COMPARISON OF PHOTOCONSISTENCY MEASURES USED IN VOXEL COLORING

Oğuz Özün[a], Ulaş Yılmaz[b]*, Volkan Atalay[a]

[a]Department of Computer Engineering, Middle East Technical University, Turkey – {oguz, volkan}@ceng.metu.edu.tr
[b]Computer Vision & Remote Sensing, Berlin University of Technology, Germany – ulas@cs.tu-berlin.de

KEY WORDS: photoconsistency, voxel coloring, three-dimensional object modeling, standard deviation, Minkowsky distance, adaptive threshold, histogram, color caching.

ABSTRACT

A framework for the comparison of photoconsistency measures used in voxel coloring algorithm is described. With this framework, the results obtained in generalized voxel coloring algorithm using certain photoconsistency measures are discussed qualitatively and quantitatively. The photoconsistency measures are based on standard deviation, Minkowsky distance, adaptive threshold, histogram, and color caching. Quantitative measurement is performed with root-mean-square-error and normalized-cross-correlation-ratio. The results show that, the photoconsistency measures which require threshold(s) may produce better/worse results depending on the selected threshold(s). Also, modeling textured objects always produce better reconstruction results.

1 INTRODUCTION

The main goal of volumetric scene modeling approaches is to find out, whether a given point is on an object's surface in the scene or not. According to the information being used in reconstruction, these approaches are classified into two groups: *shape-from-silhouette* and *shape-from-photoconsistency*. In shape-from-silhouette approaches, the model is extracted using back projections of the silhouettes onto the images: Each back projected silhouette corresponds to a volume in the space; intersection of these volumes gives the model (Mülayim et al., 2003). In shape-from-photoconsistency approaches, on the other hand, the photoconsistency of light coming from a point in the scene is taken into account: If the light coming from a point in the scene is not photoconsistent, then this point cannot be on a surface in the scene. In both approaches, the space is modeled using volume elements, *voxels*. *Voxel coloring* and its variations (*space carving (Broadhurst and Cipolla, 2000, Kutulakos and Seitz, 2000)*, *generalized voxel coloring (Culbertson et al., 1999)*, *multihypothesis voxel coloring (Steinbach et al., 2000)*, etc.) are in shape-from-photoconsistency group. These variations differ in the way they determine the visibility of a given voxel. When the same photoconsistency measure is used, a significant difference in the output is not observed. So to say, in voxel coloring algorithm and in its variations, the reconstructed model highly depends on the used photoconsistency measure. In this study, generalized voxel coloring algorithm is used to compare the effect of different photoconsistency measures. The experiments are performed on 3 image sequences. Results obtained through different photoconsistency measures are then compared using image comparison techniques, *root-mean-square-error* and *normalized-cross-correlation-ratio*.

The organization of the paper is as follows: In the following section, voxel coloring algorithm used in this study is explained in detail. This section is followed by Section 3 in which photoconsistency measures used in voxel coloring are presented. Qualitative comparison of photoconsistency measures is described in Section 4. Results obtained in the framework of this study are given in Section 5. The paper concludes with Section 6, in which the results are discussed and comments about the presented photoconsistency measures are made.

*Corresponding author.

2 SHAPE-FROM-PHOTOCONSISTENCY

Depending on surface properties, lighting conditions and viewing direction, color of light coming from a point in the scene varies. Nevertheless, different observations should be coherent. In other words, a point on a surface should be seen with similar colors when it is not occluded. This phenomenon is called *photoconsistency*. Shape-from-photoconsistency approaches of volumetric scene reconstruction are based on this property of surfaces. If the light coming from an unconcluded portion of the scene is photoconsistent, then this point should be on a surface in the scene. Otherwise, it should be empty. This claim is based on a couple of assumptions: Objects in the scene have Lambertian surfaces, and projection of any point in the scene on the images can be computed (Kutulakos and Seitz, 2000).

In this study, due to its easiness in implementation, generalized voxel coloring algorithm is used. In order to improve computational cost, convex hull of the object is computed, and it is used as the input of voxel coloring algorithm. For each view, voxel space is divided into layers according to the distance to the view point. From nearest to furthest, layers are traversed, and the voxels are checked for visibility. Initially, all pixels in all images are unmarked. At each level, visible pixels are marked, so that the visibility of voxels at the following layers can be decided. Assume that a voxel v at some further layer is visited. Compared to the other voxels which are at nearer layers, it should have less number of visible pixels in each image. Having extracted visible pixels from all images, a set of colors is obtained. If this set is photoconsistent, then the voxel is on the surface. The pixels are marked and next voxel is processed. If the set is not photoconsistent, then the voxel is removed from the model. The ordinal visibility constraint and traversal of voxels based on layers makes it possible to use pixel marking as an efficient tool to handle occlusions: Single sweep through the voxel space is enough to reduce the voxel set to a more photoconsistent state. This procedure is iterated until all voxels are photoconsistent.

3 PHOTOCONSISTENCY MEASURES

As it is mentioned in the previous section, removal or coloring decision of a voxel depends on the set of colors, which is obtained by projecting the voxel onto the images. Given n images, assume that $I_0, I_1, ..., I_{p-1}$ are the images in which voxel v is not occluded. Then, for v, a nonempty set of colors, π, is extracted from the images as shown in Equation 1 where π_j is the set of colors extracted for v from image j, and $c_0, c_1, ..., c_m$ are the extracted color values.

$$\pi = \bigcup_{j=0}^{p-1} \pi_j = \{c_0, c_1, ..., c_m\} \qquad (1)$$

Once this set is extracted, its photoconsistency can be defined in various ways. Some criteria used in the literature are as follows:

1. Standard deviation (Seitz and Dyer, 1999, Kutulakos and Seitz, 2000, Culbertson et al., 1999, Broadhurst and Cipolla, 2000),
2. adaptive threshold (Slabaugh et al., 2004),
3. Minkowsky distance (Slabaugh et al., 2001),
4. histogram (Slabaugh et al., 2004),
5. color caching (Chhabra, 2001).

3.1 Standard Deviation

Using standard deviation σ as a photoconsistency measure is proposed by Seitz and Dyer (Seitz and Dyer, 1999). If the standard deviation σ of π is less than a threshold τ, π is photoconsistent, which means the v is on the surface.

$$consistent(v) = \left\{ \begin{array}{ll} true, & \sigma < \tau \\ false, & otherwise \end{array} \right\} \qquad (2)$$

3.2 Adaptive Threshold

The consistency measure based on standard deviation works well when the object's surface color is homogeneous. Otherwise, it easily diverges. In order to handle this problem, Slabaugh et al. (Slabaugh et al., 2004) proposed a new photoconsistency measure called adaptive threshold: If a voxel is on an edge or on a textured surface, then the variation of the extracted color values is higher. Larger thresholds should be used so that the photoconsistency measure converges. Assume that a voxel v, which is on an edge or on a textured surface, is visible from p views, $I_0, I_1, ..., I_{p-1}$. Then, the color sets extracted from these images for v are $\pi_0, \pi_1, ..., \pi_{p-1}$, and the standard deviations of these sets are $\sigma_0, \sigma_1, ..., \sigma_{p-1}$, respectively. These standard deviations should be high, since v is on an edge or on a textured surface. The average of these standard deviations, which is given in Equation 3, should also be high. By using this observation, authors define a new photoconsistency measure called adaptive threshold as given in Equation 4.

$$\overline{\sigma} = \frac{1}{p-1} \sum_{j=0}^{p-1} \sigma_j \qquad (3)$$

$$consistent(v) = \left\{ \begin{array}{ll} true, & \sigma < \tau_1 + \overline{\sigma}\tau_2 \\ false, & otherwise \end{array} \right\} \qquad (4)$$

This measure brings an important advantage over the measure based on standard deviation: The value of threshold is variable according to the place of the voxel. If the voxel is on the edge or textured surface, this situation can be detected with high standard deviation in each image, and a greater threshold can be used. Adaptive threshold measure is actually superset of the measure based on standard deviation. The need for 2 thresholds is its main disadvantage.

3.3 Minkowsky Distance

Photoconsistency of a set can also be defined using Minkowsky distances, L_1, L_2 and L_∞. Minkowsky distance between two points x and y in $\Re^k$ is given in Equation 5.

$$L_p(x, y) = \left(\sum_{i=0}^{k} |x_i - y_i| \right)^{\frac{1}{p}} \qquad (5)$$

Assume that a voxel v is visible from p views, $I_0, I_1, ..., I_{p-1}$, and the color sets extracted from these images are $\pi_0, \pi_1, ..., \pi_{p-1}$. Every color entity in each of these color sets should be in a certain distance to the color entities of the other sets. Through this idea, photoconsistency of v is defined as in Equation 6. The distance between two color sets is given in Equation 7.

$$consistent(v) = \left\{ \begin{array}{ll} true, & \forall_{i,j}, consistent_{i,j}(v) \\ false, & otherwise \end{array} \right\} \qquad (6)$$

$$consistent_{i,j}(v) = \left\{ \begin{array}{ll} true, & \forall_{c_l \in \pi_i, c_m \in \pi_j}, L_p(c_l, c_m) < \tau \\ false, & otherwise \end{array} \right\} \qquad (7)$$

The most important benefit of using Minkowsky distance as a photoconsistency measure is the following. During the photoconsistency check, if the voxel is found to be inconsistent, there is no need to continue to check photoconsistency of that voxel. That means, having found a pair of colors whose difference is greater or equal to the threshold, voxel cannot be photoconsistent.

3.4 Histogram

In order to get rid of thresholds, Slabaugh et al (Slabaugh et al., 2004) proposed a new photoconsistency measure based on color histogram. Photoconsistency check is performed in two steps: histogram construction and histogram intersection. In the first step, visible pixels of the voxel v are extracted and a color a histogram is constructed for each image. Next step is photoconsistency check. To check whether v is photoconsistent or not, all pairs of histograms of v have to be compared: Two views i and j of v are photoconsistent with each other, if their histograms H_{v_i} and H_{v_j} match. A matching function, $match(H_{v_i}, H_{v_j})$, which compares two histograms, and returns a similarity value should be defined. Then the decision about the photoconsistency v is made according to the measure given in Equation 8.

$$consistent(v) = \left\{ \begin{array}{ll} true, & \forall_{i,j}, match(H_{v_i}, H_{v_j}) \neq 0 \\ false, & otherwise \end{array} \right\} \qquad (8)$$

The advantage of histogram based photoconsistency measure is that there is no need for a preset threshold. Furthermore, paired tests can be very efficient in some circumstances. For instance, if the voxel is found to be inconsistent for a pair, there is no need to test other pairs of views for photoconsistency.

3.5 Color Caching

Color caching is a photoconsistency measure proposed by Chhabra et al. (Chhabra, 2001). It brings a solution to the limitations caused by Lambertian assumption. Photoconsistency of a voxel is checked twice before it is removed: If a voxel is found inconsistent in the first step, it is passed to the second step for another check. At the first step, surface parts which show Lambertian reflectance properties are tested. Those surface parts which fail Lambertian assumptions for some reason (material properties, viewing orientation, position of the light sources, etc.) are tested at the second step. The irradiance from these parts can be inconsistent. In order to prevent carving of these parts, before carving a voxel, it should be checked with another measure which takes care of viewing orientation. Given an image, for each voxel v, a cache is constructed. Each cache holds the visible colors of v from the relevant image. Having constructed color caches for each image, these caches are checked to find out, if there is a similar or a common color in all pairs of caches. If there is a match between all pairs of caches, v is labeled as consistent. If there is any pair of views whose caches do not contain a similar or a common color, v is labeled as inconsistent. Chhabra (Chhabra, 2001) defines similarity measure between two colors $c_i = (R_i, G_i, B_i)$ and $c_j = (R_j, G_j, B_j)$ as in Equation 9, and similarity of two images I_i and I_j as in Equation 10.

$$similarity(c_i, c_j) = \left\{ \begin{array}{ll} true, & \Delta_{i,j} \leq \tau_1 \\ false, & otherwise \end{array} \right\}$$
$$\Delta_{i,j} = \sqrt{(R_i - R_j)^2 + (B_i - B_j)^2 + (B_i - B_j)^2} \qquad (9)$$

$$similarity(I_i, I_j) =$$
$$\left\{ \begin{array}{ll} true, & \exists_{c_l \in cache_i} \exists_{c_m \in cache_j} similarity(c_l, c_m) \\ false, & otherwise \end{array} \right\} \qquad (10)$$

So, the photoconsistency of voxel v is defined as in Equation 11.

$$consistent(v) = \left\{ \begin{array}{ll} true, & \forall_{i,j} consistent(I_i, I_j) \\ false, & otherwise \end{array} \right\} \qquad (11)$$

4 COMPARISON

In order to compare photoconsistency measures, one should be able to measure quantitatively the quality of the reconstructed models. In this study, similarity between captured and rendered images of model is used to obtain a quantitative quality measure. The images are compared using root-mean-square-error (RMSE) and normalized cross-correlation-ratio (NCCR). Definitions of these measures are given in Equations 12 and 13, where M and N are the image dimensions, and G is the maximum intensity value.

$$RMSE(A, B) = \frac{1}{G\sqrt{MN}} \sqrt{\sum_{i,j}^{M,N} (A_{ij} - B_{ij})^2} \qquad (12)$$

$$NCCR(A, B) = 1 - \frac{\sum_{i,j}^{M,N} A_{ij} B_{ij}}{\sqrt{(\sum_{i,j}^{M,N} A_{ij}^2)(\sum_{i,j}^{M,N} B_{ij}^2)}} \qquad (13)$$

5 EXPERIMENTAL RESULTS

Photoconsistency measures are tested using 3 objects: "cup", "star", and "box". Voxel space resolution is set to $450 \times 450 \times 450$ for all objects. Its handle makes the **"cup"** object hard to model. Furthermore, the available texture information is not adequate to obtain good results. The image sequence consists of 18 images, 16 of which are used for reconstruction and 2 for testing. Reconstructed models are shown in Figure 1 and measured quality of the reconstruction is tabulated in Table 1 and Table 2. There is no significant difference between the reconstructed models quantitatively. Qualitative results support this result. Histogram based photoconsistency measure would be the best choice for this image sequence, since there is no need for threshold tuning in this approach.

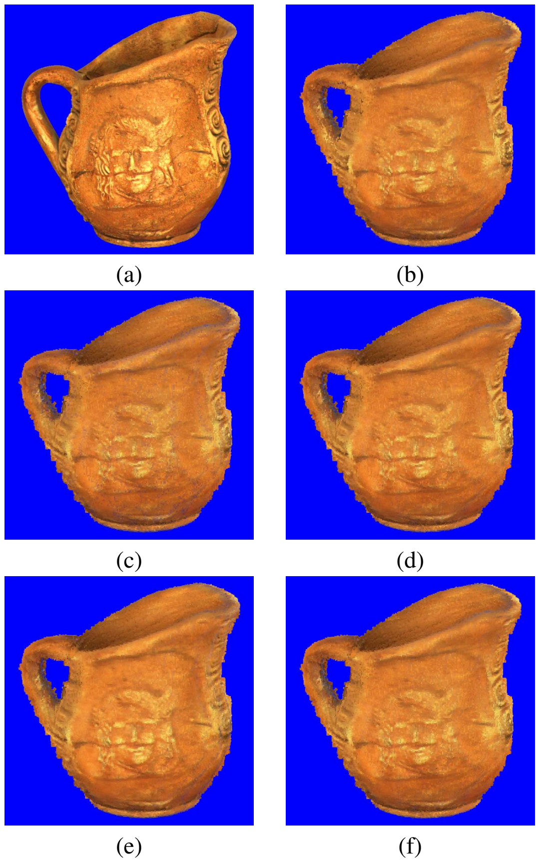


Figure 1: (a) Original image, and artificially rendered images of "cup" object obtained using (b) standard deviation, (c) histogram, (d) adaptive threshold, (e) L_1 norm, and (f) color caching.

Image No	Measure	$NCCR$ (%)	$RMSE$ (%)
08	standard deviation	2.96	10.24
	histogram	2.95	10.22
	adaptive threshold	2.95	10.22
	L_1 norm	3.00	10.31
	color caching	2.95	10.21
17	standard deviation	1.68	7.99
	histogram	1.69	8.04
	adaptive threshold	1.68	8.00
	L_1 norm	1.86	8.43
	color caching	1.67	7.97

Table 1: Error analysis using test images for "cup" object.

Measure	$NCCR$ (%)	$RMSE$ (%)
standard deviation	1.55	7.37
histogram	1.58	7.45
adaptive threshold	1.55	7.37
L_1 norm	1.63	7.56
color caching	1.56	7.38

Table 2: Average error for "cup" object.

In order to test the photoconsistency measures on an object with a simple geometry, **"box"** object is used. 12 images are taken around the object and all of them are used during the reconstruction. In Figure 2 the result obtained using L_1 norm as photoconsistency measure is illustrated. Measured quality of the reconstruction is tabulated in Table 3. Rather than its quantitative results, the visual quality of the reconstructed model gives a clue. It is about the success of shape-from-photoconsistency approaches in general: The finer the voxel space, the better is the reconstruction and the more is the computational complexity.

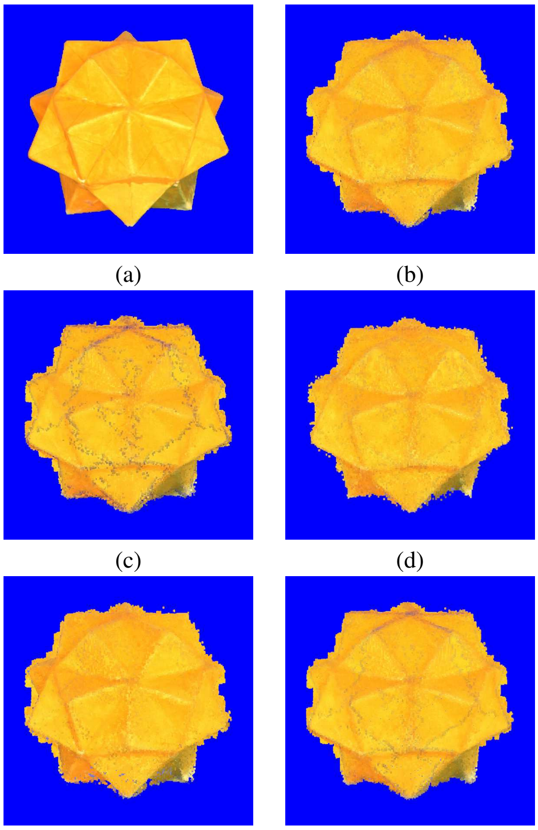
(a) (b)

Figure 2: (a) Original image, and (b) artificially rendered image of "box" object obtained using L_1 norm.

Measure	$NCCR$ (%)	$RMSE$ (%)
standard deviation	2.15	10.73
histogram	2.17	10.85
adaptive threshold	2.09	10.57
L_1 norm	2.15	10.70
color caching	2.09	10.61

Table 3: Average error for "box" object.

"star" object is a good example of objects that has not texture information but a complex geometry. The image sequence for this object consists of 18 images, 9 of which are used for reconstruction and 9 for testing. Reconstructed models are shown in Figure 3 and measured quality of the reconstruction is tabulated in Table 4. Adaptive threshold seem to produce a better result than the others. Selecting low values for the thresholds causes overcarving. On the other hand, the higher is the threshold, the coarser is the reconstructed model. There is a high dependency on selecting proper thresholds for poor-textured objects. There is no need for a threshold in histogram-based method. However, in this case, lack of texture information causes some voxels not to be carved. Uncarved voxels have the color blue, i. e. the background color. Similar effect is also observed when color caching is used as photoconsistency measure.



(a) (b) (c) (d) (e) (f)

Figure 3: (a) Original image, and artificially rendered images of "star" object obtained using (b) standard deviation, (c) histogram, (d) adaptive threshold, (e) L_1 norm, and (f) color caching.

Measure	$NCCR$ (%)	$RMSE$ (%)
standard deviation	0.85	6.39
histogram	0.91	6.54
adaptive threshold	0.74	5.89
L_1 norm	0.85	6.35
color caching	0.86	6.40

Table 4: Average error for "star" object.

6 CONCLUSIONS

A framework for the comparison of photoconsistency measures used in voxel coloring algorithm is described. Reconstruction results of 3 objects, which are obtained using generalized voxel coloring algorithm are discussed. The methods, in which thresholds are used, generally give the best results, if suitable thresholds are set. Better thresholds can be found empirically. Standard deviation gives appropriate results for the voxels, which are at the edges on the images or which are projected onto highly-textured regions of the images. Using adaptive thresholds, the threshold is changed according to the position of the voxel. But this change is controlled by another threshold, which is actually the bottleneck of the approach. The second threshold prevents carving voxels which are at the edge or highly-textured. When the second threshold is not a suitable value, it might generate cusps in the final model. When objects to be modeled are highly-textured, it is better to use histogram-based photoconsistency measure. In this approach, there is no need for a threshold. Minkowsky distance does not have a specific benefit as a photoconsistency measure. However, Minkowsky distance is a monotonically increasing function. So, if the color set is found to be inconsistent from some views, then there is no need to check the visible pixels from other views. This speeds the computation up. Color caching is an appropriate photoconsistency measures to eliminate the highlights. However, it is also possible to eliminate the specularities using background surface and selected lighting.

REFERENCES

Broadhurst, A. and Cipolla, R., 2000. A statistical consistency check for the space carving algorithm. In: British Machine Vision Conference, pp. 282–291.

Chhabra, V., 2001. Reconstructing specular objects with image based rendering using color caching. Master's thesis, Worcester Polytechnic Institute.

Culbertson, W. B., Malzbender, T. and Slabaugh, G. G., 1999. Generalized voxel coloring. In: International Conference on Computer Vision, Vision Algorithms Theory and Practice, pp. 100–115.

Kutulakos, K. N. and Seitz, S. M., 2000. A theory of shape by space carving. International Journal of Computer Vision 38(3), pp. 199–218.

Mülayim, A. Y., Yılmaz, U. and Atalay, V., 2003. Silhouette-based 3d model reconstruction from multiple images. IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics Special Issue on 3-D Image Analysis and Modeling 33(4), pp. 582–591.

Seitz, S. M. and Dyer, C. R., 1999. Photorealistic scene reconstruction by voxel coloring. International Journal of Computer Vision 35(2), pp. 157–173.

Slabaugh, G. G., Culbertson, W. B., Malzbender, T. and Schafer, R. W., 2001. A survey of volumetric scene reconstruction methods from photographs. In: Volume Graphics, pp. 81–100.

Slabaugh, G. G., Culbertson, W. B., Malzbender, T., Stevens, M. R. and Schafer, R. W., 2004. Methods for volumetric reconstruction of visual scenes. International Journal of Computer Vision 57(3), pp. 179–199.

Steinbach, E., Girod, B., Eisert, P. and Betz, A., 2000. 3-d object reconstruction using spatially extended voxels and multi-hypothesis voxel coloring. In: International Conference on Pattern Recognition, pp. 774–777.