What is Java 3D and is it for me?

Java 3D is an application programming interface (API) developed at Sun Microsystems for rendering interactive 3D graphics using the Java programming language. Java 3D is a *client-side* Java API. Other examples of Sun client-side APIs include the Abstract Windows Toolkit (AWT) and Java Foundation Classes (JFC/Swing), which are both Java class libraries for building applications with a Graphical User Interface (GUI). The client-side Java APIs are in contrast to Sun's server-side APIs such as Enterprise Java-Beans (EJB) and the other components of Java 2 Enterprise Edition (J2EE).

Making 3D graphics interactive is a long-standing problem, as evidenced by its long history of algorithms, APIs, and vendors. Sun is not a major player in the 3D graphics domain, although its hardware has long supported interactive 3D rendering. The dominant industry standard for interactive 3D graphics is OpenGL, created by Silicon Graphics (SGI). OpenGL was designed as a cross-platform rendering architecture and is supported by a variety of operating systems, graphics card vendors, and applications. The OpenGL API is written in the C programming language, and hence not directly callable from Java. A number of open source and independent programming efforts have provided simple Java wrappers over the OpenGL API that allow Java

programmers to call OpenGL functions, which are then executed in native code that interacts with the rendering hardware. One of the most popular is GL4Java, which you can find at http://www.jausoft.com/gl4java/.

However, there are few advantages to using a Java wrapper over OpenGL, as opposed to coding in C and calling OpenGL directly. Although programmers can use the more friendly Java APIs, they must incur the overhead of repeated calls through the Java Native Interface (JNI) to call the native OpenGL libraries.

Java 3D relies on OpenGL or DirectX to perform native rendering, while the 3D scene description, application logic, and scene interactions reside in Java code. When Sun set out to design Java 3D, although they did not have the resources or industry backing to replace OpenGL, they wanted to leverage more of Java's strengths as an object-oriented programming (OOP) language instead of merely delegating to a procedural language such as C. Whereas OpenGL's level of description for a 3D scene consists of lists of points, lines, and triangles, Java 3D can describe a scene as collections of objects. By raising the level of description and abstraction, Sun not only applied OOP principles to the graphics domain, but also introduced scene optimizations that can compensate for the overhead of calling through JNI.

1.1   STRENGTHS

The foremost strength of Java 3D for Java developers is that it allows them to program in 100 percent Java. In any sizeable 3D application, the rendering code will compose only a fraction of the total application. It is therefore very attractive to have all the application code, persistence, and user interface (UI) code in an easily portable language, such as Java. Although Sun's promise of Write-Once-Run-Anywhere is arguably more of a marketing dream than a reality, especially for client-side programming, Java has made important inroads toward enabling application developers to write applications that can be easily moved between platforms. The platforms of most interest today are Microsoft Windows 98/NT/2000, Sun Solaris, LINUX, and Macintosh OS X.

Java has arguably become *the* language of networked computing and the Internet. High-level support for remote method invocation (RMI), object serialization, platform independent data types, UNICODE string encoding, and the security model all provide persuasive arguments for adopting the Java language for applications that are increasingly gravitating away from a desktop-centric worldview. Many of the state-of-the-art 3D graphics applications being built with Java 3D today are leveraging the strengths of Java as a language for the Internet.

The Java 3D API itself has much to offer the application developer. By allowing the programmer to describe the 3D scene using coarser-grained graphical objects, as well as by defining objects for elements such as appearances, transforms, materials, lights, and so forth, code is more readable, maintainable, reusable, and easier to write. Java 3D uses a higher level scene description model, the *scenegraph*, which allows scenes to be easily described, transformed, and reused.

Java 3D includes a view model designed for use with head-mounted displays (HMDs) and screen projectors. By insulating the programmer from much of the complex trigonometry required for such devices, Java 3D eases the transition from a screen-centric rendering model to a projected model, where rendering in stereo allows for greater realism.

Java 3D also includes built-in support for sampling 3D input devices and rendering 3D spatial sound. By combining all of the above elements into a unified API, Java 3D benefits from a uniformity of design that few other APIs can match.

Java 3D's higher level of abstraction from the mechanics of rendering the scene have also opened the field of interactive 3D graphics to a new class of audience, people who would typically have been considered 3D content creators. Think of 3D graphics creation as a spectrum, with resources and talents distributed across a variety of tasks, as illustrated in figure 1.1.

Many new programmers have moved from Virtual Reality Modeling Language (VRML) into Java 3D. They are 3D content creation specialists; and they require the greater flexibility offered by a programming API, though they are reluctant to learn OpenGL and C. For this audience, Java 3D fills an important niche, allowing them to concentrate on content creation and application logic, without choking on the details of rendering and arcane programming syntax.

1.2   WEAKNESSES

Many of the strengths can be reversed and cited as weaknesses. For some programmers coming *from* OpenGL, there are some OpenGL features that are hard or impossible to achieve within Java 3D. Some of this audience may miss the total control they have over the scene and the rendering process. Many others, however, will quickly learn the mapping from OpenGL functions to Java 3D objects and will appreciate the productivity gains they can achieve using Java 3D.

Although Java 3D includes some clever optimizations, a skilled developer using OpenGL and native C code may be able to achieve higher performance than a Java programmer using Java 3D. If absolute rendering performance is the top-priority for your application then you may be better off using OpenGL or another native rendering API.

One particular problem, inherent in Java, which can be noticeable in performance-critical applications, is the impact of the Java garbage collector (GC). The Java runtime, the Java 3D runtime, and the application code all create objects. All these objects will eventually be garbage, and be collected by the Java Virtual Machine (JVM) GC. While the GC is running there may be an appreciable system slowdown, resulting in several rendered frames being dropped. If garbage collection occurs in the middle of a critical animation sequence, the realism of the rendered scene may be lowered for the user. However, with continued improvements in GC technology, faster hardware, and well-designed and implemented applications, such considerations are no longer prevalent.

The Java client-side APIs, and especially Java 3D, can be difficult to distribute to end users. While the biggest pool of end users run Windows, Sun has had limited success

Content Creation

VRML

```
Material  {
    ambientColor   0.200 0.200 0.200
    diffuseColor   0.800 0.400 0.500
    shininess  0.000
}
Cube {
    width  2.000
    height  2.000
    depth  2.000
}
```

Java 3D Programming

```
universe = createVirtualUniverse();
Locale locale = createLocale( universe );
BranchGroup sceneBranchGroup = createSceneBranchGroup();
Background background = createBackground();
```

OpenGL Programming

```
glClearColor(1.0, 1.0, 1.0, 1.0);
glClear(GL_COLOR_BUFFER_BIT);
glColor3f(0.0, 0.0, 0.0);
glMatrixMode(GL_PROJECTION);
glLoadIdentity();
```

Vendor Specific Programming

Figure 1.1   Java 3D fills an important gap between VRML, which is centered around describing 3D content, and OpenGL, which is a C API for rendering points, lines, and triangles

getting Java 2 (JRE 1.2) deployed on the Windows platform. Java 2 is required for Java 3D, although Microsoft's JVM does not support Java 2. This means that end users are required to download Sun's Java 2 implementation, install it, and then download Java 3D and install it, all prior to running your application. If you are deploying your application as an applet, the installation process is potentially more complex as some end users will have to manually copy or edit configuration files before they can view your applet. In addition a suitable version of OpenGL or DirectX must be installed and configured for the end user's hardware and drivers. This lengthy download and installation process can lead to frustration; I think we are some way from seeing mainstream software and games companies offering consumer-grade software products built using Java 3D, or even Java 2. Many modern end users expect the convenience of point-and-click installation and do not have the computer skills to set CLASSPATH variables or debug installation problems.

There is light at the end of the tunnel, however, as the Java WebStart project attempts to make installing and running SDK 1.2 Java applications as easy as running native applications—which may be just as well. At present it does not appear that Microsoft will be shipping *any* JVM with Windows XP.

At present, the biggest groups of Java 3D users appear to be computer scientists, businesspeople, hobbyists, game developers, and programmers. These early adopters are spearheading the deployment of Java 3D for mainstream applications.

1.3 SYSTEM REQUIREMENTS (DEVELOPER AND END USER)

Java is a resource-intensive development and deployment environment and creating interactive 3D graphics is still one of the most challenging tasks for modern PCs. Interactive 3D rendering requires hardware dedicated to 3D rendering, usually provided by third-party display hardware specially adapted for processing 3D scenes. Fortunately, 3D-display hardware has reduced in price radically over the past few years, and today's typical game PCs are able to exceed the capabilities of the expensive dedicated graphics workstations of just a few years ago.

For a realistic immersive 3D experience (first-person 3D games, for example), a consistently high frame rate is required, typically 30 frames per second (FPS) or higher. More powerful rendering hardware will be able to achieve higher frame rates, at higher screen resolutions and with higher resolution texturing, all of which contribute to the overall experience. Modern PCs could typically achieve reasonable frame rates without dedicated rendering hardware, however the processor must execute both application logic and rendering code—to the detriment of both.

Nonimmersive 3D applications (such as visualization or modeling) do not typically require as high a frame rate as immersive applications. However the application logic may become the limiting factor on frame rate, as complex calculations may be necessary prior to rendering every frame.

The frame rate that the end users see is determined by a number of factors:

- *Vertex or transform bound*—Ability of the display hardware to transform and display each vertex in the scene
- *Fill bound*—Ability of the display hardware to shade and texture the scene and push the resulting pixels to the screen
- *Application logic bound*—Ability of the application to prepare the scene for rendering

Different types of application will place different demands on those factors, and the type of application you are writing will often dictate the hardware requirements for development and end users.

The minimum requirements for most interactive 3D applications are:

- 500+ MHz main processor
- Dedicated 3D-display hardware, with at least 16 MB of texture memory. New 3D graphics cards are released regularly so you should research the latest cards within your budget. Ensure that the card has good OpenGL compatibility for use with Java 3D. The Java 3D mailing list is a good source of information on people's experiences with various graphics cards.
- 128+ MB of system RAM

An important part of designing your application should be to set your performance targets. Gather requirements from your user base on typical available hardware and ensure that your application can perform adequately on your target machine configuration. You may need to test using several popular graphics cards to ensure compatibility and performance. You may need to try several driver versions to find the best drivers for your supported cards. Unfortunately, Write-Once-Run-Anywhere does not work out too well in the world of 3D graphics!

Analyze the performance of your application using a tool such as OptimizeIt from VMGEAR (http://www.vmgear.com) to determine whether your frame rate is limited by your application logic or display hardware. Regular use of OptimizeIt will also help you to get the maximum performance from the JVM and increase garbage collection intervals.

1.4   EXPECTED PERFORMANCE

An important part of your application design is to estimate your expected performance and validate your design against your target machine configurations. Aim to build some simple prototypes that will allow you to extrapolate your finished application's performance. It is far easier to revise your designs at this stage than two weeks before completion.

For example, on my home machine—with an AMD 850 MHz processor, nVidia GeForce II Ultra (64 MB RAM) graphics card, and 256 MB RAM—I get about 35 FPS running the Java 3D Fly-Through example application (http://www.javasoft.com/

products/java-media/3D/flythrough.html). The Fly-Through city scene (figure 1.2) is composed of 195,000 triangles, 4,115 `Shape3D` instances, and 1,238 `Appearances` (uncompiled scenegraph).
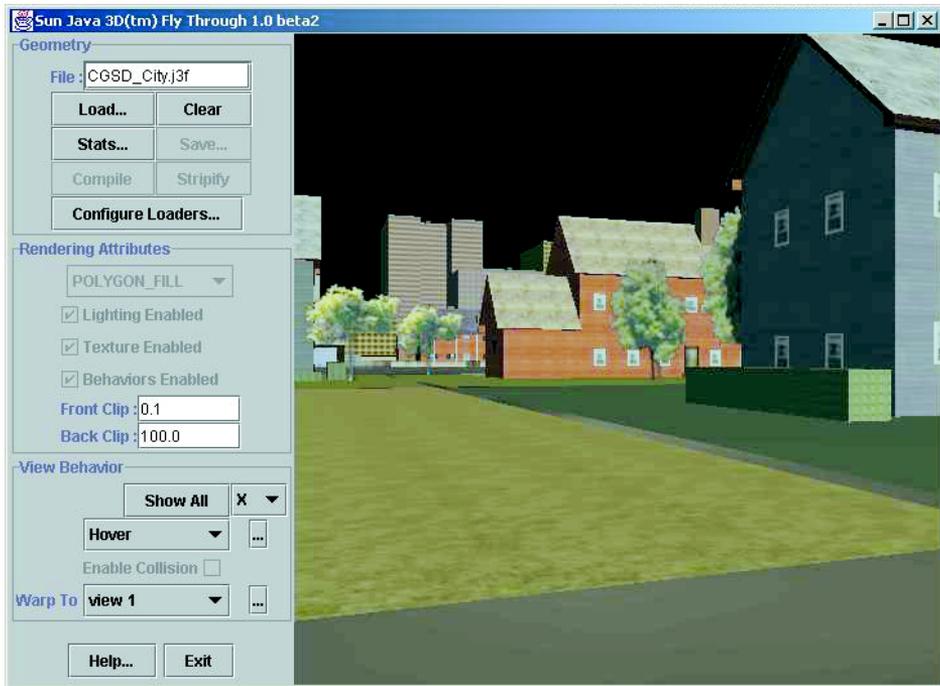

Figure 1.2   The Sun Java 3D example Fly-Through

1.4.1   Memory footprint

Java programs generally tend to require more memory than native programs. This is especially true of programs with a GUI using Swing/JFC. Java 3D can also have high memory requirements, especially if your application loads lots of large bitmaps for texture mapping objects, or defines complex geometry composed of many thousands of vertices.

To give you some idea of Java 3D's memory requirements, table 1.1 shows the total memory required for the Java 3D Fly-Through application. As you can see, bringing up the Swing application requires 25 MB, while opening the city scene pushes memory usage up to over 100 MB.

Table 1.1   Java 3D Fly-Through statistics

Working set	25 MB (no scene loaded)
Working set	108 MB (city scene loaded)

Memory usage will be an important component of your application performance. Performance will be extremely poor if your target users have less physical RAM available than the working set for your application. In this case, the operating system will have to page virtual memory to and from disk.

Another performance criterion that can be important for some applications is startup time. You should set targets for the startup time for your application. The JVM can take a considerable time to start, especially on slower machines with limited RAM. In addition, if you are loading large texture files or 3D object models, your startup time can become very significant. The RAM footprint of your application (including the JVM) and the available system RAM of the end user's computer are the most significant elements affecting startup time. You should take regular startup time measurements while you are in development to ensure that your end users are not frustrated every time they launch your application.

If you are deploying an applet, you should also be aware of the time required for it to download, as well as the graphics resources the applet requires for rendering. Texture images and 3D models can quickly become very large, so some download time targets based on typical end user bandwidth will also prove very useful.

As a reference, I measured the startup time of the Java 3D Fly-Through application. As you can see in table 1.2, launching the application took a very respectable 3 seconds, while loading the 3D content took 14 seconds. Fourteen seconds is a long time, and necessitates some form of progress indicator to reassure users that progress is occurring!

Table 1.2   Java 3D Fly-Through statistics

Start-up time	3 seconds
Loading city scene	14 seconds

1.5   RUNNING THE EXAMPLES

By now, you are probably itching to see Java 3D in action. Please refer to appendix A for a list of the example Java 3D applications and applets developed for this book, as well as detailed instructions for running the examples.

1.6   SUMMARY

Straddling the worlds of content creation and scripting on the one hand and low-level pipeline-based rendering programs on the other, the Java 3D API fills an important gap in 3D graphics APIs. With careful design and implementation, performance of Java 3D applications can rival native OpenGL applications and will exceed JNI-based Java wrappers over OpenGL.

As a Java API, Java 3D is relatively mature, first appearing at the end of 1998. But compared to OpenGL, which first appeared in the early 1990s, Java 3D is still an upstart. For example, OpenGL contains an extension facility that allows vendors to

write proprietary extensions to the API—a feature that is not yet implemented in Java 3D, though it is rumored to be appearing in Java 3D 1.4. The Architecture Review Board (ARB) controls additions to OpenGL—while Java 3D may be placed under the Java Community Process (JCP), allowing experts and vendors to influence the direction of the API.

Java 3D is the right choice if you want to program 3D applications using Java. Just as Java introduced many useful abstractions over C++ and includes a rich library of standard APIs, Java 3D introduces abstractions over OpenGL/Direct3D and includes many features that will bring your applications to market faster. Java 3D can be frustrating at times—abstraction is not always a good thing—but it will save you time as you leverage years of API development by Sun's engineers. While absolute performance is sometimes a requirement, 3D graphics hardware, processor, and memory availability are advancing so rapidly that any disparity between Java/Java3D and C/OpenGL is shrinking for all but the most memory-intensive applications.