Methods of Reducing the Visibility of Mach Bands during Gouraud Shading

Shaun Bangay
Department of Computer Science
s.bangay@ru.ac.za

Rhodes University
Grahamstown, South Africa

Abstract

Mach bands are illusory artefacts caused by the human visual system providing edge amplification in regions containing a discontinuity in the gradient of light intensity. Gouraud shading attempts to produce smooth surfaces using a polyhedral mesh as a base geometry, and is commonly considered to have visible edge effects due to Mach bands along polygon boundaries.

We investigate factors which affect the visibility of Mach bands. Experimental results indicate the intensity of Mach bands drops off with higher orders of continuity in the interpolation function. Width of the interpolation region also plays a role, although it is possible to reduce the visibility of Mach bands for very narrow interpolation regions.

An adapted Gouraud shader is created which corrects for these factors. Remaining edge effects are attributed to a number of other perceptual issues which have not been previously associated with Gouraud shading. Corrections for these effects are also described, resulting in further shader variations which can use efficient Gouraud shading strategies for the bulk of the polygon, and alternative interpolation strategies where Mach bands occur.

Keywords: Mach bands, Gouraud shading, perceptual brightness, intensity interpolation.

1 Introduction

Brightness perception phenomena affect the way we perceive the results of many of the shading models used in computer graphics. The output produced by the shader is modified by the human visual system to exaggerate edges and introduce other apparent artefacts, affecting perception of the rendered object. Mach bands in particular are a common symptom, defined as the illusory brighter or darker bands in regions where the intensity gradient changes abruptly. In this paper, we investigate factors which affect the perceived intensity of Mach bands, and devise strategies for adapting the Gouraud shading model to reduce band visibility.[1]

Many computer graphics applications perform shading of some form; from simulating the appearance of a smoothly curved surface while using an underlying polyhedral representation, to filling in missing information from lossy image compression. Gouraud shading is a well established technique for smoothly shading a polyhedral mesh. The intensity of light reflected (and thus the colour) at the vertices is calculated, and this is interpolated over the interior of the polygon. The limited number of lighting calculations (one per vertex) is conducive to fast lighting and shading. Fast hardware implementations that perform incremental linear interpolation along scan-lines are common. The resulting discontinuities in intensity gradient at polygon boundaries produce Mach bands which interfere with the illusion of a smooth surface.

The human visual system exaggerates boundaries in the shading areas, and emphasizes the faceted nature of the underlying geometry. Since these are artefacts of the shading process, they are undesirable and should be eliminated. The range of edge related artefacts which occur when using Gouraud and other shading strategies includes:

1. The Chevreul illusion[14] shown in Figure 1(a): A scalloped variation in brightness when a number of bands of uniform intensity are placed next to each other. This is commonly found during flat shading, and colour quantization.

2. Mach bands as shown in Figure 1(d): Illusory bands that appear where discontinuities occur in the intensity gradient. These occur at edges between non-coplanar polygons during Gouraud shading.

3. Cornsweet-Craik-O'Brien edges[19] as shown in Figure 1(g): If a sharp edge occurs between two regions of equal intensity then the intensities of the two regions appear different.

4. Vasarely nested square[1]: A series of nested squares of graded luminance produce exaggerated brightness down the diagonals. This can be achieved in practice, as shown in Figures 1(j) and 1(k), showing the interior of a Phong shaded cube with the light directly behind the viewer. The effect is applicable to other situations where intensity gradients meet at an angle.

5. Jagged edges: Visual sensitivity to jagged edges[17] and parallel, but slightly displaced, contours is higher than can be predicted from retinal receptor density, which suggests an active mechanism to detect these artefacts. Contrast is also enhanced when a region is surrounded by an articulated background, as opposed to a background of constant colour[1]. Reconstruction of the boundaries of the Kanizsa triangle[19] (Figure 1(l)) shows that even edge fragments are capable of reconstructing the entire polygon boundary.

While many sources describe Mach bands during Gouraud shading, little work has been done on relating it to the factors that affect perception of these bands. A number of alternative shading algorithms exist that produce little to

[1] Many of the optical illusions in this paper can be destroyed by printing. Viewing the images using the electronic version of the paper on a 24-bit display is recommended.

no Mach banding, at a cost of increased computational complexity. An improved intensity interpolation is possible: the goal is to find one compatible with the per vertex illumination requirements of Gouraud shading, through a better understanding of the cause of Mach bands.

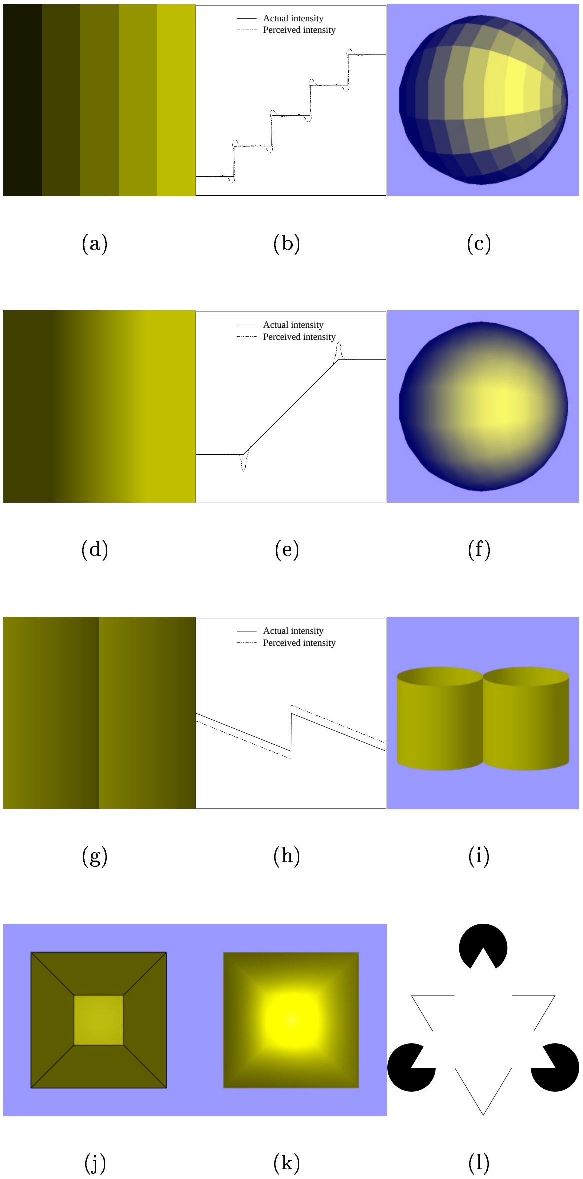


Figure 1: Shading related optical illusions. (a+b) Chevreul illusion in bands of constant colour. (c) Flat shading producing Chevreul illusion. (d+e) Mach bands during linear interpolation between regions of constant colour. (f) Gouraud shading exhibiting Mach bands. (g+h) Cornsweet-Craik-O'Brien edges (i) Cornsweet-Craik-O'Brien edges produced while lighting curved surfaces. (j+k) Vasarely nested produced by Phong shading a cube. (l) Kanizsa triangle.

2 Related Work

People are good at finding patterns in images, even where none may exist. Perceptual effects are an important factor to consider when producing effective rendered scenes[22, 6].

The visible polygon boundaries visible with Gouraud shading are attributed solely to Mach banding by almost all computer graphics texts surveyed[7, 13, 25, 26], and even by Henri Gouraud himself[8]. Phong shading is prescribed to correct this problem, although Phong[20] does mention that edges are still visible in some situations of extreme geometry.

2.1 Mach Bands in Shading algorithms

Various strategies are employed to overcome Mach bands produced by Gouraud shading. The two main approaches are either to provide finer sampling of the lighting conditions by using an alternative shading strategy, or to avoid using linear interpolation.

Shading strategies such as Phong shading[20] perform interpolation over the normal vectors at each polygon vertex, and repeat the lighting calculation for every pixel in the polygon. This has substantial performance implications, but better reproduces specular highlights and reduces edge visibility. A number of shaders of intermediate quality have been explored which offer reduced computational costs[24, 2, 5]. The shaders explored in this paper all remain true to the spirit of Gouraud shading: lighting and geometry attributes are only calculated directly for polygon vertices.

Phong[20] suggests the use of functions with continuous first order derivatives for reducing Mach bands during Gouraud shading. Duff[5] points out that discontinuities in the first derivative exist even with Phong shading, and can in some circumstances be more severe than those found in a Gouraud shader. Higher order continuity is introduced by Perlin[18] to reduce visible edge artefacts in noise.

Quadratic interpolation is necessary to avoid some geometry aliasing problems during Phong shading[23]. Mohamed et al.[15] use quadratic interpolation with Gouraud shading to produce shading behaviour that rivals output from Phong shaders. Their shader uses extra samples of the lighting conditions taken at the midpoints of polygon edges to calculate the extra weighting factors required. It also performs progressive refinement of the mesh when there sufficient variation in intensity across a polygon. However, for low thresholds, this approach becomes equivalent to standard Phong shading.

2.2 Mach Bands in the Human Visual System

Models of the human visual system exist, which emulate the human visual processes, and can predict many of the perceptual edge enhancement effects. These models offer insight into the factors controlling edge perception.

Early models[17] assume the existence of specialized neurons for detecting and enhancing boundaries, line orientation, size, depth, colour and speed. Lateral inhibition is considered by some to be the mechanism responsible for exaggerating intensity boundaries and producing Mach bands[4]. More recent studies indicate that visibility of Mach bands is more dependent on Fourier phase characteristic, and less on intensity gradients[21]. In contrast, the phenomena of *bleeding*[9], has the opposite effect - fine dark lines embedded in a region of constant colour will make it appear darker, while light lines will brighten it. More recently it is recognized that there are global effects resulting from the combined activities of many neurons, as is easily motivated by depth perception using random dot stereograms. Recent work with Mondrian patterns, where perceived brightness can be affected by colours distant from the point of interest, are further evidence of more complex neural interactions[4].

Image processing requires algorithms that behave in a similar fashion to the human visual system. Phase congruency has been proposed as a value related to perception of edge effects and of Mach bands, and has been used successfully for edge detection during image processing[12]. Intensity gradient and overall illumination are also factors which affect edge detection algorithms. Some edge detection algorithms use derivatives up to third order to reduce the effect of noise and detect false positives[11].

McCollough effects[3] introduce colour by adding fine lines of the colour next to large blocks of constant colour. The extent of these has been reported to vary with the orientation of the edges. Orientation effects can be expected, as some neurons in the primary visual cortex have been identified as each responding best to lines of specific orientation. There is also evidence that horizontal and vertical lines are favoured[9, 19]. Width of the intensity ramp is reported to affect the width and visibility of the Mach bands[14].

A range of models of the human visual system[16, 14] have been constructed which explore the effects of combining simulated biological primitive operations such as ON and OFF edge detection channels, oriented contrast detection and bounded filling-in operations. They are able to predict various optical illusions, including the occurrence of Mach bands. These models are usually presented as networks of discrete units (simulated neurons in many cases) which can correctly predict the output for a given input image, but cannot be easily reversed to describe the input that would produce a desired (Mach band free) output.

3 Identifying factors affecting Mach band visibility

Existing models of perception explore plausible neural structures that replicate the effects of known optical illusions, but do not directly explore, nor explain, the mechanisms which result in edge banding effects. They do however identify potential factors that can be used to as the basis for further experimentation.

The goal of this work is to reduce visibility of Mach bands in Gouraud shaded images. The first step involves identifying the factors that control Mach band visibility.

Mach bands have been associated with discontinuities in the first derivative of intensity[17]. Some preliminary experimentation with smooth interpolation of intensity still produces these effects, suggesting that the eye may be sensitive to changes in higher order derivatives. Phong shading manipulates aspects of the underlying surface geometry to achieve smooth blending of colours, and is able to remove most Mach bands. A sinusoidal interpolation can be used to remove Mach bands when displaying a representation of a colour space[10]. This suggests that there are interpolation strategies that will achieve the desired goal.

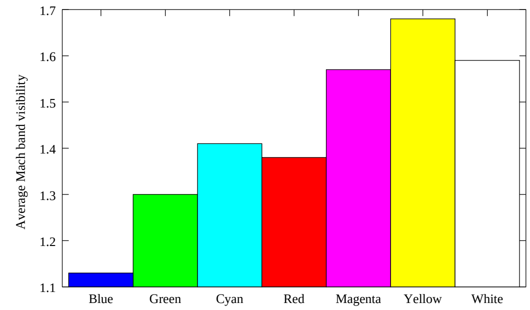


Figure 2: Relationship between Mach band visibility and colour.

3.1 Experimental design

The hypothesis to be tested is that the visibility of Mach bands can be reduced by manipulating variable factors such as those identified in section 2.2. The factors selected are chosen because they are likely to be appropriate to a shading algorithm, and because they show promise after some initial pre-testing.

For experimental verification we use an application which displays the classical scenario for producing Mach bands; specifically two regions of constant intensity, and a band between them in which intensity is interpolated from one level to another. All interpolation occurs in the RGB colour space which, while not ideal, is more likely to be used in a shading algorithm having close links to the underlying hardware.

For the purposes of this experiment, we assume that interpolation occurs between two intensity levels of the same hue. Initial experimentation shows a pronounced dependence on the choice of colour (summarized in Figure 2), although the intensity produced by the RGB monitor is more likely the dependent variable. To reduce the number of degrees of freedom in the remainder of the experiment, we fix the colour to yellow which produces particularly visible Mach bands. We make the assumption that techniques that reduce Mach banding effects for this colour will have the same effect for other colours.

The variable factors are:

Width: The width of the region over which intensity is interpolated is related to the steepness of the intensity gradient. The variable used is the proportion of the window, w, over which the interpolation occurs. The interpolation band is always centered in the window. The window occupies a visual angle of about 10° at a viewing distance of 50cm, although subjects are permitted to adjust their viewing position to best see any banding effects.

Continuity depth: Continuity depth is given by n, in the interpolation function (1), where t varies from 0 to 1 across the width, w, of the interpolation band. This function exhibits discontinuities at 0 and w only after n levels of differentiation. This value determines the

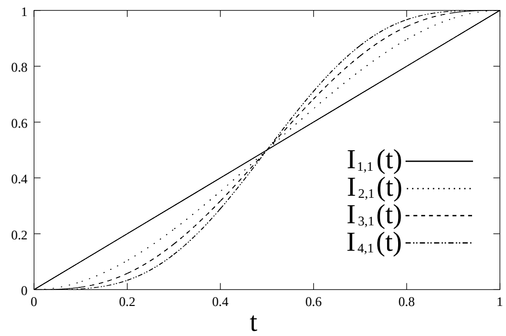


Figure 3: First four levels of the interpolation function.

"smoothness" of the intensity gradient.

$$I_{n,w}(t) = \begin{cases} 0 & t < 0 \\ \left\{ \begin{array}{ll} \frac{t}{w} & n = 1 \\ 3(\frac{t}{w})^2 - 2(\frac{n}{w})^3 & n = 2 \\ 10(\frac{t}{w})^3 - 15(\frac{t}{w})^4 + \\ \quad 6(\frac{t}{w})^5 & n = 3 \\ 35(\frac{t}{w})^4 - 84(\frac{t}{w})^5 + \\ \quad 70(\frac{t}{w})^6 - 20(\frac{t}{w})^7 & n = 4 \end{array} \right\} & 0 \leq t < w \\ 1 & w \leq t \end{cases} \quad (1)$$

The first four levels of this function are illustrated in Figure 3. These functions are derived from the normalized area under the Bezier spline of order $2n-1$: $P^{2n-1}(t)$, where $t = x$, and the y-coordinates of all control points are set to 0, except the center which is set to 1. The properties of the Bezier spline provide the required continuity.

$$I_{n,w}(t) = (2n-1) \int_0^t \begin{pmatrix} 2n-2 \\ n-1 \end{pmatrix} (\tfrac{t}{w})^{n-1} (1 - \tfrac{t}{w})^{n-1} dt \quad \text{for } 0 \leq t < w$$

The function $\tanh(kt)$ has continuous derivatives of all orders, and behaves in a manner similar to the function $I_{n,b}(t)$ although it does not pass exactly through 1 at $t = 1$. For large values of k (in practice $k \geq 5$), this difference is below the quantization error for most display devices, and so for purposes of these experiments we define:

$$I_{\infty,w}(t) \quad = \quad \frac{1 + \tanh(5(\frac{t}{w} - \frac{1}{2}))}{2}$$

Orientation: The orientation of the regions and interpolation band is set by rotating the entire image about the center of the window.

The two regions of constant intensity are filled with a shade of yellow at $\frac{3}{4}$ and $\frac{1}{4}$ of maximum brightness, using equal quantities of red and green on a gamma corrected RGB display.

Subjects are asked to quantify the visibility of the brighter Mach band according to the following scale:

0	No band visible.
1	Broad band present.
2	Narrow band with indistinct boundaries.
3	Narrow band which is clearly defined.

These categories are chosen and arranged based on the acceptability of the corresponding artefacts in a shading algorithm, and on the sequence of Mach bands produced as the independent variables vary.

Initial testing suggests that there may be persistence of vision effects affecting perception of bands in sequential viewing scenarios. We use a recovery period between tests during which the window is cleared to a uniform colour to eliminate this. The different test scenarios are all displayed in random order to minimize any other carry-over effects of previous tests. Testing is carried out with low ambient light levels, to increase sensitivity to Mach bands.

3.2 Experimental Results

By the time the final experimental results are captured, all subjects are familiar with the testing process and evaluation criteria. In total, the experiment has been carried out by 8 subjects.

Initial analysis, resulting from averaging the results over the different subjects for each of the test scenarios, clearly shows that Mach bands are always clearly visible for linear interpolation (continuity depth of 1). Within this category, band visibility is directly related the steepness of the intensity gradient and there is no clear relationship between orientation and band visibility. Results for smoother interpolation techniques show more complex relationships between the experimental factors.

With respect to each of the experimental factors:

Width: Band visibility decreases with increasing width (shallower intensity gradients) as shown in Figure 4(a). Colour quantization effects start to become problematic for very shallow gradients. There is an interesting drop in band visibility for very narrow widths as well. This is possibly due to the already very visible edge masking the visibility of the Mach band.

Continuity depth: Band visibility drops with increasing "smoothness" of the interpolation function, as shown in Figure 4(b). The fact that it is still dropping after level 3 suggests that human perceptual processes are capable of identifying discontinuities in at least the third derivative of the interpolation function.

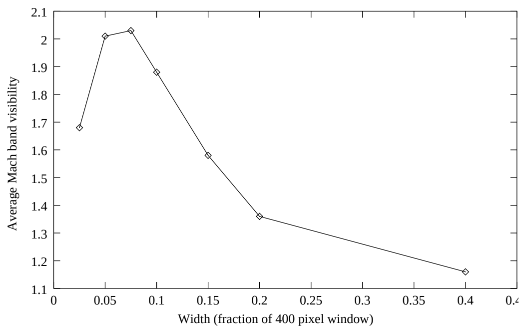
Orientation: Initial results, along with comments from the subjects, indicate a possible increased sensitivity to horizontal and vertical edges. Further experiments to confirm this have proved inconclusive.

3.3 Mach Band Reduction Strategies

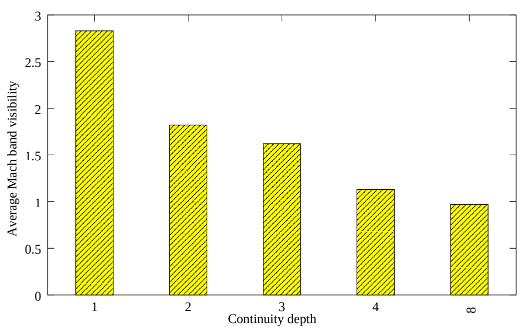
The experimental results show that the visibility of Mach bands can be reduced by controlling the smoothness of the interpolation function. Shaders that use functions with high orders of continuity will show less banding effects. Shaders that interpolate over large regions will show less visible Mach bands.

A potentially useful result is the indication of a drop in band visibility for narrow interpolation regions as well. This has potential application for flat shading, where visible edges are accepted, but where some colour interpolation across the face may be able to combat the Chevreul illusion (Figure 1(a)).

Mach bands are found during Gouraud shading at edges where two different intensity gradients meet. The use of a more complex interpolation function could be applied to a



(a)



(b)

Figure 4: Significant relationships between Mach band visibility and experimental factors. (a) Width (b) Continuity depth.

narrow region around the edge, allowing more efficient linear interpolation to be used over the remainder of the polygon. There is still a trade-off involved in the use of functions with higher orders of continuity: they tend to drop off more steeply over a smaller region resulting in a higher intensity gradient and a more visible boundary.

4 Adapting Gouraud shaders to reduce Mach bands

For convenience, all discussion assumes the use of triangular polygons, with vertices A, B and C as illustrated in Figure 5. Gouraud shading is an efficient shading technique that evaluates lighting conditions only at polygon vertices, and interpolates the resulting intensities across the entire face. In practice many implementations[13, 25] interpolate in two steps: first blending intensity vertically down the edges of the polygon, and then blending the edge intensities horizontally across each scan-line. For this discussion, a more convenient alternative (although equivalent) formulation is preferred, where the intensity of a point in the interior of the polygon is a weighted average of the intensities of all of the vertices.

For point P within the triangle as illustrated in Figure 5, the intensity of light reflected from P using Gouraud shading

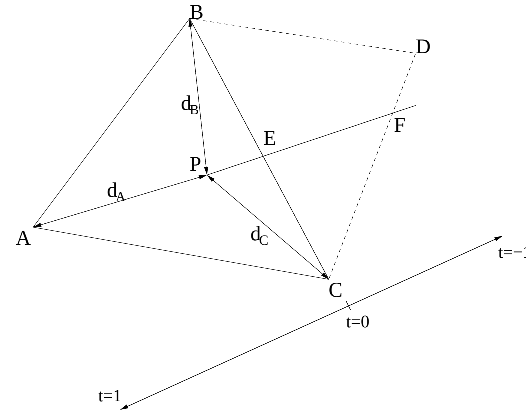


Figure 5: Calculating intensity at point P.

can be determined using the intensities at the vertices: I_A, I_B and I_C.

$$I_p = w_A I_A + w_B I_B + w_C I_C \quad (2)$$

where $w_{A,B,C}$ are based on the distances from each of the vertices to P and satisfy:

$$\begin{aligned} P &= w_A A + w_B B + w_C C \\ 1 &= w_A + w_B + w_C \end{aligned}$$

Points on one of the edges will have zero weighting from the opposite vertex. Thus the intensity of an edge point such as point E illustrated in Figure 5 can be calculated using only the intensities of the two neighbouring vertices. I_E is independent of which of the two neighbouring triangles (ABC or BDC) is currently being rendered. The intensity of the points on the line AE can be expressed in terms of the intensity of the endpoints, and a measure t of the distance from E:

$$\begin{aligned} I_p &= w_A I_A + (1 - w_A) I_E \\ &= t I_A + (1 - t) I_E \end{aligned} \quad (3)$$

where $I_E = \frac{w_B I_B + w_C I_C}{1 - w_A}$.

Although the two neighbouring triangles shown in Figure 5 have a common intensity at E, the first derivative with respect to t will have a discontinuity at E unless $I_A - I_E = \frac{I_E - I_F}{|E - F|}$. This is sufficient to produce visible Mach bands when using Gouraud shading.

4.1 Strategies for removing Mach bands during Gouraud shading

The experimental results from section 3.2 provide two factors which can influence the visibility of Mach bands. Ensuring the continuity of the higher order derivatives of the interpolation function is a powerful tool in achieving the desired goal. The dependency of band visibility on width of the interpolation region suggests two alternative choices with regard to this factor:

1. Large regions using higher order smoothing functions which may provide a better result, but at the cost of increasing shading time.

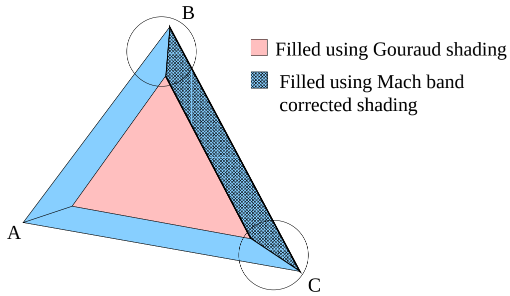


Figure 6: Correcting banding by mixing shading strategies.

2. Shade fine regions on the polygon boundary using the higher order smoothing functions and accomplish the bulk of the shading using traditional Gouraud shading.

Both alternatives can be supported by defining a boundary region around the border of the triangle which will be shaded using Mach band corrected interpolation. The interior of the triangle can then be shaded using traditional Gouraud shading, as illustrated in Figure 6. The width of the boundary region can be modified for best results.

The boundary regions can be identified by computing the minimum of the weights used in equation (2). If this minimum is below a given boundary width b, then the smallest weight corresponds to the vertex directly opposite the border region. The discussion below assumes the point to be shaded falls into the boundary opposite vertex A as illustrated in Figure 6.

4.2 Fixed Boundary Conditions

4.2.1 Strategy

The shading function across the boundary region must ensure n levels of continuity at the triangle edge, where $t = 0$, and join in a continuous fashion with the Gouraud interpolation function defined in (3) at the interior of the boundary where $t = b$. The derivatives of the Gouraud interpolation function up to the n^{th} order are known and can be matched. On the exterior of the triangle only the intensity value I_E is known. Without knowledge of the behaviour of the neighbour all higher order derivatives must be set to a canonical value of zero at $t = 0$ to ensure matching.

The function given by:

$$J_{n,b}(t) = tI_{n,b}(t)(I_A - I_E) + I_E \quad (4)$$

where, $I_{n,b}(t)$ as defined in (1), exhibits this desired behaviour. The intensity of a point P within the boundary region opposite vertex A is given by:

$$I_P = J_{n,b}(w_A)$$

Accurately matching intensity gradient at edge points on the triangle boundary would require expensive intersection calculations to determine the behaviour of the intensity interpolation in neighbouring triangles. We can approximate some measure of this behaviour while retaining the spirit of Gouraud shading. Intensity values for all vertices can, and should, be calculated prior to the onset of the interpolation process. Intensities for neighbouring vertices (such as vertex D shown in Figure 5) would then be known before any polygon filling occurs.

The value of I_F (as in Figure 5) can be determined by interpolating between the intensities of the endpoints. The weighting factor depends the relative weights along the shared edge. These can be approximated using the relationship (5) derived in Appendix A. The calculation for the situation where $w_B > w_C$ is shown in (6).

$$\begin{aligned} a &= \frac{2w_B}{w_B + w_C} - 1 \\ f &= \frac{2a}{1+a} && (5) \\ I_F &= fI_B + (1-f)I_D && (6) \end{aligned}$$

The calculation of I_F, for situations in which the neighbouring triangle is not a mirror image, is an approximation, which deteriorates as vertex D moves away from its ideal position.

The interpolation expression can be updated, to set the gradient at the boundary:

$$J_{n,b}(t) = tI_{n,b}(t)(I_A - I_E) + tI_{n,b}(b-t)\left(\frac{I_F - I_E}{2}\right) + I_E$$

The boundary regions are not of constant width, but narrow at the ends (the regions circled in Figure 6). The width of the boundary, as used in the expressions above, needs to be modified to reflect this when shading points which are in close proximity to a triangle vertex.

4.2.2 Implementation

A pseudo code description of the shading process is thus:

```
Intensity FixedBoundaryShade
   (Point P,
    LinearWeights w_A(P), w_B(P), w_C(P),
    Intensity I_A, I_B, I_C)
  if (w_A(P) < b)
    // in boundary region opposite vertex A.
    if w_C(P) > 1 - 2b
      // close to vertex C
      b = (1 - w_C(P)) / 2
    // and a similar check for proximity to B
    I_D = intensity of vertex opposite A
            across BC
    a = 2w_B(P) / (w_B(P)+w_C(P)) - 1
    if a < 0
      f = -2a / (1-a)
      I_F = f I_C + (1 - f) I_D
    else
      f = 2a / (1+a)
      I_F = f I_B + (1 - f) I_D
    return J_{n,b}(w_A(P))
  ..
    // similarly for regions opposite B, C.
  else
    // in center region.
    return I_p = w_A I_A + w_B I_B + w_C I_C
```

4.3 Vertex Shading

4.3.1 Strategy

The regions around each vertex require extra attention due to special conditions that exist in these areas. A well lit vertex often represents an extrema in terms of lighting intensity value. This serves to emphasize banding effects in this region. Particularly noticeable are effects similar to those found in the Vasarely nested square, shown in Figures 1(j) and 1(k). This can be ascribed to the interpolation gradients meeting at an angle, shown in Figure 7(a). The ameliorating effects of the other vertices are reduced in the area around each vertex.

A weighting measure based on Euclidean distance from the vertex allows the gradients in adjacent triangles to meet at the same angle, as shown in Figure 7(b). An elliptical distortion is required to cater for the difference lengths of the two triangle edges sharing the common vertex. Distortion based directly on the original linear weights produces slippage: the contours are not aligned at the common edge. These are particularly noticeable, and represent another perceptual effect caused by edge enhancement.

Slippage can be removed using the square of the original linear weights, producing the effect shown in Figure 7(c). The corrected distance based vertex weights are calculated as shown in the example of (7).

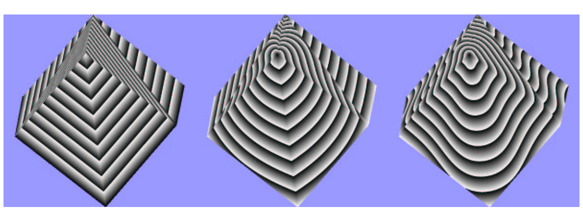
$$d_A(P) = 1 - \frac{|AP|^2}{\left[|AB|\frac{w_B^2}{w_B^2+w_C^2} + |AC|\frac{w_C^2}{w_B^2+w_C^2})\right]^2} \quad (7)$$

The distance weights are not normalized, and in fact become negative when approaching the opposite edges. This is insignificant as long as their effect is confined to the region immediately around each of the vertices.

4.3.2 Implementation

The shaders given previously can be enhanced for smoother vertex conditions by the addition of code such as the following:

```
if (w_A(P) > 1 - 2b)
  // in the region around a vertex.
  // use distance weights given in (7).
  I_Dist = d_A(P) I_A
  I_Lin = intensity from other shader.
  // confine effect to region around
  // vertex.
  g = I_{n,b}((1 - w_A(P))/2)
  return g I_Lin + (1 - g) I_Dist
// and similarly for vertices B, C
```



(a) (b) (c)

Figure 7: Illustration of weight function from a single vertex of a triangulated cube. (a) Linear weighting. (b) Distance based weighting showing slippage. (c) Corrected distance weighting.

4.4 Intensity Reweighting

Gouraud shaded pixels change intensity visibly even when at a small distance from the polygon edge. We speculate that this behaviour contributes toward the perception that edges are substantially brighter than the interior of the triangle. This effect is not as noticeable for Phong shading. Assuming a traditional local illumination model, such as that used in OpenGL[27], intensity at a point P is given by (8), where N is a normalized normal for the surface at that point, L and S are normalized vectors indicating light and viewpoint geometry, and the other terms are properties of surface and world.

$$I_P = I_{\text{amb}} + k_{\text{diff}}(N \cdot L) + k_{\text{spec}}(N \cdot S)^n \quad (8)$$

The specular term is relatively insignificant outside the specular highlight, and is ignored for the purposes of the derivation below. The relationship between the two shading techniques is:

$$\begin{aligned} I_{\text{Gouraud}} &= w_A I_A + w_B I_B + w_C I_C \\ &\simeq I_{\text{amb}} + k_{\text{diff}}(N_{\text{Phong}}) \cdot L \\ I_{\text{Phong}} &= I_{\text{amb}} + k_{\text{diff}}(\frac{N_{\text{Phong}}}{\left|N_{\text{Phong}}\right|}) \cdot L \\ &\simeq \frac{I_{\text{Gouraud}}}{d} \end{aligned}$$

where $N_{\text{Phong}} = w_A N_A + w_B N_B + w_C N_C$ is the weighted normal vector, which must be normalized before being used for illumination calculations.

The intensity distribution for Phong shading can be emulated using scaled weights dependent only on the linear weights and attributes of the vertices. This approximation applies in areas where diffuse lighting dominates. Variations for regions with significant ambient or specular effects could be similarly derived.

This approach can be integrated with the boundary shading technique by updating the intensity interpolation function as shown in (9).

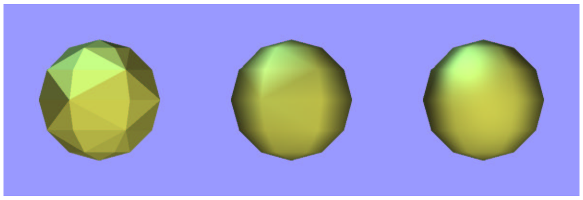
$$J_{n,b}(t) = \frac{w_A I_A + w_B I_B + w_C I_C}{\left|N_{\text{Phong}}\right| + (1 - \left|N_{\text{Phong}}\right|) I_{n,b}(t)} \quad (9)$$

5 Results

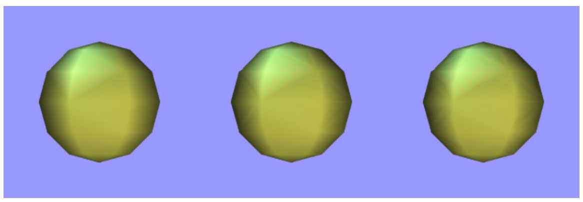
Figure 8 shows the results of the fixed boundary Gouraud shaders in action, compared to a selection of standard shaders. Bands are still present, particularly on edges where adjacent triangles have a significant change in gradient. As the continuity depth of the interpolation function increases, the bands become broader and more diffuse. The intensity of Mach bands decreases, but the edges remain visible.

The boundary requirements of this shader also introduce other visible artefacts along the edges of the triangles, particularly for large boundary regions. While Mach banding may have been reduced, the edges become visible as large regions of uniform intensity. This makes edges visible between co-planar triangles.

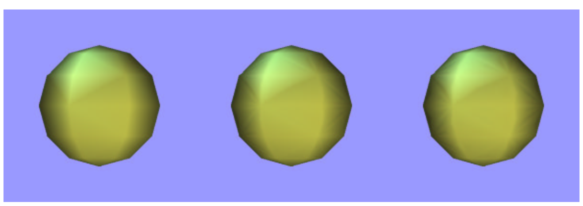
Inclusion of the vertex correction, shown in Figure 9, provides a significant contribution toward breaking up the edges



(a) Flat. (b) Gouraud. (c) Phong.

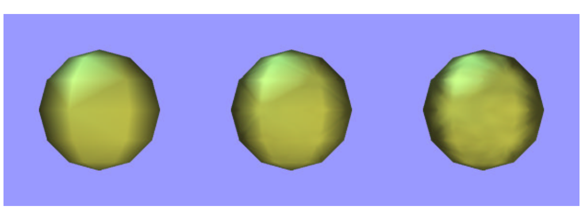


(d) (3, 0.05) (e) (3, 0.2) (f) (3, 0.33)



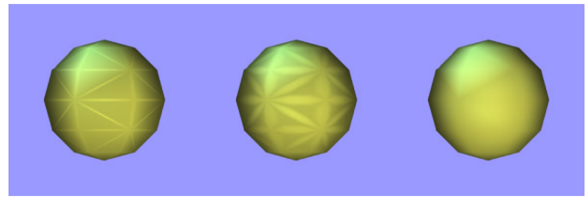
(g) (5, 0.05) (h) (5, 0.2) (i) (5, 0.33)

Figure 8: Fixed boundary shader in action. (a-c) Standard shaders. (d-i) Fixed boundary shader for given values of (n, b).



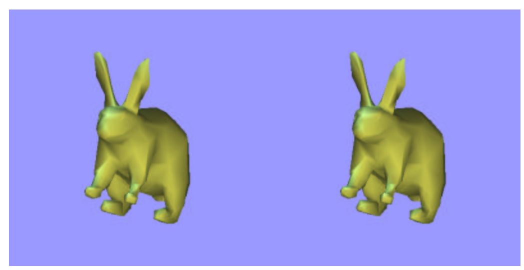
(a) (3, 0.05) (b) (3, 0.2) (c) (3, 0.33)

Figure 9: Vertex shader in action for given values of (n, b).

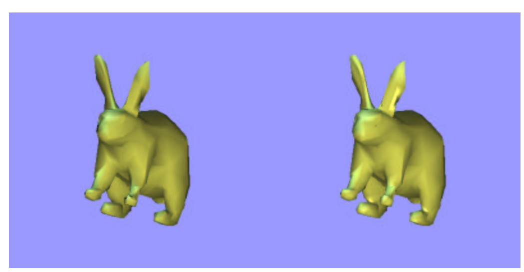


(a) (3, 0.05) (b) (3, 0.2) (c) (3, 0.33)

Figure 10: Intensity reweighting shader in action for given values of (n, b).



(a) (1, 0.05) (b) (3, 0.2)



(c) (3, 0.2) (d) (3, 0.33)

Figure 11: Shaders in action on an object consisting of 622 triangles. (a-b) Boundary shader. (c) Vertex shader (d) Intensity reweighting.

(a) (1, 0.05) (b) (3, 0.2)

(c) (3, 0.2) (d) (3, 0.33)

Figure 12: Shaders in action on an object consisting of 1538 triangles. (a-b) Boundary shader. (c) Vertex shader (d) Intensity reweighting.

and the continuity of the boundary lines. It also spreads illumination more evenly over the surface of the triangles, reducing the undesired brightness of the edges. Vertex correction over a large boundary region even goes some way toward approximating a specular highlight, although the overlap of all the vertices does provide a rather bumpy appearance to the surface.

Shading using linear interpolation is not distributing intensity evenly over the area of the triangle, and is introducing other perceptual artefacts which are also providing edge enhancement. This is confirmed by viewing the results of using the intensity reweighting shown in Figure 10. Bands along edges are eliminated. When mixed with traditional Gouraud shading, the difference between the intensity levels produced by the two shading approaches is marked.

The results of applying the shaders to more detailed objects are shown in Figures 11 and 12. Mach band reduction occurs in all cases, and a number of other edge effects are removed by the vertex and intensity reweighting shaders.

6 Conclusions

6.1 Factors controlling Mach band visibility

Steepness of the intensity gradient and smoothness of the interpolating function are identified as factors which affect the visibility of Mach bands. The relationship between band intensity and width suggests modifications that are applicable to a number of commonly used shading algorithms. Significant reduction of band visibility is possible up to fourth order continuity, at which point levels are unlikely to undergo further significant change.

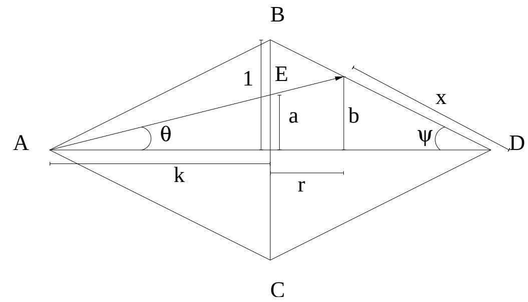


Figure 13: Deriving the relationship between edge weightings in neighbouring triangles.

6.2 Reducing banding in Gouraud shaders

It is now clear that Mach bands are not the only factors creating edge effects during Gouraud shading. The Vasarely effect plays a significant role where two adjacent triangles have different shapes, resulting in intensity gradients meeting at an angle. Any difference in interpolation along a shared edge produces visible slippage. The eye even extends small lines segments, resulting from regions containing bright vertices, to fill in the edges between these vertices. Displays with poor colour resolution will show the shaded region as containing bands of constant colour, producing further edge effects over the interior of the polygon.

Adapting the Gouraud shader as described in this paper, and using only information local to the polygon achieves the desired goal of reducing Mach bands. Strategies such as the vertex shading and intensity reweighting are effective at removing other edge effects.

6.3 Future Work

Similar studies are planned to investigate the factors producing banding phenomena in flat shading and Phong shading.

Further work is required to adapt the enhanced shading algorithms to determine if they would be suitable for implementation in hardware.

6.4 Acknowledgments

We would like to thank Lindsey Bangay for her assistance with experimental design and data analysis, as well as the members of the VRSIG for their enthusiastic cooperation throughout the many iterations of test environment.

A Derivation of edge interpolation factor for neighbouring triangles

Consider two adjacent congruent triangles ABC and DBC as shown in Figure 13. For convenience we set the scale of the shared base BC to 2, so that the absolute value of the edge point a ranges between 0 and 1. The goal is to identify the dependence of the position of the intersection of the line AE with the opposite edge, x, on the distance from the center of the common edge, represented by a.

$$\tan(\theta) = \frac{a}{k} = \frac{b}{k+r}$$
$$\tan(\Psi) = \frac{b}{k-r} = \frac{1}{k}$$

For convenience, and to simplify the derivation, we would prefer the value of the intersection point with the opposite edge to be scaled to remain in the range [0 : 1]. Thus we introduce $f = \frac{x}{\sqrt{k^2+1}}$.

$$\begin{aligned} f^2 &= \frac{x^2}{k^2+1} \\ &= \frac{b^2 + (k-r)^2}{k^2+1} \\ &= \frac{(k-r)^2(k^2+1)}{k^2(k^2+1)} \\ &= \frac{1}{k^2}\left[k(1-\frac{1-a}{1+a})\right]^2 \\ f &= +\sqrt{f^2} \\ &= \frac{2a}{1+a} \end{aligned}$$

References

[1] Adelson, E. H., Lightness Perception and Lightness Illusions, Chapter 24 in M. Gazzaniga, ed., The New Cognitive Neurosciences, 2nd Ed., MIT Press, Cambridge, MA, 339-351, 2000.

[2] Bishop, G. and Weimer, D. M., Fast Phong shading, ACM SIGGRAPH Computer Graphics, 20(4), 103-106, August 1986.

[3] Broerse, J., Viadusich, T., O'Shea, R. P., Colour at edges and colour spreading in McCollough effects, Vision Research, 39, 1305-1320, 1999.

[4] Coren, S., Ward, L. M., Enns, J. T., Sensation and perception, 5th Ed., Harcourt Brace College Publishers, Fort Worth, Texas, 1999.

[5] Duff, T., Smoothly shaded renderings of polyhedral objects on raster displays, Proceedings of SIGGRAPH '79, Chicago, Illinois, 1979.

[6] Euh, J. and Burleson, W., Exploiting Content Variation and Perception in Power-Aware 3D Graphics Rendering, Proceedings of the Power-Aware Computing Symposium (PACS), Cambridge, MA, 2000.

[7] Foley, J. D. and van Dam, A., Fundamentals of interactive computer graphics, Addison-Wesley, Reading, Mass., 1982.

[8] Gouraud, H., Continuous shading of curved surfaces, IEEE Transactions on Computers, C-20(6), 623-629, June 1971.

[9] Gregory, R. (ed), The artful eye, Oxford University Press, Oxford, 1995.

[10] Joblove, G. H. and Greenberg, D., Color spaces for computer graphics, Proceedings of SIGGRAPH '78, August 1978.

[11] Kovesi, P., Invariant Measures of Image Features from Phase Information, PhD Thesis, University of Western Australia, May 1996.

[12] Kovesi, P., Image Features from Phase Congruency, Videre: Journal of Computer Vision Research, 1(3), 1999.

[13] Hearn, D. and Baker, M. P., Computer graphics, Prentice-Hall International, Englewood Cliffs, N.J., 1986.

[14] McArthur, J. A. and Moulden, B., A two-dimensional model of brightness perception based on spatial filtering consistent with retinal processing, Vision Research, 39(6), 1199-1219 March 1999.

[15] Mohamed, A. A., Szirmay-Kalos, L., Horvath, T. and Foris, T., Quadratic Shading and its Hardware Implementation, Machine Graphics and Vision, 9(4), 825-804, 2001.

[16] Neumann, H., Pessoa, L. and Mingolla, E., A neural architecture of brightness perception: non-linear contrast detection and geometry-driven diffusion, Image and Vision Computing, 16(6-7), 423-446, May 1998.

[17] Montalvo, F. S., Human vision and computer graphics, Proceedings of SIGGRAPH '79, Chicago, Illinois, 121-125, 1979.

[18] Perlin, K., Improving Noise, ACM Transactions on Graphics, 21(3), 681-682, July 2002.

[19] Pessoa, L. and Neumann, H., Why does the brain fill-in?, Trends in Cognitive Sciences, 2, December 1998.

[20] Phong, B., Illumination for computer generated images. Communications of the ACM, 18(6), 311-317, June 1975.

[21] Ronse, C., A twofold model of edge and feature detection, Working Document WD65, Philips Research Laboratory Brussels, September 1990.

[22] Rushmeier, H., Barrett, H., Rheingans, P., Uselton, S. and Watson, A., Perceptual measures for effective visualizations, Proceedings of Visualization '97, Phoenix, Arizona, 1997.

[23] van Overveld, C. W. A. M. and Wyvill, B., Phong normal interpolation revisited, ACM Transactions on Graphics, 16(4), 397-419, October 1997.

[24] Vlasic, D., Fake Phong Shading, Master of Engineering Thesis, Massachusetts Institute of Technology, 2002.

[25] Watt, A., Fundamentals of three-dimensional computer graphics, Addison-Wesley, Wokingham, 1989.

[26] Watt, A. and Watt, M., Advanced animation and rendering techniques: theory and practice, Addison-Wesley, Reading, Mass., 1992.

[27] Woo, M., Neider, J., Davis, T. and Shreiner, D., OpenGL Programming Guide: The Official Guide to Learning OpenGL, Version 1.2 (3rd Edition), Addison-Wesley, 1999.