PHOTOREALISTIC 3D VOLUMETRIC MODEL RECONSTRUCTION BY VOXEL COLORING

Dianyong Zhang *, Zhenjiang Miao

Institute of Information Science, Beijing Jiaotong University, Beijing 100044, P. R. China
(07112064, zjmiao)@bjtu.edu.cn

Commission Ⅲ - WG Ⅲ/1

KEY WORDS: 3D reconstruction, Volumetric model ,Voxel coloring , Bounding box, Shape from photoconsistency

ABSTRACT:

Building photorealistic 3D models of object or scene from images is still a very difficult task. The built 3D models can have different application in many areas, such as industrial, virtual reality, and movies, among others. This paper shows end to end reconstruction 3d voxels model from images using voxel coloring algorithm. We define a reconstruction area by computing the minimum 3D bounding box of the object, because the size of the bounding box will affects the resolution of the reconstructed model. And background removal on the input images for voxel coloring is a necessary step, we implement the image cutout by graph cut algorithm, and have a good result. We test our method using two group of images, although we just use eight images, our reconstruction results are very good and it can match with the real object and scene in some extent.

1. INTRODUCTION

In the last decade, with the development of computer hardware and software, there have been significant computer game and animation development in computer graphics. And a realistic 3D model is playing an important role in this area. There is a growing need for 3D model reconstruction in these fields. Object reconstruction is an old and challenging problem — generating a 3D model of a object given a set of 2D images of the object.

There are two classes of object reconstruction: active methods — 3D model acquisition can be performed by laser scanners or structured light; passive methods—reconstruct 3d models by image sequences from a single camera or multiple cameras. Object scanning often demand expensive equipment and special skill to operate. As a simple and a low cost approach, 3D object reconstruction based images became more and more popular among the researchers.

At first, volumetric object reconstruction from images are based on volume intersection methods. These methods are often called as shape from silhouette algorithms[Laurentini, 1994]. The intersection of silhouette cones from multiple images defines an approximate geometry of the scene usually called the visual hull. Then, one form of reconstruction from images is voxel coloring, which began with Seitz and Dyer in [Seitz, 1999]. Voxel coloring is not like shape reconstruction approach, which is a color reconstruction problem. Rather than shape reconstruction using triangulation, voxel coloring approach traverses the scene volume voxel by voxel.

In this paper we consider the problem of carving photorealistic volume model of the real scene from a set of images. Our method similar to Rob Hess[Hess, 2006], but the definition of reconstruction area is different. We use the minimum 3D bounding box to compute the min reconstruction area. So we can save the computational time because the object locates only in the defined area. This method claims to be a simple object reconstruction method since it overcomes the feature-base and stereo methods. We test our method and get the good reconstruction result.

The rest of the paper is organized as follows. First, we describe the related papers and the systems in the following section. In section 3, we represent a algorithm of background removal on the input images for voxel coloring. Section 4 describes how to compute the bounding box of the object and how it is used in voxel coloring. Next, we will discuss our implementation of voxel coloring algorithm in detail. In Section 6, we will show our results of running this algorithm on the test datas. Finally, Section 7 conclude this paper and describe our future work.

2. RELATED WORKS

Volumetric object modelling approaches can be divided into following two classes: shape from silhouette and shape from photoconsistency.

Shape from silhouette methods using a set of binary silhouette images to get the reconstruction of volumetric scene[Laurentini, 1994]. The volume only approximates the true 3D shape since concave patches are not observable in any silhouette. Matusik et. al[Matusik, 2000] present a technique to perform the intersection of silhouette cones in image space and derive a geometry model for the visual hull. Visual hull method does not use photometric properties in each image and object concavities are not reconstructed. Furukawa et.al[Furukawa, 2009] proposed a method for acquiring high quality geometric models of complex 3D shapes by enforcing the photometric and geometric consistencies associated with multiple calibrated images of the same solid, they define the rims on the visual hull surfaces, then refine the result using energy minimization. But it is difficult to find the rims.

* Corresponding author.

Shape from photoconsistency methods use the color consistency constraint to mark object surface voxels from other voxels. Voxel coloring is an attempt to use the color information of the images to get a reconstruction. Because of the color information available in the images, voxel coloring methods can produce a better reconstruction to fit to the true scene than visual hull methods. In contrast to stereo algorithm which directly represent correspondences as disparities, image correspondences are represented by the projections of voxels to images.

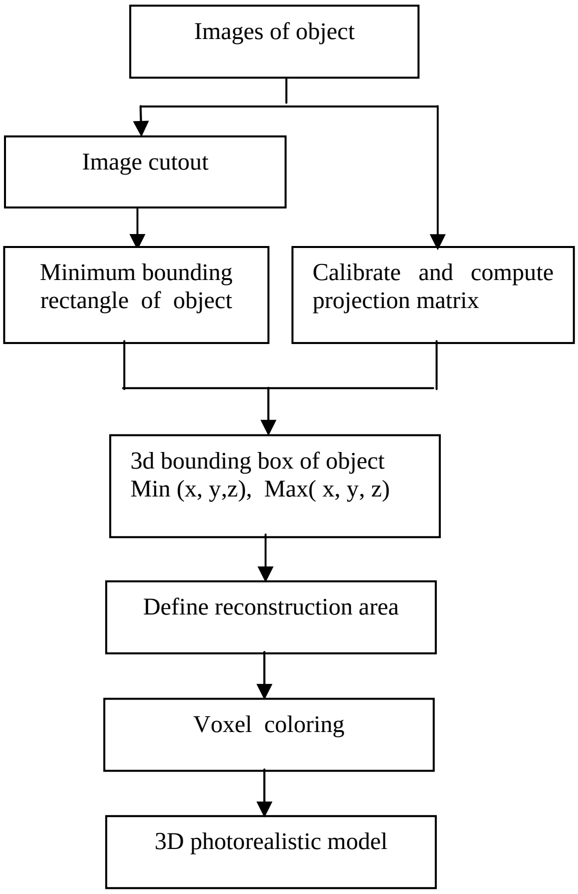


Figure 1. Framework of our system

Seitz and Dyer[[Seitz, 1999] proposed the voxel coloring algorithm, which decomposed the object or the scene to be reconstructed into small elements of volume called voxels. Color consistency can be used to distinguish surface points from other points in a scene. Voxel coloring algorithm restricts the position of cameras in order to simplify visibility information(Figure 2).

Space carving theory[kutulakos, 1998] removes the constraint of camera location. However, space carving algorithm does not use the full visibility of the scene. W.Bruce Culbertson et. al[Culbertson, 1999] present a new algorithm the extends Seitz and Dyer's voxel coloring algorithm, they called Generalized Voxel Coloring(GVC), unlike voxel coloring , it can use images from arbitrary camera locations. On the other hand it uses all images to check voxel consistency unlike space carving, which uses only a subset of images.

As the framework shown in Figure 1. We will refer to voxel coloring in this work and present some high accurate results with real object by shape from photo consistency methods..

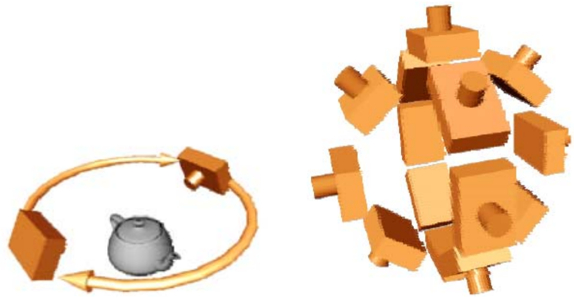


Figure 2. Camera configurations that satisfy the ordinal visibility constraint. Picture taken form [Seitz, 1999]

3. IMAGE CUTOUT

In our experiment, we use the images provided by Yasutaka Furukawa[Furukawa,2006]. For calibration, we use "Camera Calibration Toolbox for Matlab"[Bouguet,2008] by Jean-Yves Bouguet to estimate both the intrinsic(including focal length, principal point, skew and distortions coefficients) and extrinsic (translation and rotation) parameters.

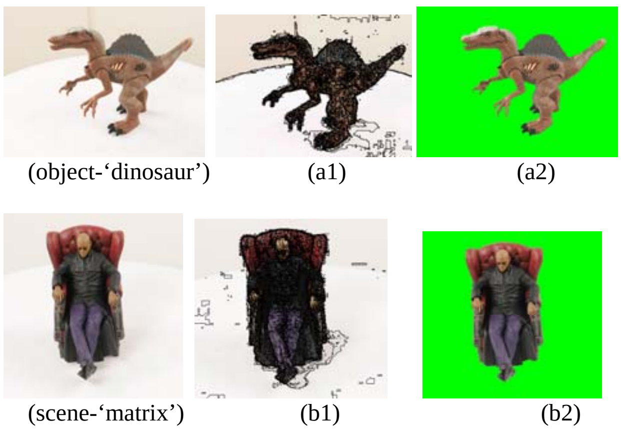


Figure 3. Background subtract of images and result.

In our paper , at first, the images are pre-segmented in regions by a propagation algorithm, then we define the foreground and background region manaually and cutout the images by a graph cut algorithm.

The similarity of two regions p_1 and p_2 can be judged by : $\|V_{p1} - V_{p2}\| < T$, where V_p is the mean color of a region in RGB color space, and T is the threshold. We set threshold = 10, the result of pre-segment show as Figure 3- a1 and b1.

Then a region based graph cut algorithm is used to segment all the pre-segmented regions into two parts, foreground and background area[Zhen.T, 2008]. The result of image segmentation show as Figure 3-a2 and b2.

4. 3D BOUNDING BOX

4.1 Minimum bounding rectangle

According the result of image segmentation, we can get the silhouette information of each image easily. Model show as white color and background show as black, show as Figure 4.

Figure 4. Binary image of object silhouette

we input the silhouette image S , based pixel row ,search point which the value of pixel are 255 form up to down, as TopMost(u,v) is the top edge point, based row , search point which the value of pixel are 255 from down to up, as BotMost(u,v) is the bottom edge point, then based column ,from left to right to get the point LeftMost(u,v)and from right to left to get the point RightMost(u,v), through the coordinate of these point, we can get the minimum bounding rectangle of the model, result as Figure 5:

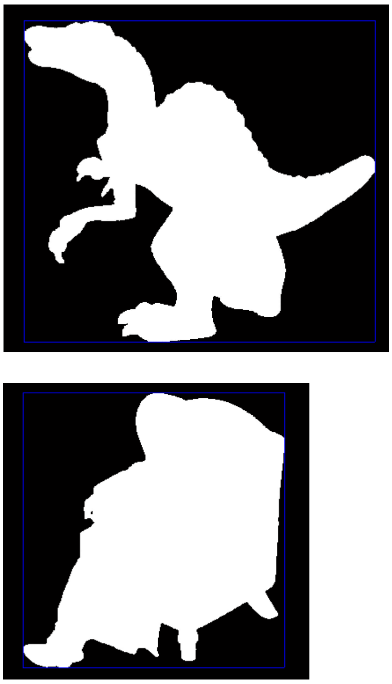


Figure 5.The blue line is the minimum bounding rectangle of the model

4.2 3D bounding box

AABB(axis aligned bounding boxes) is a bounding box aligned with the axes of the coordinate system. AABB of an object can easily be constructed from maximum and minimum global coordinates of the object along the X, Y, and Z axes. This box fully encloses the object(see Figure 6).

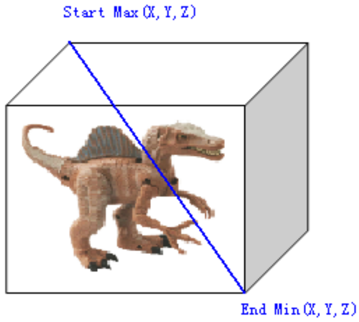


Figure 6. The sketch map of 3D bounding box

$$R=\left\{(x,y,z)\left|\begin{array}{l}\min_x \le x \le \max_x,\\ \min_y \le y \le \max_y,\\ \min_z \le z \le \max_z\end{array}\right.\right\}$$

Where $\min_x$, $\max_x$, $\min_y$, $\max_y$, $\min_z$, $\max_z$ are the minimum and maximum coordinate projected to X, Y, Z axes.

According the minimum bounding rectangle, we can get the up, down, left, right edge points, then use the projection matrix of each image, compute the points on the 2D image plane corresponding 3d points, combine with the linear restriction of the dimensional plane, we can compute the minimum 3D bounding box of the model.

3D cone of each the minimum bounding rectangle of the object can be considered as four 3D planes (Figure 7). So, 2d bounding line of n images can get the bounding box formed by 4n 3D planes.

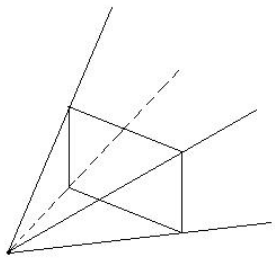


Figure 7. Four 3D plane composing minimum bounding rectangle of each image

Each 3D plane $P^i=(p_x^i, p_y^i, p_z^i, p_w^i)$ conclude a linear restriction: $P^i X \le 0$, computing six parameters of AABB bounding box , actually, set every parameter as object function of this group of linear restriction. Then compute the extremum of the function by optimization method. If we want to get the max and min coordinates, we define the next linear restriction:

$$p^i{}_x \cdot x + p^i{}_y \cdot y + p^i{}_z \cdot z + p^i{}_w \le 0, \quad i = \{1, 2, ..., 4n\}$$

According the minimum bounding rectangle of each object and projection matrix P, computing the 3D minimum bounding box of the real scene by linear restriction .

5. VOXEL COLORING

After we computed the bounding box of the object to be reconstruction, we divided the scene into n voxel layers. Then we start sweep the layer closest to the camera volume.

When the reconstruction has been initialized, each voxel is then projected onto the image plane of each input image from the image set and compute the color consistency of the set of pixels onto which each projects. Figure 8.

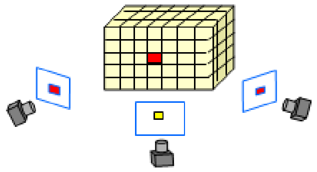


Figure 8. Carve away voxels that are not photo-consistent with images.

5.1 Voxel coloring problem

We will now define the voxel coloring problem. Given an image pixel p and scene S. We suppose each voxel v in S,which is visible in I, voxel projects to image $p \in I$. Scene S is complete with respect to a set of images if there exists a voxel $v \in S$ such that $v = S(p)$, for every image I and every pixel $p \in I$. A complete scene is said to be consistent with a set of images if, for every image I and every pixel $p \in I$,

$$color(p,I) = color(S(p),S)$$

Let N denote the set of all consistent scenes.We have the following formal definition of the voxel coloring algorithm[Seitz, 1999]: given a set of basis images $I_0, \cdots, I_n$ of a static Lambertian scene and a voxel space I,determine a subset $S \subset v$ and a coloring $color(v,S)$, such that $S \in N$. The algorithm pseudo-code is shown :

```
For  each voxel layers 1 to r
 {
    For each voxel V in each layer
        {
       For image I = 1 .. numberOfImages do
         {
          Compute footprint P of V
          }
       Compute voxel's consistency S
      If ( S < threshold )
        {
         Add V to voxel set
        }
    }
  }
```

5.2 Corners definition

Similar to [Hess,2006], a voxel is a cube, projects onto some sort of hexagon in the image plane. We define the voxel corners as follows ,

Voxel corners are projected onto the image plane. All eight corners of the voxels and their coordinates are illustrated in Figure 9.

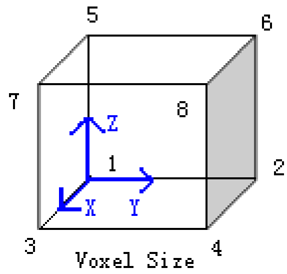


Figure 9. Voxel's corners for projection.

Each corners with their corresponding coordinates can be show as:

1: [X, Y, Z]
2: [X, Y + vsize, Z]
3: [X + vsize, Y, Z]
4: [X + vsize, Y + vsize, Z]
5: [X, Y, Z + vsize]
6: [X, Y + vsize, Z + vsize]
7: [X + vsize, Y, Z + vsize]
8: [X + vsize, Y + vsize, Z + vsize]

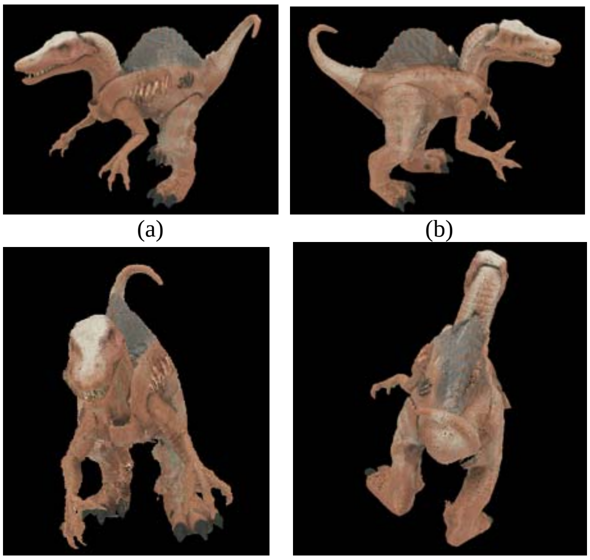
We set the min and max X and Y reprojection coordinates to create a rectangular footprint. When we get the voxel's footprint from projection, the unmarked pixels within it are found, R,G,B values are computed. We can check whether it is belong to the background voxel.

If the mean RGB value is within some user-supplied background threshold value , the voxel is considered to be a background voxel and it will not be colored. Otherwise, it will continue the voxel consistency check. Voxel consistency can be defined as the standard deviation in the color of pixels in all images, we compute each channel's standard deviation, and average the standard deviations for RGB channels and compared to a user-supplied threshold. If the mean standard deviation of the RGB channels is less than the threshold value, the voxel is considered consistent and it will be colored by the mean color of its footprints. Then, all pixels inside the voxel's footprint will be marked.

6. IMPLEMENTATION AND RESULTS

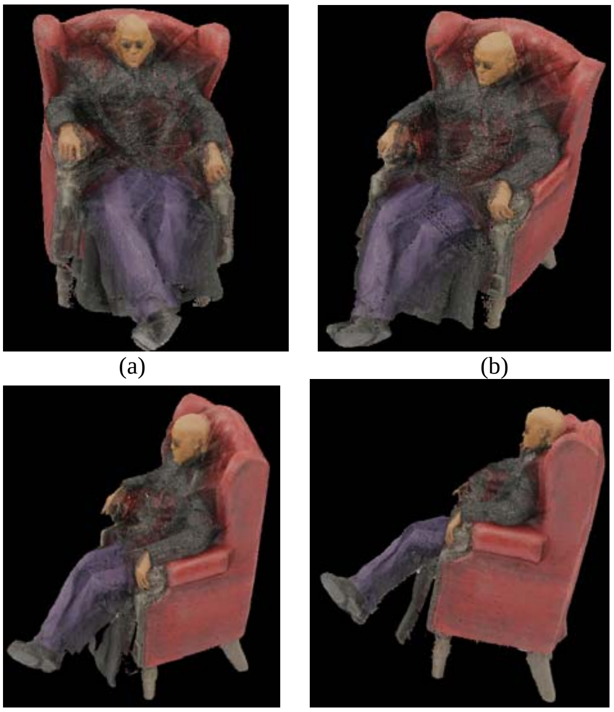
We tested our program using two set of images(an object dinosaur, a scene matrix), which are provided by Yasutaka Furukawa[Furukawa, 2006]. We just used eight images to implement the reconstruction. Results are stored in .vxl format in a file. The result were obtained by running the system on a Duo CPU 3.00GHz, the computer had 2GB of the RAM. We implement the algorithm by Matlab program, the results of running time with various voxel sizes are summarized in table1 and table2.

We input the voxel size while performing reconstruction. For example, using a voxel size of 1, the scene is broken down into voxel with sides of one unit of length. Show as Figure 10, Figure11.



(a) (b) (c) (d)

Figure 10. The reconstruction result of dinosaur from four angles.
Start Max(519.407; 790.185; 418.1958),
End Min(215.135; 212.992 ; -66.6792),
Vsize=1, Color threshold=10, Background threshold=10.



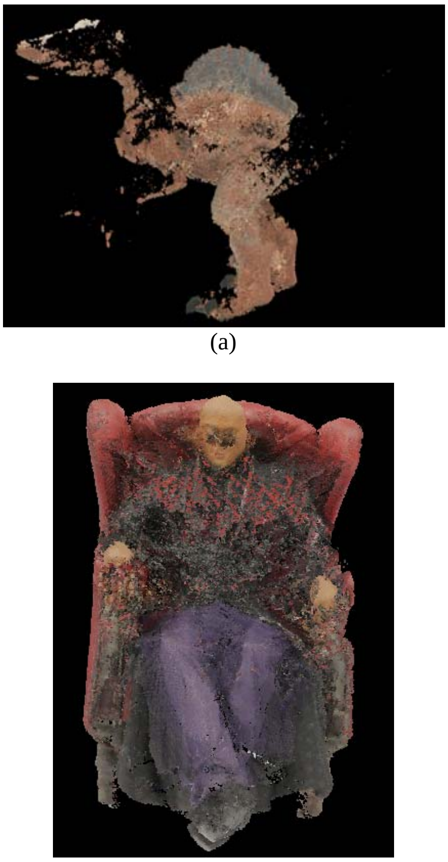
(a) (b) (c) (d)

Figure 11. The reconstruction result of matrix from four angles.
Start Max(555.475; 521.739; 358.1325),
End Min(176.925; 144.7 ; -80.0975),
Vsize=1, Color threshold=10, Backgroundthreshold =10.

Color threshold is the user input value for color variance, which controls the maximum amount of variance allowable for a voxel to be considered color invariant.

For all the models, the threshold is the higher, the reconstruction is more complete. However, the reconstruction will contain erroneous voxels(Figure.12) and it is less accurate. On the other hand, the threshold is the lower, the reconstruction result is more accuate, but it is less complete and parts of the model will be missing(Figure.13). The threshold, corresponds to the maximum allowable correlation error.

Figure 12. Color threshold = 25,part of the green color is error voxels.



(a)

(b)

Figure 13. Color threshold = 5,figure a and b.

Background threshold is also a uer-specified threshold to determine the maximum distance a voxel can be from the background color before it is considered a background voxel, different value can get the different result(Figure 14).

Figure 14. Background threshold = 40.

We have test three voxel size levels : 10 , 5 , 1. Its number of voxel colored and running time show as table1 and table2, and the reconstruction result is shown in Figure.15 and Figure.16.

Voxel size (in mm)	Voxels colored	Running time (in second)
1.0	419314	96265.23589
5.0	565	237.075930

Table 1. The running parameters of dinosaur with different voxel size

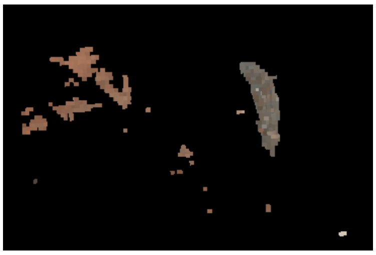


Figure 15. The dinosaur, Vsize =5

Voxel size (in mm)	Voxels colored	Running time (in second)
1.0	448298	106862.361813
5.0	8577	1154.780023
10.0	391	247.940010

Table2. The running parameters of scene matrix with different voxel size

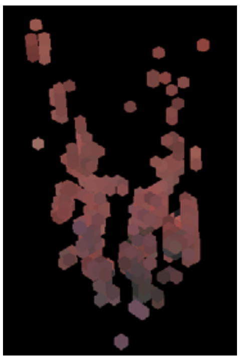


Figure 16. The scene matrix, Vsize = 10

7. CONCLUSION AND FUTURE WORK

In this paper, we have implemented a voxel coloring algorithm presented by Seitz and Dyer[Seitz, 1999] for photorealistic object or scene reconstruction. And the system can create realistic models without requiring expensive hardware and it can eliminate the drawbacks of the silhouette based reconstruction.

We test our method using images[Furukawa, 2006] an object dinosaur model and a scene matrix, similar to Rob Hess[Hess, 2006], but we compute the minimum bounding rectangle of the reconstruction area, according the calibration parameters, using the projective matrix, define the minimum 3D bounding box of the scene. Although we just use eight images, our reconstruction result is very good and it can match with the real object and scene in some extent.

The obvious open problem with this method is the handling of non-lambertian surfaces. Although the implementation of the algorithm was successful, the running time was too slower for the extra time spent calculating the footprint. Threshold selection is not easy in both the dark and bright parts of a scene when RGB color space is used. Future research will be focused on improvement of the quality of the model and the speed of the reconstruction process.

ACKNOWLEDGEMENTS

This work is supported by China National Basic Research Program (973) 2006CB303105, China National High Tech R&D program (863) 2009AA01Z334 and National Natural Science Foundation Of China 60973061.

REFERENCES

Bouguet.J.Y,2008.Online-available: http://www.vision.caltech.edu/bouguetj/calib_doc/.

Culbertson, W. B. Malzbender, T. Slabaugh, G. 1999. Generalized Voxel Coloring. Proc. Of the vision Algorithm, Workshop of ICCV, pp. 67-74.

Furukawa.Y,2006.http://www.cs.washington.edu/homes/furukawa/research/mview/index.html

Furukawa,Y.and Jean Ponce. 2009. Carved Visual Hulls for Image-based Modeling. International Journal of Computer Vision, 81(1):53-67.

Hess.R,2006, http://web.engr.oregonstate.edu/~hess/.

Kutulakos, N. and Seitz, S.M.,1998. A Theory of Shape by Space Carving. University of Rochester CS Technical Report 692.

Laurentini, A. 1994. The visual hull concept for silhouette-based image understanding. IEEE trans. Pattern Analysis and Machine Intelligence, 16(2): 150-162.

Matusik,W., Buehler,C., Raskar,R., Gortler,S., and Mcmillan,L , 2000. Image based Visual Hulls. In computer Graphics, SIGGRAPH 2000 proceedings, 369-374.

Seitz,S.M., and Charles R. Dyer. 1999. Photorealistic Scene Reconstruction by Voxel Coloring. International Journal of Computer Vision, vol. 35, no.2, pp. 151-173.

Zhen Tang, Zhenjiang Miao, Interactive Foreground Extraction for Photo and Video Editing, The 9th International Conference on Signal Processing, Proceedings, v 2,Beijing, 2008, p1483-1486.