its projections in the input images are set to 1. This strategy is used by [9] and [5]. Other more intricate data structures can also be used, such as the *layered depth images* and the *item buffers* described by [4].

3. Adaptive Space Carving

Voxel-based space carving is a simple and powerful way to reconstruct the shape and color information of a scene. Nevertheless, due to its intrinsic structure, voxel carving spends too much effort evaluating a large number of small elements that do not belong to the scene's surface, which is the real target of the 3D reconstruction process.

In order to avoid this, we use here an adaptive space-carving algorithm by progressive refinement [6] (see also Prock's work [7] for another adaptive method). In the beginning of the process, we start with a coarse representation of the reconstruction space which is defined by a bounding box of the scene associated to the cell in the root of the octree. Then, we try to classify this initial cell according to a given photo-consistent criterion. If we succeed in classifying the cell, either as photo-consistent or as non-photo-consistent (transparent), then we can finish processing that cell and attribute the adequate color to it. On the other hand, if we fail to classify it, that is, if we can not decide precisely whether the cell is photo-consistent or not, then we subdivide it in eight new cells and repeat the same classification procedure for each of the new cells created. This process is successively repeated until all regions are classified or a maximum resolution is obtained (Figure 1(a)).

In this work, we assume as in Seitz's work [9] that there is a partial ordering of the points in the reconstruction space



(a) Adaptive Space Carving (b) Selection

Figure 1. (a) - Adaptive space carving on an octree. The cells are processed in the order shown by the arrow. (b) - selection of cells in a layer at a given level of refinement (the light-colored cells are the ones to be processed)

according to their distances to the centers of projection of the cameras. Hence, at a given level of refinement, we classify the octree cells in the order defined by their distance to the cameras.

For a certain level of refinement of the octree, we can identify a voxel array structure whose elements have resolution compatible with the cells in the leaves of the octree. The layers of the corresponding voxel array structure induce virtual layers in the octree structure that define the partial order in which the cells are evaluated.

Nevertheless, differently from voxel arrays, only the octree cells that belong to the current level of refinement need to be evaluated. The cells that belong to coarser levels of refinement either have been classified or subdivided at the previous level. To determine the active octree cells in a given layer we can simply intersect a plane that approximates such layer with the whole octree. Once the intersection is determined, we select only the cells that belong to the current level of refinement. This strategy is shown in figure 1(b).

3.1. Octree cell classification and refinement

Our space-carving scheme classifies octree cells according to their photo-consistency in three groups: *consistent*, *inconsistent*, and *undefined*. In our classification scheme, consistent cells are the cells that were classified as consistent according to certain photo-consistency evaluation criteria. On the other hand, we restrict the group of inconsistent cells to those cells in the highest octree resolution level that failed the photo-consistency test and to those large cells in intermediary levels of resolution that are completely projected on the background.

The group of undefined cells includes those cells that, at a certain stage of the process, cannot be classified, neither as consistent nor as inconsistent. This may occur when a given cell contains points that belong to the empty space together with points that belong to the scene's surface, or when the photo-consistency criteria cannot precisely decide, whether the cell is consistent or inconsistent.

Whenever a cell is considered undefined, it must be subdivided in eight new octree cells that are evaluated in the next space carving iteration, which works in higher resolution.

3.2. Photo-consistency evaluation

In our approach, the photo-consistency evaluation of a cell c is done at a planar region defined by its intersection with a layer plane(Figure 1). Each layer plane here is called a *registration plane*. The evaluation itself is based on a hypothesis test that relies on information about the scene, stored in the color and opacity channels of image maps. The information corresponding to a cell is registered in scene space via OpenGL [12], by projecting the image maps as texture maps onto the registration planes. A similar approach was proposed by Sainz [8]. In the following subsections, we address in detail the issues involved in the photo-consistency evaluation of an octree cell.

3.2.1. Information used by the photo-consistency test Original space carving uses two kinds of information in the photo-consistency test: the *photometric information* that comes from the input images and the *visibility information* computed as the process is executed. Although many researchers consider this too restrictive for a space carving approach, it is also possible to use *segmentation information*. In our method, segmentation information is used, as well as a new kind of information, namely the *noise maps of the cameras*. Each noise map consists of an estimate of the noise in each color component of the sensors of a camera given by the standard deviations of the observed values in the components rgb.

For each image, we use two $rgba$ maps to store all the information used by the photo-consistency test. The first one, denominated *Photometric and Segmentation Map* (PSM) stores the photometric information in channels r, g and b , and the segmentation information in channel a. The second map, denominated *Visibility and Noise Map* (VNM) stores the estimated noise of each color component of the sensors in the components r,g and b, and the visibility information in channel a.

3.2.2. Registration To evaluate the photo-consistency of a given volumetric element v, we must sample the informa-

tion that will be used in the photo-consistency test by re-projecting v onto the projection planes of the cameras. As the projections of v may contain several elements, these elements must be registered in some way.

In our method, we register all the information on planes that approximate the virtual layers induced by the corresponding voxel array structure that also defines the cell's classification order.

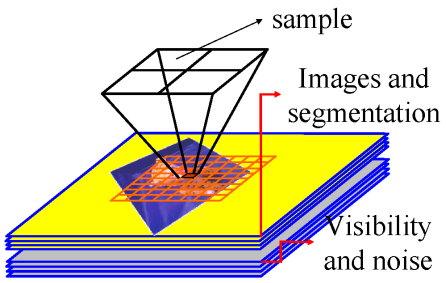


Figure 2. Registration

When testing consistency for a given virtual layer, we project the Photometric Segmentation Map and the Visibility and Noise Map for each image onto the registration plane, obtaining the corresponding *Projected Photometric and Segmentation Map* and the *Projected Visibility and Noise Map*.

Our strategy automatically solves the problem of registration and provides an adequate treatment of aliasing problems by projecting the textures as mipmaps. Moreover, fast photo-consistency evaluation is also a possibility if graphical hardware programming is used.

3.2.3. Sampling The intersection of a registration plane with an octree cell c defines a planar region pr inside c. The number of elements inside pr is determined by the resolution of the projected texture map which must be proportional to the spatial resolution of the current level of resolution of the octree. To make anti-aliasing more robust, we can generate the projected information maps with a resolution that is greater than the spatial resolution of the current octree.

Each element inside pr is addressed by indices s and t, in a coordinate system associated to the grid induced by the intersection of the registration plane with the supersampled voxels. For each element e_{st} we will take a sample $PPSM^i_{st}$ from the ith *Projected Photometric and Segmentation Map*(PPSM) at the position (s,t). Similarly, we will take a sample $PNVM^i_{st}$ from the ith *Projected Noise and Visibility Map* at the position (s,t).

3.2.4. Statistical Photo-consistency Test First of all, in order to evaluate the photo-consistency of a given planar region pr associated to a cell c we must determine the set of visible elements VE in pr composed by elements ve_{st} that are visible in some camera. If VE is empty, then pr is not seen by any camera and the cell is trivially classified as *consistent*.

In the next step we verify whether there is a visible element ve_{st} in VE that projects in the foreground of an image. When this is true, the cell is classified as *undefined* if the process is not at maximum resolution. If the process is already i maximum resolution, we classify it as *inconsistent*.

When it is not possible to take any decision based on the segmentation and visibility information, we must check the consistency of the photometric information.

In order to do this, we adopt a statistical approach in which the photometric information obtained from each camera for an element ve_{st} is considered as the realization of a random variable X_i, with normal distribution, and mean μ equal to the corresponding surface color and variance σ^2, corresponding to the level of noise introduced by the sensor. As we are using here only one camera, the likelihood test described in section 2 (also see [9]) is adequate for our purposes. When more than one camera is used, as in fixed setups, the variances of camera responses may differ significantly and this test is not the best choice anymore. In a previous work [6], we have derived a likelihood test for this more general situation.

The filtering operation applied in the registration of the images at intermediary levels of resolution causes some attenuation to the deviations observed in corresponding elements. To compensate this attenuation we propose an adjustment that is based on the fact that the mean $\overline{X}$ of a set X of independent samples $X_1, X_2, \cdots, X_n$ with the same variance σ^2, has variance σ^2/n. Hence, the thresholds associated to the hypothesis test in a given level must be multiplied by a factor that is proportional to the standard deviation of the corresponding level. For more details see our previous work [6].

This operation produces thresholds that make the photo-consistency criteria more rigid in the beginning of the process and more tolerant in the later stages.

3.3. Color attribution

In our approach, the color attributed to a photo-consistent cell is given by a weighted mean of the colors observed in its projection on the input images. This strategy may produce inadequate results when very few cameras register a region of the scene with simple geometry but with highly variable texture. In this case such region will be classified as photo-consistent and will be colored in a way that does not adequately represent the variability observed in its projection onto the images. Al-

though this problem may happen in some cases, we believe that in practice this will rarely occur. This can be explained by the intrinsic properties of the adaptive method proposed.

The adaptive refinement strategy, combined with the use of segmentation information in increasing levels of resolution, favors the subdivision of cells near the surface of the scene to be reconstructed.

Moreover, the thresholds increase as the resolution is reduced to compensate the reduction in the observed deviations due to the anti-aliasing operation. Therefore, the probability that a large heterogeneous regions be classified as photo-consistent is very low.

4. Camera Calibration

Camera calibration is one of the most important aspects related to 3D photography. In space carving, this is a fundamental process, as the knowledge of the intrinsic and extrinsic parameters is a requirement to the application of the method. When fixed camera setups are adopted, the calibration process is usually done in a pre-processing stage, before the reconstruction process takes place. In such cases, a calibration pattern is normally placed in the scene space to support the calibration of the cameras. Then, the pattern is removed and replaced by the object of interest.

In our system, this cannot be done because we use a hand-held camera. Each image captured must have sufficient information to allow the determination of the calibration parameters related to the camera that produced it. In order to produce scene images with such properties, a calibration pattern is inserted in the scene space, together with the objects of interest, in such a way that it does not affect their visibility.

One problem with this approach is that the calibration pattern will be partially occluded by the objects of interest present in scene space. Hence, we must use a calibration method that works even when the pattern is partially occluded.

The method we adopted is the one presented by Szenberg et al [10], which is based on a model-recognition approach. The main idea of this method is to automatically extract pairs of corresponding points in image space and world space from a pattern composed of line segments. In the beginning, the method tries to extract long segments that will be recognized by a model-recognition procedure based on an interpretation tree. The recognized segments produce a sufficient number of points to determine a homography between the image space and the plane that contains the model in world space. This homography enables us to find all line segments in the image of the pattern. Finally, a set of corresponding points in image space and model space is used as input data to a calibration routine (we use Tsai's method [11]) that finally obtains the intrinsic and extrinsic parameters of the desired camera. In Figure 3 we show the results of the calibration of an input camera. The cones placed on the image, positioned at the corners of the pattern, demonstrate that the calibration was performed correctly.

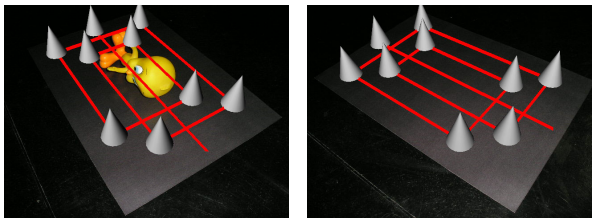


Figure 3. Input image and background image with markers verifying the calibration

5. Foreground Segmentation

In our setup, the carving process can be considerably more difficult than in other configurations because the objects of interest are surrounded by a homogeneous background defined by the calibration pattern. It is well known that photo-consistency based methods, such as space carving, produce cusps in parts of the reconstructed shape corresponding to homogeneous regions. When these regions are large, the geometry of the reconstructed object may differ considerably from the geometry of the original object. In this cases foreground segmentation must be used to produce better results.

Another reason for using foreground segmentation information is that the adaptive space-carving method used to reconstruct the objects of interest in the scene has better performance when such information is available for the input images.

Although this is considered a tricky problem, foreground segmentation can be solved in a simple way when fixed camera setups are used. In such cases, it is usual to compute an estimated model of the background for each camera based on a set of images taken from the scene without the objects of interest. Further on, a statistical test based on the computed model is used to determine which regions of the input images correspond to the foreground.

In our case, we cannot directly estimate the background model associated to each input image. We must bear in mind that previous image samples of the background will never match exactly the new input images taken from the scene. Because of this problem, we propose a strategy in which we compute an approximate model of the background from a set of images of the scene, without the target objects, taken

from arbitrary points of view. These images are registered to the input image by a warping transform so that a statistical model of the background can be computed later.

5.1. Registration of the background images

In order to compute an approximate background model for a given input image I, we must register to I all background images in the set of background images SBI .

The key observation here is that the background images are related to a given input image I by an homography, because all observed points lay on the known common plane $z = 0$. Hence, in order to register a background image BI to a given input image I we can apply a warping transformation on BI corresponding to the mapping that relates BI to I.

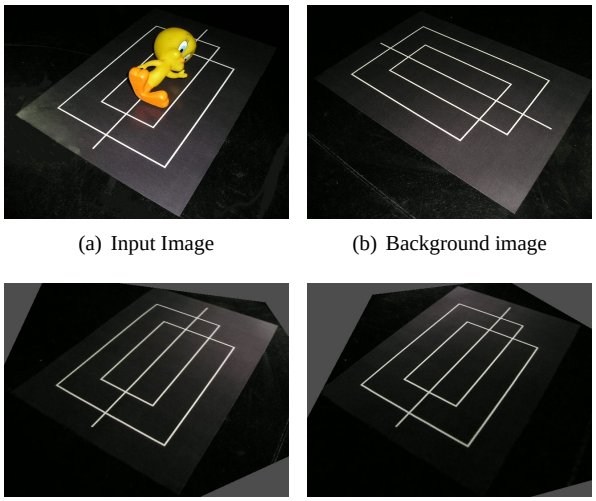
Here we assume that the intrinsic and extrinsic parameters of the cameras associated to the input image I and to the background images in SBI are known. Let Rt_1 be a 4×4 matrix that describes the intrinsic parameters corresponding to the camera c_1 that produced I, and K_1 a 3×3 matrix that describes its corresponding intrinsic parameters. Similarly, let Rt_2 be the extrinsic parameters matrix corresponding to the camera c_2 that produced a background image BI, and K_2 its intrinsic parameters matrix.

The homography between I and BI is determined by the matrix $H = K_1(Rtm_1)(Rtm_2)^{-1}K_2^{-1}$, where Rtm_1 and Rtm_2 are obtained by removing the third column of the matrices Rt_1 and Rt_2 (the entries in these columns are the coefficients of the z-coordinate which is known to be zero).

Intuitively, this matrix takes the points that define the bounds of the image in the coordinate system of camera c_2 to a polygon in the plane z=0 in world space, and then projects this polygon back to the projection plane of the camera c_1.

The warping transformation based on the homography described by H can be easily implemented in OpenGL using projective texture mapping. Let p_0, p_1, p_2, p_3 be the corner points of BI with homogenous coordinates in the coordinate system of c_2. We can compute the corresponding warped points wp_0, wp_1, wp_2, wp_3, that describe the transformed image as seen by the camera c_1, by multiplying H by the coordinates of the points p_0, p_1, p_2, p_3.

Finally, given the warped points in homogeneous coordinates, we draw the corresponding polygon with the original background image BI as texture. Figure 4 shows the registration of three background images (of a total of four) to an input image.



(a) Input Image (b) Background image

(c) Registered background image (d) Another registered background image

Figure 4. Registered images

5.2. Computing the approximated background model

The approximated background model we adopt is based on the fact that the observed color value in a pixel of a given image of the background can be seen as a realization of a random variable with approximated normal distribution with unknown mean μ and unknown variance σ^2.

As the parameters of the distribution are not known, we need to estimate them using a set of samples. This estimate is given by an interval of confidence for the mean value μ of a given pixel of the background image. The computation of this interval is based on the error between the sampling mean $\overline{X}$ and the population mean μ. According to the *Central Limit Theorem* such error $e = (\overline{X} - \mu)$ has distribution approximately equal to $N(0, \sigma_{\overline{x}}^2)$, where $\sigma_{\overline{x}}^2 = \sigma^2/n$.

Hence, the probability that e is less than a multiple of the standard deviation of the samples is

$$P(|e| < \delta\sigma_{\overline{x}}) \quad = \quad P(|\overline{X} - \mu| < \delta\sigma_{\overline{x}}) = LC,$$

where LC is the level of confidence.

Rewriting this probabilistic statement as

$$P(\overline{X} - \delta\sigma_{\overline{x}} < \mu < \overline{X} + \delta\sigma_{\overline{x}}) = NC \tag{1}$$

we have the interval with level of confidence LC

$$ico = \overline{X} \pm \delta\sigma_{\overline{x}} \tag{2}$$

Thus, our background model consists of a map that has one interval of confidence for each pixel of a hypothetical image of the background.

One problem with this simple model is that the registration is not perfect in regions of high frequency, such as near the line segments that define the calibration pattern. As a consequence, the estimated model may be inaccurate in the neighborhood of such regions. More specifically, the intervals of confidence in the elements in these regions will be larger than they should because of the displacement of high-frequency features. This may cause the classification of foreground pixels as background pixels, which cannot be acceptable.

In order to avoid this problem we propose a different way of building the maps of intervals of confidence. We do not compute the means and standard deviations for a given position (i, j) of the map by directly sampling the warped background images in exactly the same position. Instead, we sample the element in a given neighborhood of (i, j) whose observed value is closer to the observed value in the input image.

This heuristic has the effect of ignoring small nudges and shuffles that occur among the high-frequency features that were registered. As a result, the maps of intervals of confidence hereby constructed are robust enough to deal with registration errors and can be used in a correct background subtraction.

5.3. Background Subtraction

As we work in rgb space, it is necessary to define one interval of confidence for each color component using the process described above. Once these intervals are known, a simple statistical test can be used to decide whether a given pixel of the input image belongs to the background or to the foreground of the image. This is done by verifying if the observed values for each color component, in a given position (i, j), belong to the respective intervals of confidence in the corresponding position of the map. The left image of Figure 5 show the segmentation of the object based on a background model constructed without the similarity sampling in the neighborhood of a pixel. We can see in this case that parts of the object that covered the line segments of the pattern were lost. The image to the right shows the correct results obtained when the adequate sampling was done.

6. Results and analysis

The tests were executed in a 2.4 Ghz Intel Pentium IV with 1Gb of Ram memory equipped with a GeForce 4 graphic board. We used as data test a set of four 640x480 input images of a plastic toy obtained by means of a commodity digital camera. The dimensions of the reconstruction space are 25.6x25.6x25.6 cm. The background images were also taken from four different viewpoints, although more images could be used. These background images were registered to each input image so that an approximate model of the background could be computed for each of them.

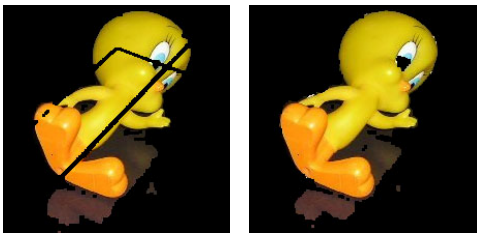


Figure 5. Wrong and correct segmentations

Figure 6 shows a magnified view of the input images and figure 7 shows the images of the object reconstructed as an octree with resolution comparable to a voxel array of 480^3 elements. We can see in the pictures that the results obtained are compatible with the input images used. The images from other viewpoints (the bottom pictures) were also correct even though very few images were used in the reconstruction process. Some parts of the object, such as the top of the head, proved quite difficult to reconstruct. This will probably require a better way to determine the threshold of the consistency test, to be investigated in future research.

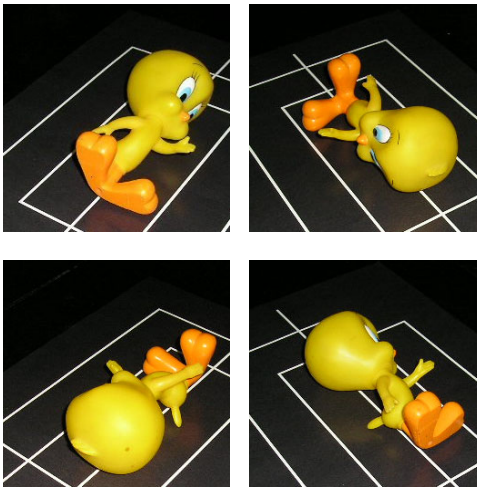


Figure 6. Magnified view of the input images

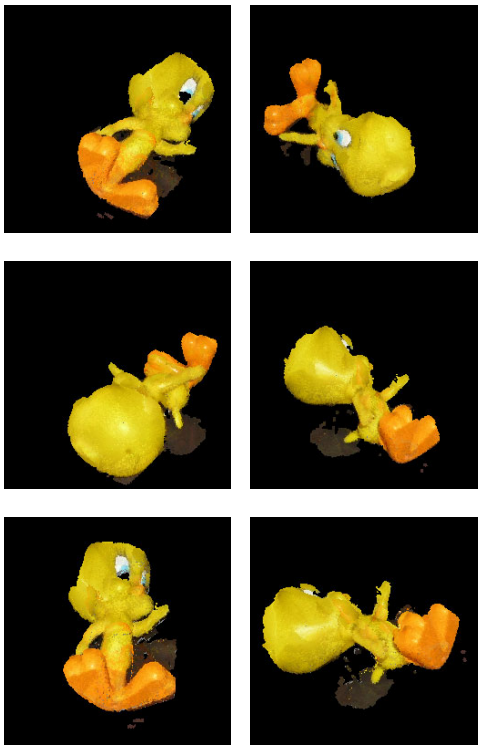


Figure 7. Reconstructed Object Images; the four on top are from the same views as the input images

7. Conclusion

This work has presented a practical 3D object-reconstruction method based on an adaptive space-carving algorithm. One of the main advantages of our approach is that it enables the user to reconstruct objects with a simple camera that is free to move, and a calibration pattern that is placed under the target object.

Easy automatic segmentation of the object in the scene is also possible using a strategy in which background images of different points of view are registered to a new input image so that an approximate background model can be computed.

Fast reconstruction is done by combining an adaptive carving strategy and a photo-consistency test that uses photometric and segmentation information in increasing levels of resolution. Memory issues are not a problem because the final solution is efficiently stored as an octree data structure that may be used as input data for other modeling procedures.

Topics that must be investigated in future works are the use of graphics hardware programming to accelerate the reconstruction process, a conservative strategy to remove large non-photo-consistent regions that do not project in the background regions of the images, and more sophisticated statistic measures that take into consideration sources of error other than camera noise.

8. Acknowledgements

This work was developed at Visgraf/Impa and was partially funded by CNPq.

References

[1] Rahul Bhotika, David J. Fleet and Kyriakos N. Kutulakos. A Probabilistic Theory of Occupancy and Emptiness. University of Rochester. Department of Computer Science. Technical Report 753. 2001

[2] A. Broadhurst and R. Cipolla. A statistical consistency check for the space carving algorithm. In M. Mirmehdi and B.Thomas, editors, Proc. 11th British Machine Vision Conference, volume I, pages 282-291, Bristol, September 2000.

[3] A. Broadhurst, T. Drummond and R. Cipolla. A probabilistic framework for the Space Carving Algorithm. In Proc. 8th International Conference of Computer Vision, pages 388-393, Vancouver, Canada, July 2001. IEEE Computer Society Press.

[4] W. B. Culbertson, T. Malzbender, and G. Slabaugh, Generalized Voxel Coloring, Proceedings of the ICCV Workshop, Vision Algorithms Theory and Practice, Springer-Verlag Lecture Notes in Computer Science 1883, September 1999, pp. 100-115.

[5] K. N. Kutulakos and S. M. Seitz, A Theory of Shape by Space Carving, International Journal of Computer Vision, Vol. 38, No. 3, July 2000, pp. 199-218.

[6] A. A. Montenegro, Reconstrução de Cenas a partir de Imagens através de Escultura do Espaço por Refinamento Adaptativo. Phd Thesis. Pontifícia Universidade Católica do Rio de Janeiro, September 2003.

[7] A. Prock and C. Dyer, Towards Real-Time Voxel Coloring, Proceedings of the DARPA Image Understanding Workshop, 1998, pp. 315-321.

[8] M. Sainz , N. Bagherzadeh and A. Susin, Hardware Accelerated Voxel Carving, Proceedings of the 1st Ibero-American Symposium on Computer Graphics, July 1-5, 2002, pp. 289-297, Guimarães, Portugal.

[9] S. Seitz and C. Dyer, Photorealistic Scene Reconstruction by Voxel Coloring, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1997, pp. 1067-1073.

[10] F. Szenberg, P. C. P. Carvalho, M. Gattass, Automatic Camera Calibration for Image Sequences of a Football Match, ICAPR 2001, March 11-14, 2001, p. 301.

[11] R. Tsai, An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, Proceedings of IEEE Conference on computer Vision and Pattern Recognition, Miami Beach, FL, 1986, pp 364-374.

[12] http://www.opengl.org