WHITE PAPER
SPMarkJava JSR 184

V 1.0.2

Sami Niemelä
Futuremark Corporation
December 14th, 2007

Table of Contents

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

1 OVERVIEW

This paper introduces SPMarkJava JSR 184—the Java benchmark for mobile devices. SPMarkJava JSR184 extends Futuremark's benchmark product lines to the Java market. It produces reliable, easy-to-understand results on the performance and power consumption of Java capable mobile devices. SPMarkJava JSR 184 was previously known as SPMarkJava06.

SPMarkJava JSR 184 is targeted for the mobile Java (J2ME) environment, measuring the device's performance and power consumption. The benchmark is developed with the same expertise as Futuremark's highly popular PC benchmarks 3DMark® and PCMark®. SPMarkJava JSR 184 has the same high quality look and feel as all other Futuremark benchmarks.

In this paper, we discuss a bit of the development of the devices and software in the handheld environment. The tests as well as the scoring methodology in SPMarkJava JSR 184 are explained in detail.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

2 HANDHELD DEVICES

Today's handheld devices have about the same processing power as an average PC had 10 years ago, and it seems that the computing power of these devices will develop extraordinarily fast. The next generation of mobile devices, that are actually partly here already, will have a significant leap on performance compared to the one in PC world 10 years ago. A new breed of sophisticated applications is possible due to the high processing power of today's handsets. Traditionally, the software for mobile devices has been, for example: calendar applications, phone books, and similar calendar/notebook usage related software. With more open interfaces to program on, the variety of the software for mobile devices has increased dramatically. Now, one can get similar kinds of software for phones as for PCs and actually be able to do a significant part of the work mobile.

Mobile devices have, until now, been primarily differentiated based on features and not performance; this is somewhat different from the PC world. The main reason for this is that the devices are still static when it comes to the used components (excluding memory amount in some cases): the user can not change components on the phone to get more processing power. However, as the features and functions of differing devices are similar, the differentiation between them can be made from the performance point of view via a benchmark. The performance of a device is useful information for operators and application developers who want to ensure the greatest performance capacity of the new services and software they introduce.

2.1 Hardware

The trend in hardware (and software) development for handhelds seems to follow PC development. Of today's (2007) mobile devices, high end smartphones may have a 300MHz CPU, over 30MB of memory and a slot for external memory card with up to gigabytes of additional storage capacity. PDAs then again seem to be a bit ahead on the performance side when looking the raw numbers: CPUs up to 600MHz, amount of memory 64MB and up and several gigabytes of storage capacity both internally and with additional memory card. The display size and resolution are, of course, smaller than in PCs. For mobile phones, screen resolution is commonly QVGA (240*320) and for PDAs VGA (640*480) or even WVGA (800*480).

A big, but so far less marketed feature in mobile multimedia capabilities is hardware accelerated 3D. The same evolution happened in PC world approximately 10 years ago and after that the gaming has really taken huge steps towards reality. Today some hardware accelerated implementations for both PDAs and mobile phones have already made available with more constantly coming up, and seemingly the point of wide adoption of accelerated 3D graphics in different kinds of mobile devices is very close.

Another development area seems to be multimedia in general: videos, sound/music and images. Again, this is similar to PCs of the last 10 years. One specific aspect can be recognized particularly in mobile phones though: the camera functionalities. The majority of todays mobile phones have a digital camera incorporated into the device and the quality of the images captured with those cameras is already similar to images captured with a 3-4 year old digital camera.

Mobility is an essential factor that brings limitations to components as well as applications. Devices' batteries may last for a week if activity is minimal; the uptime is heavily dependent on the amount of activities performed. The power consumption will increase as the CPU speeds advance, however for example a separate 3D accelerator may increase the power efficiency by handling 3D specific tasks much more efficiently than a general purpose CPU. Low power consuming implementations are heavily favored both on hardware and software and the performance or features are not always as high as they could be. The battery of a device needs to be small but carry as much power as possible. So far, the manufacturers seem to succeed in

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

this quite well; but as the need for power increases, new battery technologies will be constantly needed.

2.2 Software

Smartphones come with a variety of operating systems and software. A variety of operating systems (OS) are seen in mobile phones as opposed to virtually only three popular options for a PC. Within smartphones, Symbian OS seems to be the most popular OS at the moment, but for value and entry level mobile phones, proprietary OS's are common.

The amount of software for handhelds is increasing, and the applications available for phones are often very similar to the PC environment. The availability of additional software depends on the platform and OS of the phone, but as many OS's are open for developers, the probability of finding software for practically any purpose is ever increasing.

The software available can be both native (e.g. built for e.g. Symbian OS), or higher level, such as Java. Of these options, Java has increased in popularity especially because of the high cross-platform compatibility; the developer may not need to port the software for different platforms and OS's. J2ME is the Java application development environment developed by Sun Microsystems Inc, targeted for embedded devices, more information and the development tools can be found at: http://java.sun.com/javame/index.jsp.

Developing software for smartphones in general is different than for PCs. Phones have different usage models and more limited resources than PCs which need to be taken into account. Also, the development environment is different. Then again, as in the PC world, the development of hardware has an effect on software development and vice versa. One topic also to take into account when developing for smartphones is the power consumption; as the battery lifetime is limited, battery saving options need to be considered.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

3 SPMARKJAVA JSR 184

Figure 1: SPMarkJava JSR 184 on a Series 60 phone

As the mobile devices' capabilities today are extensive, we see that hardware performance will be of essence to users as the features are quite similar from one device to another. The evolution of the devices seems to be going towards entertainment, most prominently gaming. SPMarkJava JSR184 includes tests for measuring the 3D performance as well as other, more general mobile workloads.

SPMarkJava JSR 184 is a benchmark for the mobile Java environment developed by Futuremark Corporation. It is built to test the Java (J2ME) performance on MIDP2.0 and CLDC1.1 capable devices. The 3D tests are based on JSR 184 (also known as m3g), which seems to be most popular open API when developing Java 3D games for EU and USA markets.

The tests in SPMarkJava JSR 184 include 3D gaming, theoretical 3D, 2D gaming, image processing, video playback and JVM performance tests. The structure of SPMarkJava JSR 184 is similar to the other Futuremark benchmarks, and users of those should find using SPMarkJava JSR 184 easy. The tests are also made similar to the ones in Futuremark's PC benchmarks. SPMarkJava JSR 184 comes in two versions: Basic Edition and Professional Edition. The Basic Edition is available as a free download.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

3.1 The Tests

SPMarkJava JSR 184 consists of 8 (The Professional Edition) tests and a battery test. The tests include two subgroups of tests: Image processing testsuite and JVM testsuite. These test results are combined internally and output PNG subscore and JVM subscore.

Test Category	Test
3D	3D Game test, Low detail
	3D Game test, High detail (Professional Edition)
	3D Fill rate test
	3D Polygon throughput test
2D	2D Game test
System tests	Image processing tests
	JVM tests
	Video playback (Professional Edition)
Battery test	All (Professional Edition)

Table 1: SPMarkJava JSR 184 Tests

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

3.1.1 3D Game test

Figure 2: SPMarkJava JSR 184 3D Game Test, High detail version

The 3D Game test high detail version runs a high-end 3D game scene and measures the performance of the 3D operations. The scene contains approximately 5000 triangles rendered per frame with a total of over 21,000 triangles in the scene. The test uses bilinear filtering, the flying warship receives directional lighting (the light can optionally be switched off) and perspective correction is used (also toggleable).

The game scene is run full-screen. Different phones have different screen resolution and this has effect on the performance. Games however are usually played full-screen, so this approach was considered valid for SPMarkJava JSR 184 as well.

The graphical content of the test displays a flying warship chasing an off-road vehicle in a desolated area. The high detail version of the 3D game test is only available in the Professional Edition of SPMarkJava JSR 184.

The 3D game test is available also as a low detail version (SPMarkJava JSR 184 Basic Edition), where the amount of rendered triangles per frame is decreased to approx. 3,000, and the total amount of triangles in the scene is approx. 12,000. Mipmapping is used in the game scene (if available), no lights are used and perspective correction is set to off. The features used in the 3D game tests are shown in Table 2.

The 3D Game tests results are reported as frames per second.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

Option	High detail version	Low detail version
image filter	IMAGE_FILTER_LINEAR	IMAGE_FILTER_NEAREST
level filter	LEVEL_FILTER_LINEAR	LEVEL_FILTER_BASE_LEVEL
shading	SHADING_SMOOTH	SHADING_SMOOTH
lighting	LIGHTING_ON	LIGHTING_OFF
perspective correction	PERSPECTIVE_CORRECTION_ON	PERSPECTIVE_CORRECTION_OFF

Table 2: M3G options used in 3D game tests

3.1.2 3D – Fill rate test

Figure 3: Fill rate test

The Fill rate test measures the theoretical fill rate of the target device by drawing two 256*256 textures onto a triangle. The test uses two texture units if available. Rotation is used as a visual effect in the test. The result is reported as thousands of texels (texture elements) drawn on the screen per second (Ktexels/s).

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

3.1.3 3D Polygon throughput test

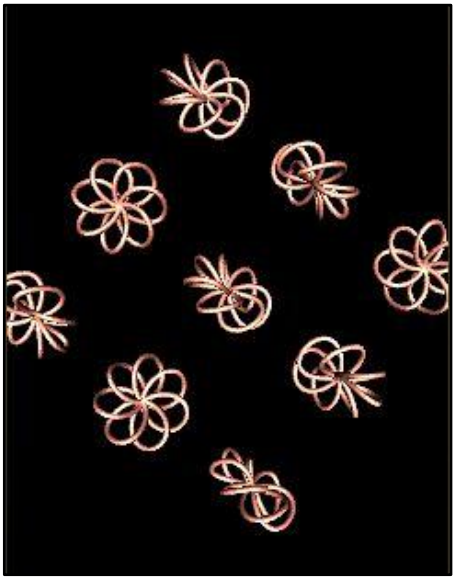

Figure 4: Polygon throughput test

The Polygon throughput test measures the theoretical polygon throughput rate by drawing rotating tori on the screen. The total triangle count in the scene is 10 000 and one directional light is used.

The result is reported as vertices processed per second.

3.1.4 2D Game test

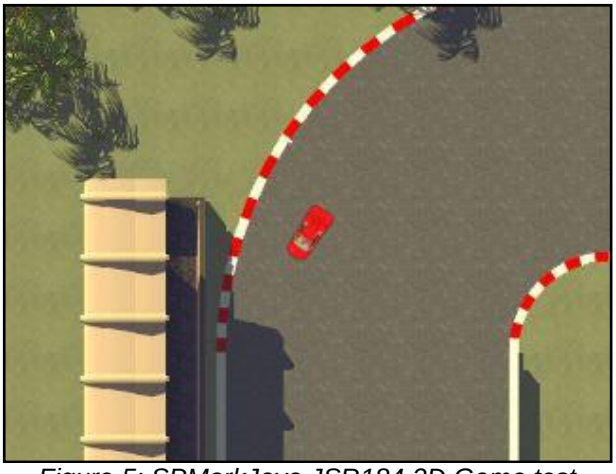

Figure 5: SPMarkJava JSR184 2D Game test

2D Gaming has been a highly popular form of gaming, especially in the handheld space. SPMarkJava JSR184 2D game test represents a racing game, which is also made playable in the Professional edition. Workload in the test is created by adding independently scrolling bitmap layers with transparency, and the race track is rendered as single, large, fully opaque

`TiledLayer` object. `javax.microedition.lcdui.game` package classes including bitmap rendering and collision detection are used heavily in this game test.

Result is reported as frames per second.

2D Game test is available as a playable game in the Professional Edition.

3.1.5 Image processing test

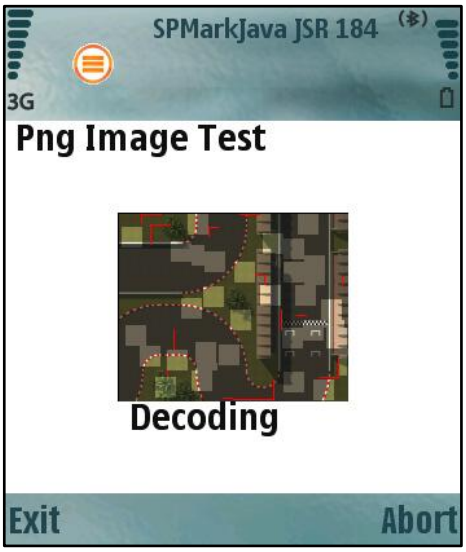

Figure 6: Image processing test

Image processing tests contain two tests: png encoding and png decoding. The test calculates a geometric mean for the tests, which is then used for the SPMarkJava JSR 184 score. The results for the individual tests are also available in the Professional Edition, reported as thousand pixels processed per second (kPixels/s).

3.1.6 JVM test

The JVM test measures the performance of the Java Virtual Machine on the device. This test can be used for e.g. comparing similar hardware solutions with differing java implementations to each other. The test suite contains a set of subtests:

- Integer addition and subtraction in various data types
- Integer multiplication and division in various data types
- conditional code (if-else-statement)
- simple object creation test
- java.lang.System.arraycopy() test
- static method point call
- floating point addition/multiplication
- sine, cosine, tangent and square root calculations
- CLDC 1.1 is required for floating point calculations.

JVM test suite results are shown as geometric mean of the tests and that subscore is used for calculating the total score. The individual JVM test results can be observed in the Professional Edition of the benchmark.

3.1.7 Video playback test

Video playback test measures whether the device is capable of running a H.263 encoded video via the JSR 135 API. The resolution of the video feed is 176*144, framerate 15fps and bit rate 40kBps. The result is reported as Pass/Fail.

The video playback test is mainly targeted for checking whether the device supports H.263 & JSR 135, as well as for use in the battery test; checking how long a video can be played through the Java API with the current battery.

The Video playback test is available in the Professional Edition of SPMarkJava JSR 184.

3.1.8 Battery Test

The Battery Test runs all the tests in a loop by default. The test loops are counted and eventually the count is reported among with the time for running the tests. The user can also select the tests to run in the Battery Test.

It should be noted that due to the nature of Java, the battery test may not always report the definite uptime of the battery, rather than the time that a Java application can be run on the device. This is because the OS may shut the Java applications down when battery level is low, but still has some power. Also, the device power management control may have effect on running the tests e.g. by switching to power save mode in the middle of running the tests. To avoid this, the power saving options should either be switched off or the timer set to maximum.

The battery test is available in the Professional Edition of SPMarkJava JSR 184.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

4 SCORING

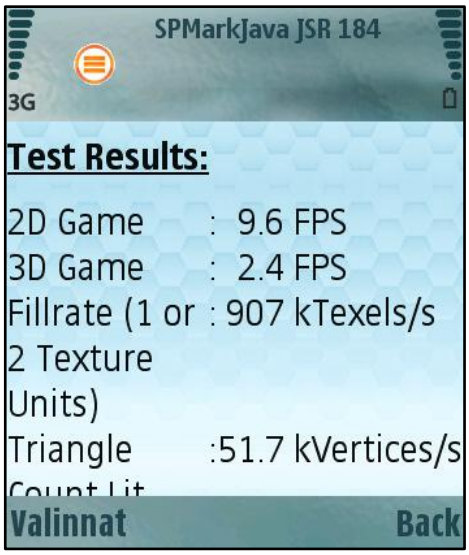

Figure 7: SPMarkJava JSR 184 results Dialogue

SPMarkJava JSR 184 provides a score for the overall performance of the device. Additional subsoccres are provided for JVM tests and image processing tests. Scores are provided only for the pre-set test suites; no scores are provided for custom test runs. The results can be seen on the screen after the test run (Professional Edition), or on the browser after submitting the results to Futuremark ORB (Basic Edition).

SPMarkJava JSR 184 Score was initially scaled to range between approximately 700 – 1500 with handsets available during the launch of SPMarkJava06 in 2005. The mobile phones of today (2007 and beyond) naturally get higher scores. Although the scale differs somewhat from Futuremark's other benchmarks, the methodology is the same: the higher the score, the better the device's performance. A couple of sample scores are shown in Table 3.

Device	Nokia® 6600(*	SonyEricsson® K750i	Samsung® U700	Nokia® N95 8GB
SPMark™Java JSR 184 Score	**N/A**	**860**	**910**	**5900**

Table 3. Example scores, SPMarkJava JSR 184 installed on phone memory.
(Nokia 6600 is not capable of running JSR 184 tests, so the score can not be calculated)*

The SPMarkJava JSR 184 Score is calculated from running all the tests. The individual test results are combined using geometric mean. The geometric mean provides a fair mechanism to combine test results as combined to assigning arbitrary weights to individual scores. The geometric mean is scaled with the multiplying factor of 17 to provide the desired range of results obtained from different reference systems. The formula for the SPMarkJava JSR 184 Score is:

SPMarkJava JSR 184 Score = 17 x (3D Game Low detail x 3D FillRate x 3D Polygon throughput x 2D game x Image processing x JVM test)$^{1/6}$.

In the SPMarkJava JSR 184 total score formula the image processing and JVM tests' results are first combined using geomean "internally", so that the weight on the total score would not depend on image processing or JVM performance too much.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

5 LIMITATIONS

While SPMarkJava JSR184 is a useful tool for measuring the performance of a Java MIDP2.0 capable device, it does have some limitations. Below is listed what we consider inappropriate uses and limitations of SPMarkJava JSR 184:

- SPMarkJava JSR 184's default test run represents a certain kind of a workload which we consider average use of a Java capable handheld device. It may not correlate to other usage models. In such cases, we recommend users create test sets of their own to see the performance figures.
- The Battery Test shows how long the workloads can be run with the current battery in the current combination of battery and OS. The power consumption of the battery is different in real world situations where the phone is not usually fully stressed during the full battery uptime. We consciously decided to fully stress the battery to avoid a test that would run for several days.

FUTUREMARK OY, KAPPELITIE 6
FIN-02200 ESPOO, FINLAND
TEL. +358 (0)20 759 8250
FAX +358 (0)20 759 8251
VAT N.O: FI1109663-4
REG.N.O: 725 396

6 SUMMARY

SPMarkJava JSR 184 is a smartphone benchmark measuring the performance of phones capable of running Java MIDP2.0 applications. It has been designed to be easy to use with clear, comparable results. SPMarkJava JSR 184 includes tests for 3D gaming, 2D gaming, image processing and JVM performance. The benchmark is available in two versions: Free Version and Professional Edition. The Free Version of the benchmark can be downloaded freely from http://www.futuremark.com/products/spmark/spmarkjavajsr184/

SPMarkJava JSR184 is created in co-operation with Futuremark BDP members including major players in the mobile industry. Development follows a neutral and transparent process to ensure the benchmark results are accurate and objective.