

# A Histogram-based Color Consistency Test for Voxel Coloring

Mark R. Stevens  
Charles River Analytics  
[mstevens@cra.com](mailto:mstevens@cra.com)

Bruce Culbertson, Tom Malzbender  
Hewlett Packard Research Labs  
{culberts,malzbend}@hpl.hp.com

## Abstract

*Voxel Coloring has become a popular technique for reconstructing a 3D scene from a set of 2D images. While many different variants of this technique exist, all rely on a test to determine if each voxel is projecting to regions of consistent color in all views of that voxel. A number of color consistency tests can be used and the specific choice has a large influence on the quality of the reconstruction. Earlier work has used variance or the  $L_1$  norm. We propose a new form of consistency test, based on histograms, that 1) is more robust at reconstructing textured surfaces, 2) deals properly with RGB color information, and 3) does not require fine tuning of parameters.*

## 1. Introduction

We present a new color consistency test for Voxel Coloring [1] that is based on histograms. Voxel Coloring has become a popular technique for reconstructing a 3D scene from a set of 2D input images. While many different variants of the technique exist, they all use some form of color consistency test. A voxel is included in the reconstruction if it projects to consistent colors, as determined by the test, in the images from which it is visible [1,2,3,4]. A number of color consistency tests can be used and the specific choice has a large influence on the quality of the reconstruction. In previous work, consistency has been defined in terms of thresholded variance or  $L_1$  norm—in other words, a set of colors is defined to be consistent if its color variance or  $L_1$  norm is less than a fixed threshold. In contrast to these forms of consistency tests, our new test 1) is more robust at reconstructing textured surfaces, 2) deals properly with RGB color information, and 3) does not require fine tuning of parameters.

Thresholded variance and  $L_1$  norm are problematic when used for Voxel Coloring. Textured regions have a higher color variance than uniform regions and most scenes have some of each. Hence, no single threshold is ideal for an entire scene. Furthermore, when a voxel is visible in an image, it is usually visible from many pixels in the image (on average about 30 pixels in our data sets). The pixels from which a voxel is visible often span regions of different color. Thus, the color distribution of the

pixels will in general be multi-modal (clustered around more than one color). Variance and  $L_1$  norm are not good consistency measures for reconstruction because they cannot distinguish between these clustered color distributions and purely random distributions; in contrast, histograms can. We performed reconstructions using both our new consistency test, as well as thresholded variance, on many data sets, and our method consistently produced reconstructions that were visually preferable.

Another disadvantage of the earlier consistency tests is that the quality of the reconstruction is very sensitive to the value of the threshold that is used. Reconstructions must be attempted with many thresholds to experimentally find the best threshold. This can be time consuming. We performed this process on two of our data sets, selected at random, using thresholded variance. On the first data set, the reconstruction failed catastrophically using a threshold that was just 5% below the threshold that yielded the best reconstruction. The best reconstruction of the second data set was obtained with a threshold 53% higher than the threshold that yielded the best reconstruction on the first data set. In contrast, we obtained good reconstructions of all fourteen of our data sets, using our histogram based consistency test, with no parameter adjustments whatsoever.

We test the color consistency of sets of pixels by comparing their histograms. There are many ways to compare histograms; we have found histogram intersection [5] to be efficient and it did a good job reconstructing all fourteen of our data sets. When the consistency of a voxel is tested, a histogram is constructed for each image that can see the voxel. The colors of all the pixels from which the voxel is visible are inserted into the histogram. Then all pairs of the histograms are compared by intersecting their corresponding bins. If any pair has no corresponding bins that are both occupied, then the voxel is considered to be inconsistent. More details follow in section 3.

## 2. Previous Work

As mentioned previously, several forms of consistency tests have been used with Voxel Coloring. Seitz and Dyer [1] give the original description of Voxel Coloring. In

their formulation, the set of pixels  $\pi_j^i$  in image  $j$  from voxel  $i$  is visible is computed. The union of all views defines the color of that voxel:

$$\Pi_i = \bigcup_{j=0}^m \pi_j^i$$

where  $m$  is the number of images, which view the voxel. To determine if the voxel is colored consistently, the likelihood ratio test of this set is computed and thresholded. Therefore, if [1]:

$$(m-1)\sigma_{\Pi_i} < \tau$$

then the voxel is labelled consistent (where  $\tau$  is the threshold parameter). Kutulakos [8] first computes the mean color of each  $\pi_j^i$  and then computes the likelihood ratio test of those means and compares them to a threshold. Culbertson *et al* [3] compare the variance of  $\pi_j^i$  to a threshold. Eisert *et al* [6] begin their voxel color consistency computation similarly to Kutulakos by first finding the mean of  $\pi_j^i$  for all the images and then computing the mean color  $\mu_j$  of each non-empty  $\pi_j^i$ . Then, unlike Kutulakos, for every pair  $\mu_j$  and  $\mu_k$  of such means, they compute the  $L_1$  norm. The  $L_1$  norm of is defined as:

$$L1_{j,k} = |\mu_j^{red} - \mu_k^{red}| + |\mu_j^{green} - \mu_k^{green}| + |\mu_j^{blue} - \mu_k^{blue}|$$

If the maximum of these norms is less than a threshold that they determine experimentally, then they consider the voxel to be color consistent. A characteristic that all of these color consistency measures have in common is that they cannot distinguish color distributions grouped in several clusters from purely random distributions.

Voxel Coloring methods differ in other ways, too. For example, most methods require metric camera calibration. However, [4] can exploit weakly calibrated cameras and [8] can tolerate some inaccuracy in camera calibration. In [1] and [2], the pixels that can see a voxel are found using a plane sweep method, whereas [3] uses item buffers or layered depth images. However, all of these variations are independent of the choice of color consistency test. Hence, they can capitalize on the advantages of our histogram based color consistency test.

### 3. Our Approach

To avoid the difficulties associated with thresholded variance and  $L_1$  norm as photo-consistency measures, we propose the use of a histogram-based method. This method is inspired by the work of [5], where histograms were used for object indexing.

Instead of pooling all of the pixels from all of the views of a given voxel, as was done in [1], we use a series

of tests on image pairs. At any point as the Voxel Coloring algorithm runs, we can find the set  $\pi_j^i$  of pixels from image  $j$  that are visible from a voxel  $i$ . We test the consistency of the voxel by comparing all pairs of such sets:

$$\forall_{k,l} \pi_k^i \approx \pi_l^i \quad k \neq l$$

where  $\pi_k^i$  and  $\pi_l^i$  are non-empty. Instead of directly comparing the sets of pixels, we say a voxel is consistent if all of the histograms of all views of the voxel intersect:

$$\forall_{k,l} Hist(\pi_k^i) \cap Hist(\pi_l^i) \neq \phi \quad k \neq l$$

Pairs of histograms intersect if at least one pair of corresponding bins has a non-zero count. Therefore, a single pair of views can cause a voxel to be declared inconsistent if the colors they see at the voxel do not overlap. We use a 3D histogram over the complete color space. Furthermore, we have found that 8 bins per channel are adequate for acceptable reconstructions; this yields 512 bins (8x8x8) for each image of each voxel.

We made several optimizations to minimize the runtime and memory requirements related to our consistency test. Notice that the histogram intersection only needs to test which histogram bins are occupied. Hence, only one bit is required per bin, or 512 bits per histogram. Histogram intersection can be tested with AND operations on computer words. Using 32-bit words, only 16 AND instructions are needed to intersect two histograms. In the worst case, the number of histogram comparisons to test the consistency of a voxel is equal to the square of the number of images that can see the voxel. Fortunately, in our data sets the average number of images that could see a voxel fell between two and three, so the number of histogram comparisons was manageable.

#### 3.1. Overlapping Histogram Bins

The voxel color consistency test, as described above, has a deficiency: it treats color values that fall near bin boundaries differently than other colors. For simplicity, consider colors that have the same green and blue values and consider the behaviour of just the red value. In our specific implementation, a pixel with red value of 15 falls in the middle of its bin. Another pixel with a red value in the range 15 to 31 will be counted in the same bin. If these two pixels are in two different visibility sets, they have intersecting histograms and are considered consistent. On the other hand, two pixels with red values 31 and 32 are close in color space, yet they are counted in different bins. These two red values would not be sufficient to pass the consistency test.

Since the choice of bin boundaries is somewhat arbitrary, the color consistency test should be insensitive to their exact placement. Furthermore, color cannot be measured exactly. For example, a camera will measure slightly different color values on successive exposures of

the identical scene. In addition, lighting usually changes slightly over time, causing minor shifts in color. Hence, for a number of reasons, we would like the consistency test to ignore small differences in color.

We corrected the problems just described by using overlapping bins. In effect, this blurs the bin boundaries. In our specific case, we enlarged the bins to overlap adjacent bins by about 20 percent. A pixel with a color falling into multiple overlapping bins is counted in each such bin. A pixel falling into just one bin is counted in just that bin, as before. This makes the consistency test insensitive to bin boundaries and small inaccuracies in color measurement.

### 3.2. Voxel Projection Size

We found histogram intersections to be a somewhat unreliable indicator of color consistency when a voxel was visible from only a small number of pixels in an image. Hence, when computing the color consistency of a voxel, if this number fell below a fixed number (we typically used 15), we treated the image the same as an image that could not see the voxel at all.

At first, it may appear that this places an unfortunate lower limit on the voxel size to insure that voxels project to enough pixels. In practice, this is very unlikely to ever be a problem. Runtime is proportional to the cube of the number of voxels; this limits the voxel size long before the requirements of the histogram consistency test. Even the earliest Voxel Coloring experiments used data sets with voxels that projected to sufficient pixels. Since then higher resolution cameras have become inexpensive, so high-resolution images are now easy to obtain. In addition, the additional information in high-resolution images is beneficial to reconstruction. Using this approach, the only way to increase the resolution of the model (the number of voxels) is to increase the resolution of the camera.

## 4. Results

We have run a comparison between the standard deviation consistency check and our histogram consistency check. For this comparison, we have used images collected of a boombox object sitting on a rotating table. Nine images of the object were collected. Eight were every 45 degrees around the object and one was looking straight down. Figure 1 shows two example images from this dataset.



**Figure 1. Two images of a boombox.**

The images were hand calibrated using the Tsai camera calibration model [7]. We hooked our histogram consistency check into our own implementation of the Generalized Voxel Coloring algorithm [3]. Our implementation of this algorithm has three steps.

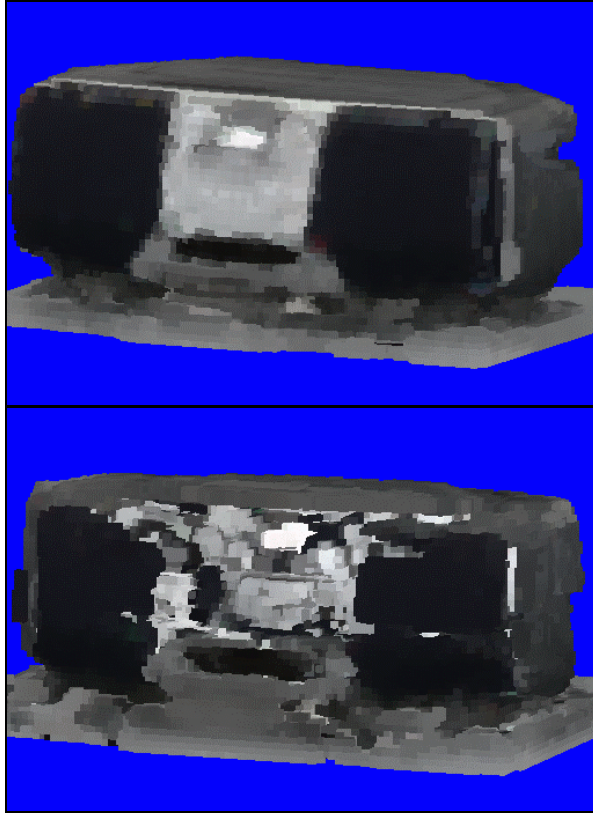
1. First, all opaque voxels (initially all voxels are opaque) are rendered, using OpenGL, into all views. The voxels are modeled as cubes and the calibration parameters are used to set up the projection matrices. The pixels under each voxel in each view are used to either update the appropriate histograms, the standard deviation, or the silhouette background counts.
2. Second, a consistency check is run on each voxel and those that do not meet the consistency criteria are discarded (marked transparent).
3. Third, the algorithm repeats the first two steps until no voxels can be carved (marked transparent).

While the histogram technique has no parameters, the variance-based algorithm requires a threshold setting. For this consistency check, numerous volumes were reconstructed using a wide range of threshold parameters. The best produced reconstruction for each parameter was used for the comparison. Figure 2 shows resulting reconstructions. As can be seen, the histogram produces a better reconstruction.

In addition, we have run our histogram reconstruction algorithm on 14 different datasets. In each case, the scenes reconstructed using the same parameter settings. Figure 3 shows several of these reconstructions.

## 5. Conclusions

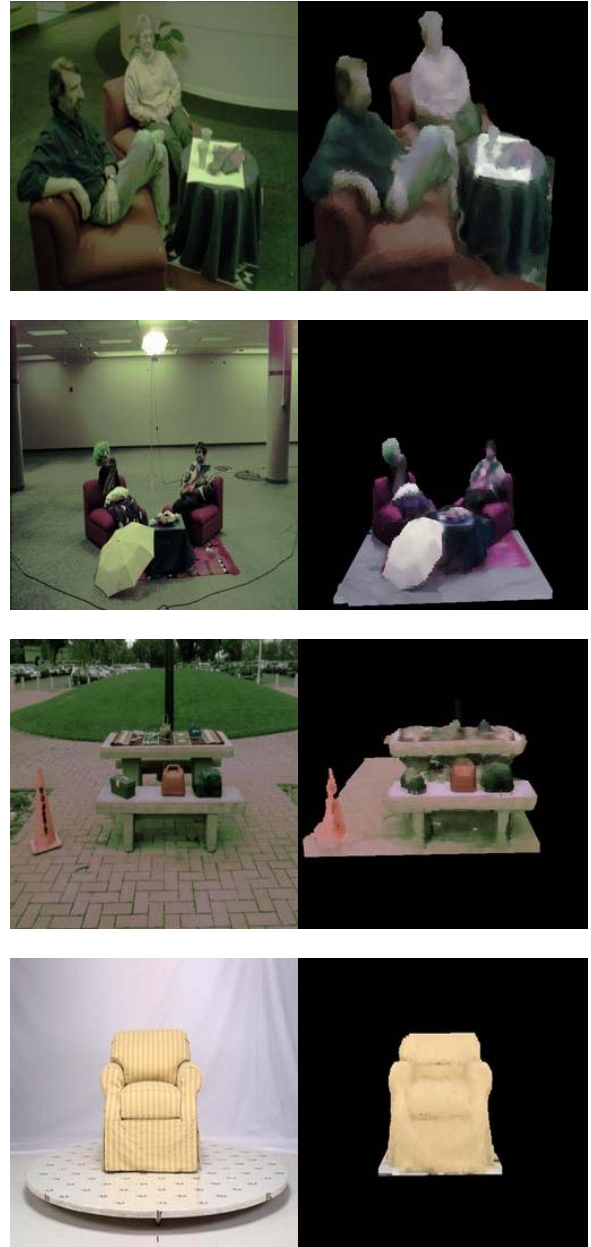
We have presented a new color consistency test for Voxel Coloring. The new technique constructs color histograms for all views of a voxel and intersects the histograms pair-wise. The voxel is considered to be consistent if all such intersections are non-empty. This contrasts with earlier methods, some of which pool all of the information from all of the views, and all of which compare color variance or  $L_1$  norm to a threshold. Our new consistency test does not require parameter tuning and can robustly reconstruct uniform and highly textured scenes.



**Figure 2. Reconstruction for histogram (top) and variance (bottom) consistency checks.**

## References

- [1] S. Seitz, C. Dyer. (1997). "Photorealistic Scene Reconstruction by Voxel Coloring", Proceedings of Computer Vision and Pattern Recognition Conference, 1067-1073.
- [2] K. Kutulakos, S. Seitz. (1998). "What Do N Photographs Tell Us about 3-D Shape?", University of Rochester, TR680.
- [3] G. Slabaugh, B. Culbertson, T. Malzbender. (1999). "Generalized Voxel Coloring". Vision Algorithms Workshop.
- [4] Saito, H. and Kanade, T. 1999. "Shape Reconstruction in Projective Grid Space from Large Number of Images", CVPR, pp. 49 - 54.
- [5] M. Swain. (1990) "Color Histograms for Object Indexing", PhD. Thesis, University of Rochester.
- [6] E. Steinbach, P. Eisert, B. Girod, A. Betz. (2000). "3D Object Reconstruction Using Spatially Extended Voxels and Multi-hypothesis Voxel Coloring", ICPR.
- [7] R. Tsai. (1987). A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. IEEE Journal of Robotics and Automation, RA-3(4): pg 323 - 344.
- [8] K. Kutulakos. (2000). "Approximate N-View Stereo", 6th European Conference on Computer Vision, pg 67-83.



**Figure 3. Several reconstructions using the histogram consistency check. The left image shows the camera view and the right is the reconstruction.**