

Multimedia in mobile phones—The ongoing revolution

Jim Rasmusson, Fredrik Dahlgren, Harald Gustafsson and Tord Nilsson

During the past couple of years, we have seen the multimedia capabilities of mobile phones advance in leaps and bounds. If this trend holds, what role will mobile phones have in coming years? How will they be used? How will services evolve? While no prediction of the future is fail-safe, Ericsson believes that a sound understanding of what mobile phones will be capable of and of the underlying technology driving the evolution gives definition and contour to the discussion.

The multimedia capabilities of mobile phones will continue to evolve for many years—memory capacity poses no obstacle; processing performance continues to advance rapidly; and the cost of adding complex multimedia functionality is declining. Large volumes of mobile phones signify that the cost of components will fall, spurring further advancement in areas such as display and camera sensor technology.

The market has accepted mobile phones as multi-functional devices—disruptive technology—that will take the place of many traditional, portable, consumer electronic devices, such as cameras and music players.

The authors describe ongoing trends in technology that affect multimedia. They also describe trends relating to display and camera technologies, algorithms and coding. Likewise, they discuss trends in memory and silicon technology along with some of the challenges associated with design.

Introduction

The introduction of third-generation mobile terminals has been accompanied by a rapid evolution in support for multimedia. Apart from video telephony, which requires special access bearers, the primary driver of this evolution is not third-generation mobile telecommunications technology. The real driver stems from audio and imaging applications made popular through other devices, and through advancements in technology that allow for cost-effective miniaturization. It comes as no surprise that the mobile phone is an accepted device for multimedia on the go. Most people carry their phones with them at all times. Several applications can thus jointly justify end-user investments in battery, display, processing performance and memory. In this context, the mobile phone represents the ideal economy of scale. Manufacturers will continue to enhance the multimedia capabilities of mobile phones, emerging cellular standards will offer substantially larger data bandwidths, and the amount of multimedia content for phones will continue to grow. Therefore, it is safe to conclude that multimedia services over the cellular networks will evolve far beyond MMS, both in terms of advanced services and data volumes.

To understand how services will evolve and how mobile phones might be used in coming years, we must first understand

- what mobile phones can be used for; and
- how underlying technology helps drive evolution.

Ericsson's two chief considerations when designing a new generation of mobile phone platforms are market trend and technology evolution. One gains an understanding of market trends from working with and listening to operators and mobile phone manufacturers, and through market analyses. Technology evolution spans silicon technology for ASICs and memory, displays, algorithms and coding formats, and much more. The challenge of design is to find the right balance between cost, functionality and performance, flexibility, and time to market. Every trade-off affects methods development and choice of hardware, software, and tools.

Understanding costs is usually fairly straightforward. Understanding functionality and performance is more complex. It entails a knowledge of the types of applications and coding formats to be supported as well

BOX A, TERMS AND ABBREVIATIONS

2G	Second-generation mobile telecommunications technology	ISO	International Standards Organization
3G	Third-generation mobile telecommunications technology	ITU-T	International Telecommunication Union – Telecommunications Standardization Sector
3GPP	Third-generation Partnership Project	JPEG	Joint Pictures Expert Group
AAC	Advanced audio coding	JVT	Joint Video Team
AMR	Adaptive multirate	LCD	Liquid crystal display
AMR-WB	AMR wideband	LED	Light-emitting diode
AMR-WB+	Enhanced AMR-WB	MIPS	Million instructions per second
API	Application program interface	MMS	Multimedia messaging service
ARM	Advanced RISC Machines	MP3	MPEG-1 layer 3
ASIC	Application-specific integrated circuit	MPEG	Motion Picture Expert Group
BMP	Bit map	NAND	Not-and
CCD	Charge-coupled device	NOR	Not-or
CMOS	Complementary metal oxide semiconductor	OLED	Organic LED
Codec	Coder/decoder	OPA	Open Platform API
CPU	Central processing unit	OSI	Open systems interconnection
DLS	Downloadable sound	PDA	Personal digital assistant
DRAM	Dynamic RAM	PPI	Pixels per inch
DSC	Digital still camera	PoC	Push to talk over cellular
DSP	Digital signal processor	QCIF	Quarter common intermediate format
DV	Digital video	QVGA	Quarter VGA
DVD	Digital versatile disc	RAM	Random access memory
EMP	Ericsson Mobile Platforms	SRAM	Static RAM
FPS	Frames per second	STN	Super-twisted neumatics
GIF	Graphics interchange format	TFT	Thin-film transistor
GPRS	General packet radio service	UMTS	Universal mobile telecommunications system
GPU	Graphics processing unit	VGA	Video graphics array
GUI	Graphical user interface	VoIP	Voice over IP
HAL	Hardware abstraction layer	XIP	Execute in place
HDTV	High-definition television		
IOT	Interoperability test		

as camera resolutions, audio quality, power consumption, graphics performance, display size and resolution, and more. When setting requirements, one must also consider what functionality will be in use simultaneously—for example, will the phone allow users to listen to music files while playing a game and downloading a file from the network? If so, can the phone also accept an incoming call and emit a polyphonic ring signal? To arrive at the right set of requirements, Ericsson develops user scenarios that describe realistic situations in which several functions are used simultaneously.

Trends in multimedia technology

Memory capacity

Two main types of memory are found in mobile phones: non-volatile program and data storage, and fast-access random access memory (RAM), for active software. By tradition, not-or (NOR) flash memory has been used for non-volatile storage. Although this type of memory is slow to update, it is quickly read, which facilitates execute-in-place (XIP) functionality. In other words, the phone's processor can fetch code directly from memory for execution. Today, however, more and more manufacturers are replacing NOR flash memory with not-and

(NAND) flash memory, which is denser and yields substantially greater capacity from the same silicon area. In addition, NAND flash memory has a substantially lower cost per megabyte—typically one-third to one-sixth that of NOR flash memory. But because the random access time associated with NAND flash memory is quite long, it is not suitable for XIP functionality. Instead, it is more characteristic of secondary storage, much like a hard disk in a PC.

Present-generation mobile phones also have substantially more RAM than their predecessors. A primary reason for this is that users are being allowed to handle and generate multimedia content in their phones. Likewise, the transition to NAND flash memory signifies that before a phone can actively use codes and data, the codes and data must first be moved into RAM. More and more manufacturers are moving away from static RAM (SRAM) to dynamic RAM (DRAM), which is substantially denser (one transistor per memory cell as opposed to some six transistors per memory cell).

Apart from built-in NAND flash memory, many phones also accommodate memory cards, which can substantially increase memory capacity. Today, many memory cards are based on NAND flash memory.

Better image resolution and increasingly complex multimedia content require greater memory bandwidth—for example, by means of high-speed memory card tech-



Figure 1
Ericsson T68, launched in 2001.

BOX B, ERICSSON T68

MULTIMEDIA CAPABILITIES IN PERSPECTIVE:

"ERICSSON T68"

When introduced in 2001, the Ericsson T68 (Figure 1) was a highly rated GPRS terminal whose most pronounced features, in terms of multimedia, were color display and imaging capabilities. In three short years, however, the multimedia capabilities and performance of mobile terminals have improved dramatically.

Display

The Ericsson T68 had a passive, 256-color, 101x80-pixel, super-twisted neumatics (STN) display. This was quite impressive in 2001. But today, most mainstream phones sport larger displays with significantly improved visual quality and resolution.

Imaging/video

Imaging was a major innovation in the T68—it handled small JPEG, GIF and BMP images. By contrast, most phones in 2004 handle more formats with higher resolution and better quality. Many also support video.

Camera

The T68 did not come with a built-in camera but was often bundled with an accessory camera that offered rudimentary functionality. The built-in

camera solutions which are so common in today's phones often yield higher resolutions and greater functionality (for example, digital zoom).

Graphics

Many contemporary phones have accelerated graphics that enable rich, animated GUIs and swift gaming. The T68 supported two-dimensional graphics, which was adequate for the then-innovative animated two-dimensional user interface.

Music player

Music players that use high-quality audio codecs, such as MP3 and AAC, are gradually making their way into mainstream phones and will soon become a standard feature. The T68 had no such capability.

Ring signals

Polyphonic ring signals are standard fare in 2004, even in many low-end phones. Some phones even accommodate ring signals in MP3 and other audio formats. The T68 was equipped with monophonic ring signals. Overall, even though the T68 was a top performer at the time of its introduction, most of today's mainstream phones outperform it many times over.

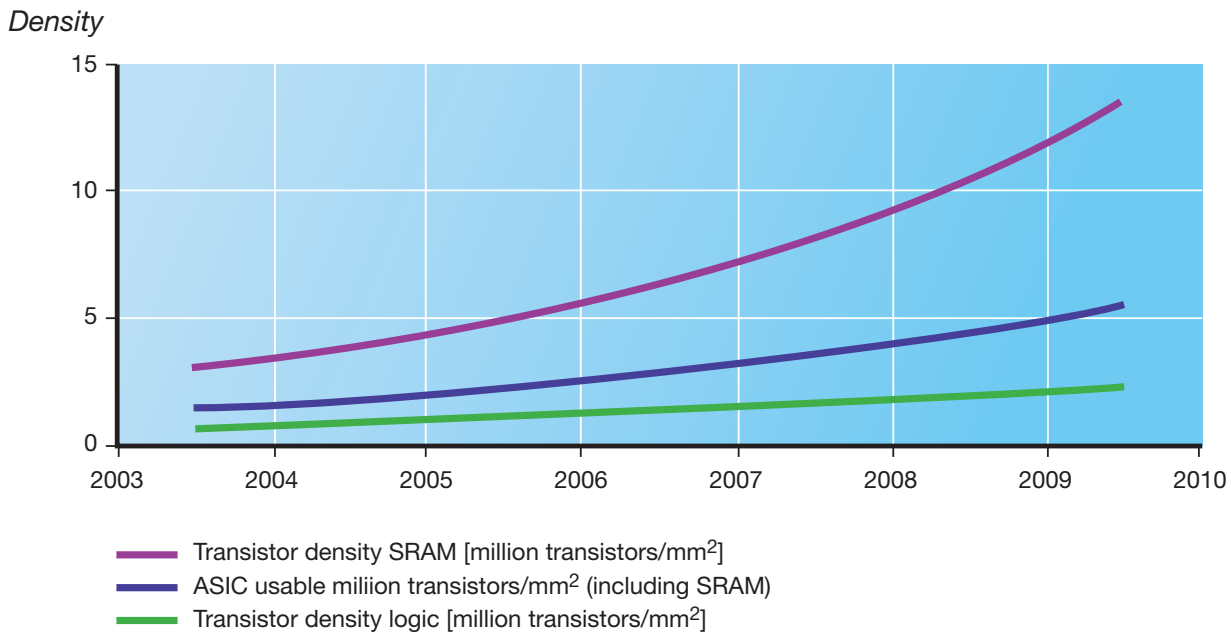


Figure 2
Expected trend in transistor density in logical circuits.

nology. Memory cards with a capacity of 2GB are now available, and 512MB memory cards cost less than USD 100. By 2007, a 2GB memory card will probably cost less than USD 100. Furthermore, microdrive technologies will soon allow for even greater memory capacity. In summary, memory capacity in mobile phones will not pose a significant obstacle for the multimedia evolution.

Processing performance

Advancements in silicon technology and processor architecture are opening the way for vastly improved CPU performance. Designers of new digital baseband ASICs for a phone platform must consider a number of important tradeoffs: cost (which often scales in terms of silicon die area), functionality, time-to-market, and power consumption.

Where flexibility and time-to-market are concerned, it makes sense to provide exceptional CPU performance. But other considerations, such as cost and power dissipation must also be weighed in. Some algorithms are demanding in terms of performance but are well suited for hardware acceleration. Typical candidates for hardware acceleration include video coding, graphics, and cryptography. Designers must thus carefully balance dedicated processing requirements

between generic CPU performance and dedicated hardware accelerators.

Advances in the area of silicon technology continue to follow Moore's Law. The International Technology Roadmap for Semiconductors (ITRS) reported that the silicon geometry of CPUs and ASICs entering into production in 1998 was 250nm; in 2000, it had shrunk to 180nm; in 2002, 130nm; in 2004, 90nm; and the projected geometry in 2007 is 65nm.¹ Ordinarily, the geometries of ultra-low-power processes appear in commercial phones one to two years after they have been perfected. Figure 2 shows the expected trend in transistor density in logical circuits. This trend points toward increasingly advanced and complex CPUs and hardware accelerators.

An additional benefit of smaller geometries is faster clock frequencies. In general, greater transistor density means greater potential for more advanced and powerful processors. There are several ways of increasing the processing performance of CPUs. Longer pipelines yield higher clock frequency, and more advanced instructions, such as DSP-like extensions (for example, the ARM9E family of processors) increase the ability to perform several operations per clock cycle (for example, multimedia extensions of the ARM11 family). Because exter-

nal memory and bus structures cannot keep up with increases in CPU speeds, more advanced cache, buffer memory, and branch predictions are used to increase effective application performance.

Display

Large and bright color displays have become a strong selling point for mobile phones, and a magnitude of features and services make good use of them—GUIs, imaging, browsing and gaming. The display is one of the most expensive components in a phone, but because it is one of the most tangible and eye-catching of all features, this cost is justified.

Display technology is evolving rapidly (Figure 3). The QVGA displays (ca 77,000-pixel resolution) introduced in phones in 2003 will become commonplace in 2005 and 2006. In Japan and Korea, for instance, QVGA displays are already standard.

The pixel density of displays in mobile phones is higher than that of displays in laptop or desktop PCs. Laptops have some 100-135 pixels per inch (PPI), whereas high-end mobile phones have between 150 and 200PPI. Some prototype displays with 300-400PPI have been developed and will arrive in the market in the form of 2- or 2.5-inch VGA (0.3 megapixel) displays. These displays will have high visual quality; graphics will appear very sharp, but most people will still be able to discern individual pixels. The resolution limit of the human eye is approximately 0.5 minutes of arc which corresponds to about 700 PPI at a viewing distance of 25cm.²⁻³ Good printers easily exceed this resolution, which is why most people prefer reading and viewing printed text and images.

Where power efficiency is concerned, the majority of dominating LCD systems leave much to be desired. Most present-day LCD systems consist of a TFT panel on top of a backlight panel. The polarizer and color filters shutter and modulate the backlight. This method of producing an image is highly inefficient, however. In fact, more than 90% of the backlight intensity is lost. Organic light-emitting diodes (OLED) take a different approach. They consist of electro-luminescent pixel elements that emit light directly. Apart from lower overall power consumption, this technology offers greater brightness and contrast and faster response times than current TFT displays. OLED display technology currently has only a small-scale presence in the market due to issues with aging, manufacturing yields and cost.

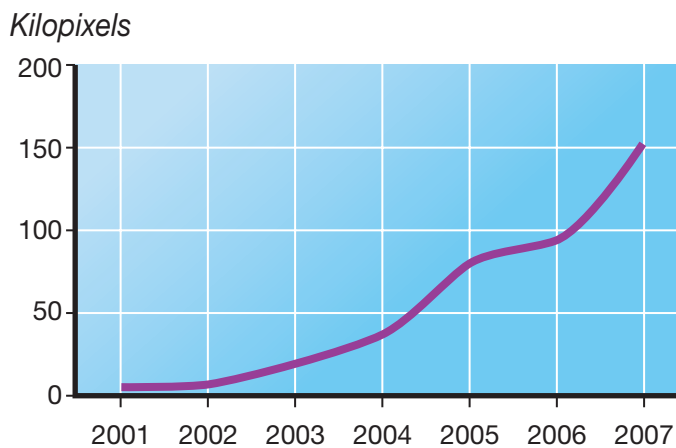


Figure 3
Trend of display resolution in mid-tier mobile phones. (Source: EMP)

Cameras and imaging

In two short years, built-in cameras have become a must-have feature of mobile phones. And, as with display technology, digital camera technology is evolving very rapidly. In 2004, the Japanese and Korean markets introduced the first mobile phones equipped with 3-megapixel cameras. Other markets are expected to follow suit in 2005 and 2006.

Camera phones usually contain a rather large suite of imaging features for enhancing and adding effects to images. These features, which include brightness, contrast, color, zooming, cropping, rotation and overlay, can be used while a still image or video clip is being shot or afterward, for instance, to spice up MMS images.

The real-time processing of megapixel resolution images is demanding and often requires hardware acceleration to assist in image compression/decompression, color space conversion, scaling and filtering. As costs continue to fall, many of the standard features associated with dedicated digital still cameras (DSC) will show up in mainstream mobile phones—for example, multi-megapixel sensors, flash, autofocus and optical zoom.

Two image-sensor technologies currently dominate: complementary metal oxide semiconductor (CMOS), and charge-coupled device (CCD). Compared to CMOS, CCD technology generally offers better sensitivity and signal-to-noise levels, but it is also more expensive and power-hungry. This technology has mainly been reserved

for high-end megapixel camera phones. CMOS technology has been more common in sub-megapixel cameras (such as popular 0.3 megapixel VGA cameras). In terms of resolution and sensitivity, however, it is fast approaching CCD, and many new CMOS-based multi-megapixel camera phones will be introduced in coming years.

Greater pixel density is a common means of achieving multi-megapixels at reasonable cost. But because pixel spacing is now well below 3 μm it is becoming increasingly difficult to maintain good performance of critical design parameters, such as sensitivity, signal-to-noise, dynamic range and geometric accuracy. The trade-off between pixel count and quality will not be trivial.

Image quality is not solely a matter of megapixels. Every aspect of digital camera imaging is being examined: small-lens systems for mobile phones will improve optical quality and offer autofocus and optical zooming functionality. Camera signal processing, such as color interpolation, white balancing, sharpening, and noise reduction, will improve with better algorithms and high-precision hardware. Flash systems will also become more efficient and powerful thanks to solutions based on white LED and Xenon discharge technologies.

Technical advances in each of these areas will yield vastly improved image quality. So much so, that many people with cost-optimized camera phones will accept the quality for their family photo albums. In all likelihood, we will also see high-end camera phones featuring full-fledged camera systems on a par with dedicated DSCs.

This rapid progress of camera technology in mobile phones might result in a case of disruptive technology where camera phones take over the role of entry-level DSCs. Indeed, citing this very scenario, one major vendor of DSCs has already dropped out of the low-end DSC market.

The standardization of imaging and camera-control functionality is underway in the Java standardization community. JSR-234, Advanced Multimedia supplements for J2ME, address this.

Video

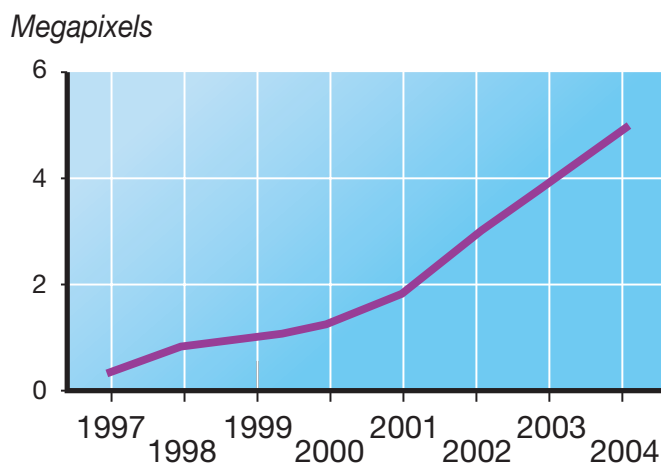
Video telephony and video streaming are two of the crowning features of 3G phones, but video capability is now also showing up in 2G mobile phones. When first introduced, the resolution and quality of video capture pretty much matched that of the phone's display. At the time, captured video images were only intended to be shown on the same phone or sent to another phone in the form of a video-telephony call or MMS. But given the rapid evolution of video technology, in a few years the video capabilities of many phones will probably be close to that of today's (2004) mainstream digital video (DV) camcorders. Notwithstanding, the requirements that video puts on computational capacity (million instructions per second, MIPS), memory size and memory bandwidth far exceed the requirements of still imaging. Therefore, to offload the CPU and save power, it pays to employ hardware acceleration. For video resolutions greater than QCIF (176x144 pixels), hardware acceleration is more or less a necessity (given current CPU performance).

Video compression

Obviously, interoperability between devices from different manufacturers as well as services provided by different operators is paramount for applications such as video telephony and MMS. The Third-generation Partnership Project (3GPP) stipulates which codecs may be used for services in UMTS: H.263 is mandatory; MPEG-4 and H.264 are optional. Because a good deal of available content has been coded in RealVideo (RV) and Microsoft Video (WMV), support is also being considered for these proprietary formats, in particular, where viewing or browsing are concerned.

Historically, two organizations have contributed toward the standardization of video codecs. The ISO Moving Pictures Expert Group developed MPEG-1, MPEG-2 and

Figure 4
The resolution evolution of image sensors for mid-end digital still cameras. (Source: EMP)



MPEG-4, which is used for VideoCD, DVD and HDTV. ITU-T developed H.261 and H.263, mainly for video-conferencing. In 2001, the two organizations formed the Joint Video Team (JVT) to develop a new recommendation or international standard targeting greater compression efficiency. In 2003, JVT announced ITU-T H.264 and Advanced Video Coding (AVC), Part 10 in the MPEG-4 suite.

More efficient compression (thanks to the H.264 codec) will improve perceived video quality, especially video telephony and streaming video at bit rates as low as 64kbps. However, these gains in compression efficiency are not free. They call for a considerable increase in computational complexity and memory, which adds to the overall cost and energy consumption of the phone. Because decoding has less affect on performance than encoding, and because the ability to consume emerging content is a top priority, we can assume that phones will initially only decode H.264. Later, support for encoding H.264 will also be added.

Graphics

The graphics subsystem of a mobile phone is involved in all display-related actions—it prepares, manipulates and blends data to be shown on the display, for example, user interface elements and windows for video and imaging.

The dominating graphics technology for GUIs and gaming in mobile phones has been two-dimensional bitmap graphics. In 2003, however, three-dimensional graphics was introduced into some high-end mobile phones. Eventually this technology will be offered in all mainstream mobile phones. Three-dimensional graphics is mainly used for gaming, screensavers and animated three-dimensional GUIs. Two-dimensional vector graphics has also been employed in some phones. This technology will be increasingly important for resolution-agnostic two-dimensional content and GUIs (defined using vectors instead of bitmaps).

Khronos, an open-standards body with more than 60 member companies from the embedded industry, is standardizing graphics APIs.⁴ One outcome of this work is the increasingly popular OpenGL ES 3D graphics low-level API, which was finalized in July 2003 and showed up in products roughly a year later. An updated version, OpenGL ES 1.1, adds functionality that better exploits hardware implementations. An OpenGL ES 2.x track, started in 2004, will

address programmability in the three-dimensional pipeline (Figure 5). Compared to fixed-function pipelines, programmability introduces a new dimension of flexibility into the graphics pipeline and enables the use of procedural algorithms for enhanced visual quality and effects. Programmability has been exploited successfully in recent computer games and will probably also be important for mobile phones. OpenVG standardizes a low-level two-directional vector graphics API. The Java standardization community is also very active and has produced high-level two- and three-dimensional graphics APIs for J2ME: JSR-184 (M3G) and JSR-226.

Three-dimensional graphics for real-time interactive gaming is very demanding. It makes extensive use of many subsystems inside the phone. Apart from the graphics subsystem, it uses the CPU, buses, and memory. The challenge, especially given increases in display resolutions, is to achieve high-performance three-dimensional gaming without consuming a lot of power. Hardware acceleration will increase performance and power efficiency many times over. At present, most three-dimensional solutions on the market are software implementations, but hardware-accelerated solutions are certain to become commonplace in a few years.

The popularity of computer games has pushed the evolution of PC graphics performance—in five years' time, pixel fill rates have increased 1000-fold and gaming is also sure to influence the evolution of graphics in mobile phones.⁵ Notwithstanding, there are fundamental differences between a personal computer and mobile phone—power consumption, size and cost, for example. These distinctions stipulate that graphics subsystems in mobile phones must be designed and optimized with emphasis on a different set of requirements.

Audio

More is the operative word in current audio trends—more codec formats, more synthetic audio formats, more audio effects, and more simultaneous audio components. New use-cases and competing codecs are behind the drive for more audio codecs. At present, the most common audio codecs are MP3, AAC, RA and WMA. The trend in audio codecs is for greater support of low bit rates (Figure 6). At the same time, voice codecs, such as AMR-WB, are evolving to provide support for general audio at bit rates that are

Figure 5
Typical three-dimensional graphics pipeline and associated stages.

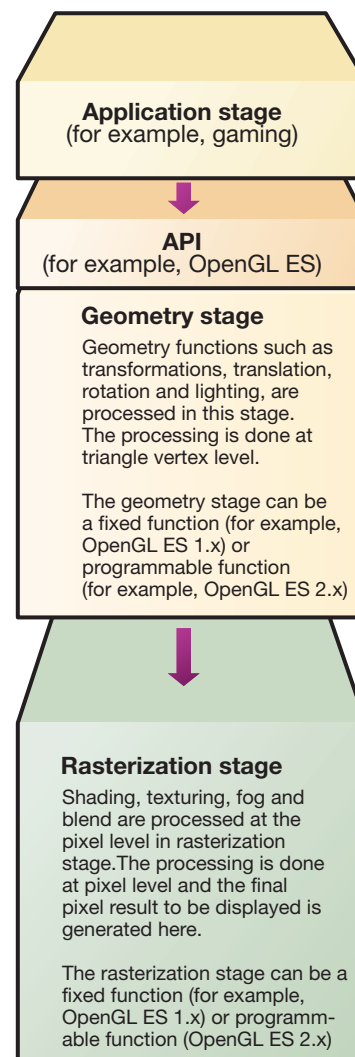
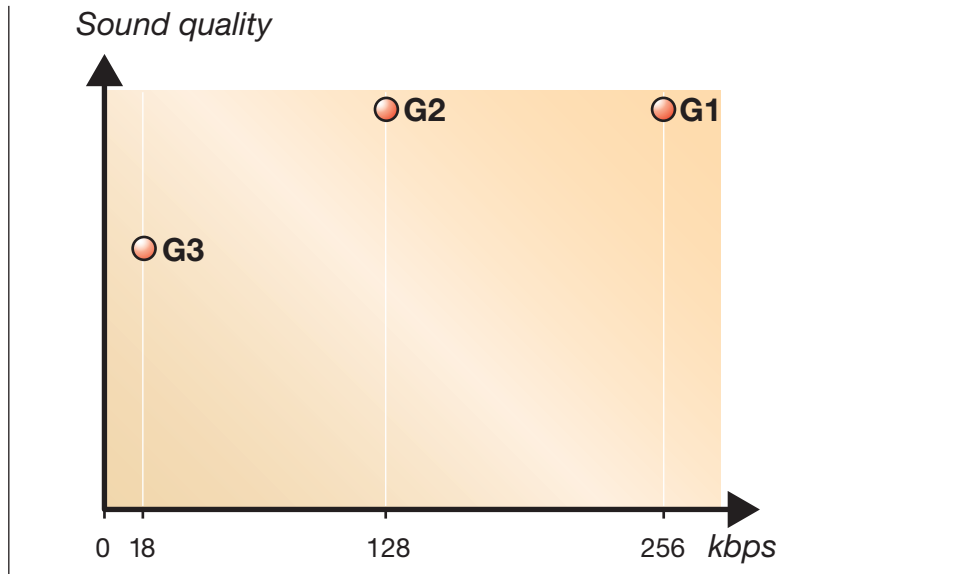


Figure 6
The current trend for audio codecs. Sound quality has remained at CD quality between first- and second-generation audio codecs, but the bit rate has dropped 50%. Third-generation audio codecs are aiming for FM radio sound quality at a low bit rate.



economically reasonable for streaming and messaging. One prominent example is AMR-WB+.⁶

A new format for synthetic polyphonic audio is mobile downloadable sound (DLS), which allows users to customize synthesized sound. For example, with DLS, users could add extra instruments with a sound that is specific to, or characteristic of, a given melody.

The current generation of mobile phones uses audio effects, such as equalizers and dynamic range compression. These effects alter the frequency curve and suppress high volume sounds to render better quality sound. New effects being introduced include chorus, reverberation, surround sound and positional three-dimensional audio. As its name implies, the chorus effect makes one sound signal sound like multiple sound signals. Likewise, the reverberation effect imitates the reflection of sounds off walls.

Most current-generation phones support stereo headsets and recently phones with stereo speakers were introduced. The surround-sound effect has been introduced to enhance the stereo listening experience. The positional three-dimensional audio effect makes it possible to move sound sources in a virtual three-dimensional space so that listeners perceive sound sources as if they are coming from a specific direction and distance.

Java standard JSR-135 enables J2ME devices to control audio players and radio tuners. JSR-234, which is projected to be finished by the end of 2004, extends the

audio functionality support for the effects described above.

The trend for more simultaneous audio components has its roots in multimedia and games. Multimedia file formats that contain multiple tracks of coded audio, synthetic audio and audio effects are becoming increasingly popular. This is especially true for ring tones. Likewise, gaming and other advanced use-cases have multiple individual audio sources. A three-dimensional audio and graphics game has many objects that emit sound from specific positions to create a virtual world. Each of these simultaneous audio sources and advanced audio effects exploits advancements in technology performance.

Voice

At best, the voice quality of current-generation mobile phones equals that of the fixed telephony network. However, the audio spectrum supported by next-generation mobile phones will be wider than the audio spectrum of fixed telephony networks. This will yield a more natural sounding voice signal and quality than that of fixed telephony. An improved voice codec, higher bit rates, and improved speech-enhancement methods have been used to accommodate the wider spectrum. The AMR-WB voice codec, which covers nearly twice the spectrum of AMR, encodes voice sounds within the spectrum of 50Hz to 7kHz (Figure 7).⁷ A prerequisite for using AMR-WB is that each of the phones involved in a given call must support it. This, in turn, requires tandem-free or

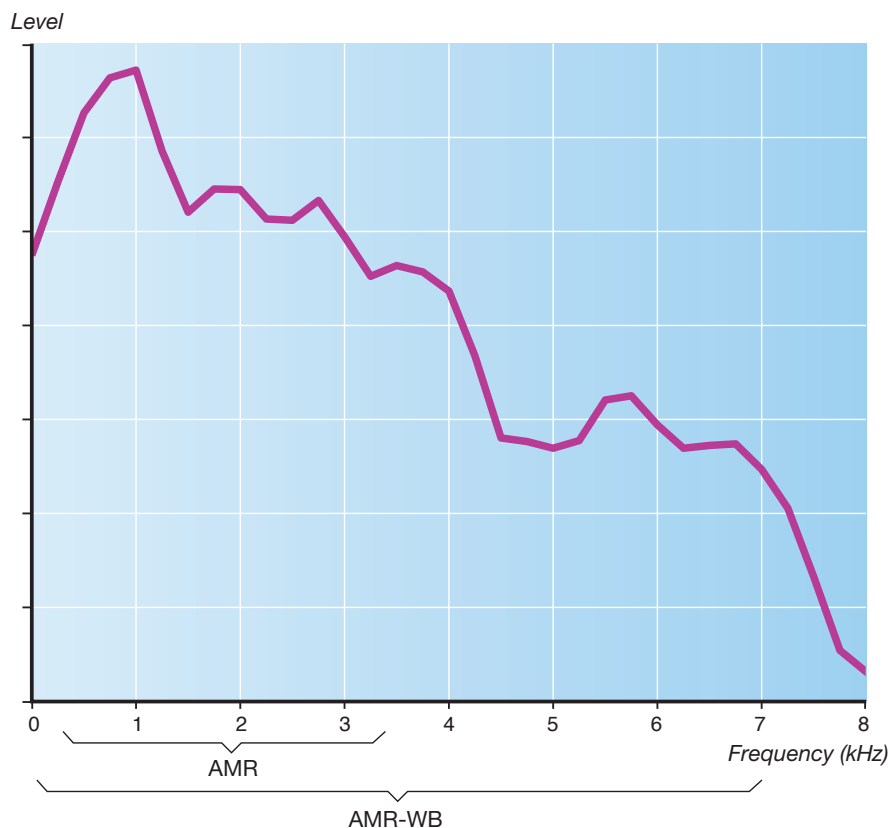


Figure 7
Speech signal spectrum. The AMR voice codec supports an audio spectrum of 300-3400Hz; AMR-WB supports a broader spectrum of 50-7000Hz.

transcoder-free operation in the network—that is, the coded voice must not be re-encoded in the network.

Network operators currently use speech-enhancement functions in the network, such as echo cancellers and noise reduction. They apply these functions to decoded voice signals before re-encoding them. Mobile phones also employ speech-enhancement functionality to reduce echo and surrounding noise for far-end listeners. Tandem-free operation limits the support of network-based speech enhancement, strengthening the case for adequate speech enhancement in phones. The more severe acoustic environment of video calls further heightens the importance of speech enhancement—users will probably increase speaker volumes and the microphone pickup of user voices will diminish.

A new voice service, Push to talk over Cellular (PoC), is based on IP transmission of the voice signal without using a circuit-switched connection.⁸ PoC is a walkie-talkie type of service that connects multiple users.

As this trend continues, we will see services like full-duplex voice over IP (VoIP) and multiple voice sessions over IP. Full-fledged VoIP services will not affect the end-user—that is, end-users will perceive them as they would any other voice call.

When VoIP is extended to support multiple simultaneous sessions, it will be possible to introduce new services with full-duplex communication to multiple users. Likewise, it will be possible to post-process each user separately. For example, with positional three-dimensional audio, one can place each participant at a distinct position in space. Doing so helps listeners to separate participants from one another.

Putting it all together—system design challenges

Mobile phones in the market today contain the features and functionality of various portable devices—phones, digital cameras, music and video players, gaming consoles, messaging clients and personal digital assis-

tants (PDA). Mobile phone hardware and software has thus become very complex. This, in turn, puts huge demands on testing and verification. For example, to guarantee compliance and stability, each software platform is verified to work in well over 10,000 test cases.

EMP's mobile phone platforms are type-approved. Considerable interoperability testing (IOT) and standards-specific tests are conducted to guarantee compliance. Increasing consumer demand for greater functionality, better performance and a larger degree of integration means that already-tough requirements for a stable, efficient and flexible implementation will become tougher still. Obviously, to succeed, manufacturers need a solid platform architecture.

Hardware architecture

The multimedia capabilities of a phone determine

- its position on the scale between entry-level and high-end; and
- what customers will pay for it.

The size and cost of the hardware needed to process multimedia has thus been allowed to swell. Indeed, it accounts for a significant part of hardware costs, especially in mid-tier and high-end phones. Besides size and complexity considerations, the rationale for completely separating multimedia processing from the radio modem and voice-processing subsystem into an independent application subsystem has grown. Doing so permits independent verification and makes it easier to support combinations of mobile telecommunications standards and multimedia capabilities.

The application subsystem consists of one or two multimedia processor subsystems and several dedicated hardware accelerators. The processor subsystems include data and program caches and tightly coupled fast memory configured to minimize cost and overcome bottlenecks that arise from a lack of available bandwidth to external memory. Dedicated hardware accelerators handle specific functions, which for reasons of power or performance, cannot be efficiently executed in general-purpose programmable CPUs or DSPs. Examples are motion estimation (video), sound synthesis (audio), and three-dimensional graphics rasterization. Complete, dedicated subsystems including memory might also be included—for example, MPEG-4 video encoders/decoders or two- and three-dimensional graphics engines. One cost of dedicated hardware ac-

celerators is additional silicon area. A flexible architecture is thus needed to support a consistent and well defined hardware abstraction layer (HAL) and associated API. That way, hardware accelerators can be added and removed with minimum impact on higher level software. An additional benefit is that a common code base can be maintained across different platforms, allowing specific application software (such as a video decoder) to run on

- an entry-tier platform without the need for hardware acceleration support; and
- a high-tier platform with dedicated hardware accelerators.

Obviously, the performance (frame size and frames per second) of phones with and without hardware acceleration will vary greatly.

Software architecture

A high level of abstraction with a clear and well-defined structure is required to handle the software complexity described above. The software architecture in platforms from EMP completely decouples the platform software from customer application software. This way, the customer software can be developed independently and reused for different platform configurations and future platforms. Application software uses the Open Platform API (OPA) to access platform functionality, such as call setup, streaming services and music playback. A hardware abstraction layer decouples hardware-dependent software from other software, which greatly facilitates the process of introducing new hardware, such as multimedia accelerators. Ericsson has designed its mobile platforms to allow native or other execution environments, such as Java, to run on top of OPA.

Figure 8 shows a schematic view of the software architecture. The software is divided into layered stacks per area of functionality. Ordinarily, each layer in each stack consists of several software components. This layered approach, which resembles open systems interconnection (OSI) layers in communication stacks, permits distinct abstraction between layers of functionality, reduces dependencies between software components, and speeds up platform reconfigurations, development, and integration of new software components.

The software architecture also introduces the Ericsson Mobile Platforms Component Model, which the platform-management software employs to integrate and register new software functionality.

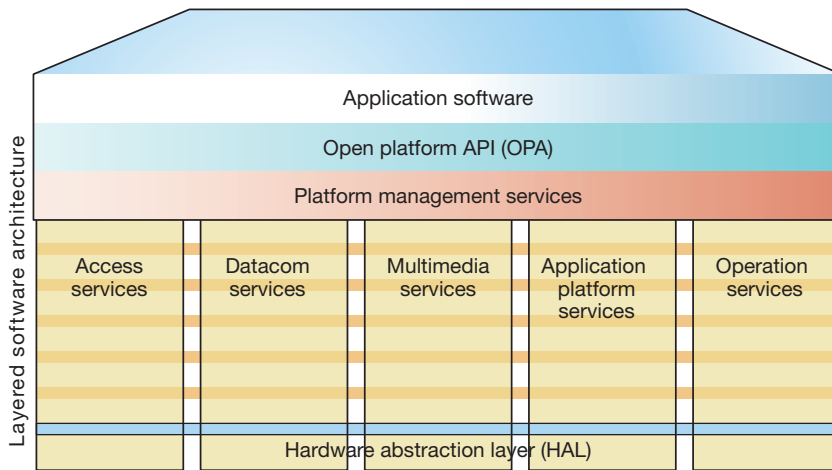


Figure 8
Ericsson Mobile Platform software architecture.

Conclusion

Since 2001, we have witnessed a rapid increase in the multimedia capabilities of mobile phones. The market has shown that the mobile phone is a natural multi-functional device, which in all likelihood, will serve as disruptive technology for many traditional, portable consumer electronic devices, such as digital cameras and music players. Current trends in technology will provide solutions for pushing multimedia capabilities forward at the same rapid pace for another four or five years.

From a hardware viewpoint, it is clear that the amount of memory required for storing and processing multimedia content will not stand in the way of mobile phones becoming rich, multimedia-centric devices.

Processing performance is certain to increase thanks to faster clock frequencies and more advanced CPUs.

Baseband processing chips will host complex hardware accelerators for dedicated, performance-demanding algorithms.

Advancements in display and camera technologies will make their way into phones, and algorithmic and coding trends will pave the way for more and richer multimedia functionality. This and coming advancements in mobile network multimedia services will open the way for enriched content sharing, more advanced streaming and broadcast services, and much more.

Ericsson has been part of the technology evolution since the beginning of the mobile phone era, driving advancements in cellular access technologies, through standardization as well as research and development. Ericsson's mobile platform products are well prepared for the upcoming multimedia evolution described in this article.

TRADEMARKS

RealVideo is a trademark or registered trademark of RealNetworks, Inc.
Windows Media is a registered trademark of Microsoft Corporation.

REFERENCES

- 1 <http://public.itrs.net>
- 2 Foundations of Vision, Brian A. Wandell, Sinauer Associates, Inc., 1995
- 3 Eye, Brain and Vision, David Hubel, Scientific American Library, 1988
- 4 <http://www.khronos.org>
- 5 David Kirk, Chief Scientist Nvidia, in "GPU Gems", Addison-Wesley, 2004.
- 6 3GPP TS 26.290: Audio codec processing functions; Extended Adaptive Multirate - Wideband (AMR-WB+) codec; Transcoding functions
- 7 3GPP TS 26.171: AMR speech codec, wideband; General description
- 8 Medman N., Svanbro K. and Synnergren P., "Ericsson Instant Talk," Ericsson Review, Vol. 81(2004):1, 16-19, 2004
- 9 JSR-135: <http://www.jcp.org/en/jsr/detail?id=135>
- 10 JSR-184: <http://www.jcp.org/en/jsr/detail?id=184>
- 11 JSR-226: <http://www.jcp.org/en/jsr/detail?id=226>
- 12 JSR-234: <http://www.jcp.org/en/jsr/detail?id=234>