PEARSON Education

# Pervasive Computing

## Technology and Architecture of Mobile Internet Applications


Jochen Burkhardt
Dr. Horst Henn
Stefan Hepper
Klaus Rintdorff
Thomas Schäck

### 1.1.2 The peace dream

Although we have been doing reasonably well in developing technology to make life easier, we did not succeed nearly as well on the implementation of the peace dream. There is only hope that as people communicate with each other, they will realize that the vine and fig tree dream can only come true if the peace dream becomes a reality for mankind. Pervasive computing infrastructure facilitates access to information and modern administration in areas that do not have a traditional wire-based communication infrastructure. Wireless infrastructure can also be implemented faster and with lower investment in the densely populated areas of the world. Thus people living there may also participate in the worldwide exchange of ideas and Internet business that may become a major driving force to let the peace dream come true.

## 1.2 Pervasive computing

> *In our time it is not only the law of the market which has its own life and rules over man, but also the development of science and technique.*
>
> Erich Fromm, The Sane Society[2]

Personal computing and Internet communication have fundamentally changed our way of working both at offices and at home. Pervasive computing, the next dimension of personal computing, will change and improve not only our work environment but also our daily life and our communication patterns with family, friends, and business partners. Small portable personal assistants combine high-speed, low-power computers, wireless communication, data storage in persistent semiconductor memory and coin-sized disk drives, small colour displays, and video and speech-processing technology. This will give users the freedom not only to communicate efficiently at any place in the world but also to access local information as well as information residing on the Internet at any place and at any time. It is typical for new technologies like the Internet and pervasive computing that they have not been accepted by the majority of people as a prophet would dream of, but are sneaking into society. Only a small number of computer experts, for example, were sending short messages a couple of years ago. Today, children and adults use pagers or their mobile phones daily to send and receive short messages. Digital cameras have begun to replace conventional cameras because they integrate easily with digital media and email. Both technologies create a new business environment and affect the communication style of people.

### 1.2.1 Pervasive computing scenarios

The personal perception, technology, and business aspects of pervasive computing are described in the following scenarios, with emphasis on technologies that are already available today and have the potential to be widely accepted in the market in the next couple of years.

#### The automobile

The best way to experience state-of-the-art pervasive computing technology is to drive a modern luxury car. When you are approaching the car, you press a button on your electronic car key to unlock the car, and to adjust the driver seat and rear mirrors to your preferences. When seated, you put your mobile phone in its cradle, and insert your key to unlock the car's electronic system and to switch off the anti-theft system. You enter your destination into the car's navigation system. You dial into your office system by pressing a button on the steering wheel and download some urgent mail to the car communication unit. The system should read the mail to you later while you are on the road.

Finally, you start the engine, and because it is a dark, rainy morning, the lights turn on automatically and the windshield wipers start at low speed. The friendly voice of the navigation system reminds you to turn left at the next traffic light. Now you are on the road, and up to 100 computers in the car make sure that the engine is running smoothly and your ride comfortable. There is some traffic on the road but you feel pretty safe because you know that the car has a computer-controlled traction and brake system that knows about the adverse weather conditions and will adjust the controls accordingly. On the highway, you switch off the navigation system because you want to listen to your urgent email message, which is read by a computerized voice. In the future, the navigation and mail system might coordinate their work, and the manufacturer might invent a pothole detector to make the ride even more comfortable. If the car is involved in an accident, you know that the automatic airbag system will protect you. One day, the automatic emergency call system may alert the carmaker support system via the wireless communication system: in case of a minor accident, it will dispatch a limousine, and in case of a heavy accident it will dispatch an ambulance and provide the emergency doctor with your personal emergency health record.

Although these functions are not available in every car today, the individual technologies already exist and we can expect this pervasive automotive environment to become a standard feature in almost every new car in a few years. The scenario demonstrates some key characteristics of pervasive computing:

- Multiple devices like car key, mobile phone, car audio system, and navigation system are integrated in the system. Devices are used for multiple purposes, e.g. the car key is used to unlock the door, to personalize settings, and to start the engine.
- A large number of different interfaces, such as keys, buttons, displays, touch screen, microphone, loudspeakers, and environmental sensors, are used to build an optimized user interface.
- There is concurrent operation of offline and temporary online systems, e.g. the motor monitoring system is working offline, but can be connected to a service center for diagnosis if required.
- A large number of specialized computer systems are interconnected via local buses and the Internet.
- Short-range (car key) and wide-area (mobile phone) wireless communication are integrated.
- Security elements, e.g. car key, theft protection system, and SIM card in the mobile phone, prevent unauthorized access.
- Personalization of functions adapts the system to the user's preferences, e.g. adjustment of the driver seat and rear mirror. User interfaces are simple, response is fast, and most of the system's complexity is hidden. No PC literacy is required to use and manage the system.

We can expect these types of functions to be integrated into networked operation and to be extended to workplace, home, and mobile environments.

### The mobile workplace

An ever-increasing percentage of traditional office workplaces will be transformed into truly mobile workplaces, with portable computers, mobile phones, portable digital assistants (PDAs), printers, and fax machines connected via wireless local area networks and wireless Internet. Totally new forms of organizations will emerge where most people can schedule their work much more flexibly than today. An indication of this ongoing transformation is the growth of mobile workplaces in some of the Western European countries, with yearly growth rates exceeding 30%. This might be a partial solution to avoid the daily traffic jams in large cities, and to overcome the lack of skilled resources in most modern economies by allowing parents to better balance their family and business life.

Mobile workplaces will not be restricted to typical office jobs. Manufacturing companies will use mobile devices to monitor manufacturing processes remotely, to schedule work more efficiently, and to improve communication with employees. Optimization of the supply chains will require a highly efficient, integrated pervasive computing infrastructure.

3

The retail and distribution industry already relies on isolated pervasive computing devices. Service terminals will become smaller and will interact with customer devices, such as smart cards, mobile phones, paging devices, and PDAs, in order to improve workflow and data integrity. Health and social services will improve information flow and quality of service. The pervasive computing workplace will become a major component of most enterprises in the future.

The massive introduction of mobile workplaces will certainly require new social skills to manage private and business life. People and organizations must learn not only to manage traditional working hours but also to cooperate in a result-oriented business environment where team members are working at different places at different time.

### The home

Pervasive computing has arrived at home mainly through mobile phones, PCs, and entertainment devices. Mobile phones have become a major infrastructure of family life in many countries by linking the busy schedules of parents and children. TV sets, video recorders, games consoles, and CD players become integrated into entertainment centres with optional access to the Internet through set-top boxes. New devices that integrate TV, DVD, and games centers, are being introduced to the market. Modern houses feature local area networks (wired or wireless) and appliance buses to control doors, windows, shutters, heating and cooling systems, home appliances, and security systems. Local area networks using data transmission via power lines or wireless technologies will provide the infrastructure to interconnect domestic devices. Through the addition of an Internet connection, the home emerges into a Web-enabled service node for home management, entertainment, security monitoring, and totally new consumer services through a domestic service gateway. Acceptance of all this new technology by users is slow because standards for integration of the different technologies are difficult to establish.

### The personal communication environment

Today, people are using multiple devices and services for personal communication, work, and entertainment. Many different devices, such as wire-line phones, mobile phones, pagers, fax machines, PCs, games consoles, digital cameras, music recorders, and video recorders, operate pretty isolated from each other. However, more and more systems are appearing on the market that combine several of these devices, creating new services such as sending a fax from a mobile phone, programming a video recorder from a PC, or controlling the heating system at home via a mobile phone. Typically, these services are not integrated and are provided via a multitude of unrelated channels and service providers. It is difficult for customers to locate a responsible party in case of a problem and to handle multiple service center addresses, personal identification numbers, user names, and passwords. In the pervasive computing environment, personal communication will evolve into an interconnected, cooperative communication landscape with a shared security and access infrastructure. Ultimately, it will work similar to the GSM environment where services are provided by a large number of service providers without forcing the users to change their interface or behaviour. Users will be able to switch off certain communication paths temporarily or route them to different media, for example incoming short messages may be routed to a voice mailbox residing on the PC or on a server. Outgoing email may be converted to voicemail for a user who does not own a PC or does not have access to a PC at the moment. Users or service providers may want to use the lowest-cost communication path to transfer large amounts of data, or may want to be alerted immediately by a short message when a certain event occurs.

The upcoming third generation of mobile devices that uses General Packet Radio Service (GPRS) or Universal Mobile Telecommunications System (UMTS)[5] will support these types of requirements, using packet-switched networks. Multiple quality-of-service options for delay, data rates and error rates are defined. Users will not be disrupted by short distortions on the communication channel but may stay online for a long period of time being charged only for the data actually transferred.

### Location-based services

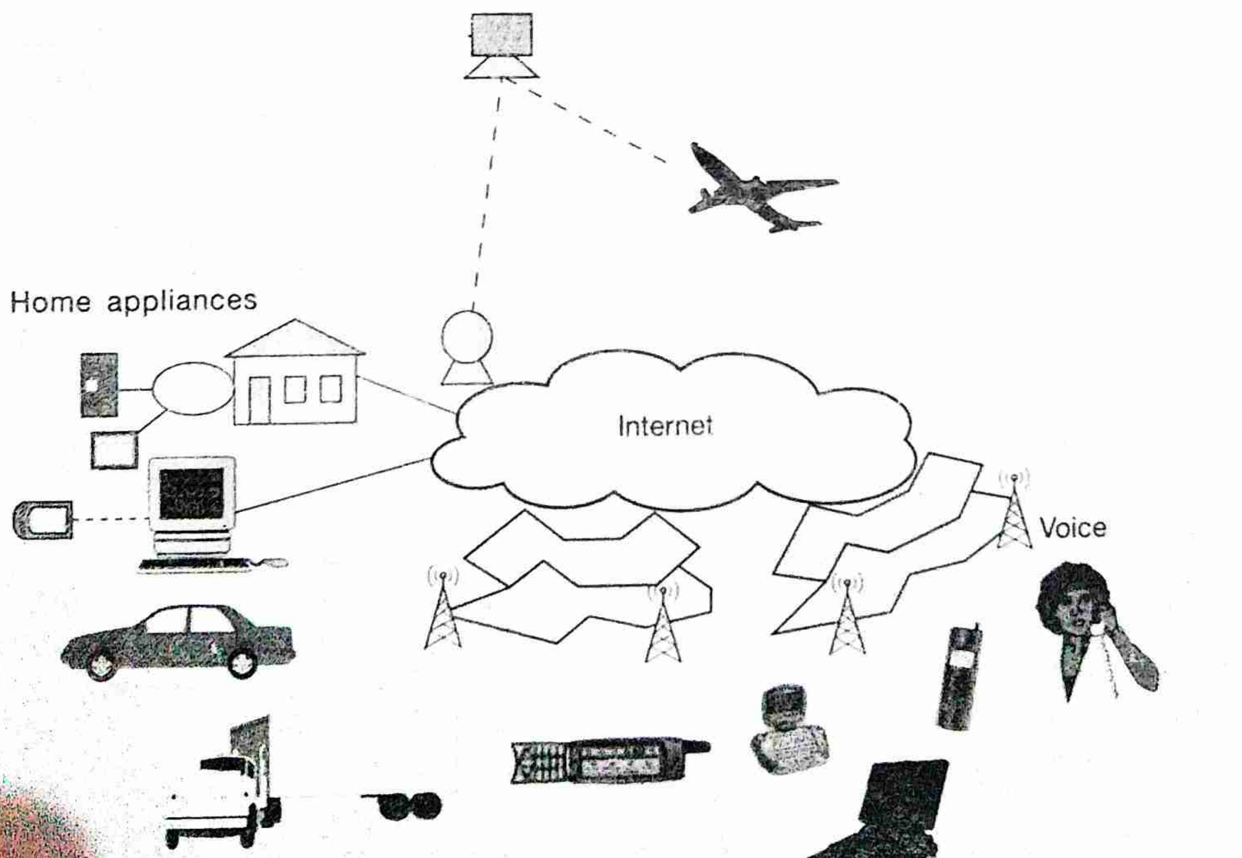
Location-based services exploit the ability of mobile phone systems to determine the position of the device by measuring signal run time from multiple transmitters to a receiver. The satellite-based global positioning system (GPS) may be used to locate the user with an accuracy of less than tens of meters at almost any place on earth. However, buildings, mountains, trees, and heavy rainfall may disrupt GPS signals. GSM operation may still be possible in some environments where GPS does not work, and vice versa. The ability to locate a caller is a strong requirement for handling of emergency calls, because the location of a mobile caller cannot be derived from the telephone extension as in a wire-line phone system. The US government has therefore enforced the introduction of emergency call location services. Caller location identification also enables all kinds of applications linking the Internet to the 'real' world, including roaming between wide area and local area networks. It also raises severe privacy issues when the user location is tracked without explicit user agreement. This problem can be addressed by a user function that allows explicit suppression of location identification, much like suppression of caller identification in GSM mobile phones.

### 1.2.2 Roaming environment

Advanced users will work with multiple devices, e.g. a multimedia PC for complex information-gathering, e-business, and office tasks; a TV set for entertainment and online shopping; and a Wireless Application Protocol (WAP) phone for mobile communication. Commuters will expect their cars to offer a personalized communication environment, and mobile workers may use a PDA during working hours. Private communication will be mostly through voice complemented, by email and short messages (SMS) (Figure 1.1). Mobile chat, virtual blackboards, auctioning, games, multicast messaging, and collaboration are some of the new communication patterns evolving in the mobile space.

Terrestrial wireless networks, satellite networks, and local area networks in offices, shops, and private homes provide access to the World Wide Web from any place and at any time. Pervasive computing devices will also be engaged in offline peer-to-peer communication or wireless networking. Users not only will roam between different mobile networks but will also use different devices and expect data and functions to be synchronized, hopefully in a totally transparent way, without much user intervention. The new, open environment must guarantee security and privacy in order to be accepted by a large user community. The GSM subscriber identity module (SIM), a secure smart card, provides a trusted

Figure 1.1



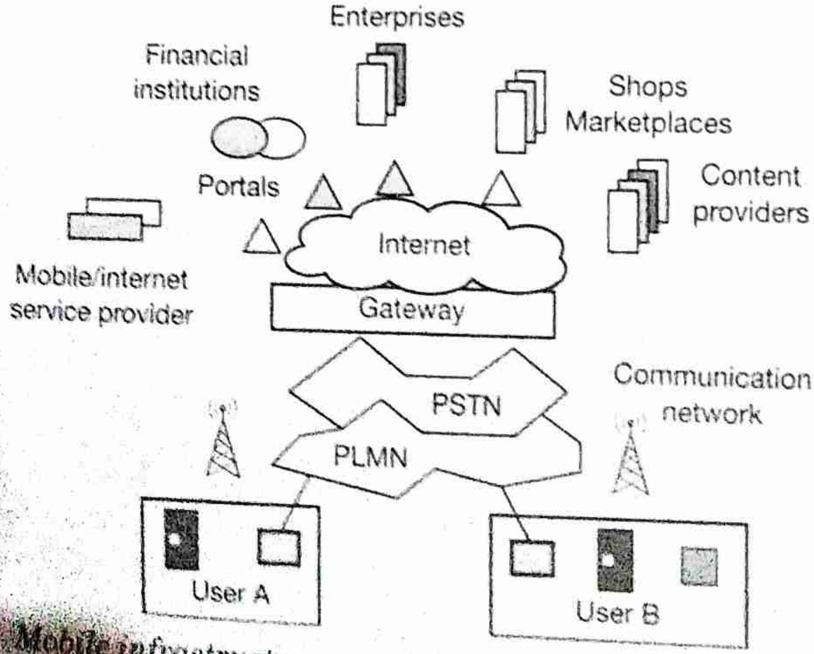
computing environment for user identification and secret key storage. The SIM will be extended to enable security functions not only for the mobile service provider but also for other parties like financial institutions and enterprises, to secure their communication with their clients. This type of global security infrastructure is not available in the PC-oriented world today. However, the introduction of public key infrastructure in PC operating systems, secure Web protocols, and the legal endorsement of digital signatures will allow us to build secure systems by merging the Internet and wireless communication into a secure, pervasive computing infrastructure suited for mobile e-business.

### 1.2.3 Pervasive computing infrastructure

As long as personal devices like organizers and PDAs were isolated from networks, the communication and data exchange with other pervasive computing devices was only possible through proprietary methods within a line of products or from the devices to a PC using special software. In the pervasive world, millions of devices, subscriber identification modules, public land mobile networks (PLMN), public switched telephony networks (PSTN), gateways (e.g. from WAP and voice to Internet Protocol, IP), servers, and applications are interconnected. Last but not least, system management and billing systems must be integrated. Content providers, shops, market places, financial services, and enterprises use this infrastructure to offer their services to their clients (Figure 1.2).

Pervasive portals provide gateways to adapt the pervasive devices to the standard Internet protocols. They also deliver a variety of functions, such as personalization, mobile device management, security, and data

**Figure 1.2**




*Mobile infrastructure*

synchronization. Portals are also able to maintain stable application interfaces between applications and various different devices with rapidly changing features.

In the traditional telecommunications environment, governments restricted the access to the telecommunications market to create secure, integrated, reliable systems. Very often there was only one state-owned organization that serviced the entire country and was very hesitant to introduce new technology. However, this was the only way to ensure interoperability and guaranteed access for users in the early days of telecommunication. In the new, deregulated communication world, multiple service providers are competing with lower rates and new functions and features based on new technology. However, interoperability and standardization are becoming crucial success factors for the entire industry. Techniques initially developed for the World Wide Web are applied to the communication sector to manage rapid evolution and interoperability. Industry committees are addressing standards issues in order to insure interoperability of their products and services. Nevertheless, standardization is still a rather slow and tedious process. Proprietary, local solutions promise to reduce time to market and to gain a competitive advantage. However, they are very often outperformed in function, performance, and price by standard solutions in a very short period of time due to the competitive environment of open markets and the economy of scale.

Two major technologies representing the two main roads to market are competing for the third generation (3G) of mobile systems.

- *GSM/UMTS* based on the harmonized 3G radio interface. This system is supported by major telecommunication standard organizations and by the Operator Harmonization Group (OHG). It provides worldwide roaming, and a smooth and compatible evolution path from existing GSM to UMTS, and supports WAP.
- *i-mode System* developed by NTT DoCoMo. This is a proprietary 3G system based on a high-speed, packet-switched network and IP. The system was developed in Japan and has now been introduced in Europe and the USA.

Whereas Europe and Asia have adopted the GSM/UMTS standard, mobile communication in the Americas is based mostly on Code Division Multiple Access (CDMA) and Time Division Multiple Access (TDMA) networks, although GSM networks are operational in the densely populated areas in the east and west. However, these networks represent less than 20% of the world market and are migrating to the mobile Internet on various paths. The strength of GSM/UMTS is certainly its support by major standards organizations, suppliers, and the large worldwide user base, whereas the strength of the i-mode system is its maturity, rich set of applications and initial affinity to Internet in respect to protocols and content representation. The i-mode system uses compact Hyper Text Markup Language (cHTML) instead of the Wireless Markup Language (WML) to ease migration from existing Internet applications to mobile applications. This strategy has worked out very well in Japan, where customers are offered a wide variety of services, with especial emphasis on entertainment.

Acceptance of WAP-based Internet services was hampered by the low data-transfer speed and low reliability of first-generation WAP systems. The introduction of high-speed packet-switched support through GPRS and UMTS will provide adequate service for 3G applications.

The evolution of pervasive computing devices with higher computing power and multimode radio-frequency (RF) circuitry will allow support for multiple communication protocols and content representations to accommodate different standards during a transition period.

### 1.2.4 Personalized services and the virtual pervasive home

*The only use of things is to be applied to the service of persons*[3]

R. H. Tawney

Pervasive computing is yet another step from the classical supply economy to a customer-driven economy[4]. Basic telephone and Internet services are transformed into highly customized and diversified products and services. Mobile users will expect services that will be easy to use and will deliver results within their expectation window of less than a few seconds. The personal phone directory of mobile phones, which allows calls to be placed much quicker than on a public phone, is an excellent example of an implementation meeting mobile users' expectations. They are not willing to enter large amounts of data using tiny keyboards or pens. They can't look up files or documents when they are using mobile devices. Only very special devices will have an ability to print on paper. Performance and process reliability must be much better than that provided by typical Web applications today. Users will demand highly personalized, easily usable services for browsing, content delivery, and transaction handling with minimum involvement in management of security, devices, and systems.

Therefore, offering typical Web applications on mobile devices is not a viable approach. Information must be condensed, hopefully presented automatically when needed, and adapted to the user's situation and device. Presentation and dialogs must be tailored to the small screen size and the limited time a mobile user is willing to spend. Mobile users also expect to work with their familiar interfaces, user profiles, applications, and data independent of device, location, and time. Users also want to control why, when, and how they are contacted. Flooding mobile users with advertising is certainly not a very good idea, whereas tailored services can create strong relationships between individuals and business.

9

A special infrastructure element, the 'virtual home', is required to support the following set of functions in a pervasive environment:

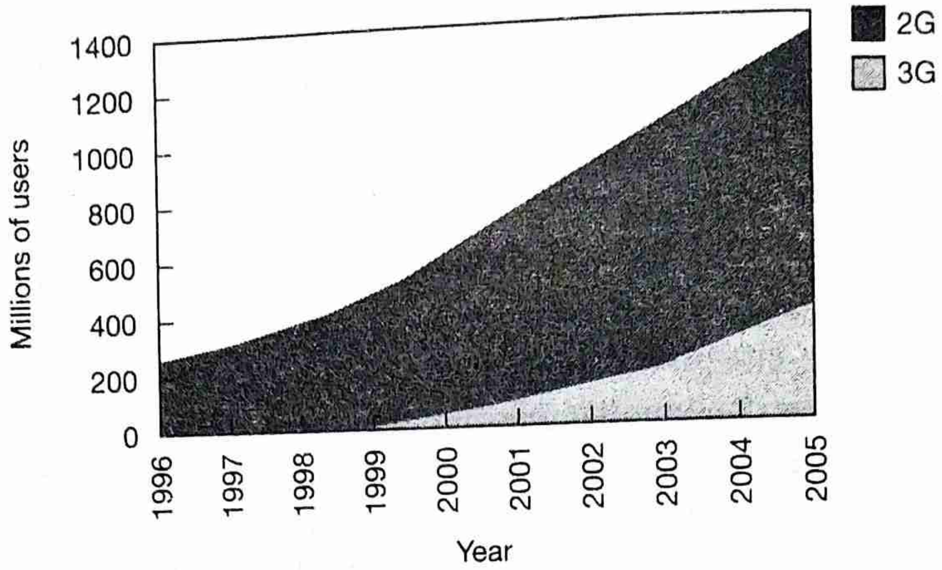
- Universal home address for users that is independent of device and location being used. It unifies, for example, phone numbers for voice and fax, addresses for email, messages, and devices.
- Secure system access, privacy (encryption), and endorsement of transactions.
- Personalization of devices, services, and communication channels.
- Server-based vault for personal and financial data, addresses, profiles, bookmarks, service directories (personal information management, payment services, etc.), and administrative data.

Typically, the virtual home will be implemented with a portal server. Users may use different virtual homes for private, business, or financial services. Today, mobile portals implement only part of the key functions or share functions with other systems, e.g. email systems. Without complete and well-managed virtual home functions, pervasive devices are difficult to use and cause major customer dissatisfaction in networked environments. Implementation is hampered not only by lack of standardized user interfaces but also by the lack of well-accepted function-distribution models in the pervasive computing space. These problems are very similar to the problems of the Internet a couple of years ago, and can be resolved by applying the Internet standardization model to the pervasive computing space.

## 1.3 The pervasive computing market

The pervasive computing market is a typical innovation market. The major driving force for initial growth in Europe and Asia is the basic mobile phone service, which provides high value to the consumer and is easy to use for anybody, whereas the driving force in USA is the extension of Internet services to the mobile space. New complementary products and services are introduced at a rapid pace, and are quickly adopted by industry and consumers. Markets are emerging in different countries with different speed and emphasis on specific products. New innovative products and services can cause major shifts in consumer spending. Therefore, the size and the structure of the market, as well as the growth within the next four to five years, are hard to predict. The number of mobile phone subscribers is probably one of the best indicators of market size and dynamics. At the end of the year 2000, about 600 million subscribers were using mobile phones. This number is estimated to grow to 1.3 billion in 2005 (Figure 1.3).

**Figure 1.3**




*Predicted growth of mobile phone subscribers worldwide*

Source: UMTS Forum[5]

In contrast to predictions for other new technologies, mobile phone sales have consistently outperformed most forecasts. The other major pervasive computing market segments are expected to grow at similar rates.

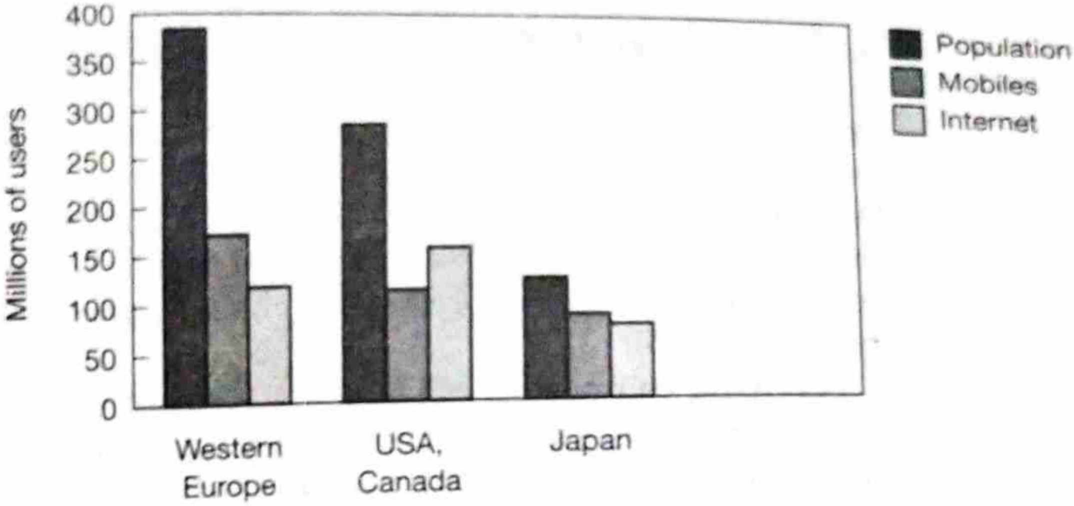
- *Embedded systems* are typically not visible to the customer, and are delivered to the consumer with other products, such as automobiles, consumer electronics, and appliances.
- *Consumer electronics* with pervasive computing elements, e.g. digital cameras or video recorders, electronic books, MP3 players/recorders with wireless modems, and direct Internet access.
- *Hand-held computers* that are typically used as stand-alone devices with optional data transfer from or to a PC or the Internet via built-in modems, mobile phones, pagers, email stations, PDAs, and mobile PCs.

Overall, pervasive computing has grown into a market with a yearly volume of several hundred billion dollars. Growth is fuelled by a combination of new technologies and innovative services. The largest part of the revenue is coming from services rather than products. The mobile service providers have developed a business model supporting rapid change. The majority of devices are not sold but are leased to the customer. Thus, consumers are moving rapidly to state-of-the-art technology independent of the condition of the overall economy.

Mobile phone coverage is approaching 100% in leading mobile countries, such as Finland, Sweden, and Italy, and growing at double-digit rates in other countries. Mobile phone users are already outnumbering the Internet users in Europe and Japan (Figure 1.4).

The ability to access the Internet via a PC seems to influence the pervasive computing market. In Japan, typical users of the i-mode service access the Internet with their mobile phones more often than with PCs.

Figure 1.4




*Mobile phone and Internet coverage*

Source: UMTS Forum[5]

This is due in part to the high speed and reliability of i-mode's packet-switched third-generation communication system, but also to the limited offerings of PC-based Internet content in Japan – not surprisingly, given the fact that most of the Internet content is presented in English. The i-mode system presents content in Japanese language and is also the first system to offer attractive mobile entertainment services that are highly valued by Japanese customers.

Most Internet users in the USA prefer solutions that combine, for example, a PDA and a PC for Internet access, or offer pocket size email systems that exploit the excellent 'wired' Internet infrastructure and can also be supported by the current mobile infrastructure.

The market in Europe is dominated by GSM phone services that support roaming and seamless GSM service throughout Europe, which is a key requirement for travelling Europeans. Most of the European manufacturers of GSM equipment and terminals have participated in the GSM standards activities and have seized the GSM market.

The WAP Internet access service is not yet widely accepted by European consumers. Volume delivery of WAP-enabled phones began in 2000, and the adaptation rate is rather low (<5%) in most countries.[7] Early WAP services seem to lack attractive applications, and seem to be neither fast nor reliable enough for the busy mobile consumer. Very often, existing HTML-based applications were 'mapped' to WML without considering the very special requirements of mobile users. It can be expected that WAP services will become more popular with the introduction of GPRS packet-switched data service providing higher reliability and bandwidth. However, this must be combined with new and exciting services for professionals and consumers in order to win back disappointed early users. Therefore, voice services are still the major source of revenue for mobile service providers. However, all providers are looking for attractive applications to increase their revenue from value-added services.

## 1.4 m-Business

Pervasive computing will play a major role in e-business especially in the consumer space and in industries like automotive and transportation. Pervasive devices will be used not only to buy travel-related products and tickets or to exchange messages and email, but will be used in almost every stage of the e-business value chain to provide better and faster service (Figure 1.5).

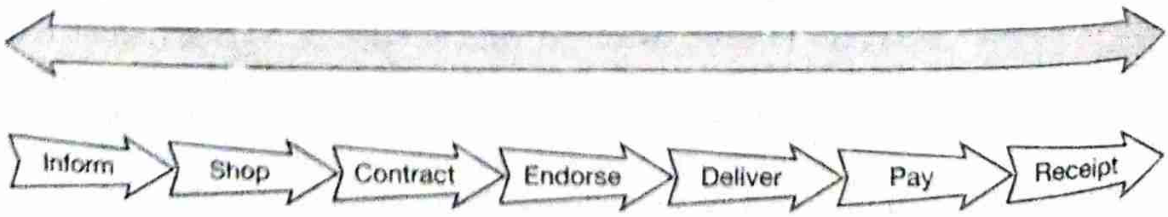
Pervasive computing will be a key element in attracting new customers and maintaining customer relationships. Some industries, such as banking, have already been hit by the overwhelming demand of customers for timely information and the ability to initiate and check transactions anywhere and at any time. Key enablers for m-business are:

- attractive offerings for the mobile user community;
- easily usable and reliable infrastructure accessible by all parties involved in the business processes;
- security through identification, authentication, privacy, and non-repudiation;
- trusted environment;
- payment system supporting payments from a couple of cents to highly sophisticated high-valued transactions;
- Business models for e-business operations.

Most of the value generated will be indirect and will help to reduce process costs for existing and new business processes. The amount of direct revenue generated by sales to mobile consumers is hard to estimate because the market is in a very early stage and many assumptions must be made in respect to product offerings and user acceptance (Figure 1.6). Development of the m-business market will also depend on the development of overall Internet trade, which still lacks common legal and tax regulations.[6]

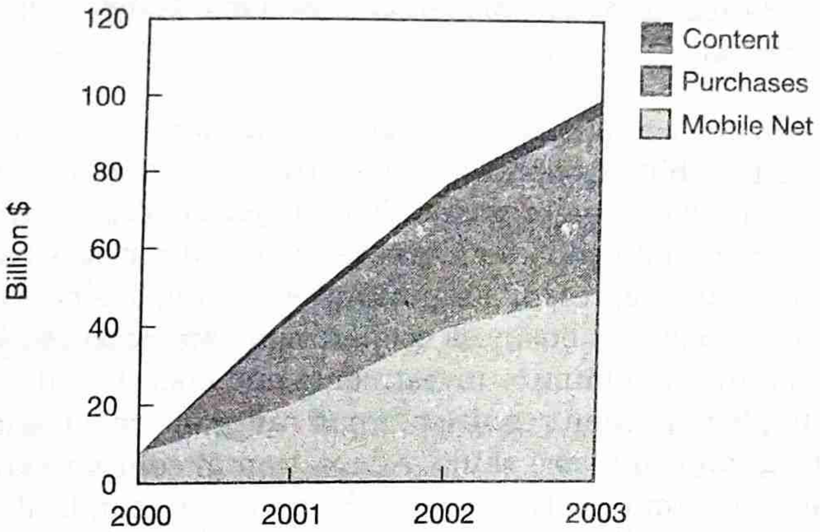
A study by the Boston Consulting Group[7] estimates the worldwide m-business market to grow to about $100 billion within the next three years. About 50 billion dollars of this revenue is assumed to flow to the

Figure 1.5




*Mobile business interactions along typical value chain*

Figure 1.6




*Worldwide m-business forecast*

Source: UMTS Forum[5]

mobile service providers. Most purchases are assumed to bé of rather low value, e.g. downloading music recordings, images, and games, and will create an estimated revenue of about 50 billion dollars. Only a very small percentage of m-business revenue will flow to content providers because consumers have adopted the Internet model and expect most services to be provided for free.

Revenue predictions at this early stage of m-business are based on early feedback from emerging markets. New product offerings, the quality of m-business support infrastructure, and mobile payment systems will impact heavily on the growth of m-business.

## 1.5 Conclusions and challenges

Pervasive computing is a new dimension of personal computing enabling communication and Internet services at any place and at any time for most of the population. It will impact our way of life as well as business processes in almost every industry. Pervasive computing is based on easy-to-use interfaces, applications, and the convergence of the following key technologies:

- very-large-scale integration of computer, communication, and consumer electronic circuits;
- wireless communication technology;
- consumer electronics and embedded controllers;
- Internet infrastructure;
- speech technology;

# 2 Application examples

There are already a lot of different pervasive devices available on the market, ranging from PDAs to WAP phones. Now the question is how to effectively use these devices in business applications. In this chapter, we explore how pervasive computing devices enable exciting new ways of conducting business. We present a series of real-life application examples showing how businesses can benefit from pervasive computing. The examples will give you an overview of what is possible and what is required, and should inspire you to think of tomorrow's applications today.

## 2.1 Retail

The retail area has seen a long development driven by the quest for faster and cheaper ways to bring goods to the consumer. Today, the consumer can select from large varieties of products, and can just buy them from stores, catalogs, and virtual shopping malls on the Internet.

Mobile devices will further improve processes for the consumers and the retail industry. Mobile computers equipped with barcode readers are used to track products during manufacturing, transportation, and in the supermarket. Smaller devices, better usability, and wireless network access will make them more adequate for the requirements of modern businesses. Standardization of the software-development platforms for these devices will help make them cheaper and ultimately more flexible.

Consumers can currently use their computers, TV sets, or mobile devices to select and order products. Tomorrow's shopping list will reside on a PDA, and may even receive input through other pervasive devices in the household. When completed, the list can be sent to the supermarket and the purchase is delivered or prepared for pick-up at a time convenient for the consumer. The availability of the orders in electronic format enables faster processing of the data. Finally, data-mining technologies allow a better understanding of the consumer's behaviour and enable direct marketing.

In 1998, Safeway UK and IBM teamed up to deliver a first-of-a-kind personalized remote shopping service.[1, 2] Selected consumers were able to use hand-held devices to build their shopping list from an electronic product

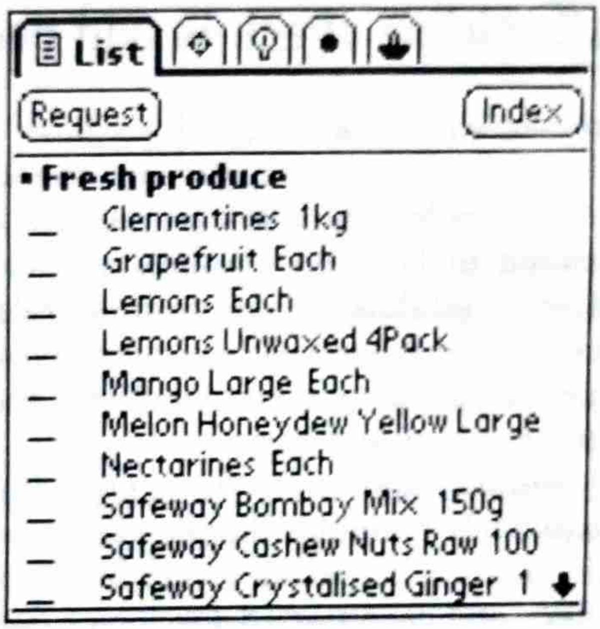


15

catalog. A preselection of products built from their past purchases simplified the creation of the electronic order. Figure 2.1 shows a screenshot of the Safeway application displaying products from one category.

The use of devices fitted with barcode readers enabled consumers to self-scan products at home or in store. The electronic order was then transmitted to the supermarket's server via a modem. Together with updates for the product catalog, the customers received feedback in the form of personalized offers. Once the order was received, the goods were collated and packed ready for collection. An overview of the system architecture is given in Figure 2.2.

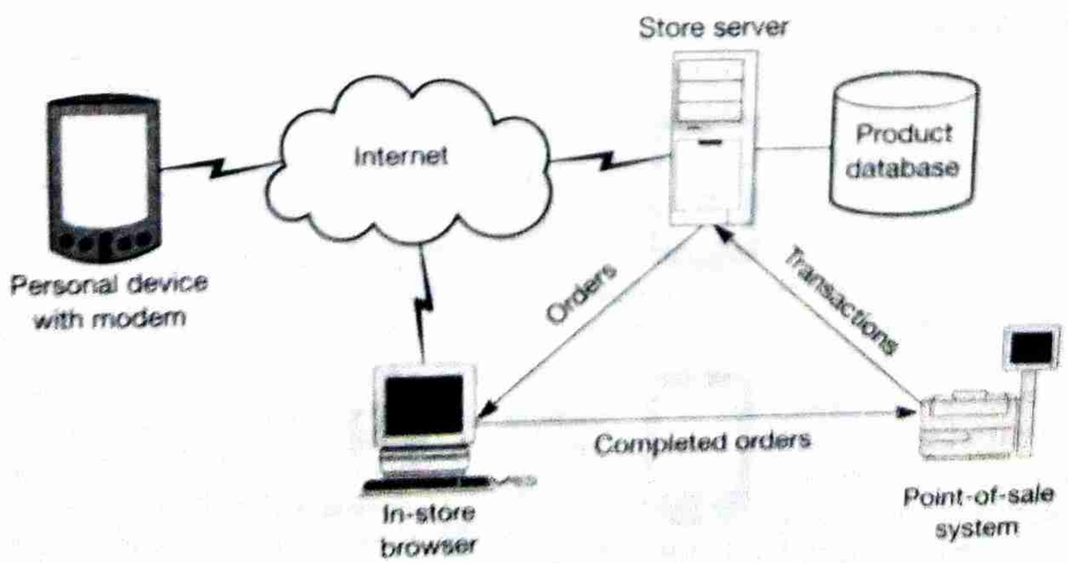
gure 2.1




*The Safeway personalized remote shopping service*

Picture courtesy of IBM Corp.

gure 2.2




*The Safeway scenario*

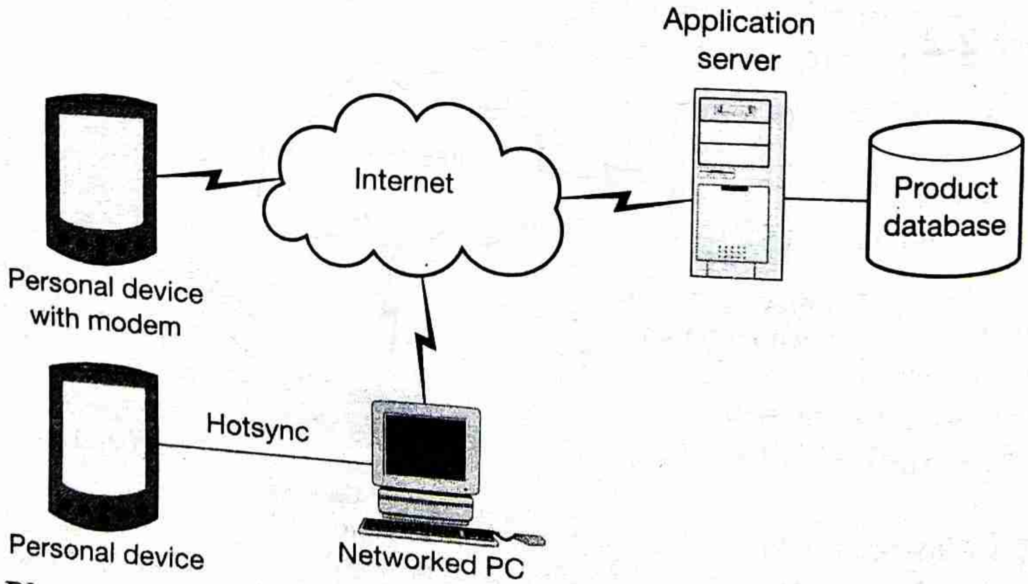
Another example for the use of hand-held devices in retail is the PlanetRx Online shop.[3] PlanetRx specializes in selling drugs online. Products can be listed and ordered from their website. As an additional feature, registered users can access the site from their Palm-connected organizer. The software is installed as an extension of the standard synchronization process, of the PDA. During the synchronization process the data stored on the device is updated, and orders are transmitted back to PlanetRx. Users can reorder products from their shopping list, or order from a weekly specials section. The complete PlanetRx scenario is shown in Figure 2.3.

## 2.2 Airline check-in and booking

For several years now, airlines have been using the Internet to publish flight schedules, let customers check their frequent-flyer accounts and claim airmiles, and allow travelers to buy tickets directly from the airline instead of booking through a travel agency. Their presence on the Internet establishes a more direct relationship between the customer and the airline. When visiting an airline's website, travelers are likely to see advertisements and special offers, and may book flights that they would not have booked otherwise.

Notably, however, by being present on the Internet, only people who are reasonably computer literate and have access to a PC can be reached. Now airlines are becoming aware of the possibilities that mobile computing offers, and have started developing new applications that enable customers to interact with airline systems via their mobile devices. With the increasing penetration of WAP phones, it will soon be possible to provide many

**Figure 2.3**




*The PlanetRx scenario*

travelers with a direct, permanently available connection to airline reservation systems. In the remainder of this section, we present two examples of mobile applications in this area: check-in and booking via WAP phones.

Usually, airlines require passengers to check in about one hour before the departure of the flight. This is so the airline can assign seats to people on the waiting list if other passengers do not show up. Early check-in results in waiting times for passengers, which is especially inconvenient for the most valuable customers of airlines – the business travelers and frequent flyers – who often travel business or first class.

By allowing check-in via WAP phones, airlines can provide increased convenience for their most important clientele. Instead of leaving for the airport early to check in, the traveller can switch on his or her WAP phone, obtain the list of booked flights, and check in for one of them with a single button click.

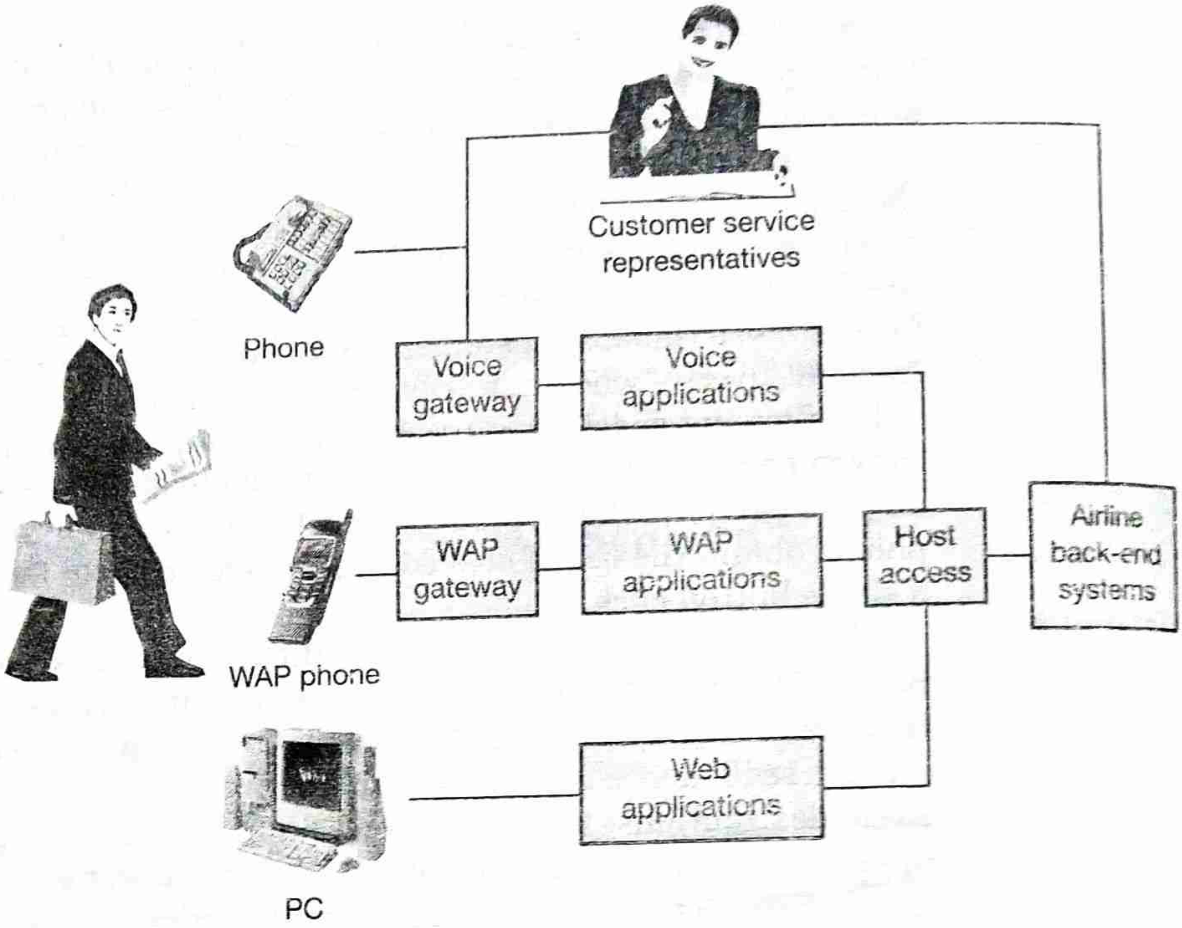
Although flight booking seems difficult and inconvenient from a WAP phone, it can be implemented in a way that provides good usability to travellers. Personalization can be used to allow users to book a flight with a short sequence of interactions. Many frequent flyers often travel to the same destinations, using the same flights again and again. Thus, using appropriate personalization data for a user, the system can make an intelligent preselection of flights that are of particular interest to the user, so that travellers can select their flights from a small list of relevant options. The data required for personalization might be specified explicitly by the user, or might be obtained from the system by analyzing the previous bookings.

To provide the kind of services described above, WAP phones have to be connected to airline back-end systems via appropriate gateways and WAP-enabled applications. Figure 2.4 shows how access via WAP phones can be added to the already existing access paths to the services of an airline.

Typically, airlines allow customers either to call a customer service representative or to log in to the airline's Web applications using the PC. Some airlines also have systems in place that allow customers to call automatic voice-driven applications. To add WAP access to existing systems, the WAP-enabled applications must be deployed on application servers and must be accessible via WAP gateways in order to be accessible from the customer's WAP phone.

The introduction of WAP technology to provide an additional access path to airline services has several advantages. Customers can effectively and quickly perform tasks such as check-in or reservations anywhere and at any time without waiting until a customer service representative becomes available. In addition, WAP applications can be implemented with little effort, and require little computing capacity. As acceptance of the new access method increases, it also reduces the number of calls to customer service representatives, saving the costs of call centers.

Figure 2.4




*Adding WAP access to airline systems*

## 2.3 Sales force automation

Until now, mobile workers relied on their portable computers in order to access and process data on the road. The availability of faster modems, and even wireless modems, has enabled them to travel and still have access to enterprise data. The complexity of the PC, however, still requires these workers to be trained in using such systems.

The availability of smaller portable devices with a relatively simple interface and the same connectivity provides a new alternative to this scenario. Not only are these devices much easier to use, but they are also much cheaper. This allows an enterprise to give out more devices to more users than before. PDAs enable the mobile professional to use the phone book and calendar while working out of the office, and to stay in contact via email. They are also used to access mission-critical data, such as contracts, technical descriptions, small databases, and graphics anywhere and at any time.

When used with a browser, PDAs can access the Internet, just like the PC. However, the quality of the rendering is limited by the capabilities of the device. Nevertheless, there is great value even in access to dedicated content from the enterprise network, or simple applications such as maps and travel information.

When used with a wireless connection, PDAs enable mobile workers to stay in touch with their company all the time. Work schedules and feedback collected during work can be exchanged with the enterprise network. This can be used to control the delivery of goods, update work assignments, submit orders, and even enter billing information while on the road.

## 2.4 Healthcare

In healthcare, there is little room for new and complicated devices unless they have an immediate benefit for the patient. Modern medicine already depends on a wide range of computerized devices, sensors, and actors. Clinical professionals have to learn not only about new methods but also about how to use new devices. Nevertheless, many clinical professionals are now using hand-held devices. The users of these devices have to learn new interfaces that are quite different to those of a PC. The new devices also fall behind PCs in terms of screen size and ease of data entry. Also, they get through several batteries when used intensively every day.

Despite all these problems, PDAs have found their way into modern healthcare. Because of their size and mobility, they are able to integrate into fast-paced hospital work as well as into a doctor's daily routine. They have turned out to be tremendously helpful in delivering up-to-the-minute patient data and medical information. Access to laboratory results and surgical reports, as well as ordering processes and physician directory look-up, can be improved by the use of PDAs.[4]

Healthcare applications are probably the most sensitive in respect to security. Integrity of data and privacy must be assured through proper hardware, software and system design. Patients must be identified correctly at all levels of treatment and medical record keeping. Patient and clinical data must be exchangeable and accessible wherever needed, but access must be restricted to professionals with the need to know. International healthcare organizations have standardized almost all data elements and structures required for documentation and data exchange between healthcare institutions, which is an excellent base for real mobile information.

Modern healthcare systems are now using smart cards for patients and professionals.[5,6] Patient cards used for identification and administrative data, e.g. Germany has issued more than 80 million chip cards to the entire population in order to simplify all administration processes. Professional cards are used for authentication of professionals and to control access to critica' data and systems. The G8 Health Card specification defines a public-key smart card architecture for an international health professional card to be used by doctors, pharmacists, administrators, nurses, and other healthcare professionals.[7] Doctors and nurses using

PDA-type devices with a smart card reader can access medical records, treatment plans, and administrative data at any time and in any place. Patients can securely transfer critical monitoring data to their doctor. The ability to work offline and to synchronize data securely on the PDA with data on the server is key in the healthcare environment, e.g. in critical areas of hospitals, the use of wireless communication is not allowed to avoid interference with electronic equipment. PDAs are also ideal to assist healthcare professionals to manage their personal work schedules, drug and International Classification of Diseases (ICD) code lists, prescriptions, treatment plans, and other critical information.[8]

The future will bring faster, smaller, and better devices that will lower the acceptance hurdle for use in healthcare environments even further. With integration into hospital communication networks, improved user interfaces, and better data input, the next generation of devices will convince professionals to replace traditional pen and paper by mobile devices. The benefit for all will then be fast and accurate access to patient data through personal mobile devices.

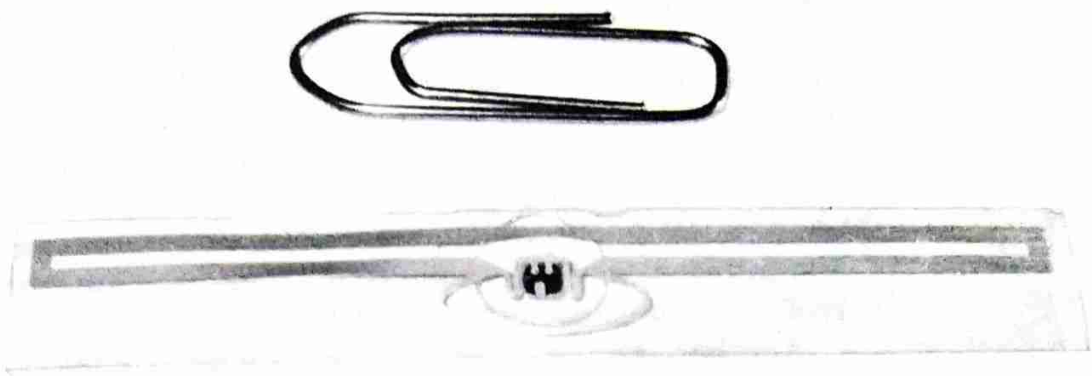
## 2.5 Tracking

The use of barcodes has revolutionized processes in many industry segments. They enable the fast and accurate identification of goods during transportation. Today they are visible on almost all products. The most common format is the one-dimensional barcode that encodes only a few characters' worth of information and is mostly used for universal product numbers. The two-dimensional bar code format allows several hundred characters of information to be stored, and eventually will even replace stamps.

The future will bring cheaper and smaller radio frequency (RF) tags that can be attached to goods such as labels. In contrast to the passive barcodes that are read optically, they will be read electronically and from larger distances. Figure 2.5 shows an RF tag consisting of a chip and an antenna enclosed in transparent plastic sheet less than a millimeter thick. Being more tolerant to environmental conditions, they will enable the tracking of more goods in more places than before. Advances in integration technology will add limited processing capability to these tags. Tomorrow's overnight parcel services will be able to determine the exact location of a parcel not just because it was registered at some place on its way but because it knows about its destination and even gives feedback in case it is routed wrongly.

An interesting example is tracking luggage in the airline industry. As tags get cheaper, it will become possible to attach to each piece of luggage a small tag that can be detected at certain points on the journey. This will enable airlines to track individual pieces of luggage from check-in to bag-

Figure 2.5



*RF tag embedded in a clear plastic strip*

gage claim. If any luggage accidentally gets separated from the passenger, the airline can immediately find out the current location of the luggage and arrange reconciliation.

## 2.6 Car information systems

Car manufacturers are becoming increasingly interested in using pervasive computing technology. A current top-of-the-range car has more than 30 different microprocessors built in, which communicate with each other over a bus. In the future, these processors will be able to not only communicate with each other, but also with the outside world. This will increase the value of the automotive system and allow for completely new services. Many car manufactures have already shown prototypes demonstrating this ability (e.g. General Motors, Daimler-Chrysler, and PSA). Analysts predict that by 2006, nearly 50% of all new cars (and 90% of all luxury vehicles) will have some kind of Internet access capability.

To connect the car with the outside world a gateway is required. The Open Service Gateway Initiative (OSGi) is an industry group that defines and promotes an open standard for networked consumer and business devices connected to the Internet. The first version of OSGi[9] is based on Java and Jini technology, and provides downloading of software, application lifecycle management, gateway security, attached device access, resource management, and remote administration. OSGi is compatible with transport systems such as Bluetooth, HAVi, HomePNA, HomeRF, USB, and WAP. It therefore provides all services needed for connecting cars to the outside world.

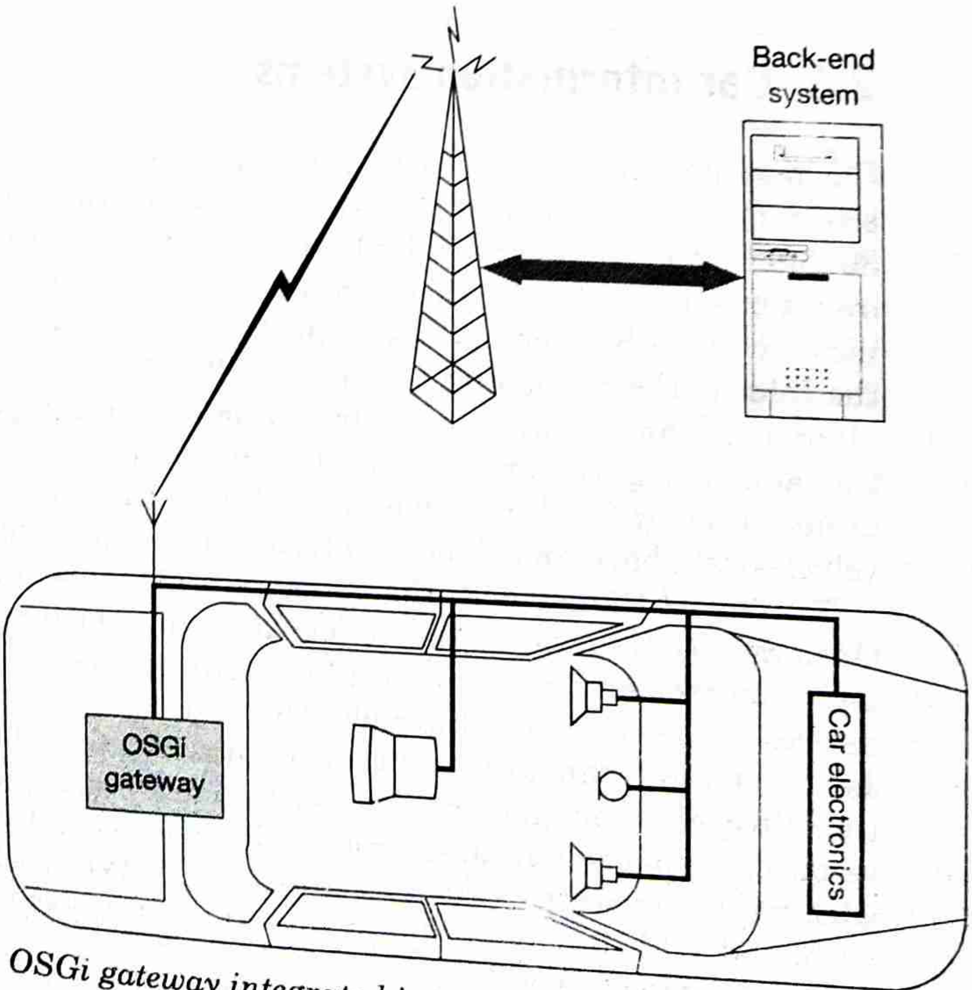
22

Figure 2.6 shows how an OSGi gateway can be integrated into a car information system. The OSGi gateway can receive and send data via a wireless connection to the back-end system. In the car, the gateway is connected to the car bus (e.g. an optical bus). It has access to the car electronics and all car sensors, and can therefore transmit all relevant car data. The back-end system can perform an analysis based on those data and inform the car holder about critical conditions. Over the gateway, the car entertainment system can receive and initiate phone calls, connect to the Internet, and access multimedia information, such as route planning, traffic announcements, and videos.

Another industry initiative with similar goals is the Internet Home Alliance, which was formed in October 2000; members include Cisco Systems, Best Buy, General Motors, Panasonic, and Sun Microsystems. The alliance's goal is to bridge data communications, telecommunications, entertainment distribution, and home-control services within the home and cars. The Internet Home Alliance tries to supplement the OSGi standard.

Connecting cars to the outside world has benefits for car owners and manufactures. Car owners can use new services such as email, Internet access, making phone calls, or automated emergency calls in case of an

**Figure 2.6**




*OSGi gateway integrated into a car information system* accident. With connected cars, the car manufactures now have the ability to get real feedback about the quality of their products. They can monitor all critical car data, such as oil temperature, inform the drivers about recalls, download software updates to the car, and more. Repair centers can receive notifications from the car information system about defects, and can order the required parts in advance. When introducing car information systems, special attention must be given to the areas of security, software updates, and management. All three areas are now being studied by the automotive industry, as they present the same problems that must be solved for all pervasive computing devices. Many solutions available for pervasive computing devices such as mobile phones or PDAs can be applied to car information systems and vice versa.

## 2.7 Email access via WAP and voice

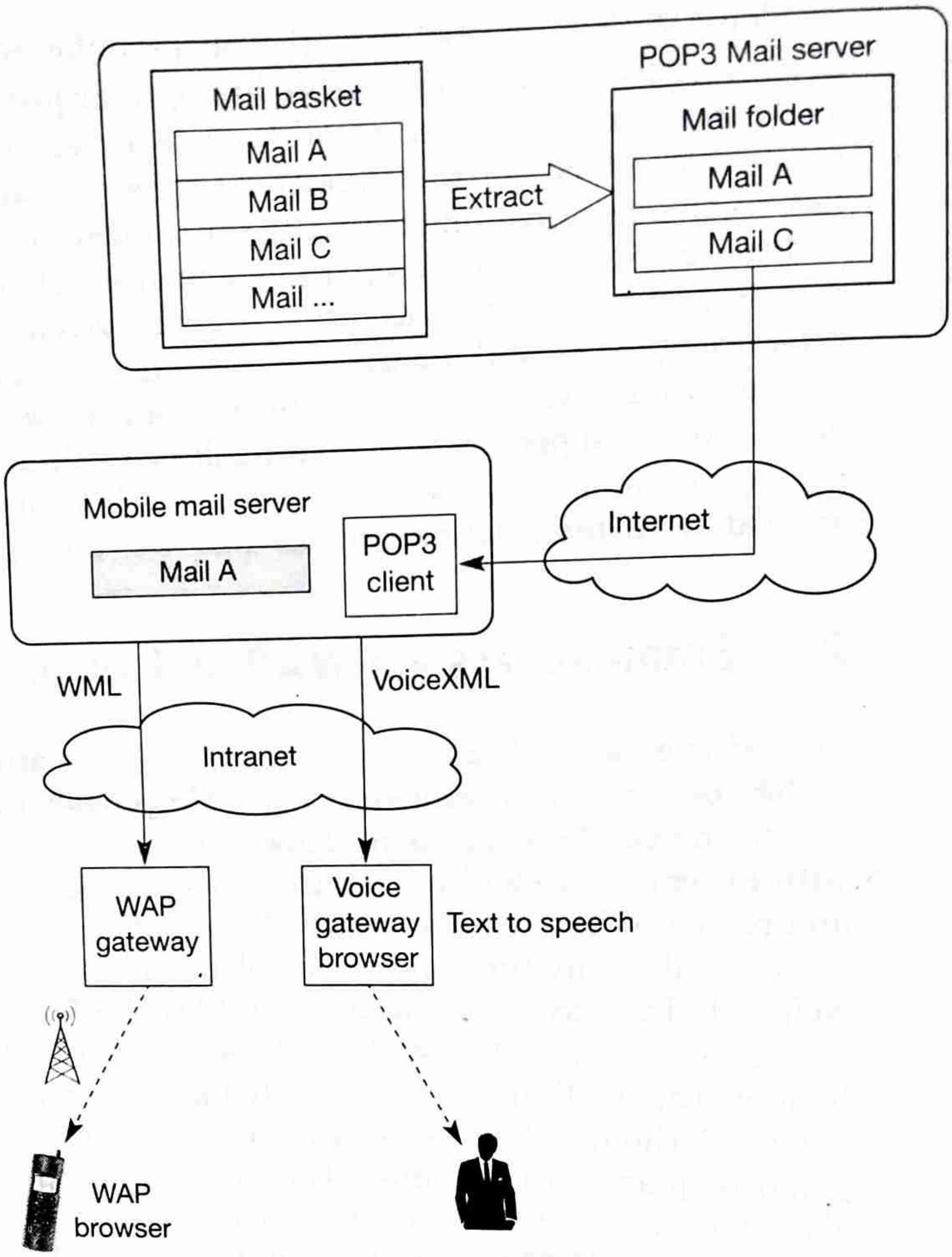
Accessing email and calendars is one of the most attractive applications of mobile devices. Voice mailboxes and SMS provide rudimentary mail services in the GSM environment. However, these services are not integrated with PC email, are limited to rather short messages, and cannot be stored and retrieved like regular email. Many mobile users would like to access their email at any time and at any place. A WAP-enabled mobile phone is well suited to accessing typical email files of a few kilobytes and to sending short responses, but WAP mail services must be tailored to the small bandwidth, small display, and limited key entry capabilities to be really useful. Although WAP phones with larger screens are already offered, small displays will dominate the market for the next years to come. Therefore, users should be able to define classes of mail to be sent to the mobile device. Mail systems like Lotus Notes and Microsoft Outlook support this type of selection through profiles and mail agents. Typically, the selected mail is stored in a folder for further processing. The content of this folder is then used as input to the WAP mail services.

An overview list of the mail should be presented to the user for selection before larger amounts of data are sent to the WAP phone in order to reduce data transfer. The content of the mail must be filtered, e.g. large attachments should be removed, and redundant text such as openings and closures should be condensed. The user should be able to mark mail as read or to delete mail, just as in standard mail systems.

Text-to-speech technology can be used to read email to customers on the move. The amount of speech that can be presented to a user is limited. Therefore, selection and filtering of email must be used to create an email system for voice, much like for WAP. Figure 2.7 shows a typical email system supporting WAP and voice access.

WAP mobile phones access the intranet of the mobile email service provider via a WAP gateway. A mobile email server hosts the application,

Figure 2.7




*WAP and voice email service*

creates WML output, and handles the incoming Hyper Text Transfer Protocol (HTTP) requests. It interfaces with the POP3 mail server where the original email is residing through a POP3 client.

A typical WAP email session involves the following steps:

1. The user accesses a mobile email server using a WAP profile and a start-page URL.
2. The mobile email server verifies whether the user is registered by checking the phone number and/or user identification and mobile password.

3. The user requests email service and provides the POP3 password for access to the POP3 mail server. The mobile email server verifies the address of the server and user identification in the user profile.
4. The server accesses a POP3 mail server on behalf of the user, and extracts the mail files for the user. A mail list is created, converted to WML, and sent to the WAP phone user for selection.
5. The user selects a mail file and sends a display request to the mobile email server.
6. The email server converts the email file to WML and forwards it to the WAP phone for display.
7. The user reads the email with a WAP browser.
8. The user responds with a request for more mail, deletion of mail, or reply.

Much the same steps are required for email access via voice. Users dial in to the voice gateway. Authentication is typically implemented using a customer identification number and a personal identification number, because numbers can easily be entered with a standard phone keypad. Additional checking of the phone number is recommended for mobile phone users. For convenience, the POP3 password will probably also be stored on the mobile email server. VoiceXML data is transmitted to the voice gateway, which hosts a voice browser and a text-to-speech engine for text rendering.

The mobile email server must be a trusted component because it has access to passwords as well as user mail. Enterprises or mobile service providers may install the POP3 server and mobile email server together in their intranet to avoid proliferation of passwords and sensitive mail.

## References

1. ACM TechNews (2000). 'Easi-Order'. http://www.acm.org/technews/articles/2000-2/0322w.html#item12
2. IBM. 'IBM and Safeway create enjoyable grocery shopping experience'. http://www.ibm.com/pvc/industry_solutions/pdf/safeway.pdf
3. PlanetRx. http://www-3.ibm.com/pvc/industry_solutions/pdf/planetrx.pdf
4. handheldmed. http://www.handheldmed.com

# 3 Device technology

This chapter gives an overview of the rapidly evolving technologies that influence the evolution of mobile devices. The key technologies are presented in three parts. The first part describes the advances in hardware technologies. The second part presents the key software technologies used for the human–machine interaction. The third part gives an introduction to some of the operating systems and Java platforms available for PDAs today. A brief discussion of each technology is given and followed by an outlook onto the future direction.

## 3.1 Hardware

There is a limit to the size of mobile devices that is imposed by the need to input and output data. The size of the input and output components, such as the keyboard and the liquid crystal display (LCD), influences the total size of a mobile device. While both are integrated into the same package, we cannot expect them to become any smaller than mobile phones are today. Once separated, the advances in integration will deliver even smaller devices. The display may be worn like a wristwatch or a head-mounted display. In the future, alternatives to the keyboard as the main input technology will appear. The devices might not even be visible any more because they will be integrated into the fabric of our clothes or hidden in glasses, pens, or jewellery.

### 3.1.1 Batteries

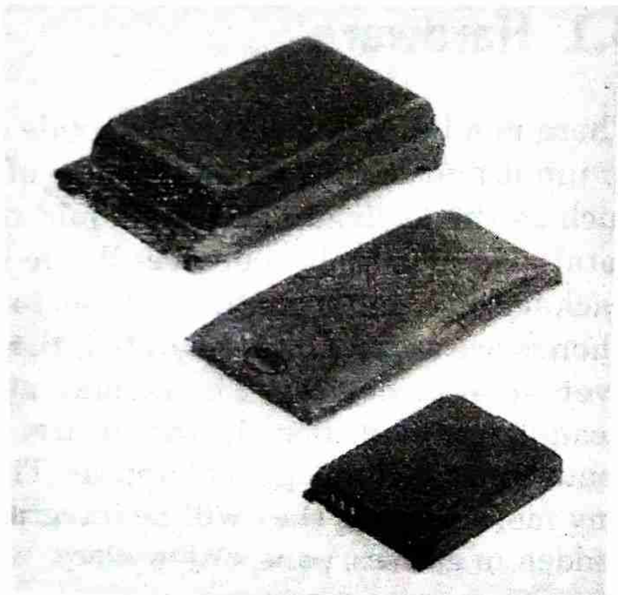
The battery technology has developed at a much slower rate than the other technologies presented here. Furthermore, most advances were immediately negated by increased power consumption from faster processors. Until recently, the nickel-cadmium (NiCad) cell was the state of the art of rechargeable batteries. These batteries were heavy and had a tendency to lose capacity through the so-called memory effect. Newer technologies, such as nickel–metal hydride (NiMH), delivered better capacity with less weight and environmentally friendlier components.

Today, lithium ion (Li ion) batteries can be found in all sorts of electronic equipment. Compared with their predecessors, these batteries are lighter and have better energy density, resulting in more power delivered by a battery with the same size. Figure 3.1 shows the various types of battery used for mobile phones.

As a result, the weight of a NiCad battery for a five-year-old mobile phone is often higher than the total weight of a modern mobile phone, including the Li ion battery. While the latter does have a lower capacity, it still offers longer talk time because of the reduced power requirements of modern devices. Table 3.1 gives an estimate of the expected standby and talk time for a mobile phone when used with typical batteries available today. The data is taken from the specification of three batteries of comparable size.

The latest in battery technology is the emergence of lithium polymer cells, which use a gel material for the electrolyte. The batteries are made from a few thin and flexible layers, and do not require a leak-proof casing. This means the batteries can be made in almost any shape or size.

Figure 3.1



*NiCad, NiMH, and Li ion batteries for mobile phones (top to bottom)*

**Table 3.1 Expected lifetime for NiCad, NiMH, and Li ion batteries**

| Chemistry | Standby time (h) | Talk time (m) |
|---|---|---|
| NiCad | 12–27 | 85–160 |
| NiMH | 16–37 | 110–210 |
| Li ion | 21–50 | 170–225 |

2

### 3.1.2 Displays

LCDs are already replacing the bulky cathode ray tubes of the past. The advances in manufacturing make them larger and more readable than before, and the dramatic weight, size, and power consumption benefits of LCD technology outweigh their relatively high cost. Today's PDAs usually feature dual-scan (DSTN) displays that control individual display elements via passive matrix addressing. This technology consumes considerably less power than the thin-film transistor (TFT) active matrix technology. This latter technology is more expensive, but is capable of significantly superior display performance and thus is generally used in portable computers.

Better and thinner displays will be available in the future based on the light-emitting organic diode (OLED) or light-emitting polymer (LEP) technologies. OLED technology was invented about 15 years ago; it only recently became commercially attractive when the initial problems with the expected life and efficiency were solved. Instead of crystalline semiconductor material, organic compounds are used. The simplified manufacturing process of smaller structures and a rich selection of organic compounds enable OLEDs to be built in almost any size and colour. This will eventually allow manufacturers to create extremely thin displays that are flexible enough to be bent and shaped as required.

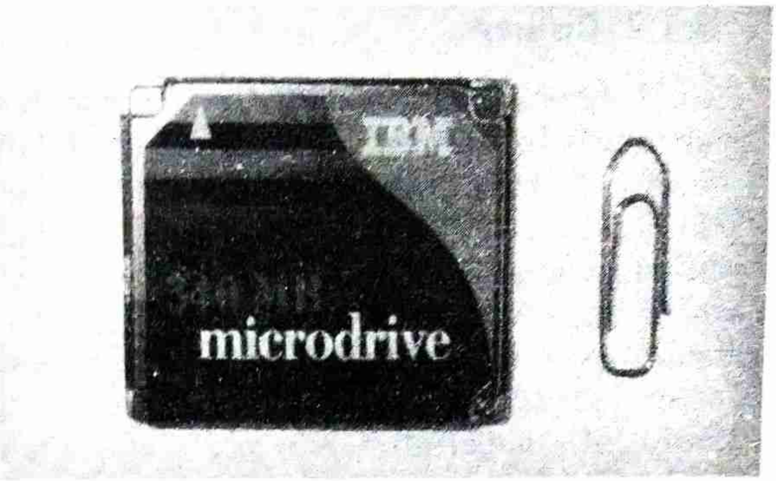
Other new display technologies, such as chip-on-glass (CoG) and liquid-crystal-on-glass (LCoG), integrate the picture elements with transistors on a layer of glass. This allows manufacturing of extremely small displays, with a pixel size of only 10 micrometers. In contrast to regular small displays like those on the back of a camcorder, the microdisplays usually require some form of magnification. They can be found, for example, in projection systems and in head-mounted displays used with wearable computers.

### 3.1.3 Memory

Memory is becoming cheaper, while the demand from applications is growing. Development is driven in part by smart phones, digital cameras, MP3 players and PDAs. For these mobile devices, the currently available technologies and their associated costs have reached a point where it is now feasible to integrate several megabytes of memory into a mobile device with an acceptable form factor.

On PCs, permanent data can be stored on hard disk drives. For mobile devices, this is often not an option because neither the space nor the power supply is available. Recently, extremely small removable disk drives like the IBM Microdrive shown in Figure 3.2 became available. Their capacity ranges between 340 MB and 1 GB, and is sufficient to store, for example, several hundred pictures when used in a digital camera.

**Figure 3.2**




*The IBM Microdrive with 340 MB capacity*

Other devices such as smart phones and PDAs store their operating system code and application data in non-volatile Flash memory and battery-backed random-access memory (RAM) instead. These semiconductor-based technologies require less power and offer faster access than disk drives. The typical capacity of built-in memory in mobile devices ranges from 2 to 16 MB. Expansion slots allow additional memory modules to be plugged into the device, which in turn allow data exchange and replace removable media such as diskettes and CD-ROM for a PC.

If a combination of Flash and low-power Static random-access memory (SRAM) memory is used, only frequently changed data is kept in SRAM. The advantages of SRAM over dynamic random-access memory (DRAM) are a simpler addressing scheme and the lack of refresh cycles. Newer uni-transistor random-access memory (Ut-RAM) tries to combine both technologies on a single die, and so delivers a higher memory capacity on smaller chips with comparably less power required. Another trend is the combination of magnetic and semiconductor memory in the magneto-resistive random access memory (MRAM) and ferroelectric random-access memory FRAM) technologies. The goal is to create a non-volatile memory with a behavior similar to that of static RAM. Because the typical refresh cycles of dynamic RAM are missing, they require very little power. This allows a single type of memory to be used in mobile devices such as PDAs and mobile phones, and replaces the common combinations of volatile and non-volatile memory in use today.

### 3.1.4 Processors

During the last couple of years, the clock rate of microprocessors and the processing power available from them has increased steadily. Rapid improvements in the CMOS manufacturing process have created ever-smaller structures and delivered higher and higher numbers of

4

transistors per chip. At the same time, the processor core voltage was lowered from the industry standard 3.3 V in 1995 to 1.35 V in 2000. This means lower heat emissions, which in turn paves the way for new improvements like larger on-die caches. This, together with advances in packaging technologies, delivers the modern Central Processing Units (CPUs) found in mobile computers and PDAs today.

### Intel's SpeedStep technology

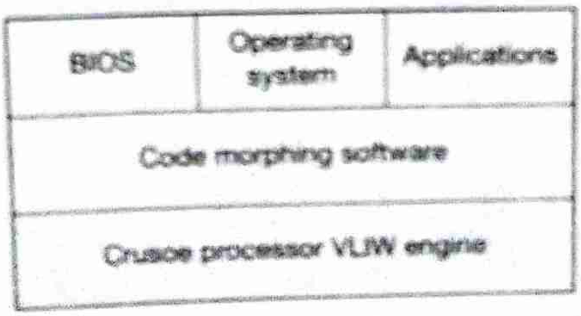
Recent processors include improvements in power management. These processors are capable of changing their internal clock frequencies and core voltage to adapt to changes in power supply. Newer designs are even capable of switching parts of the CPU on or off depending on whether the current calculations require them to be available. One such design is the SpeedStep technology from Intel. While the system is connected to an external power supply, the full clock rate and core voltage are available to the processor, resulting in the maximum performance. When running on batteries, the clock rate and core voltage of the processor are reduced, resulting in significant power savings. The transition between both modes is very fast and completely transparent to the user.

### The Crusoe processor

The latest in power-conservation technology is an attempt to reduce the total number of transistors, and to replace most of their functionality by software. The software dynamically translates the original instructions of the operating system and application software into another instruction set for the processor. Through the savings in the number of logic transistors, these processors promise to deliver the same or better performance than traditional CPUs while consuming considerably less power. The Crusoe processor by Transmeta is an example of such a design (Figure 3.3).

The Crusoe processor consists mostly of software. The relatively small processor core is designed as a 128-bit very long instruction word (VLIW) processor capable of executing up to four operations per cycle. The code morphing software on top of that emulates an x86-compatible processor.

**Figure 3.3**




*The Crusoe code morphing software*

5

In theory, this could be any other processor. During the boot cycle, the Crusoe processor loads its software into a section of the main memory. Frequently used code parts are optimized during run-time and kept in a separate cache. A technology called LongRun promises to reduce the power consumption even more by reducing the processor's voltage on the fly when the processor is idle.

The big advantage of this approach is that the Crusoe processor can be used to emulate almost any other processor and uses only a few watts, even with high clock rates. The disadvantage is the high memory requirements of the code morphing software.

## 3.2 Human–machine interfaces

Like their PC predecessors, many mobile devices also use keyboards and displays to interface with their users. However, these are usually much smaller and specialized for the application and the form factor of particular devices. Phones, for example, tend to have only number keys, plus a few extra keys for the built-in menus. This is because the size of the device is important and because users enter less text than on a PC. Other devices try to limit the number of mechanical keys to an absolute minimum, using them only to trigger the most important applications and for menu navigation. An example is the PDA. Finally, there are devices that have no means of display or keyboard whatsoever. These so-called headless devices are most often used as controllers and interface only to other devices.

### 3.2.1 Navigation

In order to operate applications in mobile devices, the user navigates through a menu structure, often using special navigation keys. An example is the integrated cursor key that delivers signals for all four directions by pressing or moving it up, down, left, or right.

**Figure 3.4**

*The Navi Roller on a Nokia 7110 phone*

Buttons that can be operated with the thumb while holding the device are especially suited for selecting entries from a menu list. These buttons can usually be turned or pressed. For each step in one direction, the menu navigation selects and highlights one of the entries. The entry is selected by pressing the button. To further improve the usability, the buttons can deliver haptic feedback to the user. Examples are the Navi Roller from Nokia (Figure 3.4) and the JogDial from Sony (Figure 3.5). Both are used in current-generation mobile phones to simplify navigation through the built-in applications.

### 3.2.2 Haptic interfaces

The programmable rotating actuator with haptic feedback is available from VDO.[1] It is basically a rotating control with force-feedback and a push button integrated into one. Sensors detect the position of the knob and an integrated motor produces feedback of torque when rotated. The way in which the motor responds when turning the knob is programmable. Haptic marks define positions of specific feedback force changes. Figure 3.6 shows a mechanical model for the haptic device to illustrate four different feedback characteristics.

When reaching a haptic mark, the user feels a resistance generated by the motor against the turning direction. This force increases until a specific position is reached. When the knob passes that position, the force gets smaller again. This can be used to create the impression of a knob that can be put into a programmable number of positions. It allows a single knob to be used for navigating through a menu structure where each menu choice is represented by one position.

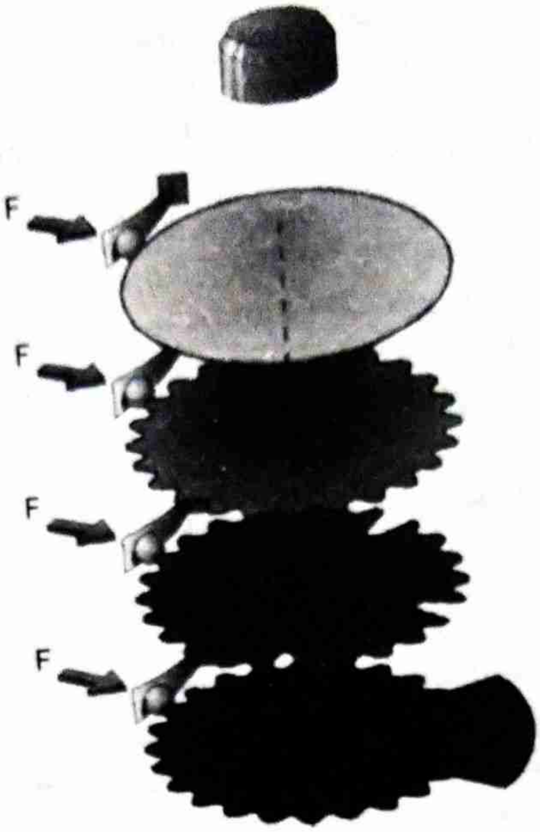
When used in a car, the knob helps the driver to feel the actual selection and to change it without even looking at the control. The programmability of the knob allows to be used it to control several settings with a different number of options through the same interface.

**Figure 3.5**

*The JogDial on a Sony CMD CD5 phone*

7

Figure 3.6




*Haptic device model*

Courtesy of VDO

### 3.2.3 Keyboards

Depending on the size of the mobile device, keyboards offer either the full set of keys or a limited set of keys for data input. Adding a full keyboard with a typewriter layout to a mobile device inevitably makes these devices larger. On the other hand, limiting the number of keys will automatically make the operation of the device more complex. Sometimes keyboards cannot be used at all because the form factor of the device simply does not offer the space for it, or the device is used in environmental conditions where a keyboard wouldn't work. Therefore, some devices completely omit keyboards in favour of other input technologies, such as handwriting or voice recognition.

#### On-screen keyboards

Devices with a reasonably large touch-sensitive display often make a compromise by replacing the mechanical keyboard with a virtual on-screen keyboard. This does not allow touch typing but still offers a convenient method for text entry. Numbers and special characters can be entered after switching into another mode, which alters the keyboard layout accordingly. Figure 3.7 shows the on-screen keyboard of a Palm PDA.

8

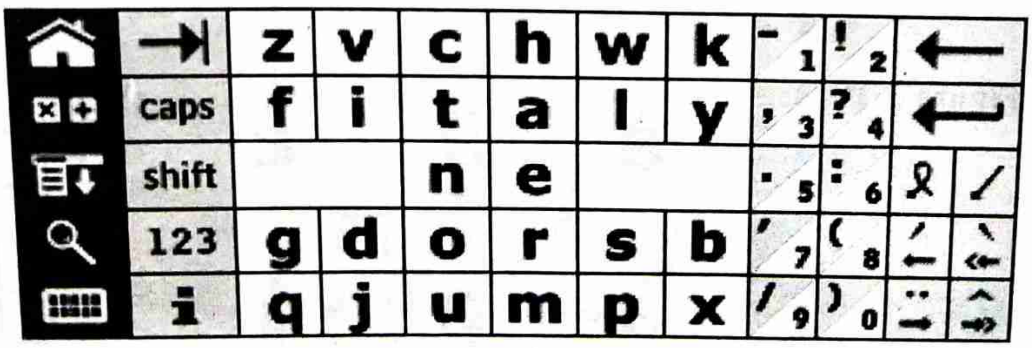
Figure 3.7


*The Palm on-screen keyboard*

Fitaly

Special keyboard layouts other than the traditional QWERTY typewriter layout may be used to speed up text input. The Fitaly keyboard,[2] for example, arranges the letters based on their individual frequency and the probability of transitions in the English language. The name is derived from the order of letters in one of the top rows of the layout, just as for the QWERTY layout (Figure 3.8). The arrangement of the letters for the Fitaly keyboard is optimized to keep the travel distance of the stylus during text input as short as possible. The letters 'i', 't', 'a', 'l', 'n', 'e', 'd', 'o', 'r', and 's' in the three center rows together with the space bar represent 73% of the keys used for normal English text. Adding the letters 'c', 'h', 'u', and 'm' brings this number to 84%. Remaining keys are never more than two positions away from the central area, and each key is placed near the keys it is most likely to follow in English text.

Figure 3.8




*The Fitaly keyboard layout*

9

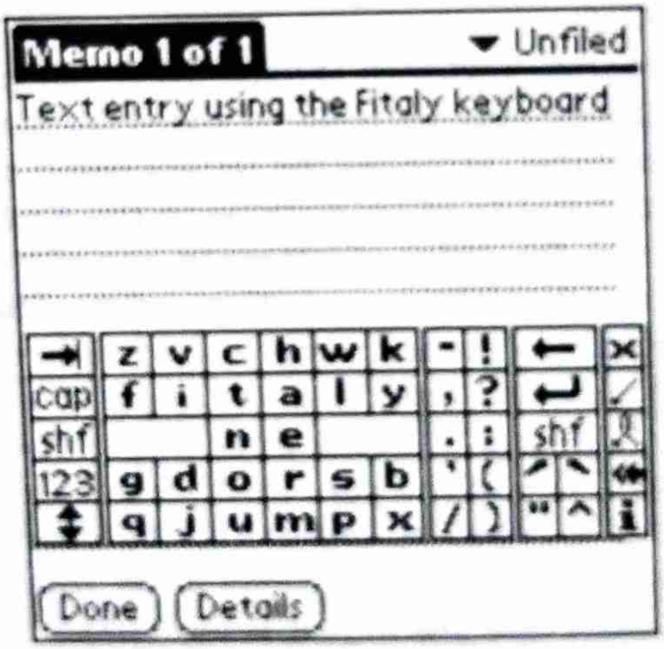
Courtesy of Fitaly

The Fitaly keyboard can be displayed like an on-screen keyboard (Figure 3.9). On the Palm and many other PDAs, where a separate input area for character recognition exists, it can also be attached to it in the form of a small overlay. This avoids losing precious display space of the application.

The Fitaly keyboard supports all 220 printable characters of the ANSI/ISO Latin-1 character set. Some less frequent characters are available by preceding the character with a special modification symbol. A technique known as sliding can be used to reach accents or umlauts. Instead of lifting the stylus after tapping a character, it is moved out of the character field in one specific direction that is suggested by the direction of the desired accent. This has been naturally extended to also work for diacritical marks (Figure 3.10).

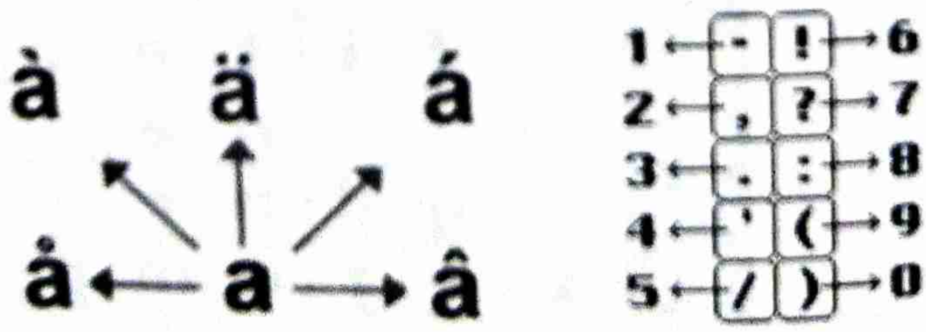
Sliding can also be used to enter a number without changing the keyboard mode. Each of the ten symbols to the right of the main area produces a digit if the stylus is moved out of the field before lifting.

Figure 3.9




*The Fitaly on-screen keyboard on a Palm*

Figure 3.10




*Inputting accented letters and numbers on the Fitaly keyboard*

Courtesy of Fitaly

### Tegic T9

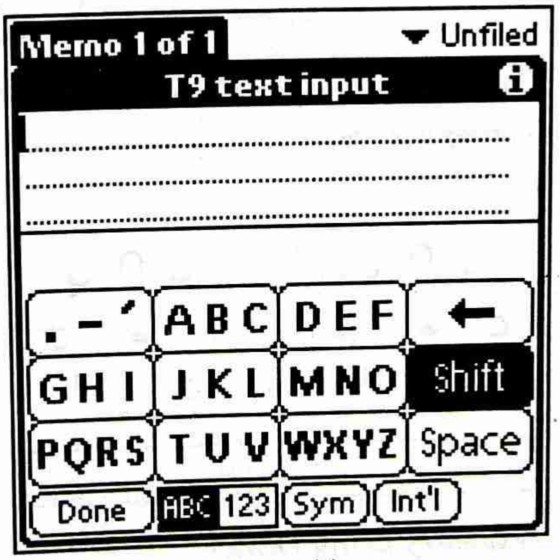
Several technologies strive to make the best use of a limited keyboard for text input. Perhaps the most popular one is the use of number keys for text input. Each number key is associated with a few letters of the alphabet. To enter words, a sequence of number keys is pressed with short pauses inserted between letters where necessary. To determine the exact letter, the number key has to be pressed several times. For example, the letters 'A', 'B' and 'C' are associated with the '2' key on the number pad of a phone. The '2' key has to be pressed once for 'A', twice for 'B' and three times for 'C'.

The T9 input system from Tegic[3] reduces the number of keystrokes by letting the user press each number key only once and working out later which of the several associated letters was meant. The software does this based on the preceding letters and with the help of linguistic rules derived from a dictionary. Ambiguities, such as the words 'Snow' and 'Pony', both represented by the same key sequence, are be resolved by prompting the user for assistance. This input system is already available on mobile phones and can be used for text-entry on PDAs. Figure 3.11 shows the text-entry process with the T9 keyboard displayed in the memo editor for a Palm device.

### Octave

Another approach to enter text without a real or on-screen keyboard is used by Octave from e-acute.[4] Octave maps each letter of the alphabet to one of eight unique strokes. The strokes are based on a common characteristic part of the letters they represent and are located around the tips

**Figure 3.11**




11

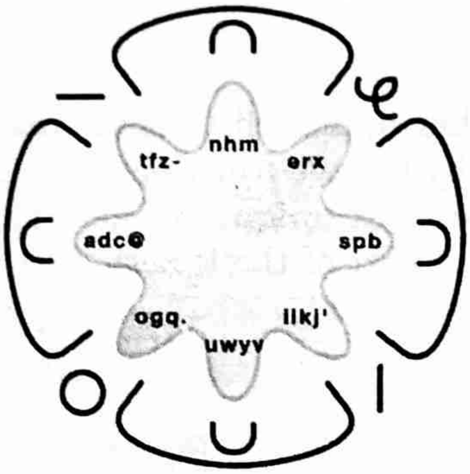
*The T9 text input on the Palm*

of a star-shaped pattern (Figure 3.12). On devices with a touch-sensitive display, a cut-out template is used to guide movement of the stylus into these eight points, each representing three or four letters of the alphabet.

Four special areas at the top, bottom, and sides outside the star are used to control the text-input process. The lower area, for example, switches into command mode, where special commands are available to insert the current date, change the language dictionary, or add a new word to the dictionary. Simple gestures within the inner area of the star are used to insert spaces or delete the last character.

During text entry, the stylus is moved from one position to the next without lifting it. To end a word, the user simply lifts the stylus from the surface. Word recognition is further augmented by dictionary-based methods. Ambiguities of stroke sequences are resolved by checking the probability of possible interpretations using the built-in dictionary. Once identified, the complete word is offered to the user for acceptance. Figure 3.13 shows the strokes needed to enter the word 'atmosphere'. After only four strokes, the complete word is identified from the dictionary and proposed for acceptance by Octave.

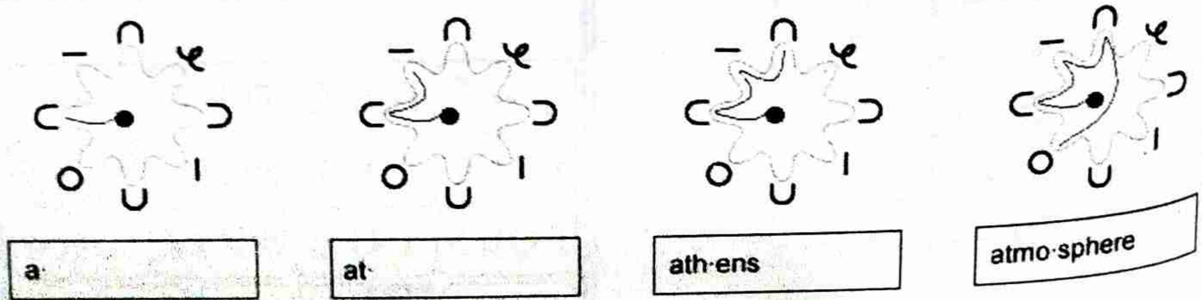
**Figure 3.12**




*The arrangement of strokes around the Octave star*

Courtesy of e-acute

**Figure 3.13**




*Text entry using Octave*

Courtesy of e-acute

The Octave input method can also be used on devices without touch-sensitive displays. In this case, a special button that can be moved into a limited number of positions generates the equivalent of the strokes. Figure 3.14 shows a prototype of a mobile phone with a single button for text entry.

### 3.2.4 Handwriting recognition

With the availability of sufficient processing power and touch-sensitive displays, handwriting recognition became feasible. The technologies available today differ widely in the amount of processing power and input precision they require. Recognition of cursive handwriting is much more complex than recognition of individually printed letters.

#### Word recognition

The most expensive approach in terms of computing power attempts to recognize complete handwritten words. This is the most natural way to write for the user, but the most difficult to recognize for a computer. The drawbacks of this approach are that it requires very precise data capture and provides only delayed feedback about the word actually recognized. Figure 3.15 shows a screenshot of text entry with the CalliGrapher[5] software on a Psion Series 5 device.

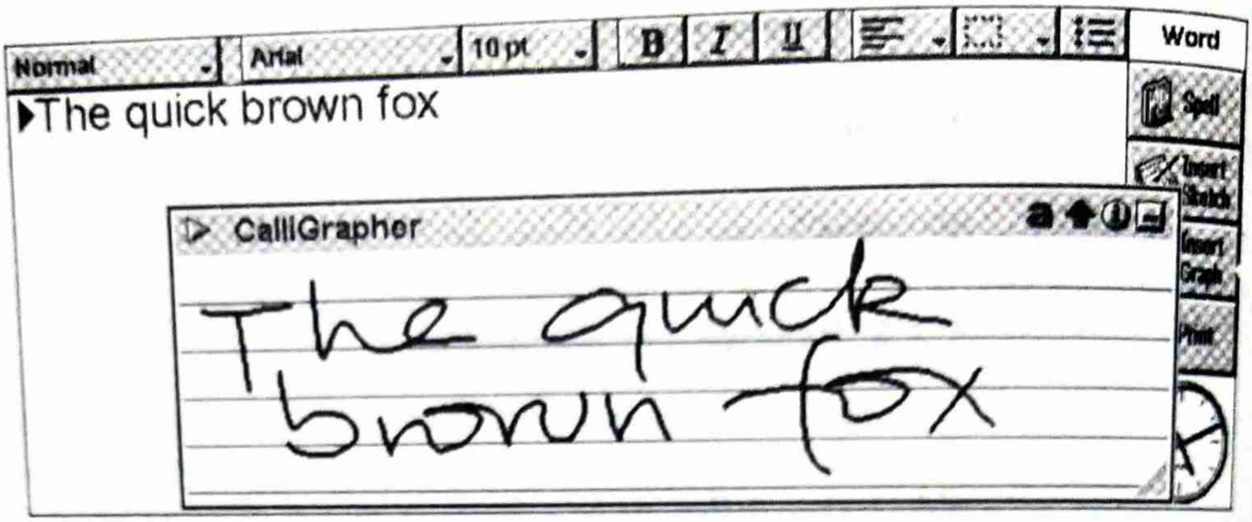
**Figure 3.14**

*The Octave button on a mobile phone*

Courtesy of e-acute

13

Figure 3.15




*CalliGrapher on a Psion Series 5*

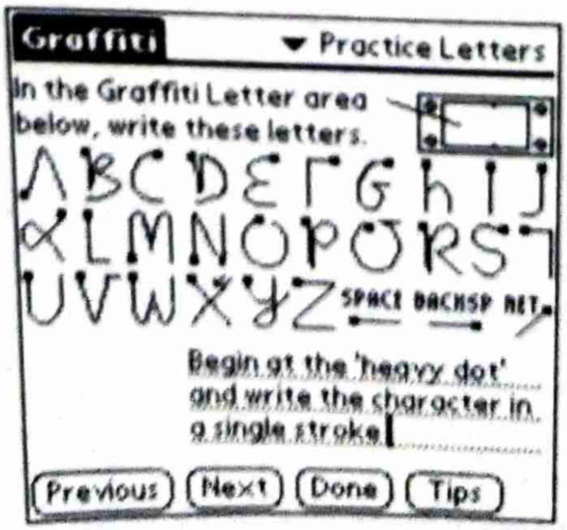
Courtesy of Paragraph

Character recognition

Other methods limit the recognition to separated characters, and require the stylus to be lifted between letters. These technologies usually achieve a very high recognition rate but require some cooperation from the user. Usually there is a limited number of ways how an individual letter has to be drawn in order to be recognized by the device. Figure 3.16 shows the Graffiti input method available on Palm OS devices. Letters have to be drawn in a single line exactly as shown, and the stylus has to be lifted after each character.

Some input methods can be trained to the user's writing habits, while others require the user to draw letters in exactly the way the device expects them to. The capability to adapt to an individual writing habit

Figure 3.16




*ffiti help on the Palm*

means that several profiles have to be managed if more than one person uses the device. Neither of these technologies can really compete with the keyboard to enter large amounts of text. However, their usability will improve with increasing computing power and memory available in mobile devices.

Localization

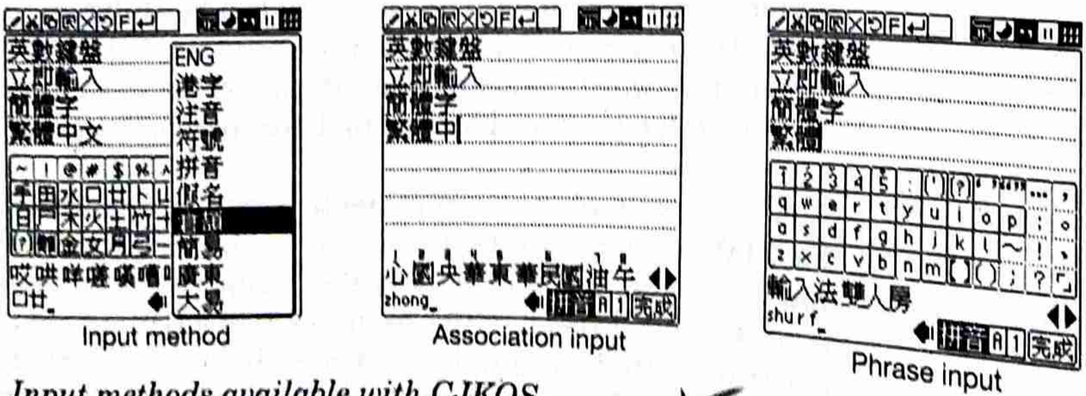
Special consideration has to be given to languages that are based not on small alphabets but on a large set of symbols, such as many Asian languages. Dedicated input methods based on standard character sets exist for these languages to allow entering tens of thousands of symbols. These input methods are typically based on the combinations of individual strokes or by using a limited set of phonics to form a single symbol. Figure 3.17 shows the available input methods of CJKOS, an operating system extension that completely adapts Palm OS for the display and input of the Chinese, Japanese, and Korean languages.

### 3.2.5 Speech recognition

Speech recognition has the advantage of being the most natural input method with only a minimum of requirements in terms of space required to integrate it into mobile devices. However, it is also the most expensive technology in terms of computing power, and the most vulnerable in extreme environments. Recognition of continuous speech is available in computers today, and will certainly become available in mobile devices too. Prototypes are currently available for research, but they still require additional hardware to support the speech processing. Chapter 7 discusses speech-recognition technology in detail.

Perhaps the most obvious devices for the integration of speech recognition are telephones. Some mobile phones already allow the selection of an

**Figure 3.17**




*Input methods available with CJKOS*

Courtesy of DYTS

15

entry from the address book by just speaking the name. In the future, they will be operated entirely by voice, understand complex queries, and may even be able to translate speech into other languages.

The flood of data generated by navigation systems, cell phones, and soon-to-arrive Internet access in cars significantly increases safety risks. Government regulations in some countries already restrict the use of phones in cars. Speech recognition seems to be able to help manage these new information sources and make them usable while driving.

In its simplest form, speech recognition can be reduced to recognition of a very small set of words. This is sufficient to control an application and to enter numeric data. Word recognition is used, for example, to operate devices while driving a car, or to check the status of a bank account via the phone.

Continuous speech recognition requires much larger vocabularies than word recognition, and has to deal with the complexities and ambiguities of the human language. It is available for text entry in PCs today. The same amount of processing power will have to be available in pervasive devices in order to use it there. This can be achieved by faster processors or by special hardware optimized for the recognition process.

## 3.3 Biometrics

Access to systems providing significant monetary value, confidential information, or critical applications must be secured against unauthorized use. User authentication is therefore a key function in any such system. Classical authentication relies on what you own, what you are, and what you know.

Typically, Internet applications rely on user identifier (UID) and a password. The UID may be in the public knowledge, whereas the password is a secret (what you know) shared by the user and the system administration. Thus, the system can check whether the user is authorized to use the system. However, any person who knows the password can perform user functions. The password can be stolen by watching the user enter their personal identification number (PIN), by capturing data during password transmission, or via access to system administration data. A stolen password is hard to detect because nothing is removed from any system.

GSM mobile phones are protected against unauthorized use through the SIM, a smart card that is issued by the mobile service provider (what you have) and a PIN chosen by the customer (what you know). The PIN is stored and checked in the secure system environment of the SIM and is not transmitted via unreliable media. Intruders must steal the SIM and the PIN in order to act like the authorized user and perform user functions. However, many users prefer not to use a PIN because it is inconvenient, and complicated procedures are required to recover a forgotten PIN and assign a new one.

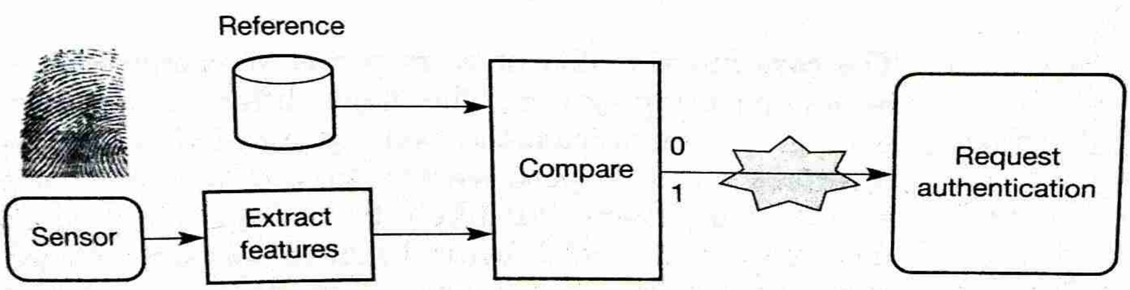
Biometrics authentication methods rely on what you are. A large number of personal characteristics, such as fingerprint, signature, hand geometry, face recognition, voice recognition, and iris scan, have been proposed. Only methods that work with small sensors, e.g. microphone, fingerprint sensor, or pen entry panels, promise near-term applicability in the pervasive computing space. Therefore, only authentication using fingerprint, speaker verification, or signature verification is discussed in the following section.

Biometrics authentication systems capture the user's characteristics with a sensor, derive characteristic values, and compare this with a known reference. The result of the comparison is either 0, if authentication was not successfully performed, or 1, if authentication was successful. In fingerprint verification, the image system extracts the end and the bifurcation points of finger lines, and uses location and direction as characteristic values for comparison with one or more stored references (Figure 3.18). Generation of reference values and actual values is subject to distortion. The comparison of the data sets is rather complex and introduces additional errors. Therefore, there is always the probability that biometrics authentication will fail, causing major inconvenience to the user. The false accept rate (FAR) is the probability that the system will accept the wrong user. The false reject rate (FRR) is the probability that the system will reject the correct user.

In the case of PIN verification with a SIM card, the FAR for any attempt by an unauthorized user is 0.01%. The system allows for a maximum of three trials before the smart card is shut off, thus the overall FAR for the SIM card system is 0.03%. The FRR is 0 because a correct PIN will always be accepted.

Biometrics authentication systems can be tuned, for example, to achieve a very small FAR. However, this will increase the FRR, and serv-

**Figure 3.18**




17

*Fingerprint authentication*

ice will be denied. Denial of service is acceptable if a reasonable exception path exists. A user of the INS-PASS issued by the US Department of Justice – Immigration and Naturalization Service may ask a security officer for entry if hand geometry or fingerprint authentication is not successful. Service denial for a mobile phone is much more difficult to handle, and will probably not be accepted by customers. In addition to the statistical FRR, every biometrics system has specific systematic problems, e.g. service may be denied if a finger is dirty or injured. Therefore, fingerprint authentication is very much limited to applications where adequate back-up processes in case of service denial are in place.

Another security disadvantage of biometrics systems is their susceptibility to fraud unless the total system is securely encapsulated and communicates with the authentication requestor via a secure channel. The SIM card system uses the smart card for secure storage of the PIN, and uses the keys for secure communication with the service provider. More than one billion smart cards are in use in Europe, and no fraud case involving professionally designed and implemented smart card has been reported. Thus, smart cards will be the security element of choice for pervasive devices. Some devices may use not smart cards but a module on a printed circuit card that serves the very same purpose.

Fingerprint verification will probably be introduced for pervasive computing devices as a convenience item, with a PIN as back-up in case of denial of service. The fingerprint authentication module must communicate with the SIM via a secure channel to maintain system integrity.

Pen-based pervasive computing devices may use signature (or handwriting) verification to secure access to the device. Modern signature-verification systems based on static images can be tailored to a FAR of 0.1% and a FRR of 5%. This level of security is comparable to a PIN-based system and reasonably convenient for the user. However, the system is prone to tampering because none of the system elements are contained in a secure housing.

## 3.4 Operating systems

The core functionality of every pervasive computing device is determined by its operating system. The major differences of operating systems for pervasive devices from the user's point of view are the human–machine interface, and the speed with which a task can be performed. For pervasive devices, there will likely be no equivalent to the Windows/Intel monopoly in the near future because pervasive devices do have a wide range of usages (from mobile phones to set-top boxes) with very constrained hardware. They will therefore deploy various chip sets (e.g. ARM, MIPS, Motorola, SH3/4, Transmeta) and multiple operating systems (e.g. Palm OS, EPOC, Windows CE).

All operating systems for pervasive devices support the basic operating system concepts, such as processes and file systems, but there are large differences between the various devices. The range is from the absolute minimum required in the Palm OS, to full-blown operating systems like Windows CE. This section will cover the most popular operating systems for the different pervasive computing areas.

Table 3.2 shows that the requirements for an operating system in PDA devices are very different from those for PCs. At the moment, there are two trends visible for pervasive computing operating systems. For personal use, the two major PDA operating systems, Palm OS and Windows CE, are becoming more similar, and can integrate phone functionality in a new device that combines a PDA with a cell phone. For home use, the development is directed towards high-performance multimedia operating systems, such as embedded Linux or BeOS.

Not covered in this chapter is VxWorks, a popular real-time operating system based on a microkernel architecture, which is quite similar to QNX Neutrino.

### 3.4.1 Palm OS

With a market share of about 80%, Palm OS is by far the most successful operation system for PDAs. PDAs with Palm operating system are available from Palm itself, Handspring, Sony, IBM, and others. The main reason for the success of Palm OS is that it is designed specifically for PDAs, and its easy-to-use approach, known as 'Zen of PDA'. It supplies only a limited number of features, but it optimizes these. This restricted functionality leads to lower memory and CPU usage, which results in longer battery life. The Palm OS is implemented on the 16-bit Motorola DragonBall CPU, which is derived from the 68k line.

At the time of this writing, the actual operating system version was 3.5, which added colour support for 256 colours. Palm OS 4.0 includes enhanced communication support (e.g. Bluetooth support) and 65 536 different colours. This will most likely lead to multimedia PDAs with integrated mobile phones.

**Table 3.2 Comparison of PDA and PC use**

| | PDA | PC |
|---|---|---|
| **Times turned on per day** | High | Low |
| **Time spent per task** | Low | High |

### Core operating system functionality

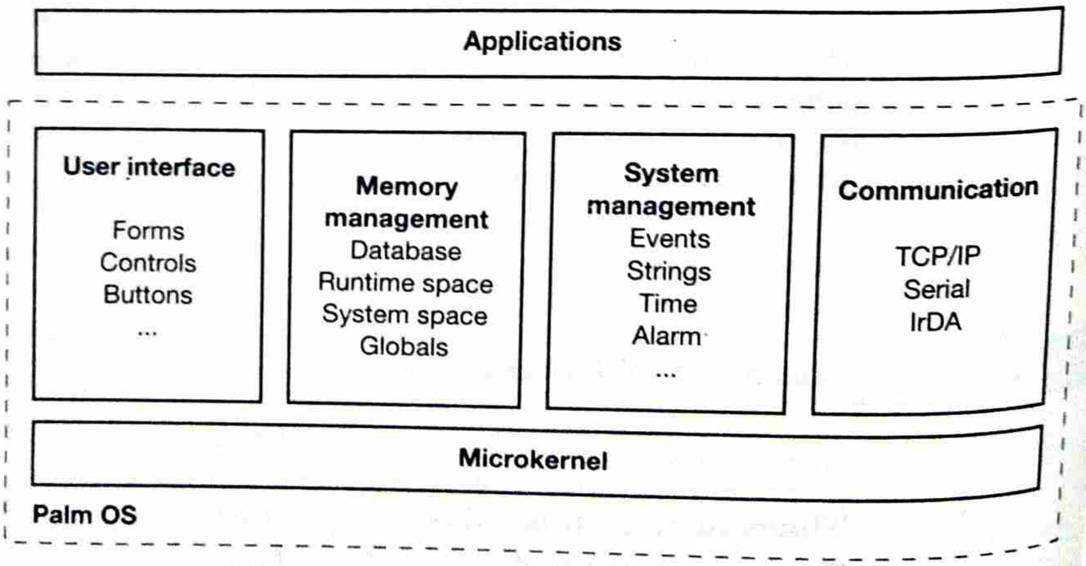
Palm OS is organized in different horizontal and vertical layers, as shown in (Figure 3.19). To make the vertical layers independent from the underlying hardware, a microkernel encapsulates the hardware-specific functionality. The vertical layers provide application programming interfaces (APIs) for the applications to access the operating system functionality. They are divided into:

- *user interface*, with the graphical input/output (I/O), e.g. buttons, forms, etc.;
- *memory management*, consisting of databases, runtime space, system space, and global variables;
- *system management*, which looks after events, alarms, date and time, strings, etc.;
- *communication layer*, which provides communication over serial I/O, Transmission Control Protocol (TCP)/IP, or Infrared Data Association (IrDA).

Palm OS has the following features:

- *User management*. Since the Palm device is considered to be a personal device, the Palm OS is a single-user operating system.
- *Task management*. Only one application can run at a time, but it is possible to call other applications from the running application. However, internally the Palm OS has the ability to use multitasking.

**Figure 3.19**




*Palm OS architecture*