EVALUATION OF IMAGE QUALITY USING NEURAL NETWORKS

R.J. CANT AND A.R. COOK

Simulation and Modelling Group, Department of Computing,
The Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK.
Tel: +44-115-848 2613. Fax: +44-115-9486518. E-mail: rcc@doc.ntu.ac.uk.

ABSTRACT

We discuss the use of neural networks as an evaluation tool for the graphical algorithms to be used in training simulator visual systems. The results of a number of trial evaluation studies are presented in summary form, including the 'calibration' exercises that have been performed using human subjects.

INTRODUCTION

Training simulators, such as those employed in vehicle and aircraft simulation, have been used for an ever-widening range of applications in recent years. Visual displays form a major part of virtually all of these systems, most of which are used to simulate normal optical views, but many (particularly in military training simulators) are used to present infra-red, sonar or radar information. The increasing diversity of applications can often make the design of new systems difficult, because there is a lack of previous experience on which to base decisions.

In graphics-based simulators, the way that images are presented is critical to the effectiveness of the system. Therefore, to achieve maximum effectiveness, the design of the graphics system which generates the displays should be determined by the requirements of the application, rather than just by what is already available commercially. Unfortunately, these requirements are initially defined in terms of the utility of the images for performing a given task and this does not translate directly into a graphics system specification.

The broad outlines of the specification may be obvious, but analysis rather than 'common sense' is needed to quantify the boundary conditions. To get the detail of the requirements into a usable form, the system really needs to be built and then tested on suitable human subjects. Whilst such research is feasible for mass-market applications, it is impractical and costly for one-off projects and cannot be repeated every time a system is modified or upgraded. This paper reviews the use of neural networks to solve this evaluation problem.

The advantage of using neural networks is that they are able to give a task-specific evaluation of an algorithm. In a practical application, a specific training and testing regime would be constructed. This could then be used to decide on the appropriate algorithm for the application. In order that this methodology itself can be tested, a number of algorithms are implemented and several different 'tasks' are defined. For each task/algorithm combination, a neural network is trained and tested. These results provide an evaluation of the algorithm for that task. Because the purpose of this paper is to investigate the neural network evaluation methodology itself, not the algorithms, the techniques chosen include some simplistic ones that have well understood deficiencies. The ability to detect these deficiencies provides a good test of the methodology.

To ensure that the results of the neural network provide a good correspondence with human performance, we have also performed similar tests on human subjects and than compared these results with the ones obtained from the neural network. The degree of correspondence between human and neural network seems to depend on the level of expertise of the human subject.

The neural network used is of a type developed by the authors and is described in detail in [Cant and Cook, 1995]. It is a three-layer, feed-forward network customised to allow both unsupervised competitive learning and supervised back-propagation training. It has been found that in some cases the results of competitive learning alone show a better correspondence with human performance than those of back-propagation. The problem here is that the back-propagation results are too good.

The following areas of investigation are presented here: resolution and sampling, shading algorithms, texture anti-aliasing techniques, and model complexity.

RESOLUTION AND SAMPLING

Many training simulators, such as periscope simulators, anti-aircraft missile simulators, and air traffic control simulators, are performance-limited by the resolution of their displays. To match the resolution of the human eye, these displays would need to be replaced with ones having over 5,000 scan lines. Unfortunately, this would not only increase the requirements of frame buffer memory and processing during rasterisation, but would also increase beyond reach the scanout bandwidth needed to refresh the display; more advanced and greatly more expensive scanout technology would have to be developed. Even with such high resolution, aliasing phenomena would still be detectable if the images were not properly sampled. Since it is not possible to match the eye's resolving power economically, training programmes must be designed around the capabilities of simulators as they are. Typically, display resolution and sampling accuracy affect training tasks such as the detection, recognition and attitude determination of vehicles at long range.

Resolution Effectiveness

This study concerned the recognition and attitude determination of three types of fighter aircraft when drawn at low resolution. The aircraft that were used are models of the Harrier, Mig 27 and F-16 that were actually employed in periscope simulators developed by Ferranti Simulation and Training during the 1980s. They are shown in Figure 1. The objectives of the study were to determine: (i) what happens to image quality when the resolution of the image is near its limit; (ii) how 'basic' (i.e. non-anti-aliased) images compare with anti-aliased images; and (iii) what the differences are between human and neural network perception. In achieving these objectives, image quality was first evaluated using neural networks. This was then compared with the evaluation of humans, as detailed in the following sections.

Neural Network Evaluation

Figure 1 Harrier, Mig 27 and F-16 aircraft models

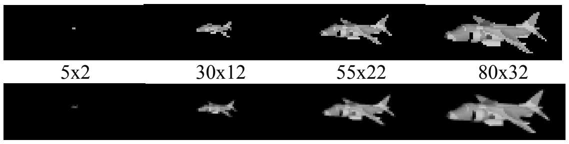


Figure 2 Basic (top) and anti-aliased (bottom) images at a range of resolutions

The neural network evaluation involved training the image analysis system to categorise images of the three aircraft according to type and orientation over a series of low resolutions. The quality of the images over the range of resolutions chosen is illustrated by Figure 2. For each resolution, the image analysis system was tested to determine how well it had learnt (i.e. how well the images could be categorised from the information they contained). By plotting network performance against image resolution, the effect of resolution on image quality could then be determined. To see if anti-aliasing provided any significant advantage, the evaluation was carried out first with basic images and then with anti-aliased images.

During the evaluation, the image analysis system was configured as follows:

Output Categories. The network's output layer was set up with 24 output categories in a two-dimensional arrangement: the three types of aircraft 'Harrier', 'Mig 27' and 'F-16', each with the general heading directions '0°', '45°', '90°', '135°', '180°', '225°', '270°' and '315°'. To match the dimensionality of the output layer, a two-dimensional hidden layer was set up with 6 nodes across by 8 nodes down giving a total of 48 hidden nodes.

Image Generation. Gouraud-shaded images of the aircraft were generated using ambient and diffuse illumination only with a directional light source 45° from the right. For each image, the type of aircraft was selected at random, as was its heading (any of 360° in yaw angle accurate to 0.01° with up to ±15° roll). In order to evaluate particular resolutions, images were generated with specific numbers of pixels horizontally and vertically. The aircraft models were suitably scaled to ensure that they would always fit into this width and height for all the required orientations. The width-height ratio was always 5 to 2.

Training and Testing. The image analysis system was trained to categorise images of the three aircraft by presenting it with 115,000 randomly-generated images. When tested, the image analysis system's task for each of a further 10,000 randomly-generated images was to identify the type of aircraft and to determine its heading to the nearest 45° as per the output categories.

Rather than training a separate network for each resolution, a single 'combined' network was used for which images of random resolution were generated and scaled up to fill out the input layer. Two independent combined networks were employed: one for basic images and one for anti-aliased images. This meant that the network could not simply memorise the training data but had to generalise between resolutions. Furthermore, instead of training the networks specifically on the discrete resolutions 5 by 2, 10 by 4, etc., an almost continuous range of resolutions was used, implemented by using steps of less than one hundredth of a pixel. However, the networks were still tested on the integer resolutions to simplify the statistical analysis and comparison with human tests.

Neural network performance against image resolution for the basic and anti-aliased images is presented in Figures 3 and 4. The graphs show that, for the lower resolutions, the image quality increases with resolution and anti-aliasing has a clear advantage. However, the relative improvements in image quality diminish rapidly when the resolution exceeds 20 by 8 pixels.

Comparing Human Evaluation

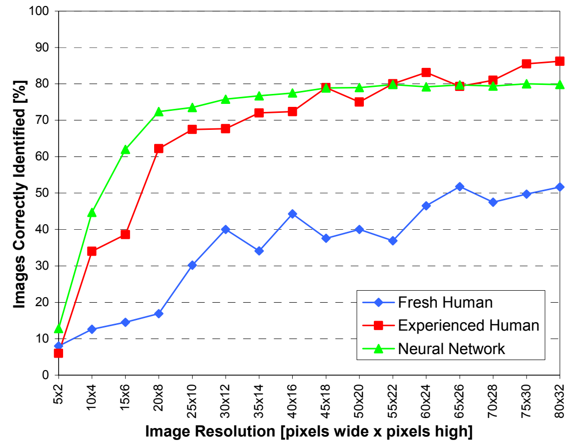


Figure 3 Neural network and human results for basic images

Figure 4 Neural network and human results for anti-aliased images

In order to compare human performance with network performance, a corresponding human test was devised in which the human subject was presented with images of aircraft similar to those presented to the combined networks. Each test was divided into two halves: a series of 50 'test' images, followed by a series of 50 'control' images:

Part 1: Test Images. The test images were presented on a 640 by 480 pixel display. Just as for the combined networks, the aircraft type, orientation and image resolution, ranging between 5 by 2 and 80 by 32 pixels, were chosen randomly. For each test, the images were either all basic non-anti-aliased images or all 16-sample-per-pixel anti-aliased images, selected at random. (A random mixture of basic and anti-aliased images was originally used, but the results indicated that the human subjects were distracted by the change in image type.)

Part 2: Control Images. The control images were always presented in anti-aliased form on a 1,024 by 768 pixel display. For each test, the set of control images was equivalent to the set of test images, but the order in which they were presented was scrambled (needless to say, the human subject was not made aware of this). The resolutions (and perspective) of the aircraft were scaled so that each control image was rendered to appear physically the same size as it did during the test images, thus ranging from 8 by 3.2 to 128 by 51.2 pixels. The control images were included in the test mainly to allow the results of the test images to be verified, but the higher resolution, anti-aliased control images also provided the opportunity to see if any 'training transfer' occurred from the lower resolution, basic or anti-aliased test images.

For all images, the human subject had to use the mouse to select the most appropriate of 24 image categories (equivalent to the network's output categories). The test was set to a suitable level (such that scoring as low as 0% or as high as 100% was possible but unlikely). The test was used to measure the performance of: (i) a large number of 'fresh' people (around 120), each having taken the test just once or twice, who were neither familiar with the test itself, nor the aircraft, nor analysing computer-generated images of three-dimensional objects; and (ii) a much smaller number of 'experienced' people (two in particular), each having sat the test many times over, who were familiar with the test, the aircraft, and the analysis of 3D scenes. Approximately 150 results files were collected for each of these two groups of human subjects.

Results and Analysis

For each of the two groups of human subjects (i.e. fresh and experienced), the set of results files collected were processed to provide overall performances for the human at each of the 16 different image resolutions. Graphs showing the overall performances of fresh human, experienced human and neural network against image resolution, for basic and anti-aliased test images, are presented in Figures 3 and 4. For both human and network, the scores for the anti-aliased test images were generally better than those for the basic test images, and the scores for the higher resolution, anti-aliased control images (not plotted) were better than those for the anti-aliased test images, although to a lesser extent. For some people, there seemed to be a positive influence when the control images in the human test were preceded by anti-aliased test images (indicating that lower resolution observation could be transferred to the corresponding higher resolution images).

The shape of the curves show that network performance is very similar to experienced human performance. The individual test scores for the experienced subjects fell consistently around the 70% to 80% mark whereas the overall network scores were only about 5% to 6% higher. The network performed much better at the lower resolutions than the human, but marginally worse at the higher resolutions. This was due to the human's bias towards the higher resolution images (relating more readily to the real world) as opposed to the network's objective approach. In contrast, network performance is substantially better than fresh human performance. The individual test scores for the fresh humans varied wildly between 10% and 60%, and there was very little coherence between the results at individual resolutions from one person to the next. It is clear that the quality of the images presented is what underlies both human performance and the learning ability and subsequent performance of the network. The fact that the network scored significantly better than the human at low resolution shows that information to correctly classify the type and orientation of the aircraft was present in the images, and was therefore potentially available to the human. Furthermore, it was found that the network was confused in a similar way to the human. For example, it was sometimes difficult to tell whether the aircraft were pointing towards or away from you. However, as might be expected, it was near the boundaries between the direction categories where the majority of misclassifications took place.

SHADING ALGORITHMS

In the real world, the way surfaces are shaded depends on the position, orientation and characteristics of the surfaces and the light sources illuminating them. Shading provides essential depth cues to the 3D structure of objects as well as indicating the materials from which the objects are made. During the evolution of 3D computer graphics, to allow objects to be represented realistically in computer-generated images, a number of illumination, reflection and shading techniques have been developed. So far, however, only the simpler methods have been applied to real-time systems. Space does not permit the detailed description of these algorithms in the present paper. The reader is referred to any good textbook on computer graphics for details, for example [Foley et al., 1990].

Algorithm Effectiveness

There are three shading algorithms which are commonly used in real-time graphics. In ascending order of complexity they are: flat shading, which evaluates the required reflection model only once per polygon; Gouraud shading, which evaluates it at a comparable once per vertex but approximates the evaluation at each pixel; and Phong shading, which evaluates the reflection model fully at each pixel. Phong shading itself can be implemented in more than one way, the most critical factor being the normalisation of the interpolated vectors. The study was designed to determine how well these algorithms really represent shading information, and whether the algorithms are helped or hindered by the inclusion of specular highlights.

In the following studies, image quality or shading quality is based on the determination of lighting direction. When generating images of objects using a particular shading technique, the proportion of images that can be correctly categorised is used as a measure of how accurately the direction of illumination is represented in the images. In this study, the direction of illumination was determined from images of aircraft. The same three aircraft models as the models employed in the resolution and sampling study were used. The objectives of the study were to determine: (i) what the differences are in image quality between the different shading algorithms; (ii) how including specular highlights compares with diffuse-only images; (iii) how important the normalisation operations in Phong shading actually are; and (iv) how neural network perception compares with human perception. The same approach was taken here as for the resolution and sampling studies, the neural network evaluation being benchmarked against a study using human subjects. These results are also presented in [Cook and Cant, 1996].

Neural Network Evaluation

In the neural network part of the study, the algorithms flat shading, Gouraud shading, 'unnormalised' Phong shading, and full Phong shading were used to train eight separate neural networks. Four networks were trained with 'diffuse' images in which diffuse reflection only was used with each of the algorithms, and four were trained with 'specular' images in which both diffuse and specular reflection were used. The networks were trained to categorise which direction the light was coming from. Each network was then tested to determine how well the lighting direction had been represented by the images. By comparing the relative performances of the eight networks, the relative performances of the different shading algorithms, and the effect of including specular highlights, could be determined.

During the evaluation, the image analysis system had the following configuration:

Output Categories. For this study, the lighting direction was divided into 15 categories: five across (225° split into five 45°-wide sectors) by three down (90° split into three 30°-high sectors). The output layer of each neural network was set up in a corresponding two-dimensional 5 by 3 arrangement, and each network's hidden layer was set up with 30 hidden nodes in a two-dimensional 5 by 6 arrangement.

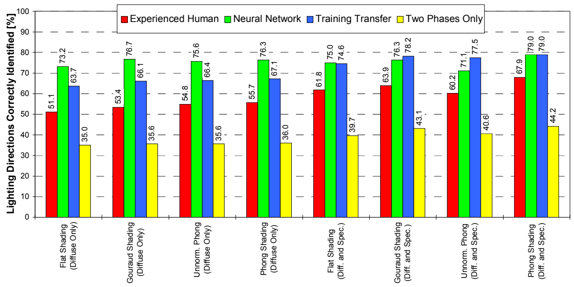


Figure 5 Results for shading evaluation

Image Generation. Anti-aliased images of the aircraft were generated at a resolution of 80 by 32 pixels using 16 supersamples per pixel. For each image, the lighting direction was chosen randomly (any angle up to 112.5° to the left or right accurate to 0.01°, and any angle up to 45° up or down). The type of aircraft was also chosen randomly, but the aircraft were always drawn pointing to the right.

Training and Testing. Each neural network was trained to categorise the illumination direction of aircraft by presenting them with 115,000 randomly-generated images. For network testing, for each of a further 10,000 randomly-generated images, the task was to categories the direction of illumination horizontally and vertically according to the 15 output categories.

Comparing Human Evaluation

In the corresponding human test, the human subject was presented with images of aircraft similar to those presented to the neural networks. Again, each test was split into two parts: a series of 50 diffuse images, followed by a series of 50 specular images. The shading algorithm for each aircraft in the human test was selected randomly. The images were always anti-aliased but drawn at a resolution of 320 by 128 pixels, four times those presented to the neural networks.

The subject had to use the mouse to select the most appropriate of 15 lighting directions which corresponded to the network's output categories. In this case, only experienced human subjects were used because of the greater difficulty of the task.

Results and Analysis

The human and neural network performances for all shading techniques are presented in Figure 5. The level of performance reached by the experienced human was substantially lower than that of the back-propagation neural network, illustrating the difficulty of the exercise. Pure competitive learning, which results from the first two stages of training, are therefore presented to supplement the fully trained network data. The results for the human show a steady improvement as the shading algorithm becomes more refined, with a significant rise in performance when specular reflection is introduced. The only exception to this trend is the relatively poor performance when unnormalised Phong shading was used.

The neural network results objectively show that all the shading techniques performed to a similar level and are therefore all more or less equally effective. When the reflection model consisted only of diffuse reflection, Gouraud shading, unnormalised Phong shading and full Phong shading were all approximately equal, and were all marginally better than flat shading. When specular reflection was included, the network results show that there were small improvements for flat shading and full Phong shading, no significant difference for Gouraud shading, and a clear drop in performance for unnormalised Phong shading. However, these results indicate only that the different shading techniques provide similar amounts of information about lighting direction, not that the information is presented in a form suitable for human use. To overcome this difficulty, a 'training transfer' exercise was added.

As a training transfer exercise, the network that was trained using the most realistic of the shading techniques (full Phong shading with diffuse and specular reflection) was tested on images rendered using the less realistic shading techniques. These results are also presented in Figure 5. The relative differences between them show a strong resemblance to the human results and therefore shows that such a training transfer exercise is a valid measure of performance. This exercise was a valid measure of performance because, in a similar way to the human, the network used experience of realistic shading to categorise images of simpler shading.

Summary

The study presented here clearly shows that, for both network and human, the more sophisticated (and more expensive) shading algorithms are generally more effective and including specular reflection is generally better still. The relatively poor results for specular Gouraud and unnormalised Phong shading are most likely caused by the specular highlights being confined to the vertices of the polygons. Furthermore, the reduced network result for unnormalised Phong shading shows that such inaccurate highlights can actually cause some confusion. To improve the performance of unnormalised Phong or Gouraud shading, objects would need to be made up of a larger number of smaller polygons to enable the highlights to be positioned more accurately.

MODELLING COMPLEXITY

Much of the task of creating effective computer-generated images lies in how the objects in the environment are modelled. However, deciding how and to what level of detail the objects should be modelled is often highly subjective. Objects should be defined such that they are appropriate for the particular application. For example, a military simulator may need trees and bushes to be represented with enough detail that camouflaged tanks and other vehicles can be hidden amongst them. In contrast, trees and bushes in a flight simulator may be used to provide basic visual cues, such as position and orientation, and can thus be represented much more simply. In other applications, object recognition may be the fundamental requirement. Missile crews may need to distinguish between friendly and enemy aircraft, for example. When the aircraft are modelled, attention must therefore be paid to the particular features of the aircraft that make them distinctive.

Object Modelling

In order that storage and processing requirements are minimised, objects are usually modelled with simplified geometry, and small details such as bolts, grooves, and surface roughness are often represented with texture. Graphics systems often employ some form of detail elision. When distant objects are drawn, detail elision allows the processing requirements to be decreased but at the expense of increased storage. In detail elision, each object has several models, each with a different level of detail. The model with the most detail is used when the object is close to the observer while the less detailed models are selected when the object is farther away. Ideally, the level of detail would be chosen so that

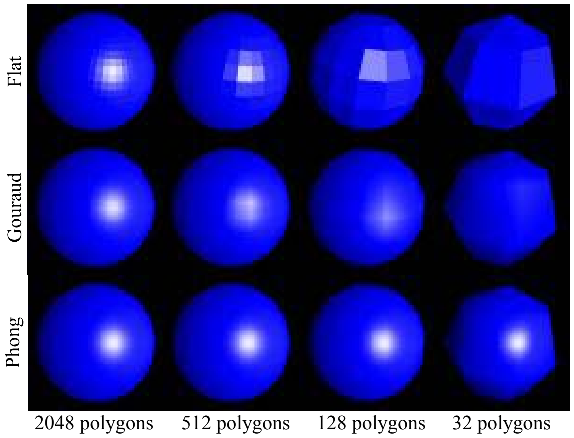


Figure 6 Modelling a sphere with different numbers of polygons

the loss in detail from fewer polygons would coincide with the loss in detail from the display's resolution.

Modelling Effectiveness

Essentially, the level of detail to which objects are modelled can affect how well the objects are recognised. In addition, because Gouraud and Phong shading is based on the vertices of polygons, the way objects are modelled can also affect how well they are shaded, especially where curved surfaces and specular highlights are concerned.

Figure 6 shows how a sphere can be represented using different numbers of sides. It also demonstrates how modelling complexity can affect the quality of shading. A sphere has been modelled with different numbers of polygons and rendered using flat, Gouraud and Phong shading. Diffuse and specular reflection have been used in all cases. Flat shading allows the sphere's curved surface and specular highlight to be represented reasonably well only when a very large number of polygons are used. Whilst the impression of surface curvature can be improved with Gouraud shading, an effective highlight still requires a large number of polygons. Phong shading, however, can simulate the specular highlight well no matter how few polygons are used. To back up the observations described above, experiments were performed using the image analysis system. In a similar way to the previous study, the neural network was trained and tested on the direction of illumination. Between 5 and 10 output categories across by the same number of categories down were employed. The lighting algorithm combined diffuse reflection with specular reflection. For each image, the sphere was randomly orientated and presented in anti-aliased form. The experiments allowed a curve to be plotted for each shading algorithm. Each graph shows how well lighting direction is represented for particular modelling complexities of the sphere. The results shown in Figure 7 relate to a 'training transfer' exercise similar to the one performed in the previous study. The results for training and testing on the same algorithm are qualitatively similar but slightly less differentiated.

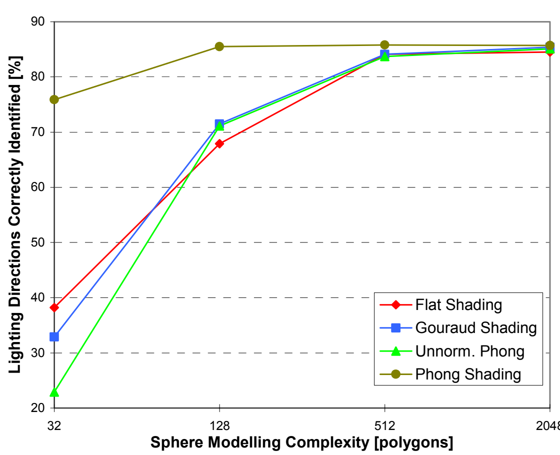


Figure 7 Results of modelling complexity evaluation

Results and Analysis

The results show clearly that there is a critical cutoff point where the performance falls off rapidly. In a realistic simulator application we would expect that the cutoff point would vary from object to object and from task to task. The neural network technique would be a convenient way to discover where this critical point is and thus to avoid waste that would result from using an unnecessarily large number of polygons.

An interesting side issue is the way in which Phong shading removes the cutoff, allowing a much smaller number of polygons to suffice. This emphasises the way in which a change of rendering algorithm can reduce the number of polygons needed. The fact that the neural network can detect this so easily shows the potential power that this method has.

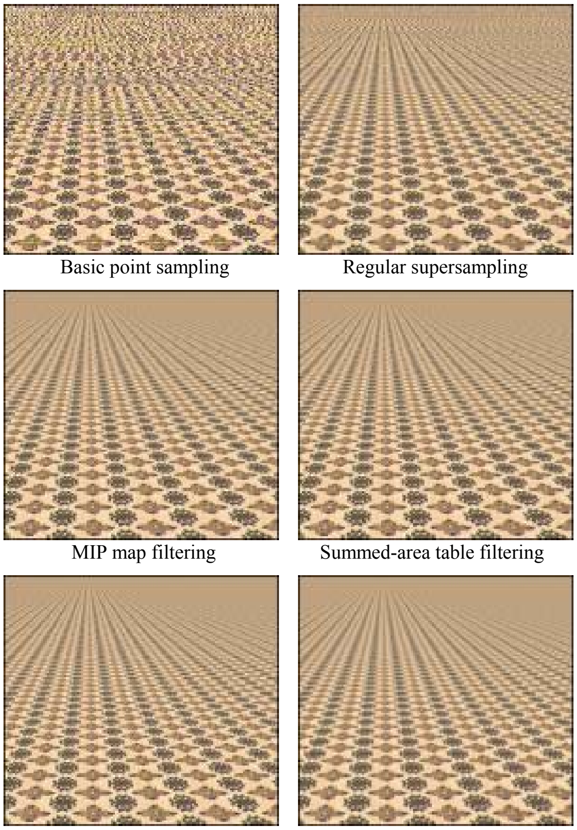
TEXTURE ANTI-ALIASING

The scenes generated by early real-time graphics systems contained surfaces that were plain and simple. These surfaces were not only unrealistic and visually uninteresting, but also lacked essential cues for depth, orientation and motion—cues which were unavailable to pilots training on flight simulators, for example, when flying at low altitude. The demand for greater realism and the availability of greater processing power, however, encouraged a number of texture mapping algorithms to be developed. Texture mapping has now become a vital component in most real-time graphics systems. Unfortunately, mapping texture onto surfaces accurately requires significant computing power and simplified algorithms designed for real-time performance can often suffer from severe aliasing or excessive blurring. The range of algorithms that can be used are described in [Williams, 1983; Crow, 1984; Glassner, 1986; Heckbert, 1988; Schilling et al., 1996; Cant and Shrubsole, 1997]. Images that were created using a selection of these algorithms are shown in Figure 8.

In this study, the real-time texture mapping algorithms—basic point sampling, regular supersampling, MIP map filtering, summed-area table filtering, and potential map filtering—are evaluated. To provide an 'ideal' comparison, it was also necessary to implement a very high quality algorithm that used Gaussian filters. This algorithm is not at present even remotely feasible for real-time implementation but provides a useful benchmark. The study is also presented in [Cook, 1998].

Neural Network Evaluation

The effectiveness of the texture mapping algorithms is based on how well orientation information is represented by a textured plane at a number of different depths. During evaluation, texture with aliasing or blurring problems was expected to cause confusion and was therefore expected to perform worse than texture of higher



Basic point sampling — Regular supersampling

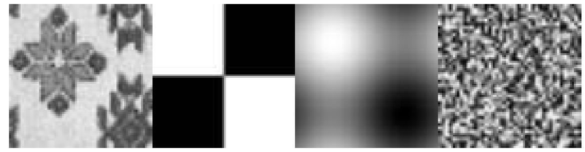
MIP map filtering — Summed-area table filtering

Potential map filtering — Gaussian-weighted filtering

Figure 8 Different algorithms for texture anti-aliasing

quality. The evaluation study had two parts: an objective 'own algorithm' evaluation in which neural networks were tested on the same texture mapping algorithms on which they were trained, and a 'training transfer' evaluation in which neural networks were trained on the 'ideal' brute force Gaussian-weighted algorithm and tested on each of the real-time algorithms.

During the study, images of texture were generated at different orientations over a range of different depths. For each image, the texture was mapped onto a finite plane. Variations in horizontal and vertical orientation were achieved with rotations about the z and x axes, and different depths of the texture were simulated by modifying the scaling of the texture across the plane. A separate neural network was trained and tested for each texture mapping algorithm at each depth of the texture. Each network was defined with a 10 by 10 node input layer. Each output layer was set up with 36 output categories: nine 5°-wide categories horizontally ranging from 22.5° to 67.5° for z angle by four 5°-wide categories vertically ranging from 65° to 85° for x angle. The hidden layers were configured as 9 nodes by 8. As input, texture images 10 pixels square were generated at the centre of projection and transferred directly to the network's input layer.
Each network was trained using a total of 32,500 images. In order to test the network, a further 5,000 texture orientations were randomly generated. The network's task was to determine which of the 36 output categories the orientation belonged. The study was separately performed using four texture maps: first with a 41 by 41 wallpaper texture, then with a 'high frequency' 40 by 40 checkered texture, then with a 'low frequency' 40 by 40 sinusoidal



Wallpaper — Checkered — Sinusoidal — Noise

Figure 9 Selection of textures used in the study

texture, and finally with a 40 by 40 noise texture, as shown in Figure 9.

Graphs showing a training transfer evaluation of each of the real-time algorithms against the Gaussian-weighted algorithm are presented in Figures 10 to 13. For completeness, the Gaussian-weighted algorithm's own performances (against itself) are also included. For each algorithm, the percentage of how well the network could deduce the orientation of the plane from the texture is plotted against a number of different depths of the texture. The graphs show that the differences in algorithm performance are clearly detectable, as are the differences in behaviour between the different texture maps.

Note that the lack of high frequencies in the sinusoidal texture makes the algorithm performance indistinguishable. In contrast, the best indication of the relative differences in performance between the different texture mapping algorithms is provided with the noise texture. This is because the noise texture has no repeating features and the only way the orientation of the textured plane could be determined accurately was if the texture was mapped accurately. In order of increasing quality, the algorithms are basic point sampling, regular supersampling, then summed-area table filtering and MIP map filtering and, best of all, potential map filtering. This trend is also generally true for the other texture patterns.

Summary

This section has shown that neural network analysis can provide useful information about how the relative performances of the real-time texture mapping algorithms can be affected. As well as by the effects of aliasing and blurring, the algorithms can also be affected by the actual texture.

CONCLUSIONS

The studies presented here show that neural network analysis in general gives the result that might have been expected intuitively. This is reassuring, although the superficial conclusion might be that such analysis can be left to experience and educated guesswork. There are, however, one or two slight surprises. Although, on reflection, these can be understood, they might easily be missed had less thorough methods been relied on.

The full potential of the method can only be realised when it is applied to real life requirements analysis. The facility to define the limits of performance in such a case, and to flush out any subtle problems, should prove invaluable.

REFERENCES

Cant, R.J. and A.R. Cook. 1995. "The Evaluation of Image Quality for Training Simulators Using Artificial Intelligence." In *Proceedings of the 1995 European Simulation Multi-conference* (Prague, June 5-7). SCS Publication, 812-816.

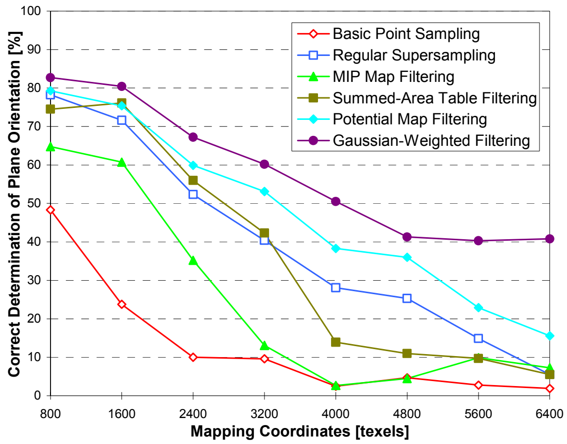


Figure 10 Results for wallpaper texture

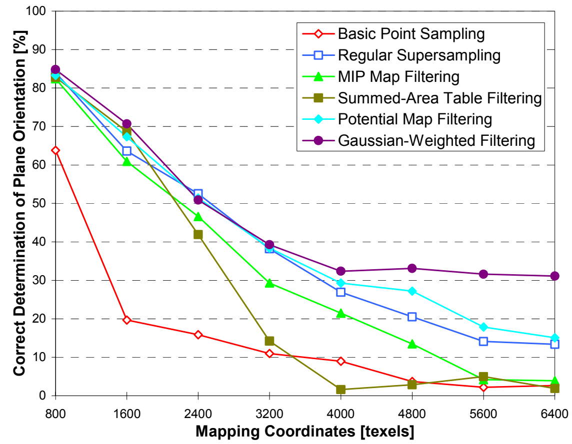


Figure 11 Results for checkered texture

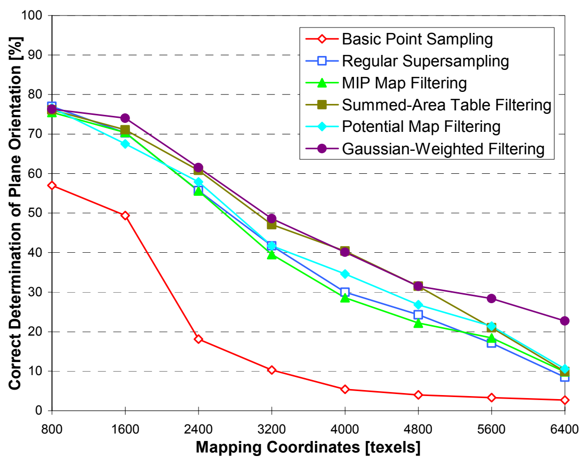


Figure 12 Results for sinusoidal texture

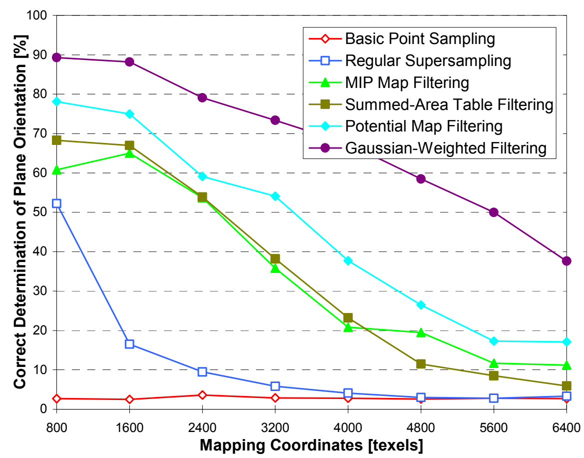


Figure 13 Results for noise texture

Cant, R.J. and P.A. Shrubsole. 1997. "Texture Potential Mapping: A Way to Provide Anti-aliased Texture Without Blurring." In *Visualization & Modeling*, R. Earnshaw, J. Vince, and H. Jones, eds. Academic Press, 1997, Chapter 12.

Cook, A.R. and R.J. Cant. 1996. "Human Perception Versus Artificial Intelligence in the Evaluation of Training Simulator Image Quality." In *Proceedings of the 1996 European Simulation Symposium* (Genoa, October 24-26). SCS Publication, 19-23.

Cook, A.R. 1998. "Neural Network Evaluation of Real-Time Texture Mapping Algorithms." In *Proceedings of the 1998 European Simulation Symposium* (Nottingham, October 26-28). SCS Publication, October 1998, 643-647.

Crow, F.C. 1984. "Summed-Area Tables for Texture Mapping." In *Proceedings of SIGGRAPH '84* (Minneapolis, Minnesota, July 23-27). Computer Graphics, Vol. 18, No. 3, July 1984, ACM SIGGRAPH, New York, 207-212.

Foley, J.D., A. van Dam, S.K. Feiner, and J.F. Hughes, 1990. *Computer Graphics: Principles and Practice* (Second Edition). Addison-Wesley.

Glassner, A.S. 1986. "Adaptive Precision in Texture Mapping." In *Proceedings of SIGGRAPH '86* (Dallas, Texas, August 18-22). Computer Graphics, Vol. 20, No. 4, August 1986, ACM SIGGRAPH, New York, 297-306.

Heckbert, P. 1988. "Survey of Texture Mapping." In *Computer Graphics: Image Synthesis*, K.I. Joy, ed. Computer Society Press, 1988, 321-322.

Schilling, A., G. Knittel, and W. Strasser. 1996. "Texram: A Smart Memory for Texturing." In *IEEE Computer Graphics & Applications*, Vol. 16, No. 3, May 1996, IEEE Computer Society Press, 32-41.

Williams, L. 1983. "Pyramidal Parametrics." In *Proceedings of SIGGRAPH '83* (Detroit, Michigan, July 25-29). Computer Graphics, Vol. 17, No. 3, July 1983, ACM SIGGRAPH, New York, 1-11.