An Empirical Comparison of Progressive and Wavelet Radiosity - Summary and Critical Review

Erik Cedheim, Ramzi Ferchichi, Dan Lind, Henrik Nyman, Linus Lundberg

Department of Information Technology, Uppsala University
P.O. Box 337, S-751 05 Uppsala, Sweden.
Email: {erce4281,rafe5725,dali7481,heny2875}@student.it.uu.se

1 Introduction

This is a summary of *An Empirical Comparison of Progressive and Wavelet Radiosity* by Andrew J. Willmott and Paul S. Heckbert.

The paper is about a comparison between two of many methods for calculating radiosity in a scene. These two methods are *progressive radiosity* and *wavelet radiosity*, both algorithms are inherently iterative so you get the convergence of the algorithms visualised. The algorithms were implemented from scratch in a common framework so that no algorithm would benefit from differences in coding-style, compiler or computer.

To save time and keep the project manageable not all algorithmic options were considered. Discontinuity meshing, importance-driven radiosity and distribution raytracing for example were ignored. Besides the paper summarised here the authors has also produced a broader study where 3 different radiosity methods across seven test scenes with on average 6 different settings.

As mentioned above two different algorithms were tested progressive radiosity and wavelet radiosity. The progressive radiosity iteratively calculates one column of the form factor matrix at a time. This method can be made adaptive by the introduction of a two-level mesh where the coarser patches shoot light to a finer set of adaptively refined elements. This is known as *sub structuring*. In wavelet radiosity the choice of basis function and the authors of the report chose to include five of the most commonly discussed in their tests: The Haar basis, flatlets of order 2 and 3 (F2, F3) and multi wavelets also of order 2 and 3 (M2, M3).

Progressive radiosity is performed until the residual error drops below a preset faction of the original residual error. Wavelet radiosity is performed until no more refinement is needed on the *links*.

When meshing in progressive radiosity the patches are subdivided into meshes where each edge is shorter than a preset maximum. For wavelet radiosity with Haar basis as well as progressive radiosity, post-process smoothing is applied. In both multiwavelets and flatlet the authors render directly from the basis functions.

Formfactors were calculated by approximating the formfactor by its point to point counterpart which works well as long as the two patches are sufficiently far apart. To avoid inaccuracies at close distances the authors use polygon-to-point calculations when the patches are too close together.

Three tests where performed to test:

- *Inter-reflection*, a tube with a lightsource in one end tests the algorithm's ability to handle inter-reflection between long chains of patches.
- *occlusion*, tests the algorithm's ability to cast soft to very hard shadows by moving an occluder between a surface and a receiver.
- *complexity*, a scene with a varying amount of chairs each chair has 27 polygons and the room consist of 5 polygons.

The tests were performed by inserting checkpoints in the algorithm. At each checkpoint the algorithm writes the entire scene rather than the image produced. To measure the error in a scene the authors had to pick some sort of reference model. Where analytical solutions were possible it was used otherwise the matrix radiosity method was used with analytical patch-to-point form factors and a mesh at least four times as big as any other implementation.

Since the radiosity is view-independent a view-independent method of error measurement was used. Since the compared tests share the same set of base polygons the error was set to the difference of radiosity of each polygon and then summed up to an area-weighted mean value.

The results are displayed in plots. For each run the resulting values are represented as *time equalised* plots. To produce time equalised plots all methods are rolled back to the time of the first converged method. In a time equalised plot each method is represented with the error of the first converged value over the parameter to the method.

2 Results

2.1 Time and Error

In the end of the error-measurement tests, when the higher-order methods ran out of memory, the wavelet methods showed a trend toward lower errors, whereas the trend for the progressive methods were toward higher errors. Haar did best of the wavelet methods because the M2 basis have poor visibility handling and the M3 basis is hampered by it's extra time and space overhead.

The convergence time of the progressive methods increased quadratically whereas the convergence time of the wavelet methods were linear with scene complexity. The difference was that as the complexity increased, the time which the progressive methods spent shooting light from the walls of the scene increased, which were not the case with the wavelet methods as they used almost a constant number of links between a wall and other scene parts.

Note that for the unscaled complex scene, where room size did not increase with the number of chairs, the progressive radiosity method outperformed the hierarchical radiosity.

2.2 Wavelet Memory Use

The experiments has shown that wavelet radiosity methods consume excessive memory, especially compared to the memory consumption of the progressive algorithms. In a particular run, the M3 basis reached 550Mb while the progressive method with sub-structuring only reached 1Mb. Clustering has been shown to solve this issue but was not implemented for these experiments.

2.3 Empirical Complexity of Wavelet Radiosity

Strong effects on time and memory use in the experiments was observed for the wavelet radiosity methods. The most interesting of these relations are the following:

- $\#links = \sigma k^2 + \beta k$
  In the wavelet radiosity methods, the number of initial links is σk^2, where k is the number of polygons in the scene and σ is the average fraction of scene polygons visible from a particular polygon. The term βk is the number of links created during refinement. For the Haar wavelet method the extra links term dominated, yielding linear time cost, while the initial links term was dominating for higher order methods which resulted in near-quadratic time cost.
- $\#links = \theta(1/epsilon)$
  The number of links generated by the wavelet methods for a fixed scene is proportional to 1/epsilon for small enough epsilon. As the total number of links approaches the minimum, σk^2, when epsilon is increased, the relationship becomes sublinear and approaches O(1). This relationship is very useful when setting epsilon for wavelet radiosity because halving epsilon will double the time and memory cost of the algorithm.

2.4 Progressive and Hierarchical Shadow Handling

Simple tests of shadow handling were done using the blocker scene. The experiments shows that sub-structuring achieves the lowest error when the blocker is closer to receiver while hierarchical radiosity achieved the lowest error when the blocker approaches the light source. Although hierarchical radiosity achieved the best result when shadows are softer it took much longer to complete. This is due to the fact that hierarchical radiosity subdivides both the receiver and the light source, whereas sub-structuring only subdivides the receiver.

2.5 Wavelet Shadow Handling

Two methods were tested for wavelet shadow handling, *fractional visibility*, which scales the result of quadrature with estimated visibility between two patches, and *visibility-in-quadrature*, which leaves visibility in the integrand. None of the methods performed well the complex experiment. The M2 basis with fractional-visibility method resulted in overly blocky shadows, and when switching to the visibility-in-quadrature method the shadow became smoother but still suffered from large discontinuities between elements.

The problems with the fractional-visibility approach is that it assumes constant visibility across the element even when the radiosity is no longer constant with the higher-order methods. The visibility-in-quadrature method handled variation of visibility across an element better but suffered from aliasing problems as it sampled visibility from a 2x2 grid rather than 4x4 grid.

2.6 Reflectance

The tube experiment showed how well the radiosity methods perfomed with high reflectance scenes. Substructing were slow and the Haar had large errors for high reflectance values. The experiment hightlights how close-proximity surfaces are a problem for wavelet methods because they create many links in these cases which slows them down markedly and significant accumulated projection error produces bright strips of light which worsen as the link hierarchy deepens. Several solutions to the projection errors exists but were not tested.

The sub-structuring method's problems were caused by *overrefinement*, which occurs with high reflectance scenes and lights that touch another surface at a reflex corner. The multigrid approach can solve this problem by first illuminating the scene with a coarse mesh with sub-structuring turned off and when the sub-structuring method is run using this mesh as a starting point, patches that are shot early will not overrefine areas where the final radiosity gradient is smooth.

3 Conclusions

The main conclusions are:

- **Wavelet radiosity using the Haar bassis is often best for moderatly complex scenes.**
  Haar was best of the tested methods for the 'complex' experiment. It's time cost grew linearly while the progressive radiosity with substructuring grew quadaratically. Haar often produced the most accurate result compared to the reference method.

- **Wavelet methods consume excessive memory.**
  For very complex scenes the wavelet methods became impractical because their memory requirement grew quadratically with scene complexity. Wavelet methods of order M use $O(M^2)$ memory per element and $O(M^4)$ memory per link. **Substructuring is a better choice when the scene is too complex for wavelet methods**.

- **Higher order wavelet methods has poor shadow handling.**
  Shadows are very blocky when using the fractional visibility method or the visibility-in-quadrature methods unless one uses heavy subdivision. Currently the higher-order wavelets handle visibily poorly which prevents them from achieving higher accuracy.

- **Substructuring is recommended for hard shadows.**
  On the blocker experiment it was observed that substructuring is most accurate method when a shadow contains an umbra. Wavelet methods typically refine less than substructuring near shadow boundaries and therefore their shadows appear blurrier and blockier.

All conclusions are based on the experiment scenes and on one specific implementation. The use of clustering and other advanced techniques are expected to alter these conclusions.

4 Critical Review

An Empirical Comparison of Progressive and Wavelet Radiosity is a very ambitious project that have taken the authors 2 years to complete. You can tell that they are proud of it as the report has many mentions of the size of their code and resulting data.

It is nice to see an empirical comparison of algorithms in a field that usually only has theoretical comparisons. This is also, according to the authors, the reason for writing the paper as they feel that theoretical results are not enough and things like memory usage and constant time does matter in a real life implementation. This is something we agree with.

The authors have implemented all the algorithms themselves and motivate this with that they feel that code written by different people can greatly affect how well it performs. They also wanted a common framework to base all their different radiosity methods on so that it was the methods efficiency and accuracy that was tested and not the implementation of the algorithms.

The paper we have been given to review is in fact just a short summary of a much larger paper by the authors. This is probably why there are many things that are left unexplained in the paper we got, and why statements are made without proper proof. This is generally not the case in the larger paper where the authors have taken their time to explain things in much greater detail. Perhaps this is unavoidable when trying to compress a 90 page paper to 10 pages, but we still feel that the short paper could have been structured in a better way. At the moment its really only useful for people who already have a good knowledge of the subject. A very good point with the report is that the authors have taken care to always note down when they are deviating from known methods.

Since the large report is so much longer we are here reviewing mainly the short report (although, we have read the longer report to get a better understanding of some things).

4.1 Error Measurement

In the paper they use a view-independent error measure. This is done by calculating the intensity of each patch for all the different radiosity methods and comparing it to a reference solution patch by patch. This reference was where possible calculated using analytical methods, or with the matrix radiosity method using a high detailed mesh.

This approach has both its advantages and disadvantages. On the positive side this method will give an error that is easy to compare and measure. The problem is that a view independent error measure might not really reflect how good the image will look to the user (which is usually the main concern in computer graphics). As an example of this, a methods inability to produce hard shadow will be heavy penalized since there will be a great difference between a very dark patch and a light patch. A much greater difference than if we were comparing two method's ability to do soft shadows against each other.

Another concern is the choice of reference solutions. In one of the more complex scenes the matrix radiosity method proved too memory intensive for the test platform and the progressive radiosity method was used instead - the very same method that they later tested against the reference. It would seem possible that because of this progressive radiosity got an unfair advantage in this part of the tests.

4.2 Conclusions of the Paper

While the conclusions made in the paper are sound they depend heavily on the authors definition of a "complex scene". They start out with a simple scene: a few chairs in a room, and then increase complexity by inserting more chairs. What they also do is increase the size of the room along with the number of chairs. This means that the distance between chairs stays the same instead of filling the space with more and more chairs.

Doing this instead of keeping the same room and inserting more objects results in dramatically different results than if the earlier approach would have been used, even though the both methods are dealing with the same number of polygons (usually otherwise a good measure of scene complexity). In the first method (expanding room size as complexity increases) wavelet methods perform much faster than progressive radiosity. If on the other hand the authors had just put more chairs into the same size room progressive radiosity would have been fastest. It is interesting to note how important such a simple definition such as "complex scene" has on the results.

4.3 The Paper's Validity Today

The paper is from 1997, but we feel that it is still valid and interesting now. However, we are not sure the results from the experiments would have been the same on todays machines. The authors have problems with many of the algorithms being too memory hungry. The wavelet methods in particular suffer from this. The thing is that the authors used a computer with 128 MB of primary memory and must have had to do a lot of time-consuming page swapping to disk in the memory hungrier methods. Nowadays large amounts RAM is not very expensive and we think that the wavelet methods would have gotten a decent performance boost compared to the other methods on a computer with more memory.