7

Photorealistic 3D Model Reconstruction based on the Consistency of Object Surface Reflectance Measured by the Voxel Mask

K. K. Chiang, K. L. Chan and H. Y. Ip
City University of Hong Kong
China

1. Introduction

To create three-dimensional (3D) models of real scenes and objects is an old and challenging computer vision problem. Systems that can reconstruct the 3D model of object, for instances the human head or cultural artefacts, have found many applications such as virtual character animation and interactive museum exhibition. Real object models can be reconstructed automatically using active and passive methods. Object range scanning by laser or structured light are typical examples of the active methods. They often demand expensive equipment and special skill to operate. Moreover, they are not very good in modeling very glossy objects. The passive methods can acquire images of the object at different viewpoints using off-the-shelf CCD cameras (Chang & Chen, 2002). The camera is usually calibrated by taking pictures of a specially designed calibration pattern or object. The camera viewpoints can be arbitrarily selected and the camera model is adjustable. For instance, Niem (Niem, 1999) proposes a 3D object reconstruction method using a mobile camera to capture image of the object and calibration pattern simultaneously.

Our system of 3D object model reconstruction consists of four major steps: camera calibration, volumetric model reconstruction, polygonal model formation and texture mapping. An overview of our system is shown in Figure 1. The camera calibration is to obtain the intrinsic and extrinsic parameters defining the internal camera properties and the viewpoint orientation with respect to the object. The object and the calibration patterns can be captured simultaneously. Therefore, the camera can be placed anywhere and each view can be calibrated independently. One of the popular approaches for volumetric modeling is shape from silhouette (SFS), which is to recover the shape of object from its contours. However, reconstruction of a complex rigid object from its images is a challenging computer vision problem, especially when the object exhibits large textureless surface or concave surface. Previously, we enhance the SFS-based volumetric modeling algorithm by imposing the photo-consistency in neighboring views and the aggregation of evidence in volume space via the use of voxel mask (Wong & Chan, 2004; Chiang & Chan, 2006). Although the algorithm is very good in tackling textureless as well as concave surface, it is still unable to model non-Lambertian object surface accurately. In the present investigation, we propose a novel volumetric modeling algorithm that further improves the shape reconstruction by explicitly taking into account the object surface specularity. Then the polygonal model is

Open Access Database www.i-techonline.com

Source: Scene Reconstruction, Pose Estimation and Tracking, Book edited by: Rustam Stolkin,
ISBN 978-3-902613-06-6, pp.530, I-Tech, Vienna, Austria, June 2007

formed by the marching cubes algorithm (Lorensen & Cline, 1987). Various operations are applied to refine the polygonal model. Finally, a texture map is created from the original image sequence and mapped onto the polygonal model to give it a realistic appearance. Computer graphics techniques are adopted for the synthesis of missing or unseen texture.

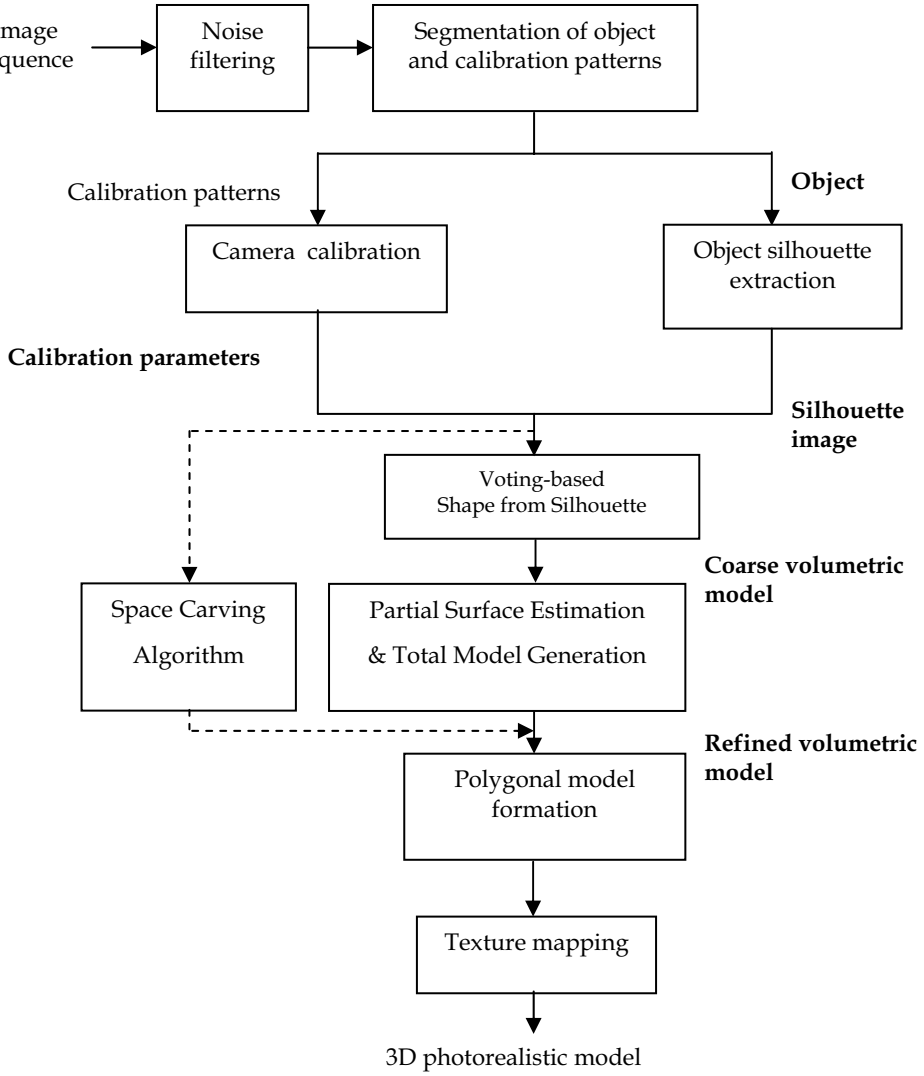


Figure 1. Overview of 3D model reconstruction system

2. Volumetric model reconstruction

Volumetric modeling is an important part of 3D model reconstruction. Conceptually, it is based on computations in 3D volume space in order to construct the object volume in the

world coordinate system that is consistent with the input images. Image-based volumetric modeling methods generally assume that the object or scene is Lambertian. Ideally, only matte surface exhibiting purely diffuse reflection satisfies Lambert's Law. Many real objects are moderately glossy, e.g. plastic and ceramic objects. They exhibit both diffuse and specular reflections. The model reconstructed by conventional volumetric modeling algorithms contains many errors due to violation of this assumption. Therefore a lot of post-processing effort is needed before the 3D model can be used in practical applications.
In our system, the topology of the object is first obtained by a voting-based SFS technique. Detail object shape is reconstructed in two steps: partial surface estimation and total model generation. In our first attempt, partial surface is estimated by exploiting color-consistency and the aggregation of evidence in volume space. A novel 3D voxel mask is used for measuring the color dispersion, instead of using the conventional image block matching technique. To tackle real glossy objects, we further enhance the 3D model reconstruction system by proposing a novel partial surface estimation algorithm that can handle the co-existence of both diffuse and specular surfaces. Based on the dichromatic reflection model, we derive an explicit relation characterizing the specular reflection. Voxel mask is employed for measuring either photo-consistency or surface specularity during the volumetric modeling.

2.1 Shape from silhouette/photo-consistency

The volumetric modeling assumes that there is a bounded volume (V) within which contains the object of interest. This volume is often assumed to be a cube and the most common approach to representing it is a regular tessellation of cubes called voxels (v). SFS reconstructs the volumetric model using a sequence of images. Each viewpoint location can be estimated from the camera calibration process. Our first volumetric model reconstruction method, called shape from silhouette/photo-consistency (SFSPC), uses the voting-based SFS firstly to reproject the object silhouettes onto the 3D voxel space to obtain the topology of object. Partial surface estimation and total model generation are then used to refine the volumetric model which will be explained later.
In our system, the contour of the object is simply extracted from each of the input images with the use of a monochromatic background during image acquisition. The reason why we set up the system in front of a monochromatic background is to save the computation time. Object silhouette can be extracted easily and subsequent computation can be confined to the object region. It should be mentioned that a relaxation of the acquisition environment is possible at the cost of longer computation time. That will not affect the reconstructed model as the background voxels can still be carved away by the partial surface estimation algorithm, no matter it exploits photo-consistency or surface specularity.
Each silhouette image is represented in binary form. SFS recovers the volumetric description of the object from multiple silhouette images by volume intersection. Firstly, a bounding cone is constructed, using the camera focal point and the corresponding object silhouette. If a point of the voxel is back-projected onto the inside the object silhouette, that voxel may be occupied by the object. Otherwise, the voxel is outside the object. To make the estimation of the object in 3D space more accurate, multiple silhouette images taken at various viewpoints can be used. Then, a visual hull (VH) is constructed by intersecting all the bounding cones formed by the object silhouettes (Laurentini, 1994).

Many SFS-based methods assume that the silhouette images are errorless. However, in real situation, silhouette images may contain errors which can affect the accuracy of the reconstructed model. Therefore, an extension of conventional SFS, such as the voting-based SFS, is implemented such that silhouette image error does not seriously affect the reconstructed object shape. First, the volume space V is set up and the score of each voxel is initialized to zero. Based on the voting-localizing scheme, a score of the voxel is incremented each time if the voxel is projected onto the silhouette image region, that is the projected color is not the background color. An accumulated value is obtained until all images are visited. The voxel is considered as visible only and is included in the visual hull VH if its accumulated score is larger than a pre-defined threshold as shown in the following pseudo-code.

```
For each voxel v in V
   v.score = 0
End For
VH = ϕ
For each voxel v in V
   For each silhouette image i, where i = 1…NumOfImages
      Color = ProjectVoxelOnImage(v, i)
      If Color != BackgroundColor
      Then v.score = v.score + 1
      End If
   End For
End For
For each voxel v in V
   If v.score ≥ Threshold
   Then VH = VH ∪ v
   End If
End For
```

The voting-based SFS can be treated as a generalization of SFS, by which setting the threshold as the total number of views becomes the conventional SFS. The advantage of the voting-based SFS over the conventional SFS is to minimize the adverse effect of the artefact in silhouette images to the resultant model.

It is well known that SFS is not guaranteed to reconstruct a correct 3D shape of the target object, especially when the object has concave surface. To cope with this problem, it is necessary to incorporate other technique to detect and recover the shape concavity. Shape from photo-consistency is one of the choices. For a greyscale or color image, the photometric

information does give improvement in volumetric reconstruction. If a voxel corresponds to the surface of an object, the texture and color properties of its projection points on the images, from which it can be seen, must be nearly the same. There is an assumption made in any color-consistency method: surface of the object is assumed to satisfy the Lambertian reflectance model, so that every surface appears equally bright in all directions regardless of the illumination. The shape from photo-consistency algorithm consists of two steps: (i) partial surface estimation, and (ii) total model generation. The basic idea of partial surface estimation is that if a voxel corresponds to the surface of the object, the color of its projection point on each image is similar to each other. Due to lighting fluctuation and image noise, the ideal color-consistency (or zero color dispersion) is not possibly achieved using the whole set of images. However, it is much easier to estimate a fairly good partial surface from a small set of neighboring images by relaxing the color dispersion to a small value. As there are as many partial surfaces as the images and each partial surface may still have error, a merging step that can integrate the partial surfaces and at the same time reduce error is needed.

For the estimation of partial surface for each viewpoint, the color dispersion for each voxel must be calculated first. Color dispersion is to measure the difference of color values for a voxel projected to a number of consecutive images. The visible voxels with minimum color dispersion are chosen and regarded as the partial surface for that viewpoint. Equation (1) is used to calculate the color dispersion:

$$D(v,i) = \sum_{j=i-m}^{i+m} \sum_{block\ size} \{C(p(v,i)) - C(p(v,j))\}^2 \tag{1}$$

where the number of consecutive images is $2m+1$, *block size* is the block area in the image for pixel-by-pixel comparison, and $C(p(v, i))$ is the color value of the projection point of voxel v on image i. The partial surface for the i^{th} viewpoint is referred as V_{sp}^{i}. Then, for each viewpoint i, v in the visual hull VH is projected onto the consecutive images from view i-m to i+m for calculating the color dispersion. After the color dispersion calculation is finished, a set of voxels V_r^i along a ray from the focal point of i through v is obtained. Assign v to V_{sp}^{i} if its color dispersion is the minimum among the voxels in V_r^i. These operations are performed for all voxels and repeatedly for all viewpoints. To ensure that the voxels of each partial surface are photo-consistent, only a small number of consecutive viewpoints in the neighborhood of the partial surface viewpoint are examined instead of examining the whole sequence of camera views. If the voxel is occluded in one or more examining viewpoints, the color dispersion is high and that voxel will very unlikely be selected as partial surface.

To generate the complete 3D volumetric model, the partial surfaces have to be integrated. However, error may still occur in each partial surface. For instance, variation of the color value of a surface point between two captured images due to environmental lighting can lead to false detection of the surface voxel. Therefore, a relaxation is allowed in integrating the partial surfaces by the voting-localizing scheme such that the error is rounded to construct a high precision model. The total model generation algorithm is as follow.

1. Set the score of all voxels in VH as zero.

2. For each voxel v, increment the score using the following scheme:
    a. Calculate the distance between the focal point of i and v, $L(v, i)$, and the distance between the focal point of i and voxel v', $L(v', i)$, where v' is the intersection of a set of voxels along a ray (from the focal point of i to v) and the partial surface V_{sp}^{i}.
    b. Both distances are measured in the Euclidean space. If $L(v, i)$ is greater than or equal to $L(v', i)$, that means v is behind or lying on V_{sp}^{i}. Then increment the score of v by 1.
3. After looping all the viewpoints, if the total score is greater than a given threshold, v is appended to the final volumetric model.
4. The process is repeated for all voxels.

2.2 3D voxel mask

The common practice of conventional shape from photo-consistency, or methods based on the color-consistency, is to project 'one' voxel into consecutive image views and threshold the variance of the color values. This makes the algorithm heavily rely on the image quality. Even in the multi-hypothesis testing (Eisert et al. 1999), only one voxel is used to measure the photo-consistency which is insufficient. On the other hand, the use of image-based pixel-by-pixel comparison in a block area for measuring the color dispersion cannot guarantee a correct volumetric model. Therefore, the formulation of color dispersion should be re-defined. Instead of using image-based comparison and multi-hypothesis testing of one-voxel projection, a novel 3D voxel mask for evaluating the photo-consistency is proposed. Voxels in the mask are used to measure the photo-consistency among the consecutive images.

The idea of voxel mask is somewhat similar to the image filter kernel. The voxel mask is oriented in three orthogonal directions with respect to the camera viewpoint. The line joining the camera and the examining voxel - centre of voxel mask, is called the axial direction. The other two directions are called coronal and sagittal. Each voxel in the 3D voxel mask is denoted as v_{vm}. The structure of the voxel mask and the generation procedure are shown in Figures 2 and 3 respectively.

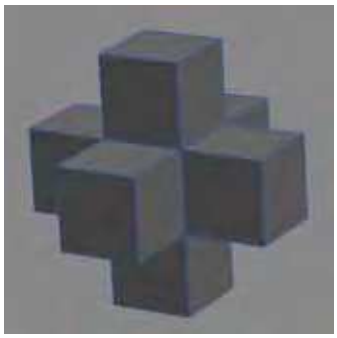


Figure 2. 3D voxel mask

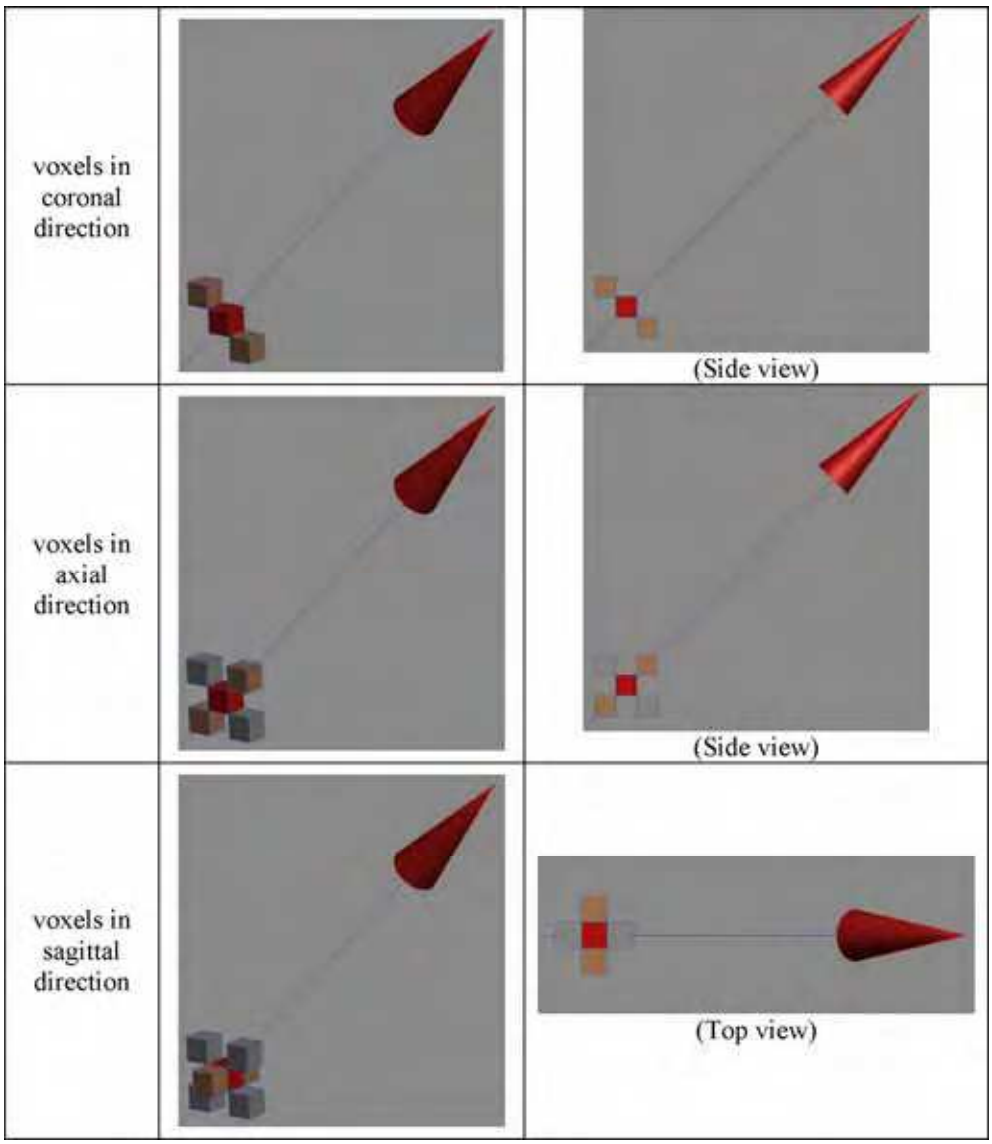


Figure 3. Generation of 3D voxel mask (cone: camera viewpoint, red voxel: centre of mask, brown voxels: voxels in a particular direction)

The mask for each examining voxel is generated independently. The structure of mask is symmetric and invariant in the voxel space. This ensures that the same mask structure is used to evaluate the color dispersion among the neighboring images. However, not all voxels in the mask will be used to evaluate the color dispersion. In each direction, if any component v_{vm} is not in the visual hull, all voxels in this direction are excluded from the color dispersion calculation. This mechanism can prevent the surface voxels from

erroneously removed from the volumetric model. As the number of voxels in the mask used to evaluate the color dispersion may be different, the accumulated color dispersion value is averaged by the actual number of comparisons. The new function for color dispersion evaluation is re-defined as:

$$D(v_{vm}, i) = \sum_{j=i-m}^{i+m} \{C(p(v_{vm}, i)) - C(p(v_{vm}, j))\}^2 \quad (2)$$

The modified partial surface estimation using 3D voxel mask in our SFSPC algorithm is shown in the following pseudo-code.

```
For each camera viewpoint i
    V_sp^i = ϕ
    For each voxel v in VH
        v.NoOfComparison = 0
        v.ColorDispersion = 0
        Form voxel mask for v
        For each voxel v_vm in voxel mask
            v.ColorDispersion += D(v_vm, i)
            v.NoOfComparison++
        End For
        v.ColorDispersion = v.ColorDispersion / v.NoOfComparison
    End For
    For each voxel v in VH
        If v.ColorDispersion = MIN_{V_r^i}(v.ColorDispersion)
        Then V_sp^i = V_sp^i ∪ v
        EndIf
    End For
End For
```

2.3 Shape from specularity consistency

Volumetric modeling methods generally assume that the object surface is Lambertian. Real objects and scenes may have specular surfaces. Today, the visual quality of models becomes the main point of attention due to the increasing demand for 3D models in various areas such as virtual reality and product design. The presence of specular reflection hinders the accurate reconstruction of the model but it is this phenomenon that gives the 3D model a true sense of realism. It is therefore important to have sophisticated volumetric modeling methods that can handle non-Lambertian surfaces.

The image pixel represents the reflection of light incident on a microfacet of the surface. The dichromatic reflection model (Shafer, 1985) describes the surface reflection as the sum of diffuse and specular reflections. The diffuse reflection is characterized by subsurface scattering and represents the shape of surface. The specular reflection occurs at the air/material interface and is only observed at some locations on the surface. We propose to model the non-Lambertian surface in our multi-view volumetric modeling framework.

Figure 4 shows the relationship between incident light, specular reflection, and camera viewpoints. The surface reflection R is modeled as $R = R_d + R_s$, where R_d and R_s are the diffuse and specular reflections respectively. R_d is proportional to $cos(\theta_i)$ where θ_i is the angle between the surface normal and the incident light direction. R_s can be decomposed into $R_{sc} + R_{sv}$. R_{sc} depends on surface material properties and θ_i. R_{sv} is modeled as a cosine function $R_{sv} = k \cdot cos(\theta_v - \theta_r)$ where θ_r is the angle between the surface normal and the specular reflection, θ_v is the angle between the surface normal and the viewpoint direction, and k depends on surface material properties. For a surface point illuminated by a fixed light source, R_d and R_{sc} are both constant and is represented as $R_c = R_d + R_{sc}$. Without using the polarizer, we cannot separate R_d and R_{sc}. But we do not need to eliminate R_{sc} as we have already stated that specularity gives the model a true sense of realism. We only need to exploit R_{sv} which varies with the change in viewpoint orientation, and devise volumetric modeling method that can accurately identify the specular surface voxels. Assume that R_i is the color of the surface voxel observed in viewpoint i. R_{i+1} is the color of the same surface voxel observed in adjacent viewpoint $i+1$. These two adjacent viewpoints are separated by an angle $\Delta\theta$. Therefore,

$$R_i = R_c + R_{svi} = R_c + k \cdot \cos(\theta_{vi} - \theta_r) = R_c + k \cdot \cos(\theta) , \quad (3)$$

$$R_{i+1} = R_c + k \cdot \cos(\theta - \Delta\theta) . \quad (4)$$

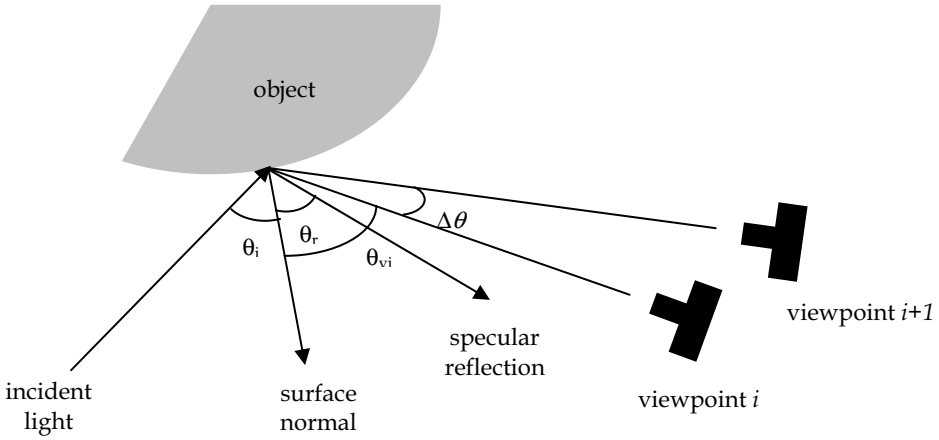


Figure 4. Relationship between incident light, specular reflection and camera viewpoints

If $\Delta\theta$ is small,

$$R_{i+1} = R_c + k\,[\cos(\theta) \cdot \cos(\Delta\theta) + \sin(\theta) \cdot \sin(\Delta\theta)] \cong R_c + k \cdot \cos(\theta) \cdot \cos(\Delta\theta) , \quad (5)$$

$$R_{i+1} - R_i = k \cdot \cos(\theta)[\cos(\Delta\theta) - 1] = R_{svi}[\cos(\Delta\theta) - 1] . \quad (6)$$

This relation still holds for positive or negative $\Delta\theta$ or θ. To evaluate the similarity in specularity among neighboring voxels, a planar voxel mask is used. A small size voxel mask is good enough to analyze the surface specularity in a local object surface. Voxels in the mask are used to measure the R_{sv} in consecutive images. The generation of a voxel mask for each examining voxel is based on the camera viewpoint. The mask is oriented in three

orthogonal directions with respect to the viewpoint as shown in Figure 5. The mask for each examining voxel is generated independently. The structure of mask is invariant in the voxel space. This ensures that the same mask structure is used to evaluate the specularity among the neighboring images. Again, the size of the voxel mask can be reduced when any voxel in the voxel mask (v_{vm}) is not in the visual hull. As the number of voxels in the mask used in the evaluation may be different, the accumulated surface specularity (v.SpecularityDispersion) is averaged by the actual number of comparisons involved. If the examining voxel is a surface voxel, all the specularities measured by the voxel mask are small and similar. The accumulated (normalized) surface specularity is minimum.

When there is sufficient texture information or the object surface exhibits inhomogeneity, the partial surface estimation can easily define the surface voxels. However, if the object contains a large, flat area in homogeneous color, accurate partial surface is difficult to obtain as the true color of that surface is almost the same in neighboring views. Considering the existing of image noise, the voxel with the color or specularity very similar to neighboring voxels may not be a real surface voxel. Based on our observation, the dispersion between the voxel near to the camera along the ray v_{near} and the minimum dispersion is checked. If the difference is smaller than the pre-defined threshold, v_{near} is assigned as the surface voxel. Otherwise, the examining voxel is the surface voxel. This measure ensures the reconstructed volumetric model is conservative rather than carves away true object voxels. The partial surface estimation, denoted as shape from specularity consistency (SFSC), is shown in the following pseudo-code.

```
For each camera viewpoint i
    V_sp^i = ϕ
    For each voxel v in VH
        v.TotalNumOfComparison = 0
        v.SpecularityDispersion = 0
        Form planar voxel mask for v
        For each voxel v_vm in voxel mask
            v.SpecularityDispersion += R_svi
            v.TotalNumOfComparison += 1
        End For
        v.SpecularityDispersion = v.SpecularityDispersion / v.TotalNumOfComparison
    End For
    For each voxel v in VH
        If v.SpecularityDispersion = MIN_{V_r^i} (v.SpecularityDispersion)
            If v_near.SpecularityDispersion – v.SpecularityDispersion ≤ Threshold
            Then V_sp^i = V_sp^i ∪ v_near
            Else V_sp^i = V_sp^i ∪ v
            End If
        End If
    End For
End For
```

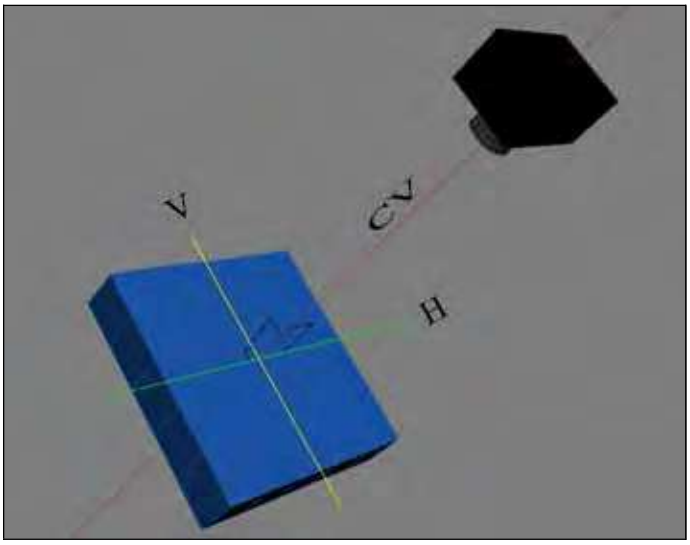


Figure 5. Planar voxel mask and its orientation – V: vertical axis, H: horizontal axis, CV: camera viewpoint axis

3. Result

Three objects (tiger, flower, square rubber) have been used to test our first volumetric model reconstruction algorithm SFSPC. The images are captured by a consumer-type digital camera at the resolution of 1024 × 768 pixel. The program is run on a 1.2GHz PC with Microsoft Windows 2000 and 512 MB RAM. 44 images are used to reconstruct the tiger model while models of flower and square rubber use 36 images. The resolution of the voxel space is 256 × 256 × 256 for all models. For the generation of partial surfaces, the number of consecutive views is set to be five (2 preceding views, the current view and 2 successive views). Figure 6 shows image views of the objects (top row), the corresponding partial surfaces (middle row), and partial surfaces at another view (bottom row). Figure 7 shows the volumetric models of the three objects. The voting-based SFS takes from 5 minutes to 12 minutes. The partial surface estimation plus total model generation takes from 32 minutes to 129 minutes.

The square rubber is used to test the robustness of the SFSPC algorithm. Comparison is made between SFSPC and shape from silhouette/stereo (SFS^2) (Matsumoto et al., 1999). We select SFS^2 as the reference because SFSPC follows the concept of SFS^2 while we enhance the shape from photo-consistency step by the use of 3D voxel mask. Also, SFS^2 has demonstrated a better shape reconstruction than various existing techniques. The square rubber is used as it exhibits convex, flat, as well as deeply concave surfaces. The number of consecutive images is set to be five and the threshold for total model generation is 0.85. Different views of the reconstructed model are shown in Figure 8. It can be seen that SFSPC always generates a better volumetric model than SFS^2. Figure 9 shows several rendered views of texture mapped models at arbitrary viewing positions.

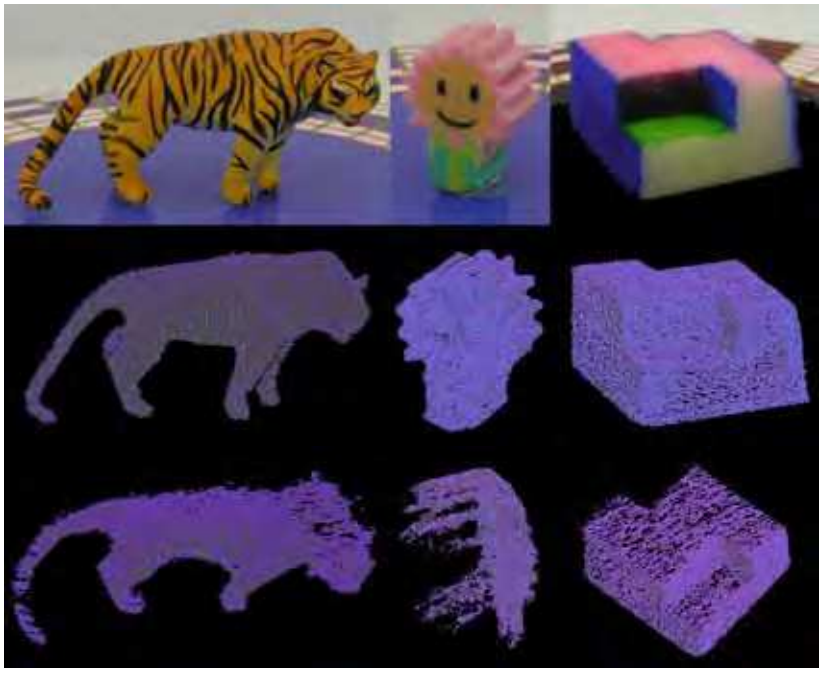


Figure 6. Image views and partial surfaces of three test objects reconstructed using SFSPC

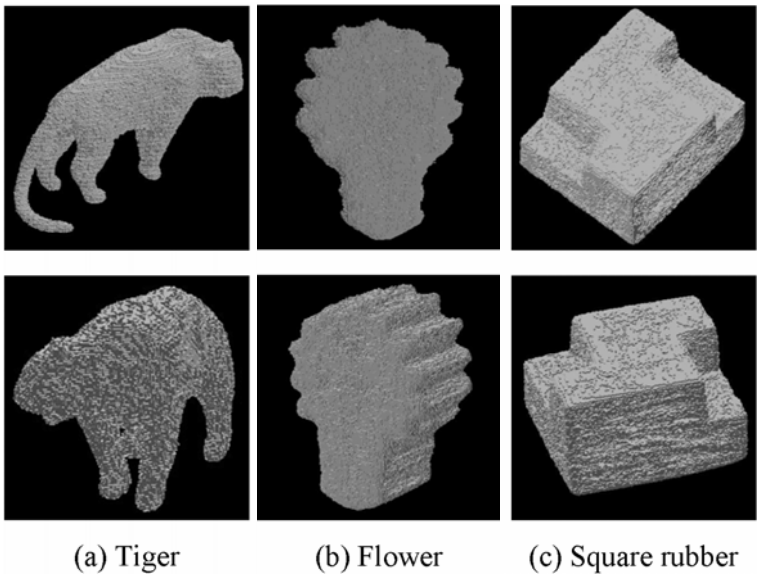


Figure 7. Different views of volumetric models of three test objects

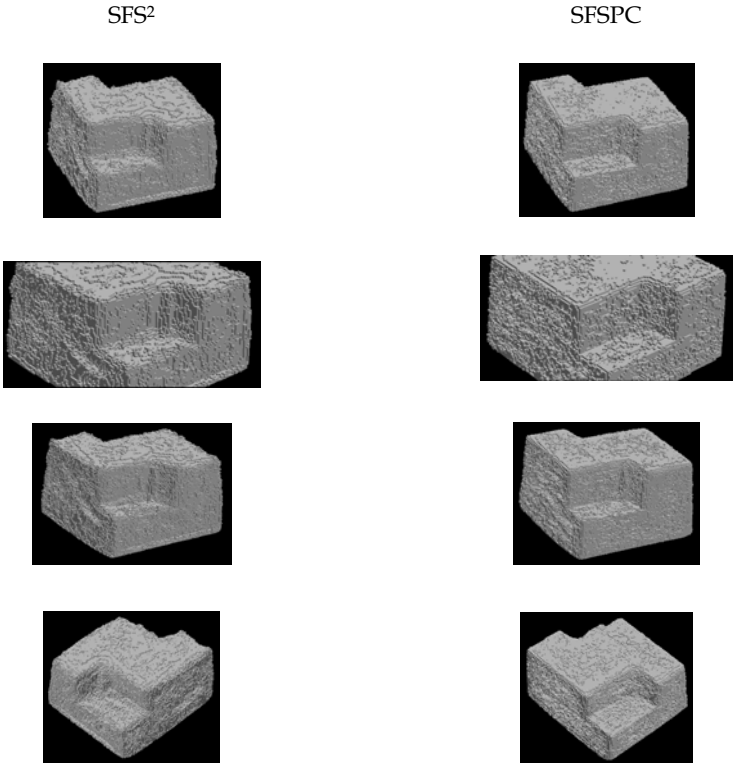


Figure 8. Comparison of SFS² and SFSPC

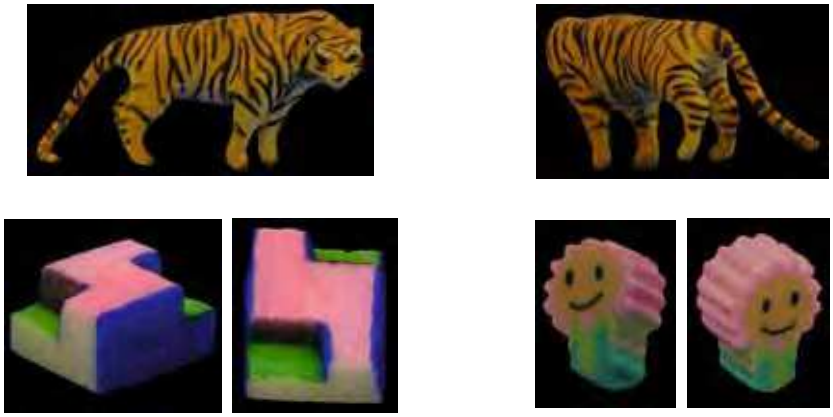


Figure 9. Synthetic views of texture mapped models of tiger, flower and square rubber

Next, we test our second volumetric model reconstruction algorithm SFSC using three plastic head objects and two ceramic teapots. The objects exhibit variety in topology (there is a hole in teapot handle) and also different degrees of concavity. They all exhibit moderately specular reflection. Each image sequence is captured by a consumer-type digital camera with a 2,048 x 1,536 CCD sensor. The target object and the calibration box are placed on a computer controlled turntable in front of the stationary camera. With neighboring views separated by 10 degrees, each image sequence contains 36 images. Photographs of the head objects and teapots are shown in Figure 10. The 3D model reconstruction system is run on an ACPI Multiprocessor x64-based PC with two Intel Xeon CPUs running at 3.6 GHz and 2 GB RAM. Each volumetric model is reconstructed in a volume space of 128 x 128 x 128.

(a) Head object 1

(b) Head object 2

(c) Head object 3

(d) Teapot 1

(e) Teapot 2

Figure 10. Photographs of the head objects and teapots

We also implement the well-known Space Carving algorithm (Kutulakos & Seitz, 2000) for comparison purpose. The volume space (128 x 128 x 128) is carved in four directions, with a group of 9 views allocated for each carving direction. Variance of pure color region is estimated in the background of the scene. Figure 11 shows the reconstructed models of head object 1. Figure 12 shows the reconstructed models of teapot 2. Figure 13 shows the texture mapped models of the objects which are reconstructed by SFSC.

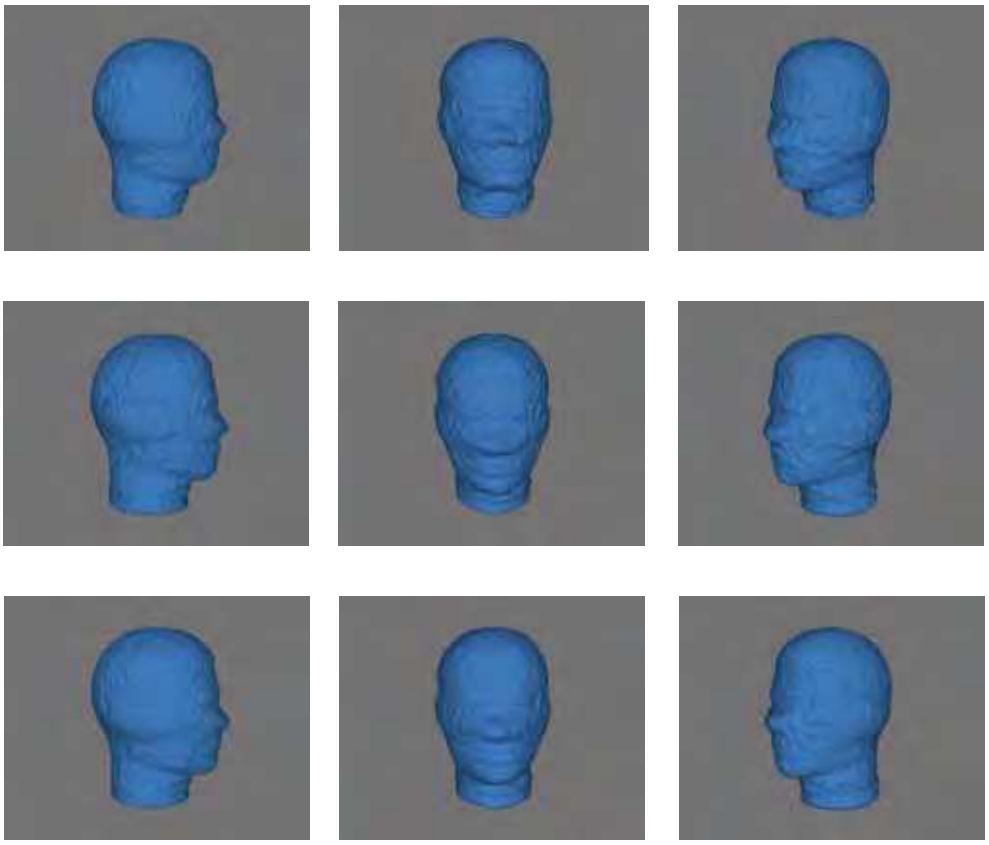


Figure 11. Reconstructed models of head object 1 using: (top row) SFSPC; (middle row) Space Carving Algorithm; (bottom row) SFSC

We also compare the performance of these volumetric model reconstruction methods quantitatively. The reprojection correctness is the ratio of the number of pixels correctly reprojected from the volume space into the object and background regions to the total number of pixels of the acquired image. In principle, this measure only indicates the accuracy of the model silhouette in the acquisition viewpoints. The quality of the reconstructed model in arbitrary viewpoints should better be judged visually than comparing the reprojection accuracy. Ideally, the reprojection correctness should be 100%.

The accuracy is lowered due to erroneous removal of model voxels and non-removal of background voxels in the refinement of the volumetric model. Therefore, the reprojection accuracy values among the three comparing algorithms are very close. Theoretically, the closer the reprojection accuracy to 100%, the better is the volumetric modeling algorithm. Figure 14 shows the reprojection correctness of head object 1. Figure 15 shows the reprojection correctness of teapot 2. For head object 1, SFSC achieves higher reprojection correctness than the Space Carving Algorithm in all image views, while the average reprojection correctness is comparably with SFSPC. For teapot 2, SFSC scores higher reprojection correctness than the Space Carving Algorithm in 34 image views, while the average reprojection correctness is slightly lower than SFSPC.

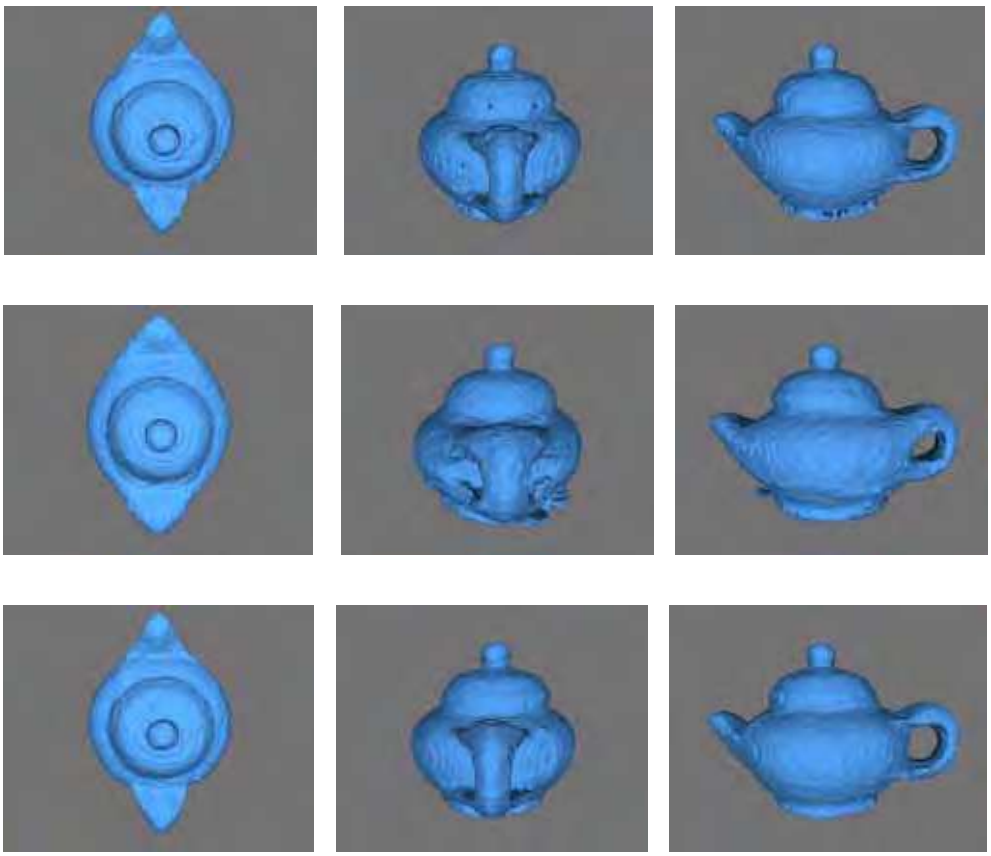
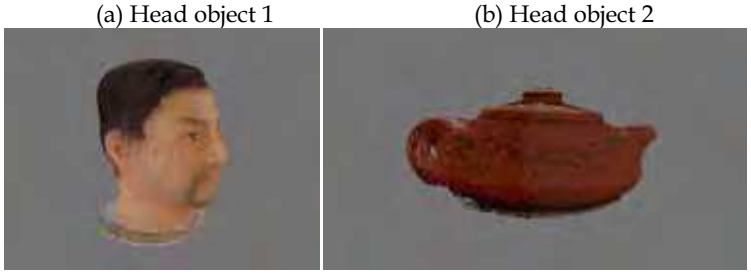


Figure 12. Reconstructed models of teapot 2 using: (top row) SFSPC; (middle row) Space Carving Algorithm; (bottom row) SFSC

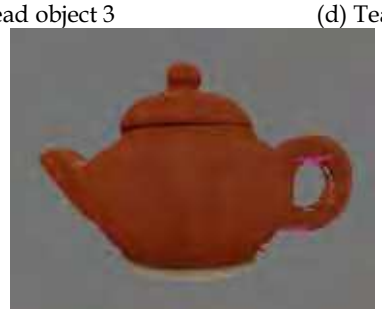
4. Discussion

SFSPC works well to model objects with more color inhomogeneity or sufficient texture information. It can quite accurately carve the concave regions. However, if the object contains a large, flat area in homogeneous color, partial surface estimation may not truly identify the real surface voxels. Besides, the computational time is still a problem. The higher the resolution is set, the more accurate and finer reconstructed model is generated. However, as the number of voxels in the volumetric model increases, more computation is needed to process the entire model. As for SFSPC, most of the time is spent on the partial surface estimation process. The computational time is directly proportional to the number of remaining voxels after the voting-based SFS. It is the trade-off between the computational time and quality of the model. Methods based on color-consistency evaluation depend on the quality and consistency of the input images, which are in turn sensitive to the lighting condition in the image capturing environment. This factor also affects SFSPC. It is the reason why the Lambertian reflectance model is assumed during the generation of the volumetric model. However, this assumption does not always apply in practical situations.

(a) Head object 1 (b) Head object 2



(c) Head object 3 (d) Teapot 1



(e) Teapot 2

Figure 13. Texture mapped models of head objects and teapots

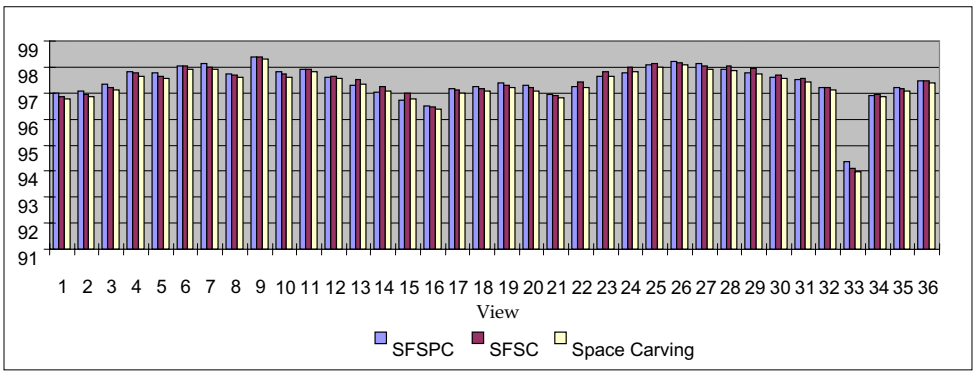


Figure 14. Reprojection correctness of head object 1

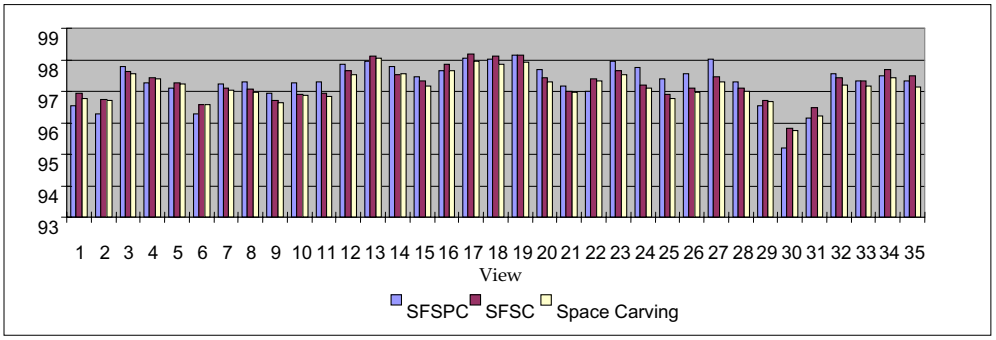


Figure 15. Reprojection correctness of teapot 2

Our second volumetric modeling method SFSC works well on the objects not only with sufficient texture information but also with color homogeneity. Moreover, it can nicely reconstruct the concavity regions. It can be seen that the new algorithm, which explicitly takes into account surface specularity, can produce better results particularly in the reconstruction of frontal face, ears and top of the head object. The SFSPC can erroneously carve away many voxels (see the volumetric model of teapot 2). This is due to the Lambertian surface reflection assumption adopted in this method. The Space Carving Algorithm produces models which are very smooth with insufficient detail structure (see the facial features of head object 1). This is a very conservative method and many background voxels are still preserved in the volumetric model as shown in the volumetric model of teapot 2. Although there is no top view in our image sequence, SFSC can produce more accurate model than the other two methods (see the first column of Figure 12).

5. Conclusion

In summary, we develop a system that can reconstruct the photorealistic 3D object model from a set of photographs. This image-based modeling system can create realistic models without requiring expensive hardware. The volumetric modeling is an important step in the system. We adopt the shape from silhouette approach to obtain the topology of the object. Detail object shape is reconstructed based on the constraint of photo-consistency. Meanwhile, a 3D voxel mask is introduced to check the photo-consistency of voxel, instead of using the conventional pixel-by-pixel block matching technique. However, the problem caused by non-Lambertian object surface is outstanding. To solve this problem, we propose a new algorithm that explicitly measures surface specularity during partial surface estimation. Meanwhile, a planar voxel mask is introduced for checking the consistency of specularities obtained from neighboring voxels in order to confirm the validity of a surface voxel. All these changes can enhance the performance of the volumetric model generation and solve the problem caused by non-Lambertian object surface. Our results show that the new volumetric modeling algorithm can produce better models than the Space Carving Algorithm and the concept of surface specularity is significant in generating high quality object model.

Future research will be focused on further improvement of the quality of the model and the speed of the reconstruction process. New calibration pattern can be designed that can facilitate more accurate camera calibration. The research in volumetric modeling is still an on-going problem. The accuracy of the volumetric modeling also depends on the volume space resolution. However, the computational time is increased with higher volumetric resolution. Future work should be done in optimizing the modeling algorithm.

6. Acknowledgement

The work described was fully supported by a grant from City University of Hong Kong (Project No. 7001857).

7. References

Chang, Y.J. & Chen, Y.C. (2002). Facial model adaptation from a monocular image sequence using a textured polygonal model. *Signal Processing: Image Communication*, Vol. 17, 373-392

Chiang, K.K. & Chan, K.L. (2006). Volumetric model reconstruction from unrestricted camera views based on the photo-consistency of 3D voxel mask. *Machine Vision and Applications*, Vol. 17, No. 4, 229-250

Eisert, P.; Steinbach, E. & Girod, B. (1999). Multi-hypothesis, volumetric reconstruction of 3-D objects from multiple calibrated camera views. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp. 3509–3512, ISBN 0-7803-5041-3, Phoenix, USA, March 1999, IEEE

Kutulakos, K.N. & Seitz, S.M. (2000). A theory of shape by space carving. *International Journal of Computer Vision*, Vol. 38, No. 3, 199-218

Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 2, 150-162

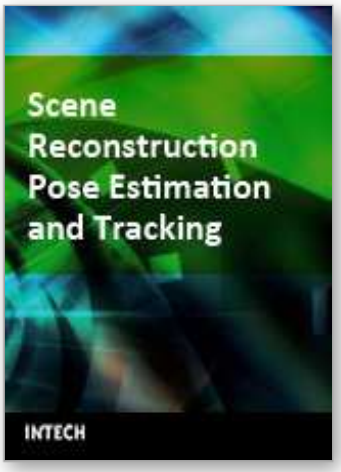
Lorensen, W.E. & Cline, H.E. (1987). Marching cubes: a high resolution 3D surface construction algorithm. *Computer Graphics*, Vol. 21, No. 4, 163-169

Matsumoto, Y.; Fujimura, K. & Kitamura, T. (1999). Shape-from-silhouette/stereo and its application to 3-D digitizer. *Proceedings of the 8th International Conference on Discrete Geometry for Computer Imagery*, pp. 177-188, ISSN 0302-9743, Marne-la-Vallée, France, March 1999, Springer-Verlag

Niem, W. (1999). Automatic reconstruction of 3D object using a mobile camera. *Image and Vision Computing*, Vol. 17, 125-134

Shafer, S. (1985). Using color to separate reflection components. *Color Research and Applications*, Vol. 10, 210-218

Wong, S.S. & Chan, K.L. (2004). Multi-view 3D model reconstruction: exploitation of color homogeneity in voxel mask. *Proceedings of the 3rd International Conference on Image and Graphics*, pp. 142-145, ISBN 0-7695-2244-0, Hong Kong, December 2004, IEEE



Scene Reconstruction Pose Estimation and Tracking

Edited by Rustam Stolkin

ISBN 978-3-902613-06-6

Hard cover, 530 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

This book reports recent advances in the use of pattern recognition techniques for computer and robot vision. The sciences of pattern recognition and computational vision have been inextricably intertwined since their early days, some four decades ago with the emergence of fast digital computing. All computer vision techniques could be regarded as a form of pattern recognition, in the broadest sense of the term. Conversely, if one looks through the contents of a typical international pattern recognition conference proceedings, it appears that the large majority (perhaps 70-80%) of all pattern recognition papers are concerned with the analysis of images. In particular, these sciences overlap in areas of low level vision such as segmentation, edge detection and other kinds of feature extraction and region identification, which are the focus of this book.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

K. K. Chiang, K. L. Chan and H. Y. Ip (2007). Photorealistic 3D Model Reconstruction based on the Consistency of Object Surface Reflectance Measured by the Voxel Mask, Scene Reconstruction Pose Estimation and Tracking, Rustam Stolkin (Ed.), ISBN: 978-3-902613-06-6, InTech, Available from: http://www.intechopen.com/books/scene_reconstruction_pose_estimation_and_tracking/photorealistic_3d_model_reconstruction_based_on_the_consistency_of_object_surface_reflectance_measur

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.