Enhancing the Effects You Can Create with Computer Graphics by Using the OpenGL Shading Language (GLSL)

Mike Bailey
mjb@cs.oregonstate.edu

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

The Basic Computer Graphics Pipeline, OpenGL-style

MC = Model Vertex Coordinates
WC = World Vertex Coordinates
EC = Eye Vertex Coordinates
NDC = Normalized Device Coordinates (-1.→1.)
SC = Screen Coordinates (pixels)

Rasterization

Framebuffer

MC

NDC

SC

RGBA pixels

Fragments

NDC

Vertices

Fragment

Shader

Fragment

Shader

Vertices

NDC

SC

RGBA pixels

Fragments

MC

NDC

SC

RGBA pixels

Fragments
GLSL Variable Types

**uniform** These are “global” values, assigned and left alone for a group of primitives. They are read-only accessible from all of your shaders. They cannot be written to from a shader.

**out / in** These are passed from one shader stage to the next shader stage. In our case, out variables come from the vertex shader, are interpolated in the rasterizer, and go in to the fragment shader.

GLSL is like C With Extensions for Graphics:

- Types include int, ivec2, ivec3, ivec4
- Types include float, vec2, vec3, vec4
- Types include bool, bvec2, bvec3, bvec4
- Vector components are accessed with .rgba, .xyzw, or .stpq
- Types include mat4
- Types include sampler1D, sampler2D, sampler3D to access textures
- You can use parallel SIMD operations (doesn’t necessarily get implemented in hardware):
  
  ```
  vec4 a = vec4( 1., 2., 3., 4. );
  vec4 b = vec4( 5., 6., 7., 8. );
  ...
  vec4 c = a + b;
  ```
- Vector components can be “swizzled” (to abgr = from rgba)
- Type qualifiers: const, uniform, in, out
- Variables can have “layout qualifiers” to describe how data is stored
- The discard operator is used in fragment shaders to get rid of the current fragment

The discard Operator Halts Production of the Current Fragment

```
if( random_number < 0.5 )
    discard;
```

GLSL is Missing Some C-isms:

- No type casts – use constructors instead:
  ```
  float x = 3.14;
  int i = int( x );
  ```
- No pointers, strings, or enums

**Warning:** integer division is still integer division!

```
int n = 2;
int d = 4;
float f = float( n / d );  // still gives 0, just like C, C++, Python, and Java do
```

**To get the correct result:**

```
float f = float( n ) / float( d );
```
Our Shaders’ View of the Basic Computer Graphics Pipeline

- Fixed Function (non-you-programmable)
- You-Programmable

There are actually four more GLSL shader types we won’t be covering here. In CS 457/557, we will cover all of them.

A GLSL Vertex Shader Takes Over These Operations:
- Vertex transformations
- Normal Vector transformations
- Computing per-vertex lighting (although, if you are using shaders anyway, per-fragment lighting looks better)
- Taking per-vertex texture coordinates (s,t) and interpolating them through the rasterizer into the fragment shader

A GLSL Fragment Shader Takes Over These Operations:
- Color computation
- Texture lookup
- Blending colors with textures (like GL_REPLACE and GL_MODULATE used to do)
- Discarding fragments

**Built-in Vertex Shader Variables You Will Use a Lot:**
- `vec4 gl_Vertex`
- `vec3 gl_Normal`
- `vec4 gl_Color`
- `vec4 gl_MultiTexCoord0`
- `mat4 gl_ModelViewMatrix`
- `mat4 gl_ProjectionMatrix`
- `mat4 gl_ModelViewProjectionMatrix (= gl_ModelViewMatrix * gl_ProjectionMatrix)`
- `mat3 gl_NormalMatrix` (this is the transpose of the inverse of the MV matrix)

**Output built-in**
- `vec4 gl_Position`

Note: while this all still works, OpenGL now prefers that you pass in all the above input variables as user-defined in variables. We can talk about this later. For now, we are going to use the most straightforward approach possible.
**Built-in Fragment Shader Variables You Will Use a Lot:**

Output built-in \[
\text{vec4 } \text{gl\_FragColor} \quad = \text{the RGBA being sent to the framebuffer}
\]

Note: while this all still works, OpenGL now prefers that you pass the RGBA out as a user-defined out variable. We can talk about this later. For now, we are going to use the most straightforward approach possible.

---

**My Own Variable Naming Convention**

With 7 different places that GLSL variables can be created from, I decided to adopt a naming convention to help me recognize what program-defined variables came from what sources:

<table>
<thead>
<tr>
<th>Beginning letter(s)</th>
<th>Means that the variable ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>is a per-vertex in (attribute) from the application</td>
</tr>
<tr>
<td>u</td>
<td>is a uniform variable from the application</td>
</tr>
<tr>
<td>v</td>
<td>came from the vertex shader</td>
</tr>
<tr>
<td>tc</td>
<td>came from the tessellation control shader</td>
</tr>
<tr>
<td>te</td>
<td>came from the tessellation evaluation shader</td>
</tr>
<tr>
<td>g</td>
<td>came from the geometry shader</td>
</tr>
<tr>
<td>f</td>
<td>came from the fragment shader</td>
</tr>
</tbody>
</table>

This isn't part of "official" GLSL – it is just my way of handling the chaos.

---

**The Minimal Vertex and Fragment Shader**

**Vertex shader:**

```glsl
#version 330 compatibility
void main()
{
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

**Fragment shader:**

```glsl
#version 330 compatibility
void main()
{
    gl_FragColor = vec4(.5, 1., 0., 1.);
}
```

**A Reminder of what a Rasterizer does**

There is a piece of hardware called the Rasterizer. Its job is to interpolate a line or polygon, defined by vertices, into a collection of fragments. Think of it as filling in squares on graph paper.

Rasterizers interpolate built-in variables, such as the (x,y) position where the pixel will live and the pixel’s z-coordinate. They also interpolate the normal vector (nx,ny,nz) and the texture coordinates (s,t). They can also interpolate user-defined variables as well.

A fragment is a "pixel-to-be". In computer graphics, "pixel" is defined as having its full RGBA already computed and is headed to be stored in the framebuffer. A fragment does not yet have a computed RGBA, but all of the information needed to compute the RGBA is available.

A fragment is turned into an RGBA pixel by the fragment processing operation.
This pattern example is defined by three uniform variables: \(uS0\), \(uT0\), and \(uD\), all in texture coordinates \((0..-1)\). \((uS0,uT0)\) are the center of the pattern. \(uD\) is the length of each edge of the pattern. The \(s\) and \(t\) boundaries of the pattern are like this:

A Little More Interesting, I:

What if we Want to Color in a Pattern?

\[
\begin{align*}
\text{s} &= uS0 - uD/2. \\
\text{t} &= uT0 + uD/2. \\
\text{s} &= uS0 + uD/2. \\
\text{t} &= uT0 - uD/2.
\end{align*}
\]

\(uS0, uT0\) are the center of the pattern in texture coordinates

\(uD\) is the size of the pattern in texture coordinates

The vertex shader needs to pass the texture coordinates to the rasterizer so that each fragment shader gets it:

The vertex shader:

```cpp
#version 330 compatibility
out vec2 vST;
void main()
{
    vST = gl_MultiTexCoord0.st;
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

The fragment shader answers the question: “Am I (the current fragment) inside the pattern or outside it?”

The fragment shader:

```cpp
uniform float uS0, uT0, uD; // from your program
void main()
{
    float s = vST.s; // the s coordinate of where this fragment is
    float t = vST.t; // the t coordinate of where this fragment is
    vec3 myColor = vec3(1., 0.5, 0.); // default color
    if (uS0 - uD/2. <= s && s <= uS0 + uD/2. &&
        uT0 - uD/2. <= t && t <= uT0 + uD/2.)
    {
        myColor = vec3(1., 0., 0.); // change the pattern color if inside the pattern boundaries
    }
    gl_FragColor = myColor with lighting applied;
}
```

All 4 of these must be true to conclude this fragment is inside the pattern!
Here's the cool part: It doesn't matter (up to the limits of 32-bit floating-point precision) how far you zoom in. You still get a crisp edge. This is an advantage of procedural (equation-based) textures, as opposed to image-based textures.

It's because, in a sphere, the s coordinate encompasses twice as much angle (-180° → +180°) as the t coordinate does (-90° → +90°). So the same amount of "s" produces twice the distance as the same amount of "t". If you care, you can fix it like this:

```cpp
float s = vST.s;
float t = vST.t;
s = 2.*s;
```

If the Equation Defines a Square, Why Does the Pattern Look Like a Rectangle?

```
s = \frac{\Theta - (-\pi)}{2\pi} \\
t = \frac{\Phi - (-\pi)}{\pi}
```

Vertex shader:

```cpp
#version 330 compatibility
out vec2 vST; // texture cords
out vec3 vN; // normal vector
out vec3 vL; // vector from point to light
out vec3 vE; // vector from point to eye
const vec3 LIGHTPOSITION = vec3(5., 5., 0.);
void main() {
    vST = gl_MultiTexCoord0.st;
    vec4 ECposition = gl_ModelViewMatrix * gl_Vertex; // eye coordinate position
    vN = normalize(gl_NormalMatrix * gl_Normal); // normal vector
    vL = LIGHTPOSITION - ECposition.xyz; // vector from the point to the light position
    vE = vec3(0., 0., 0.) - ECposition.xyz; // vector from the point to the eye position
    vec3 Normal = normalize(vN);
    vec3 Light = normalize(vL);
    vec3 Eye = normalize(vE);
    vec3 myColor = vec3(1.0, 0.5, 0.0); // default color
    vec3 mySpecularColor = vec3(1.0, 1.0, 1.0); // specular highlight color
    if( dot(Normal,Light) > 0. ) // only do specular if the light can see the point
    {
        float d = dot(Normal,Light);
        vec3 ref = normalize(reflect(-Light, Normal)); // reflection vector
        vec3 diffuse = uKd * d * myColor;
        vec3 specular = uKs * pow( max( dot(Eye,ref),0. ), uShininess ) * mySpecularColor;
        vec3 ambient = uKa * myColor;
        vec3 ambient += diffuse + specular;
        gl_FragColor = vec4(ambient + diffuse + specular, 1.);
    }
    else
    {
        vec3 Normal = normalize(vN);
        vec3 Light = normalize(vL);
        vec3 Eye = normalize(vE);
        vec3 myColor = vec3(1.0, 0.0, 0.0); // default color
        vec3 mySpecularColor = vec3(1.0, 1.0, 1.0); // specular highlight color
        vec3 ambient = uKa * myColor;
        vec3 ambient += diffuse + specular;
        gl_FragColor = vec4(ambient + diffuse + specular, 1.);
    }
}
```
Applying Per-Fragment Lighting

Per-fragment Lighting is Good, Even Without a Pattern!

First, follow these steps:

1. You will see two files that are already in your Sample folder: glslprogram.h and glslprogram.cpp
2. In your sample.cpp file, un-comment the line:
   \#include "glslprogram.cpp"

These two files have been reduced to have just the shader features you need for Project #6.

If you are not working on Project #6, but are working on something bigger, I have more complete versions of glslprogram.h and glslprogram.cpp – just ask me.
void Animate( )
{
    int ms = glutGet( GLUT_ELAPSED_TIME );  // milliseconds
    ms %= MS_IN_THE_ANIMATION_CYCLE;
    Time = (float)ms / (float)MS_IN_THE_ANIMATION_CYCLE;        // [ 0., 1. )
}

Do this in Animate() like you've always done:

Here is our C++ Class to Simplify the Shader Setup for You

Do this in InitGraphics() somewhere after where the window has been created and GLEW has been setup:

Setting up a Shader is somewhat Involved:

Here is our C++ Class to Simplify the Shader Setup for You

Setting up a Shader is somewhat Involved:

Here is our C++ Class to Simplify the Shader Setup for You

We cover the full GLSL API in CS 457/557

Pattern.Init( );
bool valid = Pattern.Create( "pattern.vert", "pattern.frag" );
if( ! valid )
{
    fprintf( stderr, "Yuch!  The shader did not compile.
" );
}
else
{
    fprintf( stderr, "Woo-Hoo! The shader compiled.
" );
}

This attempts to load, compile, and link the shader program. If something goes wrong, Pattern.Create() prints error messages into the console window and returns a value of valid=false.

In C/C++, the exclamation point (!) is pronounced "not".

In C/C++, the exclamation point (!) is pronounced "not".

Graphics chips have functionality on them called Texture Units. Each Texture Unit is identified by an integer number, typically 0-15, but oftentimes more.

To tell a shader how to get to a specific texture image, assign that texture into a specific Texture Unit number and then tell your shader what Texture Unit number to use. Your C/C++ code will look like this:

```
glActiveTexture( GL_TEXTURE5 );                 // use texture unit 5
glBindTexture( GL_TEXTURE_2D, TexName );
```

The file gl.h has these lines:

```
#define GL_TEXTURE0                             0x84C0
#define GL_TEXTURE1                             0x84C1
#define GL_TEXTURE2                             0x84C2
#define GL_TEXTURE3                             0x84C3
#define GL_TEXTURE4                             0x84C4
#define GL_TEXTURE5                             0x84C5
#define GL_TEXTURE6                             0x84C6
#define GL_TEXTURE7                             0x84C7
```

We cover the full GLSL API in CS 457/557

Graphics chips have functionality on them called Texture Units. Each Texture Unit is identified by an integer number, typically 0-15, but oftentimes more.

To tell a shader how to get to a specific texture image, assign that texture into a specific Texture Unit number and then tell your shader what Texture Unit number to use. Your C/C++ code will look like this:

```
glActiveTexture( GL_TEXTURE5 );                 // use texture unit 5
glBindTexture( GL_TEXTURE_2D, TexName );
```

The file gl.h has these lines:

```
#define GL_TEXTURE0                             0x84C0
#define GL_TEXTURE1                             0x84C1
#define GL_TEXTURE2                             0x84C2
#define GL_TEXTURE3                             0x84C3
#define GL_TEXTURE4                             0x84C4
#define GL_TEXTURE5                             0x84C5
#define GL_TEXTURE6                             0x84C6
#define GL_TEXTURE7                             0x84C7
```
The Whole Process Looks Like This, I:

```c
// globals:
unsigned char * Texture;
GLuint TexName;
GLSLProgram Pattern;

// In InitGraphics( ):
glGenTextures( 1, &TexName );
int nums, numt;
Texture = BmpToTexture("filename.bmp", &nums, &numt );
gTexParameterf(GL_TEXTURE_2D, GL_TEXTURE_WRAP_S, GL_REPEAT );
gTexParameterf(GL_TEXTURE_2D, GL_TEXTURE_WRAP_T, GL_REPEAT );
gTexParameterf(GL_TEXTURE_2D, GL_TEXTURE_MAG_FILTER, GL_LINEAR );
gTexParameterf(GL_TEXTURE_2D, GL_TEXTURE_MIN_FILTER, GL_LINEAR );
gTexImage2D(GL_TEXTURE_2D, 0, 3, nums, numt, 0, 3, GL_RGB, GL_UNSIGNED_BYTE, Texture);
Pattern.Init();
bool valid = Pattern.Create("pattern.vert", "pattern.frag");
if( !valid )
{
    //...
}

The Whole Process Looks Like This, II:

This is the hardware Texture Unit Number. It can be 0-15 (and often higher depending on the graphics card).

```
What if You Want to Use Two Textures in a Shader?

// In Display()
Pattern.Use();
glActiveTexture(GL_TEXTURE5);
glBindTexture(GL_TEXTURE_2D, TexName0);

Pattern.Use();
glActiveTexture(GL_TEXTURE6);
glBindTexture(GL_TEXTURE_2D, TexName1);

Pattern.SetUniformVariable(“uTexUnit0”, 5);
Pattern.SetUniformVariable(“uTexUnit1”, 6);

<< draw something >>
Pattern.UnUse();

Why Would You Want to Use More Than One Texture in a Shader?

Once the RGBs have been read from a texture, they are just numbers. You can do any arithmetic you want with the texture RGBs, other colors, lighting, etc. Here is an example of blending two textures at once:

Fragment shader:

#version 330 compatibility
in vec2 vST;
uniform sampler2D uTexUnit0;
uniform sampler2D uTexUnit1;
void main() {
    vec3 newcolor = texture(
        ...
    ) * vec3(0.5, 0.5, 0.5);
    gl_FragColor = ...
}
Setting rgb from the Untransformed xyz, I

Fragment shader:

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_ModelViewMatrix * gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

Vertex shader:

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_ModelViewMatrix * gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

What's Different About These Two?

Set the color from the **untransformed** (MC) xyz:

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

Set the color from the **transformed** (WC/EC) xyz:

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_ModelViewMatrix * gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

Turning XYZs into RGBs in Eye (World) Coordinates

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_ModelViewMatrix * gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

Setting rgb from the Transformed xyz, II

```
#version 330 compatibility
out vec3 vColor;
void main()
{
    vec4 pos = gl_ModelViewMatrix * gl_Vertex;
    vColor = pos.xyz; // set rgb from xyz!
    gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```
Hints on Running Shaders on Your Own System

- You need a graphics system that is OpenGL 2.0 or later. Basically, if you got your graphics system in the last 5 years, you should be OK, unless it came from Apple. In that case, who knows how much OpenGL support it has? (The most recent OpenGL level is 4.6)
- Update your graphics driver to the most recent version!
- Do the GLEW setup. It looks like this in the sample code:
  ```
  GLenum err = glewInit();
  if (err != GLEW_OK)
  {
    fprintf(stderr, "glewInit Error\n");
  }
  else
    fprintf(stderr, "GLEW initialized OK\n");
  ```
  This must come after you've created a graphics window. (It is this way in the sample code, but I'm saying this because I know some of you go in and "simplify" my sample code by deleting everything you don't think you need.)
- You use the GLSL C++ class you've been given only after a window has been created and GLEW has been setup. Only then can you initialize your shader program:
  ```
  Pattern.Init();
  bool valid = Pattern.Create("pattern.vert", "pattern.frag");
  ```

A Common Error to Look Out For

Here is a piece of code:
```
#version 330 compatibility
out vec3 vColor;

void main()
{
  vec4 pos = gl_Vertex;
  vec3 vColor = pos.xyz; // set rgb from xyz!
  gl_Position = gl_ModelViewProjectionMatrix * gl_Vertex;
}
```

Wrong! By re-declaring vColor in "vec3 vColor = pos.xyz", you are making a local copy of it and writing into the local copy, not the out variable! The out version of vColor is never getting written to, and so the vColor in the fragment shader will have no value. (BTW, C/C++ works the same way.)

Don't ever re-declare in, out, or uniform variables!

Trust me, you will do this sometime. It's an easy mistake to make mindlessly. I do it every so often myself.

Differences if You are on a Mac

Apple froze their GLSL support at version 1.20 – here is how to adapt to that:
- Your shader version number should be 120 (at the top of the .vert and .frag files):
  ```
  #version 120 compatibility
  ```
- Instead of the keywords in and out, use varying
- Your OpenGL includes will need to look like this:
  ```
  #include <OpenGL/glu.h>
  ```
- You don't need to do anything with GLEW
- Your compile sequence will look like this:
  ```
  g++ -framework OpenGL -framework GLUT sample.cpp -o sample -Wno-deprecated
  ```
1. Declare the GLSLProgram above the main program (i.e., as a global):
   ```
   GLSLProgram Pattern;
   ```
2. At the end of InitGraphics(), create the shader program and setup your shaders:
   ```
   Pattern.Init();
   bool valid = Pattern.Create("pattern.vert", "pattern.frag");
   if( ! valid ) { . . . }
   ```
3. Turn on the shader program in Display(), set shader variables, draw the objects, then turn
   off the shader program:
   ```
   Pattern.Use();
   Pattern.SetUniformVariable( ...
   OsuSphere();
   Pattern.UnUse(); // return to the fixed function pipeline
   ```
4. When you run your program, be sure to check the console window for shader compilation errors!

Tips on drawing the object:
- If you want to key off of s and t coordinates in your shaders, the object must have s and t coordinates
  (vt) assigned to its vertices – not all OBJ files do!
- If you want to use surface normals in your shaders, the object must have surface normals (vn) assigned
  to its vertices – not all OBJ files do!
- Be sure you explicitly assign all of your uniform variables – no error messages occur if you forget to do
  this – it just quietly screws up.
- The glutSolidTeapot() has been textured in patches, like a quilt – cute, but weird
- The OsuSphere() function from the texturing project will give you a very good sphere. Use it, not the
  GLUT sphere.