1.3 Cellular Phones as Embedded Systems

Yrjö Neuvo

Professor; Executive Vice-President, CTO, Nokia Mobile Phones; Member of Nokia Group Executive Board, Finland

I. INTRODUCTION

A. Market development

In the 1990s wireless technologies developed in a way that can best be described as 'unbelievable'. Consequently, the number of mobile subscribers has grown over 100-fold in the last ten years, as shown in Figure 1. Currently, there are over one billion subscribers globally and the number is growing at a healthy pace. Current scenarios predict another 400 million subscribers by 2005, leading to a total of some 1.6 billion subscribers. Continuing with scenarios for mobile voice, there is also a big opportunity in the growth of mobile traffic as fixed voice calls are increasingly being replaced by mobile calls. By 2008, there is a good chance that the current number of global mobile subscriptions will double.

The production volumes of cellular phones are remarkable when compared against other consumer products. In 2002 approximately 405 million cellular phones were sold worldwide, whereas during the same time the worldwide shipment of personal computers (PC) was 136 million units and personal digital assistants (PDA) 12 million. Even the total production of vehicles, 57 million units, is almost a decade behind that of cellular phones.

B. Cellular generations

Radio systems continue to develop from the current second generation (2G) and evolved 2G systems toward higher data rates and better mobility. While the third-generation ([1]-[6]) cellular system is being rolled out and the first terminals have reached the shops in Japan and Europe, standardization work continues towards even higher data rates. Several access methods complement 3G, e.g. the wireless local area network for high-speed hot spot data connection and Bluetooth for short-range radio connection. Furthermore, discussions and research on the next-generation cellular system, 4G, have already started [7]. 4G is expected to enable versatile mobile broadband services by providing wide coverage, ultra-high bit-rate radio access. In the multiradio environment, the services are required to be radio-agnostic. This is best achieved by utilizing the Internet protocol (IP) convergence as the unified connection layer. The evolution of radio systems toward higher data rates and mobility is illustrated in Figure 2.

C. Digital convergence

The digital industry is experiencing rapid convergence of parts of consumer electronics, communication, information technology, media and entertainment industries. The convergence enables people to create, share, and consume digital content using interoperable devices. The cellular phone, the main device that people always carry with them, is the key platform for the mobile convergence applications, e.g. web browsing, imaging, and high bit-rate video streaming. Mobility means that the mobile Internet will bring us new features to facilitate the way we live and work [8].

With the convergence, cellular phones will evolve from the traditional cost-optimized handhelds to multifunctional terminals in a variety of form-factors. The clear trend is toward smaller terminals with new applications and user interfaces providing a continuously improving end-user experience. The objective of this article is to discuss the key challenges of future cellular terminals. The architecture development and opportunities are presented for both the wireless access and application engines. The main challenge, however, lies in the management of the growing system complexity. The most potential methods and technologies to tackle the complexity challenge are presented.

II. TODAY'S CELLULAR PHONE

Cellular phones are extremely complex embedded systems where all functional blocks are custom-made for mobility. The combination of miniaturization and functionality is unprecedented compared to other consumer products. A block diagram of today's typical classic category triple-band cellular phone is shown in Figure 3. The engine consists of three main chips: one for RF, one for baseband, and one for mixed-signal and energy management. The chips are either proprietary designs or based on available chip sets. The product includes a total of 345 parts. The characteristics of the cellular phone shown in Figure 3 are represented in Table I.

III. ARCHITECTURE CHALLENGES FOR WIRELESS ACCESS

New implementation architectures enable the adoption of new technologies. The new technologies, then, bring improvements in product requirements, such as performance, miniaturization, power consumption, reliability and cost. Several successful development phases in the past laid the foundation for improvements also in the future.

A. Past architecture development for wireless access

In wireless access implementation there have been several major technology steps during the last 15 years of cellular phone development. The steps have provided clear improvement in miniaturization and functionality, as shown in the wiring board-level engine development illustrated in Figure 4. In the first phase, when RF and baseband circuitry were implemented on separate printed wiring boards (PWB), the main challenge was to embed the whole engine on the same wiring board. The successful integration based on careful design of the wiring board and shieldings reduced the size and cost of the cellular phones dramatically and enabled the first real handportable devices.

The next major step was the integration that took place within both the baseband and RF. DSP and MPU functionalities were integrated together with logic circuitry on a large baseband ASIC. In the RF side, the discrete superheterodyne RF was mapped onto Silicon RF ICs. Again, significant size and cost benefits were achieved.

But in the RF front-end, it was the next architecture, the direct conversion, that really benefited from the silicon implementation. Although the direct conversion receiver architecture is very old, it was not feasible for mass-production for a long time due to DC offset and signal self-mixing problems. The architecture became practical only in mid 1990s when RF IC technology had matured to include sufficiently fast and homogeneous transistors and the quality of design models and tools had improved significantly.

Direct conversion architecture provided huge benefits in component count reduction: several filters and synthesizers were removed and the total silicon area was reduced. Furthermore, when dual- and triple-band transceivers became a requirement, the direct conversion architecture proved to be a flexible platform for the multi-band operation: frequency planning is easy, and several functional blocks can easily be reused for many bands. Figure 3 shows a typical direct conversion architecture of a multi-band receiver.

B. Current baseband and RF challenges

With the high data rates, the processing power requirement and the complexity continue to grow. Today, the baseband ASICs that include the processor core or cores are not only limited in performance by the computational power, but also by the on-chip communication, line delays and clock distribution. In order to reduce the load from the processing cores, decentralized architectures are being considered. Then, small controller processors

0-7803-8267-6/04

or units running at low frequencies can be distributed within the engine. With multiple processors, management of the processing power resources becomes one of the key challenges.

An alternative approach to optimize the baseband is to utilize configurable logic, configurable processors or processor generators. They provide more freedom in optimizing the trade-off between performance and flexibility. Still, due to the exploding baseband complexity of cellular platforms, the main challenges have to do with design methodology, verification, and testability. In the RF domain, digital CMOS technology has been the hottest topic in recent years. The technology has proven to be applicable also for the RF, but it suffers from several disadvantages compared to the widely used RF BiCMOS. In particular, substrate-coupling effects are stronger and the quality of transistor models is not at a sufficiently mature level for first-time success in RF IC design. The cost calculation of RF CMOS reveals that a one-to-one replacement of the RF circuit with a CMOS version provides only marginal, if any, cost benefit. The calculation changes immediately when large digital content, e.g. digital filtering and control logic, is integrated on the same chip.

C. Single chip?

To fully exploit the speed of the latest digital CMOS technology in RF, completely new circuits and architectures need to be invented. In contrast to just mapping the traditional RF circuits to CMOS, the new architectures should be based on fast sampling and time-discrete signal processing. Then, the interface between the RF and baseband becomes blurrier and the portability of the RF front end to the future CMOS process nodes improves dramatically. Whether this will be the next cornerstone in the architectural development paving the way to true single-chip radios is not yet known. In any case, these new digital radios will establish even more pressure for developing better system and behavioral modeling tools for complete engine optimization.

D. Multiradio

With the wide variety of radio systems, terminals offering access to several systems are required. Access to all available systems—multi-band GSM, WCDMA, CDMA2000, Bluetooth, FM radio, GPS, WLAN—is technically feasible. In practice, there are two main questions about multiradios yet to be answered: (a) For what combinations of systems can the radios be integrated with a reasonable cost and size? (b) What combinations of systems make sense in the market place?

In the baseband domain, the required functions for the different radio standards are so much alike that today's baseband already has the capability for multiradio processing. Due to the different air interface specifications, the RF section of a multiradio can easily become very complex with a large number of parallel circuit blocks. In multi-standard or multi-band RF ICs, programmability of circuit blocks can be used for complexity and silicon area reduction.

Multiradios introduce also miniaturization challenges to the antennas and front-end filters. The antennas need to be designed as compact multi-system antenna modules and filter miniaturization is crucial. Bulk-acoustic-wave technology (BAW) is one key enabler for filter miniaturization. Micro-electro-mechanical systems (MEMS) technology still requires proof for RF applications.

IV. APPLICATION ENGINE

Considering applications, three different computing platforms have had remarkable success during recent years. The most powerful platform has been used in PCs, and the lighter one in PDAs. Typical cellular communication products have provided the lightest application set, as the cost, application needs and available technologies of the cellular engine have dominated the entire concept. However, the Nokia Communicator launched in 1996 can be regarded as a pioneering product in combining communication and PDA functionalities. That evolution is proceeding well and the application space is becoming richer all the time in all cellular product categories.

A. Application engine implementation

As shown in Figure 3, GSM/GPRS computation or the whole protocol stack and application functions have been implemented with one ASIC having a DSP, MPU, logic, on-chip memories, and external memory and control interfaces. In the future, due to multi-standard communication and multimedia applications, platforms will be more complicated. Part of the complexity comes from higher data rates within wireless access devices, displays and cameras, for instance. Standardizing interfaces is necessary to ease the handling of complexity and the enabling of a large variety of HW extensions. For example, MIPI alliance has been launched to orchestrate the interfaces between application core and wireless access and user interface components [9]. In general, standardization must be organized to focus on the core areas of the platform.

The clock speed rate of an MPU is not equal to the overall performance in a cellular device. When the MHz and power consumption trends of the processors used in PDAs or especially PCs and cellular devices are compared, growing absolute gaps in both parameters can be seen, as shown in Figure 5. When communication services with 3G or WLAN are used, cellular devices with a volume of about 100 cc cannot withstand too high power levels caused by the application processor. On the other hand, HW acceleration strategy plays a significant role. Which functions will be implemented with HW must be carefully considered; in this respect, embedding techniques and solutions play a key role. With the improvements in display resolution and color space, 3D graphic accelerators could be one way to optimize overall performance, for instance.

B. Memory Challenge

With the convergence terminals, in particular, the total memory requirement is increasing rapidly. Several different memory chips and their interconnections consume large areas on the PWB and a lot of data transfer is taking place between the embedded memories and multiple memory chips. In the future, nonvolatile RAM (NVRAM) may challenge the flash memory chips. NVRAM is a universal technology that is expected to require fewer memory chips, less PWB, fewer I/O, and lower power consumption. The most mature NVRAM technologies are ferroelectric RAM (FeRAM), magnetoresistive RAM (MRAM), and ovonics unified memory (OUM). In addition to the technological challenges, the new NVRAM technology introduces challenges in memory management and usability issues.

The focus in memory technologies will be on denser mass memories. There is a trend from communication-centric memory architectures to larger mass memories supporting e.g. data downloading or local storage. Applications like 1-Gbyte game or movie downloads, or one-hour of high-quality video recording would require more flash memory. At the moment, memory cards like MMC or SD can be considered good mass memory solutions, as e.g. gigabyte mass memories are still too expensive. Large memory space in a cellular phone provides immediate access to the user's data, music or games.

C. Application platform

Everyday life is becoming more and more mobile. In part, this means that the mobile Internet will bring us new features that will help us in the way we live and work. For example, web browsing, video calls or high bit-rate streaming are starting to be available. Data traffic will be handled more with IP packets, and IPv6 enables Internet devices to move from place to place while offering seamless or transparent Internet connectivity to all applications.

Java is the desired platform that enables the deployment of mobile handset value-added services for service and content providers. The specific version of Java has been developed to meet the inherent resource constraints of cellular handsets. In addition, support for multimedia and many other features have been added. Java is in the primary platform for deploying third-party mobile applications when the applications are targeted to run on a plethora of devices from different suppliers. The developer communities around Java and Symbian have grown significantly. In this environment, coordination between different manufacturers using Java is very beneficial for compatibility of cellular applications.

The run-time environment part of Java or virtual machines can be implemented with HW or SW; there are pros and cons with both approaches. SW–based solutions provide flexibility and modularity, and development can be independent of the HW. However, this approach leads to challenges with memory requirements in long-term evolution, as the number of optional packages in SW will grow. With HW-based solutions, power can be saved and less RAM is needed, but SW becomes more dependent on inflexible HW implementation. Nevertheless, Java should be taken into account when optimizing processor and cache architectures, and more research is still needed with the HW and SW approaches.

Rich applications are based on evolution in technologies, like displays, cameras, wireless access and data processing power. Many of these applications are part of multimedia and require diverse support from the HW and SW platforms. For example, video codec is needed to support video applications, meaning significant additional data processing requirement for the baseband platforms. The codecs must be standardized to guarantee compatibility between different products. Multimedia evolution comprises a multitude of new applications, like snapping pictures, video streaming or video clips, and real-time video calls. MMS plays a key role in user-experience evolution. In Table II the data rates, picture sizes and frame rates of cellular generations for video streaming are presented.

The main target with multimedia is to provide features that help people to share experiences, business data, and moments of life in a convenient and easy way. One important aspect is the quality of the features. For example, casual photo snapping will evolve towards photographic and printable high-quality pictures, and video clip resolution will eventually achieve a level that makes large-screen viewing a reality. All these applications require large memory storage capacities; therefore, mass memories will be needed in handheld products in the future.

V. TACKLING SYSTEM COMPLEXITY

Besides communication services, new PDA applications – starting with the simple calculator or calendar type of applications – have continuously evolved. Multimedia applications and games, for instance, will continue this evolution and put plenty of demands on the processing power of handheld devices. This evolution of communication and application functions has substantially increased the system level complexity or the amount of functionality, resulting in the following trends:

- Number of functional blocks increases ⇒ Number of interfaces increases
- Power consumption increases ⇒ Overall heating increases
- Amount of data increases (megapixel color displays etc.) ⇒ Required memory space and data traffic increases

A. Power challenge

Since middle of the 1990s, many 2G phones have provided users with good operation times. As an example, the first member of Nokia 6000 cellular phone family was introduced in 1997. It provided 3.3 hours of talk-time, and 180 hours of standby time. Using that as a reference and focusing on talk- or application-times, the general trends in power consumption and Lithium-Ion battery capacity can be represented, as shown in Figure 6. Constant annual growth of approximately 10% in battery capacity has enabled continuous battery volume shrinkage, while having the absolute milliAmperehour level constant over the years. It has also enabled new features like multislot transmission in GSM uplink without significant compromises in operation time. Figure 6 describes the highest output power levels in the transmitter, which in real life happens only very infrequently. Most of the time, the total power consumption is significantly lower. The good news is that technology evolutions in which battery capacity is one element enable good talk-time also for the first WCDMA phones, and even at the maximum transmitter power level. However, careful design is needed for the highest power consumption peaks, when 3G or WLAN communication is run simultaneously with multimedia applications. Therefore, it is important that the power consumption of the wireless access engines, meaning the cellular and non-cellular accesses, must be reduced in the future to release power for the applications.

In Figure 7, the power partitioning in a 3G phone is represented. Power consumption can be divided into four significant portions: transmitter, including power amplifier, receiver, digital and mixed-signal ASICs, and user interface HW, including display, speaker etc. Compared to the power partitioning in a 3G terminal, the power amplifier more noticeably dominates the power consumption in 2G terminals. Typical power-added-efficiencies of the GSM power amplifiers are in the range of 40-55%. In WCDMA, a non-constant envelope modulation is used; therefore, the power amplifier has to operate in linear mode, resulting in lower power-added-efficiencies – typically in the range of 30-40%. Since the transmitters are rarely used at the maximum power level in cellular systems, not only should the maximum efficiencies be compared, but the overall power consumption over the probability-density function of the transmitted power of the terminal should also be evaluated.

Future challenges include maintaining 10% annual growth in battery capacity and handling hot spots in highly integrated engines. As illustrated in Figure 8, the number of components will decrease continuously, as traditional chips turn to SoCs (System on Chip), or components become modules or SiPs (System in Package). Finally, the system complexity is increasing so much that better system level design tools are needed to optimize and verify the designs, and to test the hardware.

B. Means for complexity management

There are many potential methods and technologies to tackle the increasing complexity of the handheld systems. One of them is standardization of the core areas of the platforms. Symbian as an operating system is one example. It is specifically designed for PDAs, cellular phones, and other wireless information devices offering efficient and adaptive usage of all resources and instant access to user data, for instance. As a standardized OS, Symbian guarantees wide compatibility among handheld devices.

Another example is related to linkage between wireless access, application engine and user-interface HW. So far, this linkage has been based on dedicated solutions and one example is represented in Figure 3. As the variety of components that can be used to build the products in the future will significantly grow, only from the logistics point of view it is not possible to go on with dedicated implementations. A more generic interface must be standardized to easily attach new components into the system. The MIPI alliance has been launched to solve this challenge.

As complex systems comprise an enormous amount of parameters and dependences to be optimized, advanced EDA tools are needed to support product development. Tools are enablers for faster time-to-market and boosters of evolution. Achievements in embedding, miniaturization, power savings or design verification depend strongly on the tools. The top-down design flows currently available have weak spots that require too much effort and orientation from designers.

Today's multilevel requirements include, e.g., C-code, VHDL, and transistor-level descriptions, not only traditional mixed-signal based environments. A better framework is needed to fully enable utilization of intellectual property (IP) blocks in integrated circuit design. Standardized data transmission formats (cf. GDS II) need to be created to transfer information from one tool to another. Ease of use is a must. For example, library translations should not be visible to a designer. Collaboration between technology and EDA tool houses is needed to provide seamless, reliable and tested design environments. To make significant strides in embedding, more effort is still needed to develop practical substrate noise modeling tools. Cross-technology modeling like 3D package modeling combined with RF IC simulations must be supported, which is extremely important for chip modules or stacking, e.g. Removing overheads from the digital domain designs requires better support for optimizing the HW/SW partitioning. In some areas, like antenna design, the tools are in good shape, but new, more powerful and flexible simulation algorithms should be developed.

To achieve the best performance in handheld devices, any overhead in the system must be eliminated. Therefore, adaptive solutions are required. For example, in baseband platforms, dynamic voltage and frequency scaling, digital technologies to tackle leakage current, digital architectures applying asynchronous structures, and usage of reconfigurable circuits should be applied. All possible means for preventing unnecessary switching in digital circuits and overheads in latencies need to be applied.

Flexibility will be key in the ability to implement all those numerous product categories in the future, since they cannot be based on one or two platforms. As a part of this strategy, enormous ASICs may not be the most probable choice, because they can be too expensive and more difficult to test. In addition to flexibility, power consumption, cost and time-to-market are also important drivers. Whatever techniques are applied, the total cost of the product must be reasonable.

VI. CONCLUSIONS

A cellular phone is a paradigm of an embedded system having highly optimized cost, size, efficiency and performance. In terms of production volumes, the cellular phone is in a class of its own. Development has been based on advances, either evolutionary or disruptive, in implementation architectures and technologies. In particular, cellular phones have benefited from the early exploitation of leading-edge semiconductor technologies.

With the digital convergence already having crossed the threshold, the future of cellular terminals is very exciting. In wireless access, the architectures and technologies continue to be optimized towards versatile multiradios. And even more radical development is taking place in the application side, with 3G and rich applications like multimedia services and fancy games being established simultaneously. The evolution of system complexity is taking a large step ahead; consequently, system complexity management becomes the main challenge. To meet the time-to-market requirements while applying new technologies and techniques, advanced system level design tools are needed. In particular, power consumption, top-down design flow, flexibility, HW reconfigurability, programmability, and embedding are the key features that should be mastered in advanced system level design tools. As complexity increases, more standardized solutions are needed to guarantee compatibility. Standardization must be organized to focus on the core areas of the platforms, e.g., the operating systems and system interfaces.

References:

[1] T. Ojanperä, R. Prasad (editors), Wideband CDMA for Third Generation Mobile Communications, Artech House, 2001.
[2] H. Kaaranen, A. Ahtiainen, L. Laitinen, S. Naghian, V. Niemi, UMTS Networks, Architecture, Mobility and Services, Wiley and Sons Ltd., 2001.
[3] T. Halonen, J. Romero, J. Melero (editors), GSM, GPRS and EDGE Performance, Evolution towards 3G, Wiley and Sons Ltd., 2002.
[4] H. Holma, A. Toskala, WCDMA for UMTS Radio Access for Third Generation Mobile Communications, Wiley and Sons Ltd., 2002.
[5] G. D. Mandyam, J. Lai, Third Generation CDMA Systems for Enhanced Data Services, Academic Press, 2002.
[6] A. Hottinen, O. Tirkkonen, R. Wichman, Multiantenna Transceiver Techniques for 3G and Beyond, Wiley and Sons Ltd., 2003.
[7] www.wireless-world-research.org
[8] Mobile Internet Technical Architecture, IT Press, 2002.
[9] www.mipi.org
[10] C. Lindholm, T. Keinonen, H. Kiljander, Mobile Usability: How Nokia Changed the Face of the Mobile Phone, McGraw-Hill Companies, 2003.

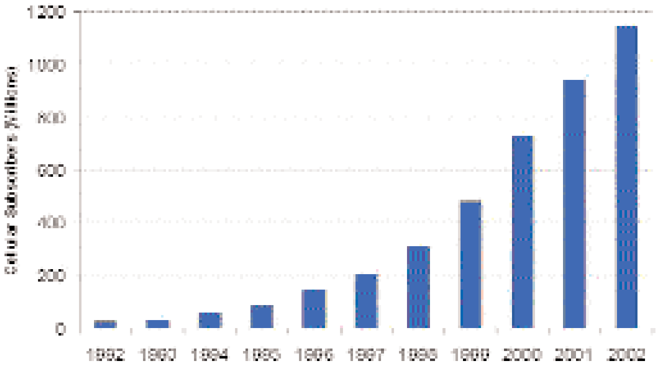


Figure 1.3.1: Mobile subscriber development.

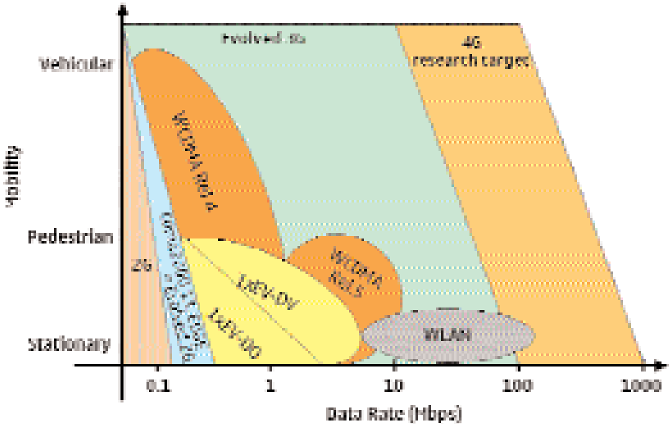


Figure 1.3.2: Mobility and data rate for radio systems.

0-7803-8267-6/04

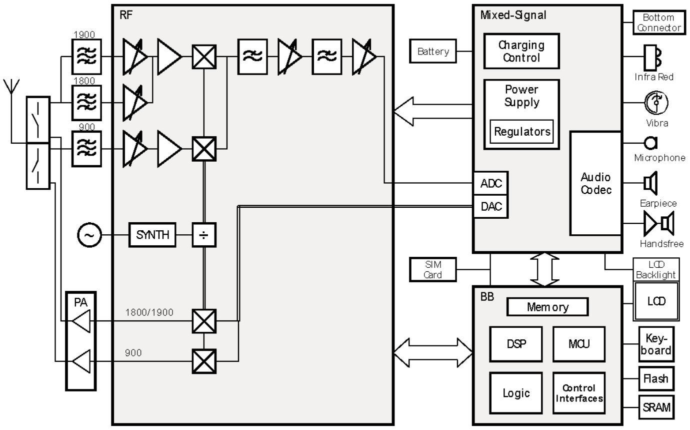


Figure 3. Block diagram of a cellular phone

Figure 1.3.3: Block diagram of a cellular phone.

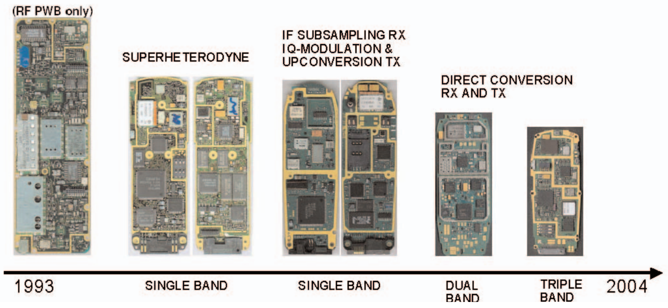


Figure 1.3.4: Engine development for cellular phones..

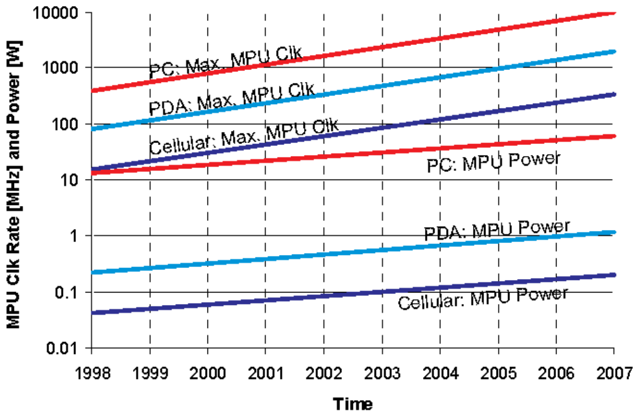


Figure 1.3.5: Maximum clock speed and power consumption trends of the application processors in PCs, PDAs and cellular phones.

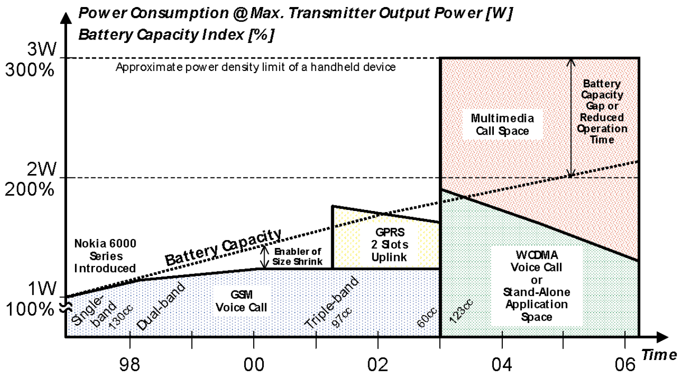


Figure 1.3.6: Battery capacity and power consumption indexes with the maximum output power level in cellular transmitters.

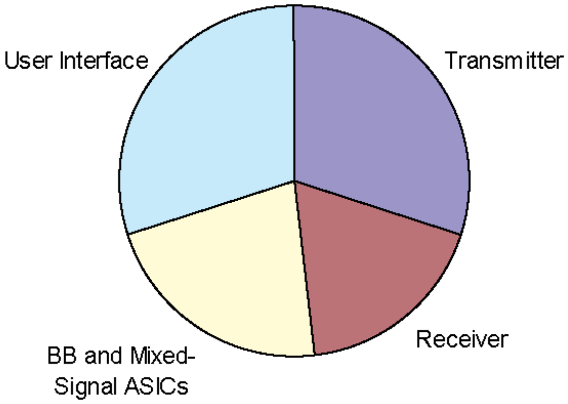


Figure 1.3.7: Power consumption break down in video streaming in a 3G phone.

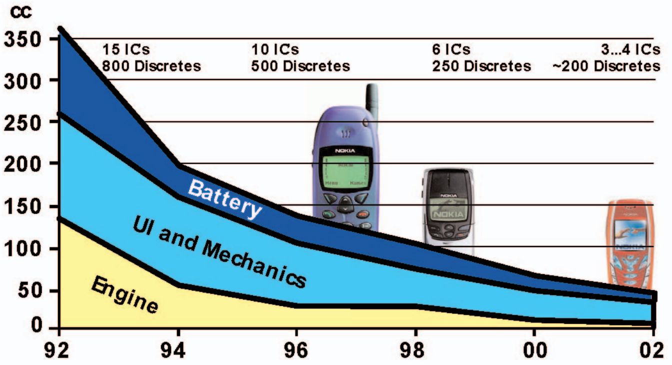


Figure 1.3.8: Miniaturization of cellular phones.

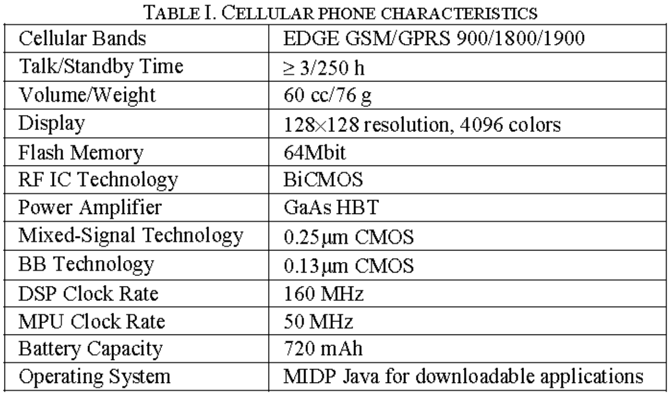
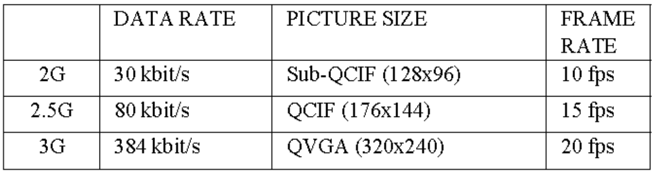


TABLE I. CELLULAR PHONE CHARACTERISTICS

Cellular Bands	EDGE GSM/GPRS 900/1800/1900
Talk/Standby Time	≥ 3/250 h
Volume/Weight	60 cc/76 g
Display	128×128 resolution, 4096 colors
Flash Memory	64Mbit
RF IC Technology	BiCMOS
Power Amplifier	GaAs HBT
Mixed-Signal Technology	0.25μm CMOS
BB Technology	0.13μm CMOS
DSP Clock Rate	160 MHz
MPU Clock Rate	50 MHz
Battery Capacity	720 mAh
Operating System	MIDP Java for downloadable applications

Table 1.3.1: Cellular phone characteristics.



	DATA RATE	PICTURE SIZE	FRAME RATE
2G	30 kbit/s	Sub-QCIF (128x96)	10 fps
2.5G	80 kbit/s	QCIF (176x144)	15 fps
3G	384 kbit/s	QVGA (320x240)	20 fps

Table 1.3.2: Practical data rates, picture sizes and frame rates of cellular generations for video streaming.

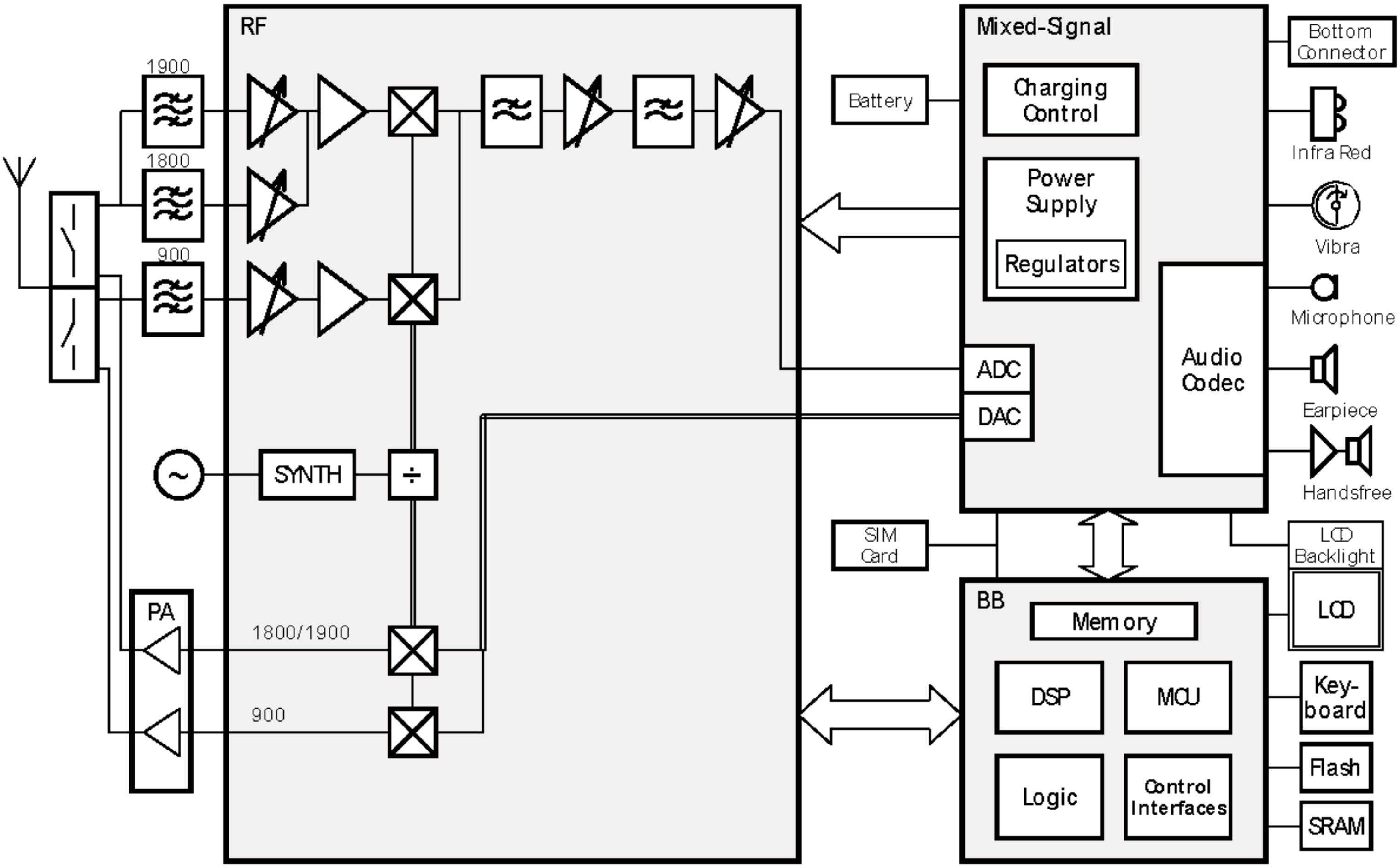


Figure 3. Block diagram of a cellular phone

Figure 1.3.3: Block diagram of a cellular phone.

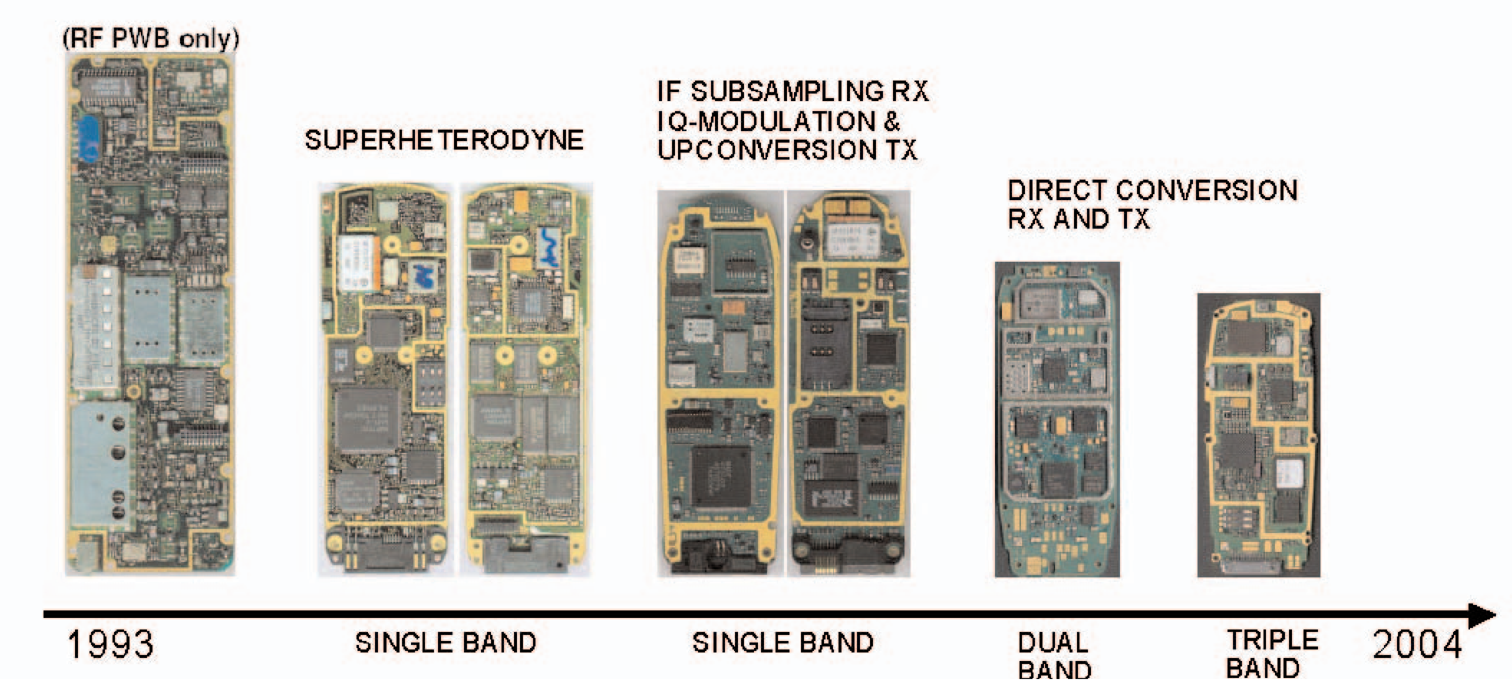


Figure 1.3.4: Engine development for cellular phones..

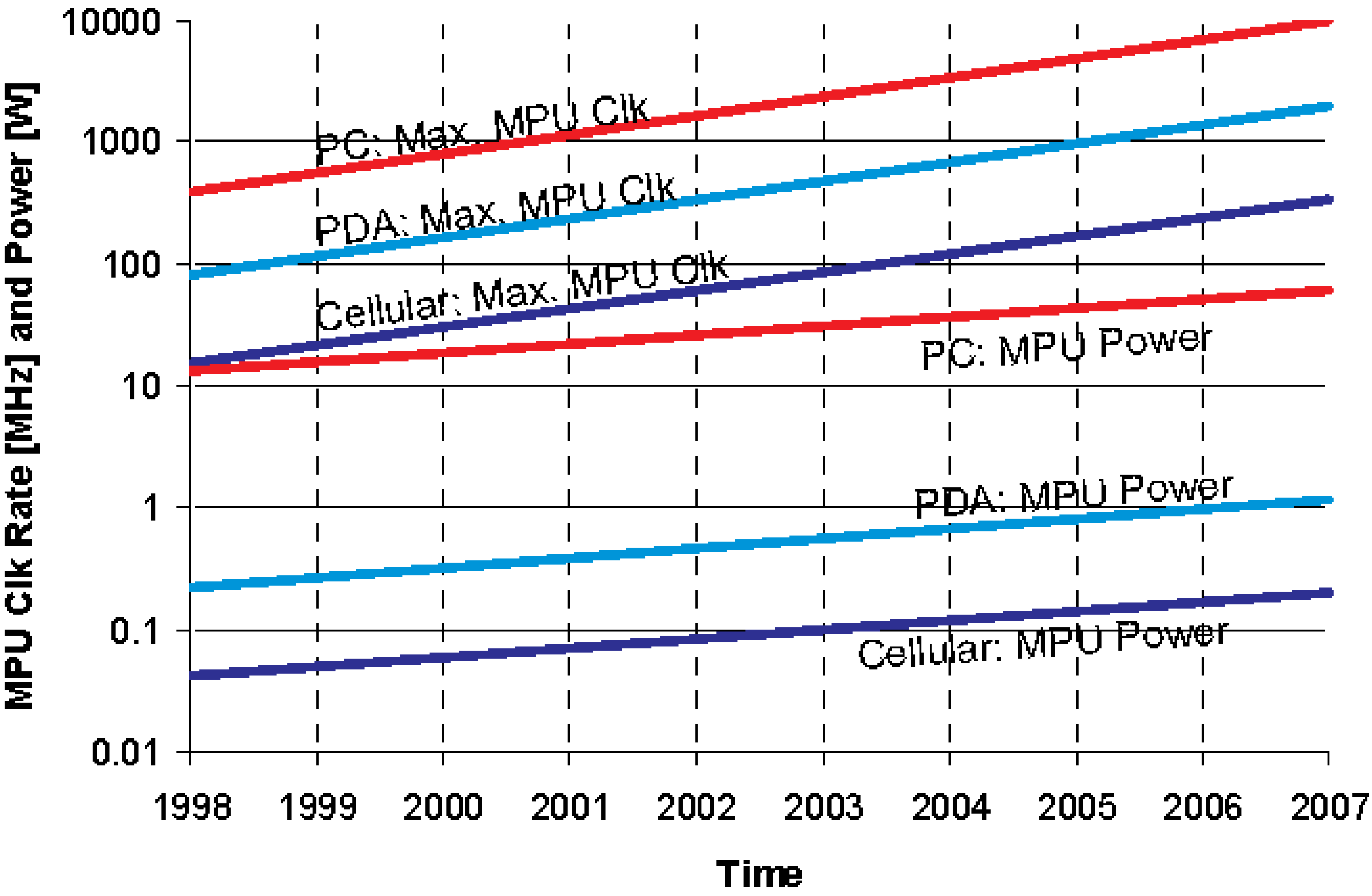


Figure 1.3.5: Maximum clock speed and power consumption trends of the application processors in PCs, PDAs and cellular phones.

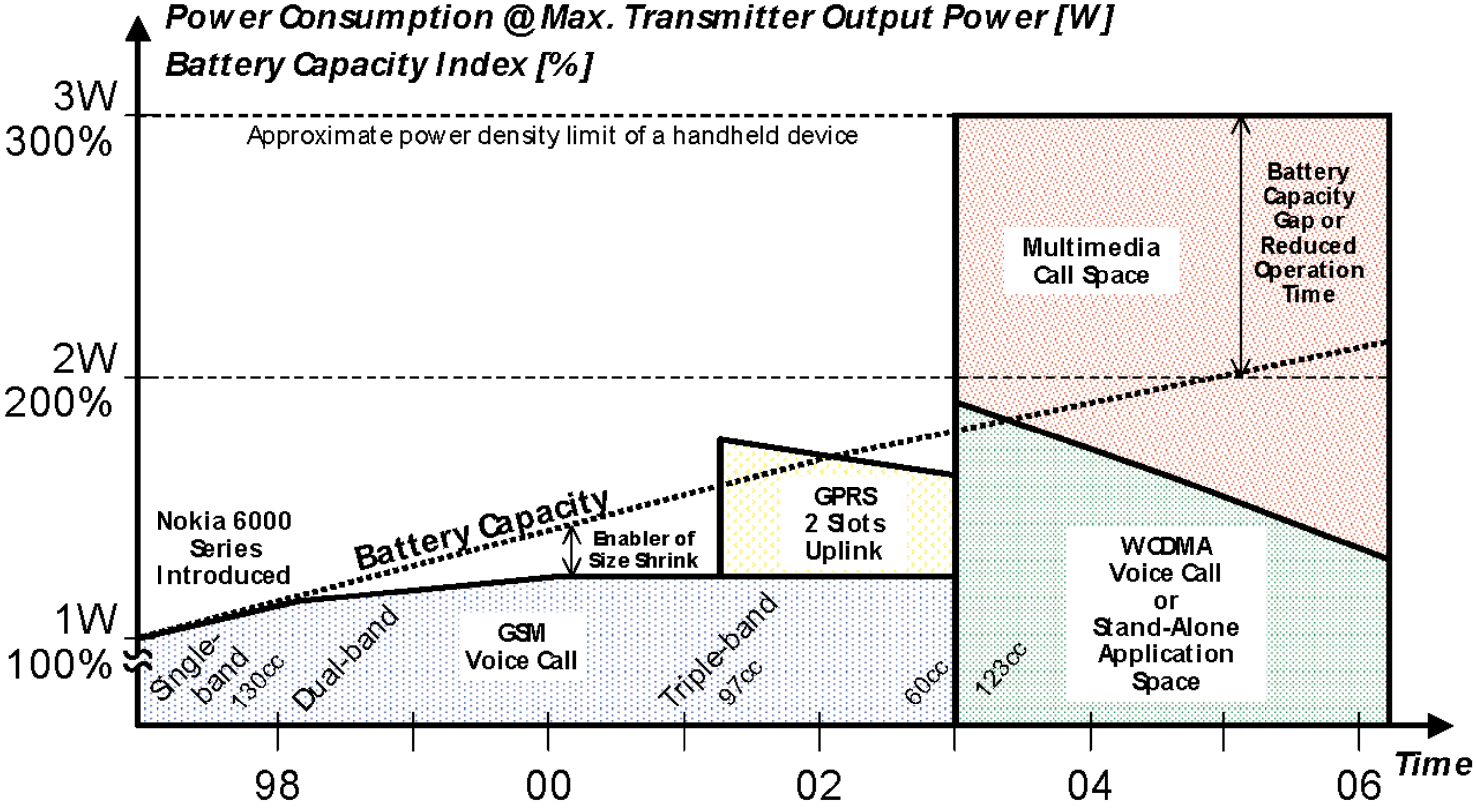


Figure 1.3.6: Battery capacity and power consumption indexes with the maximum output power level in cellular transmitters.

0-7803-8267-6/04

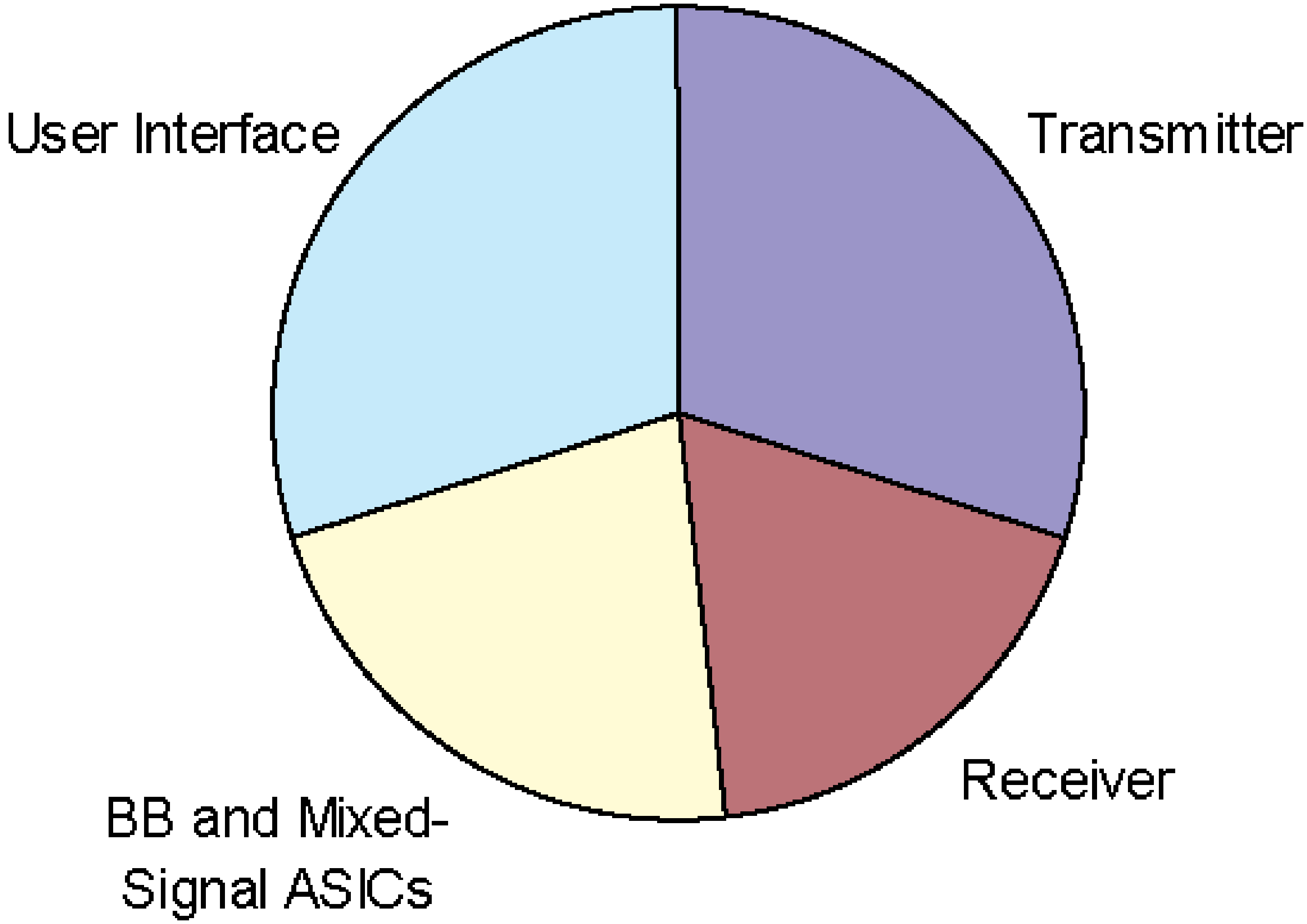


Figure 1.3.7: Power consumption break down in video streaming in a 3G phone.

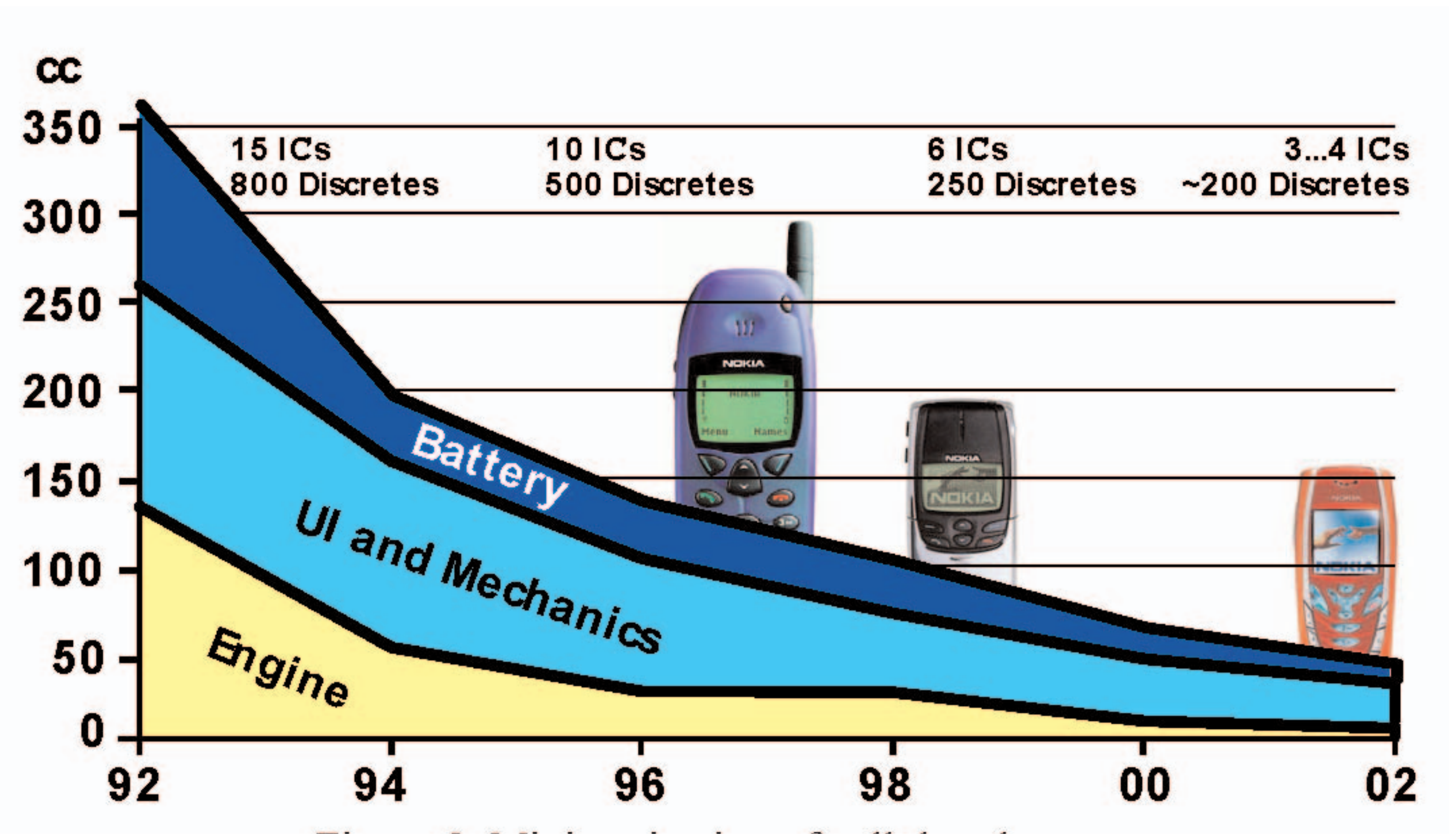


Figure 1.3.8: Miniaturization of cellular phones.

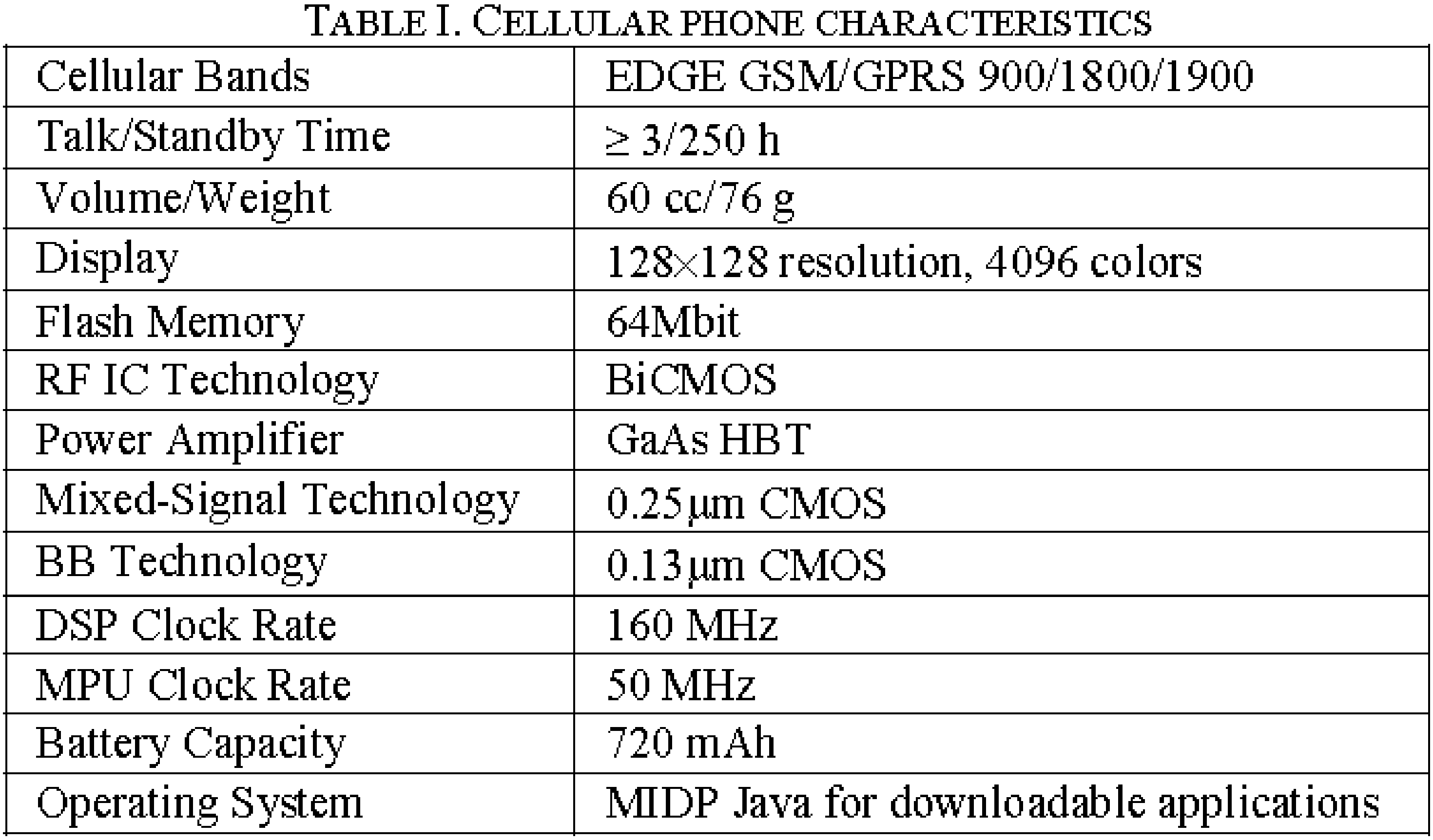


TABLE I. CELLULAR PHONE CHARACTERISTICS

Cellular Bands	EDGE GSM/GPRS 900/1800/1900
Talk/Standby Time	≥ 3/250 h
Volume/Weight	60 cc/76 g
Display	128×128 resolution, 4096 colors
Flash Memory	64Mbit
RF IC Technology	BiCMOS
Power Amplifier	GaAs HBT
Mixed-Signal Technology	0.25µm CMOS
BB Technology	0.13µm CMOS
DSP Clock Rate	160 MHz
MPU Clock Rate	50 MHz
Battery Capacity	720 mAh
Operating System	MIDP Java for downloadable applications

Table 1.3.1: Cellular phone characteristics.

	DATA RATE	PICTURE SIZE	FRAME RATE
2G	30 kbit/s	Sub-QCIF (128x96)	10 fps
2.5G	80 kbit/s	QCIF (176x144)	15 fps
3G	384 kbit/s	QVGA (320x240)	20 fps

Table 1.3.2: Practical data rates, picture sizes and frame rates of cellular generations for video streaming.