APPENDIX

Sensing and Sensors for Context-Aware Computing

Huadong Wu (whd@cs.cmu.edu)

APPENDIX

Sensing and Sensors for Context-Aware Computing

1. Introduction

Two major research themes in computer science that elaborate on having computer technology better serve human beings are ubiquitous computing and context-aware computing: ubiquitous computing makes context-aware computing possible and context-ware computing makes ubiquitous computing valuable. According to Dey et al [5], "*context* is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves" and "a system is context-aware if it uses context to provide relevant information and/or services to the users, where relevancy depends on the user's task".

Surely, it is very difficult to have computers understand our human being's — the user's — contexts. According to Pentland (Alex Pentland, academic head of MIT Media Lab [7]), this is first a cognition problem. "In the language of cognitive science, perceptual intelligence is the ability to solve the frame problem: it is being able to classify the current situation, so that you know what variables are important, and thus can take appropriate action. Once a computer has the perceptual intelligence to know who, what, when, where, and why, then simple statistical learning methods are probably sufficient for the computer to determine what aspects of the situation are significant, and to answer a wide variety of useful questions".

We human beings use our sensing organs – eyes, ears, nose, skin, nose and tongue – to sense the outside world and use all kinds of knowledge to interpret what we feel. Naturally, sensing/sensor technology is the first step towards the perceptual intelligence in context-ware computing.

As *context* is a very broad concept, its relevant sensing and sensor technology can include various components. For the analysis of general context, it can be described in three domains: real world (space), time line, and spiritual world (including social relationship). However, from a specific user-application's perspective, for a given time, the context situation can be simplified by characteristics of: (1) outside environment (2) his/her own activity, and (3) his/her own basic information and current physiological and psychological states.

Today's technology is far away from enabling computers to sense all general contexts. With regard to context-aware computing research status, we list the following context information sources roughly according to their importance, accuracy, or the implement difficulties. This list is much more tightly related to the current sensing and sensors' technology development status, instead of a taxonomy elaboration.

1. **<u>User's explicit instructions, known facts, and direct inferences</u>**
   Assuming the system is for personal usage (single user, or several pre-registered users with each one has his/her unique system-behavior profile associated with the user ID), the user's identification needs be guaranteed all the time[1]. With the ID authentication

[1] If this is in doubt, the commonly used password mechanism will be used as the last resort.

granted, the current user's explicit instructions and schedules are regarded as the most accurate and reliable information sources for further contexts extraction. Because these kinds of information pieces exactly reflect the user's tastes and preferences for the given situations, these information pieces as well as other known facts (such as a line number being associated with a caller's ID) should be among the most important parts of contexts to be used. They can be used either as context components for applications, or as examples for system training to improve its future behavior.

2. **Sound**

   There are occasions when a camera visual system cannot provide useful context cues due to its awkward viewing perspective or object occlusion, but a sound detection system — though less sophisticated than the visual system — can still provide contexts, because of sound's more pervasiveness. Noise level and noise pattern can help to infer contexts such as "whether it is in a noisy subway station or in a quite library"; and if a human voice pattern is detected, it is then possible to infer "with whom the user is talking", or even further, "whether user is in an important conversation or in a casual chatting". Microphones as sound sensors are technically mature and economically cheap. The key to the acoustic context inference is in artificial intelligence research, or more specifically, in signal processing and pattern recognition.

3. **Vision**

   Vision sensing provides more than 90% information in our human's ordinary life: we can catch and infer a great deal of context information by just a glimpse of a scenario. The machine-vision sensor for context-ware computing can range from a simple photodiode to a most advanced APS (Active Pixel Sensor) CMOS camera system. Vision sensors' (mostly, CCD and CMOS cameras) hardware is now quite cheap and technically mature, but the corresponding image understanding technology is far behind — there is still a long way to go before a computer image-understanding system can provide us abundant and reliable context information as our human vision does.

4. **Vibration and motion**

   Motion and vibration information can be a very good clue to infer the user's activity and the appliance usage status, such as whether the user is driving a car or sitting at his office desk, whether the appliance is held in the user's hand or put on a desk etc.. Theoretically the motion information and position (location and orientation) information can be mutually derived as in the situation of submarine navigation systems, but the implement technology seems too complicated to be a practical solution under our application's cost and weight/size constraints.

5. **Location, proximity and Orientation**

   Obviously, user's physical location and orientation often define an important aspect of his/her environment context. For example, at large scale, a user's expectation will be completely different when he/she is at home versus when he/she is out-of-town; whereas at small scale, when the user is in his/her office versus when he/she is in a near-by cafeteria. Currently there isn't a mature location and orientation sensing technology for general applications. For outdoor applications, GPS or DGPS is the best choice for most cases; but for indoor proximity and orientation sensing, there are still many competing researches going on with many implementation schemes — almost all schemes need some kinds of infrastructure support, with some require more than others.

6. **User mental activity state**
Even if a computer can reliably sense the user's mood like a smart human secretary, to what extent the system should adjusts its behavior according to the sensed information is still a big challenge. On the other hand, although a computer cannot be as smart as a human secretary, the user's personal physiological state information will be greatly appreciated even if it is reduced to just for daily health status monitoring purposes. The user's affective state information that can be sensed relatively unobtrusively includes: heart rate, respiration rate, blood flow and volume pulse variation, Galvanic skin response (GSR which depends on the sweat gland activity and believed to be an very important indicator of mood stress), body temperature, skin complexion and Electromyogram (ESG which measures muscle electrical activity) etc. Technically these measurements can be implemented without great difficulty (it only requires meticulous efforts to make it less obtrusive and make the users willing to wear it), but whether and how these activities' pattern can be mapped onto a user's mood remains almost a mystery to be unveiled.
7. **Outside environment**
Environment information includes: ambient temperature, humidity, barometry, oxygen-richness and perhaps smell detection. With the Micro-Electro-Mechanical System (MEMS) manufacture technology advancement, it is now possible to incorporate local environment information with only a marginal extra cost, but how to effectively use these kinds of contexts is an interesting research subject to be addressed.
8. **People interaction and social relationship**
This is the most difficult, yet a very important, part of context-ware computing research. Firstly, there is still a very long way to go before we can reliably sense the presence of other people, and even more to do to recognize their identities and states (activity, expression, etc.). Secondly, people still don't know how to effectively describe the very complicated social relationship with a computer-understandable language yet, let alone how to detect such subtle relationship automatically.

In the rest of this document, we will analyze all aspects of the context-ware computing related sensing and sensor technologies, describe the physical principles behind them, the implementation infrastructure requirements, researches to improve performances and possible breakthroughs.

2. Sound

2.1. Acoustic Sensing

Sound information exists as patterns of ambient air pressure variation — in the form of oscillation of ambient pressure in the human auditory frequency range, roughly from 20 Hz to 20 kHz[2][3].

As vocal sound is the major communication means in current wireless cellular phone usage, the first step in our context aware computing project perhaps should be something that can

[2] Ultrasonic signal, sound waves whose frequency is greater than 20 KHz, is often used for object detection and distance measurement application, and will be discussed later.

[3] Human hearings on average are most acute to 2 ~ 5 kHz frequency components – lower than 0 db (20 μPa).

improve this basic service. Some examples of better cellular phone service include: to decide when and how much the conversation voice volume should be raised due to noisy environment (or if possible, when to activate some kinds of actively noise-cancellation techniques); to identify speakers so that it can always be sure that some user-specific features are activated only when those users are actually using the phone; to identify the speaking content to some extent so that it can take speaking orders from the user, or to guess the relative importance of the current conversation.

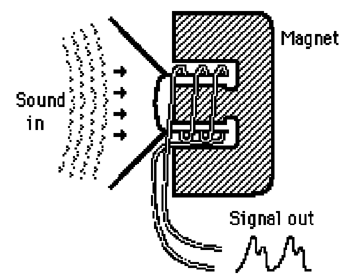
Acoustic signal can be sensed by microphones. The microphone sensors' hardware technology is very mature and can be manufactured in mass-product at very cheap costs, as described before, it is the artificial intelligence technology in sound signal processing that makes a big difference for providing valuable context cues.

2.2. Principles of Microphone sensors

Sound sensors or microphones are designed to pick up acoustic vibrations, convert the ambient pressure variation patterns into the form of electric current variation patterns. There are many different types of microphones: carbon, dynamic, ribbon, Piezo-electric, and condenser microphones, whereas the capacitive condenser type is the most commonly used because of its excellent overall frequency response. [13], [15], and [14] (page 295~301)

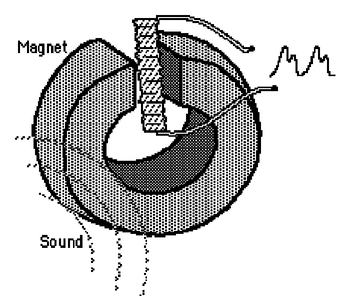
Carbon microphones were the first microphones widely used in the very early stage of telephone industry, and are now no longer used because of their poor frequency response and bad signal-to-noise ratios. A carbon microphone consists of a sealed small "bath" of carbon powder, with a metal diaphragm on the top of the bath. A battery is applied across the "bath" button, and as sound flexes the diaphragm the diaphragm in turn squeezes the carbon powder in the button, the carbon grains change their electrical resistance with the pressure thus a varying current is produced.

Dynamic microphones are essentially headphone drivers in reverse in their structures: the diaphragm (otherwise the speaker cone but typically being a thinner plastic film) is attached to a dense coil of wire, which is suspended in a magnetic field. The wire coil movement in the magnetic field, coupled from the diaphragm cone movement and that in turn is due to the sound pressure variation, generates a voltage in the coil.

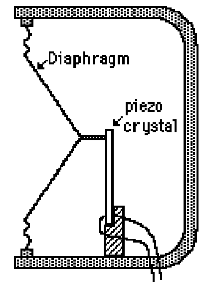


Dynamic microphones were once the basic workhorses of film and video industries, but they are now gradually replaced by condenser microphones – though they are often recommended for recording location narration as their less sensitivity and narrower frequency response range make them especially suitable to isolate the vocal voice from background sounds.

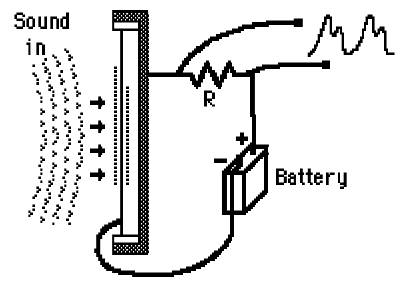
Ribbon microphones are variations on the dynamic microphone themes. It suspends a metallic foil (ribbon) in a magnetic field whose surface area also serves as a diaphragm. The movement of the ribbon in the magnetic field generates a voltage proportional to the velocity of the ribbon very much analogy to the dynamic microphone situation, but with less mass and inertia than a dynamic microphone, it has a very much-improved transient response.



Ribbon microphones are very rich in sounding; they're especially great on horn instruments and are often used in vocals to “fatten and warm the sounds up” [15]. Ribbon microphones are so sensitive and fragile that they are usually used only for special purpose recording in professional studio environments.



Crystal microphones use a thin strip of piezoelectric crystal material such as barium titanate and lead zirconate attached to a diaphragm. The two sides of the crystal acquire opposite charges when the crystal is deflected by the diaphragm, which moves according to the sound pressure. The charges are proportional to the amount of deformation and disappear when the stress on the crystal disappears. The electric output of crystal microphones is comparatively large, but the frequency response is not comparable to a good dynamic microphone.



Condenser microphones typically use a metallic diaphragm as the moving electrode and a closely placed back plate as the stationary opposite electrode in a capacitive measuring circuit. The sound wave pressure pushes and pulls the diaphragm and the change of the opposite electrode spacing will cause the change of electrical charge in the capacitive measurement circuit, which will in turn force a small current in the capacitance measurement circuit.

To obviate the need of a stabilized high polarization voltage in the capacitive measurement circuits, some condenser microphones use a permanently pre-polarized dielectric material for either their diaphragms or their back-plates. These kinds of condenser microphones are also called pre-polarized condenser microphones or electret condenser microphones. This technology makes the condenser microphones very cheap and sometimes very light and small — but at the cost of some frequency response characteristics degradation, as they usually cannot compete with the best regularly polarized condenser microphones.

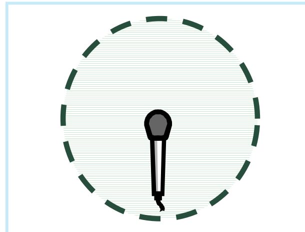
The metallic diaphragm is a very light membrane optimized to couple the ambient acoustic pressure vibration so that the charge on the membrane depends only upon the spacing — there is no appreciable resonance to skew the frequency response. As the condenser microphones' sensitivity is usually very high, and their overall frequency response is perhaps the most faithful and flat among all microphone types, condenser microphones are now the most popularly used type in all kinds of applications.

One thing regarding the usage of condenser-type microphones is that special care should be taken to avoid pop and crack noise when the microphone is too close to the sound source.

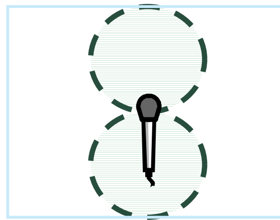
2.3. Microphone Sensor Implementation

Like any other sensors, the performance of microphone transducers can be described by specifications such as: sensitivity, overload capacity, linearity or distortion, frequency response, instrumentation-noise, signal-to-noise ratio, etc. Besides these specifications, there is a special characteristic regarding microphone design, i.e., signal pickup pattern, — the way that a microphone picks up sound from various directions. This sound signal pickup pattern originates from how much “pressure” versus how much “pressure-gradient (pressure difference between front and back side of the diaphragm)” is imposed on the microphone diaphragm in the sound field.

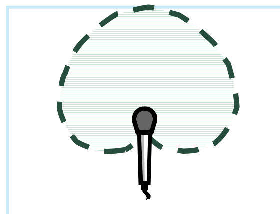
To simplify from application perspectives, the microphone signal-pickup pattern can be thought of as the counterpart of energy dispersion pattern for loud speakers, it is nothing more than the area over which a microphone will pick up a sound effectively. Microphone signal-pickup patterns are specified by the microphone output versus the angle of the sound source for sounds of a given amplitude and various frequencies, they are shown as circular curves in polar diagrams where the output is represented by the radius length at the incident angle. For any a microphone type, its capsule and housing can be designed to have one of the four typical sound-pickup patterns listed in the following paragraphs.



Omni-directional: this is the simplest microphone design – it picks up sound all around regardless of its origin. Omni-directional microphones are very easy to use and generally have very good to outstanding frequency responses (their frequency-sensitivity is flatter).

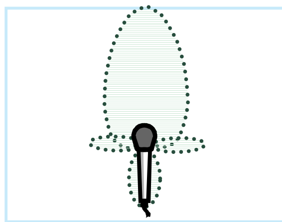


Bi-directional (figure of eight, figure-8): it is not very difficult to produce a pickup pattern that accepts sounds both in the front of and at the rear of the diaphragm but with little responds to those from the sides. This is the way any diaphragm actually behaves – if sound can strike in the front and the back as well. The frequency response of bi-directional microphones generally is almost as good as that of omni-directional microphones, at least for sounds that are not too close to the microphone.



Cardioid: this type of sound-pickup pattern takes the shape of a heart, picking lots of sound in front of the microphone but not much from back. They are probably the most common microphones, widely used for sound reinforcement in hand-held vocal microphones, or for stage recording such as concerts or ensembles where audience noise is a possible problem.

However, the cardioid sound-pickup pattern has some problems with it also. First, the actual polar curve graph varies with frequency components, for lower frequencies it is more like the omni-directional sound pickup pattern; second, the frequency response for signals arriving from the back and sides is uneven; and the third[4], a cardioid microphone will emphasize the low frequency components of any source that is very close to the diaphragm.



Hyper/super-cardioid, Unidirectional, and shotgun: it is possible to exaggerate the directionality of cardioid type microphones (at the cost of exaggerating some of the problems also), the elongated heart shaped pattern will pick up sound from a long narrow cone-shaped area in front of the microphone. The hypo-cardioid sound pickup pattern is often seen as a good compromise between the cardioid and bi-directional patterns as it gives a better overall rejection and flatter frequency response at the cost of small back pickup lobe. The "shotgun" sound pickup pattern pushes the techniques to extremes by mounting the diaphragm in the middle of a pipe: a shotgun microphone is extremely sensitive along the main axis, but possesses pronounced extra lobes which vary drastically with frequency. In fact, the frequency response is so bad for the shotgun pattern that it is usually

[4] The third one is popularly known as "proximity effects", which is sometimes deliberately used as a feature: as many singers and radio announcers rely on it to add "chest" to a light voice.

electronically restricted to the vocal voice range, where it is used to record dialogue in film and video industry.

As microphones cannot effectively sort out desired sound (in the situation of speaking to a cellular phone, for example) from undesired noise (reflected sound, ambient noise etc.), besides the microphone gain settings, the even more important aspect regarding how to use microphones properly is where the microphone should be placed. For a given sound source (a speaker for example) in any room, in any desired direction there is always a “critical distance” where the direct sound and the reflected sound are equal in intensity. In general, an omni-directional microphone should be placed no farther from the talker than 30% of the critical distance; and a unidirectional (cardioid, super/hyper-cardioid, or shotgun) microphone should be positioned no farther than 50% of the critical distance. The recorded speech will not sound articulate if the microphone is put too near or too far away to the speaker’s mouth.

2.4. Microphone Sensor Research

Although today’s microphone for audible sound pickup is mature and good enough to be cheaply used, there are still some interesting researches going-on to further reduce size, improve performance, and minimize costs. (http://www.darpa.mil/MTO/sono/microphone.html)

2.4.1. Microphone with CMOS and MEMS technology

MEMS technology makes it possible to build the sensing diaphragm directly onto a silicon chip; combined with the on-chip amplifier circuits, the whole MEMS microphone can be made very sensitive, small, cheap, and reliable. By arranging and electrically parallel connecting micro-machined capacitor transducers, the resulting system will achieve not only a wide frequency response range but also a large dynamic range (in excess of 100 dB).

Towards an even further performance enhancement, Eun Sok Kim from USC explored CMOS-compatible piezoelectric microphone design, trying to utilize piezoelectric microphones’ intrinsically simple fabrication requirement (no polarization voltage) and wider dynamic range properties (compared with condenser microphones) but to overcome their low-sensitive disadvantages by using corrugated MEMS diaphragm design to reduce residual stress in it. (http://www.darpa.mil/MTO/sono/presentations/UofHawaii.pdf)

2.4.2. Pickup enhancement and noise cancellation

A unique application of the condenser microphone is the pressure zone microphone (PZM) such as the Crown PZM (http://www.crownaudio.com/crownaudio/mic_htm/pzm.htm), which has a hemispherical pickup pattern. A miniature condenser microphone is mounted face-down next to a sound-reflecting plate or boundary. The microphone was developed especially for on-surface mounting to avoid the tonal coloration of other types of microphones which are placed close to a surface. The microphone's diaphragm is placed in the thin “pressure zone” just above the mounting surface where the incident and reflected sounds combine effectively "in phase" over the audible range. The benefits are a wide, smooth frequency response, excellent clarity, and little off-axis coloration.

The “pressure zone” concept comes from the idea that the reflected waves experience a 180° phase change upon reflection and therefore add to the incoming wave. This has the potential for a doubling of effective sound amplitude, which gives four times the sound intensity or a 6 decibels increase in sensitivity. (http://hyperphysics.phy-astr.gsu.edu/hbase/audio/pzm.html)

For personal communicators and teleconferencing sound pickup to combat the detrimental effect of background noise and the sound reverberation that serious degrades microphone reception, Elko at Bell Labs explored a differential microphone array solution. The designed structure uses sensor elements' array that spaced very closely compared to the acoustic wavelength, and the elements are combined in an alternating sign fashion, such that the differential effects results a super-directional microphone system. [41]

2.5. Contexts from Sound

2.5.1. Signal processing

Once microphone sensors detects sound signal, the first task in signal processing is to measure sound loudness level and check for some specific patterns such as human voice, music, engine noise etc..

There are some research reports for noise pattern such as vehicle identification, mechanical vibration diagnosis, and even music understanding etc. showing in academic journals, but the most intensively research are in speech recognition and speaker recognition, and there are some remarkable improvements over the last ~3 decades.

2.5.1.1. Noise level

Sound loudness measurement is simple yet a very important aspect of sound characteristics for context understanding. For example in a noisy environment you would like the system to raise the voice of your calling partner, but you may feel embarrassed when your cellular rings loudly while the music is very soft in a concert.

2.5.1.2. Microphone deployment and gain setting

As our human hearing system has such an amazing dynamic range, often a piece of sound signal that sounds very clear to our human being may easily run out of the dynamic measurement range of a microphone. In general human vocal recording or music recording practices, for high quality recording, it often needs a dedicated attention to set up microphones to an appropriate distance and to adjust the microphone signal pickup gain.

Low quality recording method (the commonly used AGC— Automatic Gain Control method) seems not adequate for general context-aware computing sound-sensing purposes. Because very sophisticated AI computation is needed to extract different levels of contexts, the "garbage in garbage out" situation definitely needs be avoided. However, for many applications, we will have to deal with the problem of sensing various sound sources of large dynamic range, and it is often too tedious or even impossible for the users to adjust the microphones' deployment and their gain settings.

This problem may be solved to some extent when the MEMS large dynamic microphones become readily available (perhaps in 3~5 years). On the other hand, with hardware price dropping down and more computation power cheaply available, an alternative possible solution is to deploy several microphones of the same sound pickup patterns closely in an array and set their gain different (specifically from small to large). With this arrangement, a simple signal-processing algorithm and a sensor fusion scheme can always provide suitable raw sound data for further context extraction analysis.

2.5.1.3. Signal preprocessing

In voice recognition research, the term “signal processing” usually refers to the signal preprocessing — preparing for the further pattern recognition work (speech recognition, speaker recognition, music understanding, etc.). The signal preprocessing is much the same for all those following pattern recognitions, and because of the similarity, some of such dedicated DSP chips are designed to perform the tasks.

The first step in sound signal preprocessing is to divide a piece of long continuous sound signal into small pieces or frames (usually with large overlaps) that can be regarded as an invariant signal. With speech recognition, the frame lengths usually are selected as 10~30 ms.

Then, after applying some kind of smoothing windows, the frames of signal are undergoing a fast Fourier analysis. The resulted spectra – spectrograms – are the raw data for checking specific patterns.

2.5.2. Speaker recognition

Speaker recognition deals with verifying someone who says she/he is (and assigning a new identity to the voice of an unknown speaker). Today, speaker recognition technology has not successfully found the key features to reliably distinguish registered speakers yet, so computers are generally not smart enough to recognize speaker’s identity with randomly spoken contexts as our humans do in daily life. Text independent speaker identification is still an in-progress research topic in the voice recognition research field.

However, for given speaking contents, the speaker recognition technology is much easier and becomes technically more mature — it is close to actual application deployment. BACOB bank (one of the top five retail banks in Belgium) has been considering offering speaker verification security features to its customers for “banking through telephone” services since early 2000, based on the success of a pilot system [31].

In middle 1999, BACOB initiated a pilot telephone-banking project using speaker verification from Keyware (http://www.keyware.com/). The pilot is with BACOB employees who have accounts with BACOB. Each employee in experiments enrolls three items: full name, address, and date of birth. During verification, the system selects one of the enrolled items for verification and asks the caller to say it. If the response is not satisfactory, a second item is selected, and second and third responses are not satisfactory the caller is rejected and the call is terminated. The system performed very well in providing secured and easy-to-use services: it rejected identical twins of two employees who attempted transfer money between the employees’ accounts, and in another case an employee was unable to fool the system by making a tape recording.

2.5.3. Speech recognition and contents understanding

“Speech recognition” refers to identifying the words and phrases that a person is saying; and “contents understanding” refers to the system’s understanding what the person is trying to communicate.

There are two general types of speech recognition: discrete and continuous. The discrete speech recognition requires its user to speak one world at a time — with a noticeable pause between every two spoken words; whereas continuous speech recognition allows the user speak to it more naturally at more or less normal rate. Modern speech recognition research projects and products are mostly continuous speech recognition, though there is a disadvantage in it — a continuous system does not make ongoing adaptation to the user's speech whereas a discrete system usually does.

Some speech products and some research prototypes claim to be "speaker independent", but usually their recognition correctness rate is much lower than that of "speaker dependent" systems. For speaker dependent systems, as the user trains and speaks to the system, the software creates a user-specific voice file that contains a lot of information about the user's voice qualities, pronunciations, and patterns of word usage[5]. The software uses this acoustic and linguistic information to make its best guess at each word or phrase as it is dictated.

There are now four major companies vie for the speech market: Dragon Systems, IBM, Learnout & Hauspie, and Phillips. For the biggest user group, the Windows users, their currently available speech recognition commercial products are: Dragon Systems' Dragon Systems (NaturallySpeaking, DragonDictate), IBM's ViaVoice, Learnout & Hauspie's Voice Express, and Phillips' Free Speech.

Nevertheless, overall, for speech recognition to be widely used as the major human-computer interface, there is still a long way to go. Besides recognition accuracy, the two biggest hindrances of its wide-acceptance are the difficulty to train a recognition system and the difference between public expectation and reality of the "intelligence" in spoken language content understanding.

A speech recognition system alone is still not so smart after all: it does not understand what the user tries to communicate. As context-aware computers do not exist yet, today's computer speech recognition systems are far from chattering with human users. However, for some given situation, as in the telephony call forwarding, the so-called IVR (Interactive Voice Response) research makes remarkable progress recently and there are many impressive pilot products using a conversational style in telephone service. [32]

2.5.4. Sound source localization (phased-array microphones)

The sound waveform propagation speed in atmosphere is known. Theoretically, with two microphones by correlation calculation we can easily detect the time-delay or phase difference between the two sound channels so that the sound source can be inferred to be somewhere on a parabola curve; and with three microphones the sound source can be inferred to be at one of the intersection points of the two parabola curves.

This method is a simple and effective way to localize sound sources, but in practice it works only under the conditions that the being detected sound source is the only one or the dominant one and the sound reflection is relatively weak.

Another big potential to localize a sound source is to use phased array microphones, similar to long-wave phase radar detection in principle and perhaps to navy submarine sonar detection applications. It is believed that the there is a large amount of research (and perhaps a bunch of successful methods) for sound source localization classified as advanced military technologies, so there are not many research reports publicly available though it is mentioned as a very interesting research topic occasionally.

---

[5] Both discrete and continuous speech recognition software captures the user's preferred vocabulary, but the voice file in discrete speech recognition software is built primarily on the user's pronunciation of individual words whereas the voice file in continuous speech recognition also contains information about the user's grammar and word usage (i.e., which words/phrases tend to be used in what order).

3. Vision

Vision comprises one of the very important parts in a user's contexts. The contextual cues from vision can range from brightness, object/people recognition, to object/people spatial dimensional and location estimation or measurement.

3.1. Vision sensors

Vision sensing now is much matured technology; there are many cameras available with all ranges of performance specifications and price.

3.1.1. Environment brightness, infrared radiation

Brightness is a very import aspect of environment, and it can be very easily measured by cheap photosensitive resistors, photosensitive diodes or photosensitive transistors. The photosensitive resistor's resistance changes as the amount of the light that falls onto the resistor changes. The amount of current a photosensitive diode conducts varies as the level of light varies. For a photosensitive transistor, the amount of energy passing from the transistor's collector to the emitter is controlled by the amount of light falling on the base, which is sealed in the transistor.

Light is a kind of electromagnetic waveform whose color changes from red to violet with the wavelength changing from ~0.7 μm to ~0.3 μm. When the wavelength is larger than 0.7 μm, it is outside our visual range, and called infrared light. The very important characteristic of all heat objects is that they all emit some amount of infrared light: a human body is constantly emitting infrared light and it can be detected by measuring the emitted infrared light.

The photosensitive sensing element in a photosensitive diode or a photosensitive transistor can be made to especially sensitive to infrared light. Thus, the cheap infrared diode sensors can be very preliminary sensors for detecting human existence in context-aware applications. Another example of infrared proximity sensor applications can be found in public restrooms, detecting (e.g.) your hands underneath the faucet to provide an accurate measurement to the nearest object within a range of 15 cm to nearly a meter.

3.1.2. Image sensing: CCD video cameras

For video image sensing, CCD (Charge Coupled Device) video cameras are already cheap and become the standard sensors with various products of different performance specifications available. A CCD video camera consists of three parts: a lens with house to focus the object scene on the image plan; a CCD chip with its photosensitive top surface aligned at the image plan; and signal reading-out accessories. The CCD chip is the center to convert photo image into electric charge image.

A CCD is a semiconductor device similar to the integrated circuits found in televisions and computers. It's usually a small (less than 1" square) silicon chip subdivided into as many as 4 million picture elements, commonly known as pixels. The pixels are arranged in a matrix, or rectangle. When light (in the form of a photon) hits a particular pixel, an electron is generated. The CCD chip will store this electron at the location of the pixel; for example, if a photon of light hits a pixel at the upper right of the CCD chip, the electron will be stored at the upper right of the CCD chip. More than one electron can be stored at a particular pixel location – if a number of photons hit a specific pixel, a corresponding number of electrons will be stored at that location. [40]

By scanning the matrix of electrons stored at the pixel locations on the CCD chip sequentially with the reading-out accessory circuits, the image is sent out from the camera at given video frame rate.

3.1.3. Image sensing: CMOS APS cameras

A CCD chip requires the use of many other support chips to provide a system solution, so the CCD system consumes a lot of power. CMOS technology advancement provides cheaper and denser diodes and transistors that now can be used for imaging. In 1994, the Jet Propulsion Laboratory re-invented the CMOS APS (Active Pixel Sensor) technology, which can integrate those supporting circuits into or close to the pixel component so that it has the potential of CCD image quality with lower power and system costs.

Right now CMOS APS technology has poorer image quality, because of leakage noise in the photodiodes. However, this situation is being drastically improved, it is believed that the CMOS APS technology will soon displace the CCD in the low end of the market — for toys, biometrics, and PC videoconferencing.

It is predicted that CMOS APS systems will displace more than half of the mainstream CCD markets by the year of 2003 (http://www.hpl.hp.com/features/wayne_greene_interview.html). Next generation video cameras have the trend to package everything on a single chip: CCD, conversion circuitry, even lens — to make cheap eyes in a myriad of applications such as surveillance, security, party games, etc.

3.2. Vision sensing and context extraction

3.2.1. Variant factors for image sensing

Vision perception or image understanding as part of our human being intelligence is by now far more can computers can understand: many aspects that we take it for granted really matter in artificial intelligence computation.

In order for a camera to sense a high-quality object images, the object needs be appropriately focused onto an image plane. For a given object and at given distance from the object, the camera lens basically decides how big the image will be on the image plan: a lens with a longer focus length will result in a bigger object-image with smaller view angle and verse versa. A camera with fixed lens focus length is of course not convenient to use, in reality, most cameras have a combined lens with adjustable focus length to some extent. To simplify calculation in image understanding, it often assumes the object-image mapping as a linear one-to-one mapping like a pinhole camera system. However, we need exert a careful consideration in choosing camera lens system, because the out-of-focus images will often cause extra difficulties in object recognition when the objects are not well focused.

Overall, today's computer image understanding system is still pretty preliminary: it is very sensitive to view perspective variation — for given application tasks the artificial intelligent algorithms will work only with very strict perspective constraints, and this situation seems wont change for a very long time.

Our human vision system has an amazing capability to adjust to a very large dynamic lighting range, much larger than any camera CCD or film imaging system. For an artificial intelligent image understanding system to work properly, we have to carefully set suitable camera aperture as well as control scenery lighting conditions.

3.2.2. Stereovision

We human being vision system has two eyes to look at an object from slightly different perspectives, and meanwhile our brain synthesizes the observed stereo images so as to enable us to perceive 3-D space information.

Similarly, with two cameras pointing at an object from slightly different perspectives, and under the condition that we know the two cameras' position as well their object image position information, then theoretically we ought to be able to derive the object's position in 3-D space with some triangular geometry calculation. This is what computational stereovision does. However, in reality, the main difficulty to implement this simple concept stems from a fact that a computer is hardly smart enough to match corresponding points in the two images. Around this problem, there are many researches going on and some have come out impressive results, but in general, their recovered 3-D object models are far from satisfaction.

3.2.3. Object recognition

The very important part of our intelligence — imagination — is not quite understood yet. Yet, perhaps this is the part that plays the key role in our daily object tasks. There are some research activities address partly occluded objects' recognition, but most successful recognition systems do not use much "imagined" information to facilitate their tasks.

The most preliminary yet often quite effective object recognition method is to "guess" based on some color block existence. For example, if we know that this is a picture of a pastoral scenery image, then we have a big chance to be correct if we assign a block of upper blue pixels as "sky" and a block of lower white pixels as "sheep".

Another commonly used attacking approach, beside the neural network method, which can best be described as another "guess" method, is trying to geometrically model the observed objects. The so-called "structure from motion" problem that has eluded computer science vision researchers for years is: "give you a videotape of a room that one has made by walking around with a handheld camcorder, can you create a three-dimension model of the room as well as determine the camera trajectory?" Researcher in the Robotics Institute Carnegie Mellon claimed to have found a robust solution to this problem. [42][43]

Nevertheless, we are now still too far away for context-aware sensing system to generally infer objects and their function information from vision image. Yet, for applications of very constrained conditions, it is possible that a context-aware computing system can derive its location and proximity information by recognizing important landmarks, and in the near future to recognize some devices through vision.

3.2.4. People Tracking

There has been much work in the area of people tracking using computer vision techniques. Several real-time systems have been presented that work reliably under a few established conditions that do not change over time. Color-based tracking with a static background is the basis of the widely used people-tracking system. Several other systems have also concentrated on similar color metrics for tracking faces in real-time. [17]

The increase in computational power has made it possible to further extend the tracking methods to allow using multiple cameras and combing their images and other visual cues (like motion, for example). This has resulted in a variety of real-time systems that have been used by thousands of people at various exhibitions.

There has also been a great progress in locating people in a scene using methods to look for faces. These methods, though much more reliable than the color-tracking methods, are computationally quite expensive making them somewhat infeasible for real-time systems yet. These methods require higher resolution images than the color-tracking methods. Some attempts have been made to incorporate these methods with real-time systems to aid in visual tracking system initialization and registration.

A major issue with most of the real-time systems is that each method works under specific conditions. Methods for merging of color, motion and depth measurements towards reliably finding and tracking a person under varying lighting conditions with large camera movements are still under research and perhaps there is a very long way to go before the people identification and activity contexts can be readily available from computer vision systems.

4. Vibration and Motion

"Motion" refers to position change related to some world reference coordination system. Roughly, as we are not dealing long-term measurement such as in marine navigation, the earth and sun rotational inertial effects can be neglected: we are living in an inertial world where position is a simple integration of speed and speed in turn is a simple integration of acceleration.

Motion, like location/proximity, comprises a very important aspect of contexts in most context-aware computing applications, but unfortunately, it cannot be easily measured without a supporting infrastructure. However, one good thing about motion detection is that in almost all circumstances, we need not deal with steady motion but rather we are more concern with intermittent regular or irregular motion, which is tightly coupled with acceleration — and acceleration is much easier to measure.

In the earth inertial reference coordination system (i.e., the ground reference system), a simple way of acceleration measurement[6] is to use a "proof" mass plus some force measurement means (such as, a spring plus a position meter, a pendulum structure, a piezoelectric gauge, etc). With the MEMS technology drastic advancement in the last few years, now the whole measurement system can be integrated onto a single silicon chip.

As MEMS accelerometers with one to three sensitive measuring axes (e.g. Anolog's ADXL202) are now so cheap and easy to use — it does not require any supporting infrastructure to implement, one should not be surprised to find them widely used in context-aware sensing systems. A good such example is the Esprit TEA-II context-aware research project prototype, where accelerometer measurement is one of the primary sensor information to derive the hand-held device's usage status. (http://www.tea.starlab.net)

5. Location, Proximity, and Orientation

A sensor system that can provide people or objects' location information with accuracy within a few centimeters, orientation with accuracy within a few degrees, everywhere without the need for a large number of base-stations would be ideal for general purposes context-aware

[6] A more accurate acceleration measuring method is to use a gyroscope accelerometer, where the torque generated from attaching a proof mass to the rim of a gimbaled gyroscope is measured and counted through the gyroscopic precession mechanism. A gyroscopic accelerometer can be very accurate as used in submarine and ballistic missile navigation system, but it tends to be too expensive to use also. [46]], [44]], and (http://www.fas.org/nuke/hew/Uas/Weapons/Airs.htm]

computing. Such an ideal location system doesn’t exist and perhaps won’t be feasible in the near future. However there do exist many practical sensing technology candidates for different context-aware tasks’ non-intrusive location, proximity and orientation detection, ranging from inertial gyroscope, stereovision, to object-tagging systems.

Except for inertial gyroscope navigation system, almost all location and orientation systems need some kind of infrastructure to support them to work properly, with the object-tagging technology perhaps being the least complicated to implement but most dependent on infrastructure support. Object location and proximity, relative position and orientation resolving, object registering and tracking, all the sensing technologies essentially use the same or similar physical principles. In the following sections, we are going to analyze the most commonly used implementation technologies and researches, roughly from coarse to fine granularity.

5.1. Location sensing implementation technologies

5.1.1. GPS, DGPS

GPS was a United States Department of Defense initiative seeking to accurately pinpoint troop positions in military scenarios. A robust system was needed that could accurately ascertain a soldier's location on a global scale.

To meet these criteria a satellite based system was designed whereby a receiver on the planet surface would be able to ascertain its location from computations based on listening to clock signals sent from a subset of a constellation of 24 orbiting satellites.

In a simplified explanation, the receiver computes its position by calculating the distance of the satellites using the clock signals they transmit. Now a notional sphere can be constructed around the satellite with a radius of this distance. If an additional sphere is calculated from another satellite, the receivers position must then lie somewhere along the intersection of the two spheres. Using a further third sphere reduces the location to two points of intersection from which the actual location can be deduced (one is easily discarded as it is normally way out in space). In practice, the distance to a fourth satellite is also required so that the clock error between the receiver and satellite clocks may be calculated. It is possible to locate oneself anywhere in the world within an average of 30-meter accuracy.

Satellites broadcast on two frequencies and provide two sets of positioning data, one available to the public and the other an encrypted signal restricted to military use. For non-military users, the accuracy is not as high as it potentially could be as the DOD employ a selective availability policy to reduce accuracy and, supposedly, to prevent military rivals from using the technology. However, the civilian uses of GPS, especially in maritime environments, have expanded rapidly and eventually the selective availability policy is expected to end. However, measurement errors will still limit the accuracy of a single receiver to about 15~20 meters.

Accuracy can be improved to a few meters by using differential GPS (DGPS): +/- 2 meters horizontal accuracy and +/- 3 meters vertical accuracy with a differential base station within 100 kilometers [30]. The idea behind the differential positioning is to correct bias errors at one location with measured bias errors at a known position. A reference receiver, or base station, is used to compute corrections for each satellite signal. Further improvements to sub-meter levels are possible using techniques based on carrier phase measurements on the satellite signals and, by applying these methods to both broadcast frequencies, modern survey quality GPS receivers can achieve sub-centimeter positioning.

There are many GPS and DGPS receiver products commercially available now, they are not very expensive (prices in the range of $100~$250) and they usually have a RS232 cable interface ready so that a GPS receiver can be easily incorporated into a map system running on a laptop or a PDA for vehicle traveling guidance applications.

For example, Hertz NeverLost In-Car Navigation System uses GPS with smart sensors (perhaps plus a digital compass and odometer readings) to achieve turn-by-turn guidance accuracy throughout the U.S. and in selected cities in Canada at nominal daily fee of $6 (http://www.hertz.com/serv/us/prod_lost.html).

5.1.2. Gyroscope, MEMS inertial navigation system

A typical gyroscopic localization or navigation system consists of a gyroscope to measure and maintain inertial orientation and a set of three linear accelerometers to measure acceleration in three orthogonal axes.

The central mechanism of a gyroscope is a wheel similar to a bicycle wheel. Its outer rim has a heavy mass and rotates at high speed on very low friction bearings. When it is rotating normally, it resists changes in direction. With a gimbaled structure where the gyroscope is set in the center of the structure such that the whole structure can move in 3-dimension space without interfering the gyro rotation, any surface attached to the gyro assembly will remain rigid in space even though the case of the gyro turns. This rigidity in space property can provide a reference for any change of heading, pitch and roll of a moving structure.

With the inertial orientation reference, monitoring the acceleration in two orthogonal axes in horizontal plane (or with an additional accelerometer in vertical direction if we need measure 3-D space movement), and first integrating the acceleration into velocity then further integrating the velocity into distance, we can provide system location information without any outside infrastructure support.

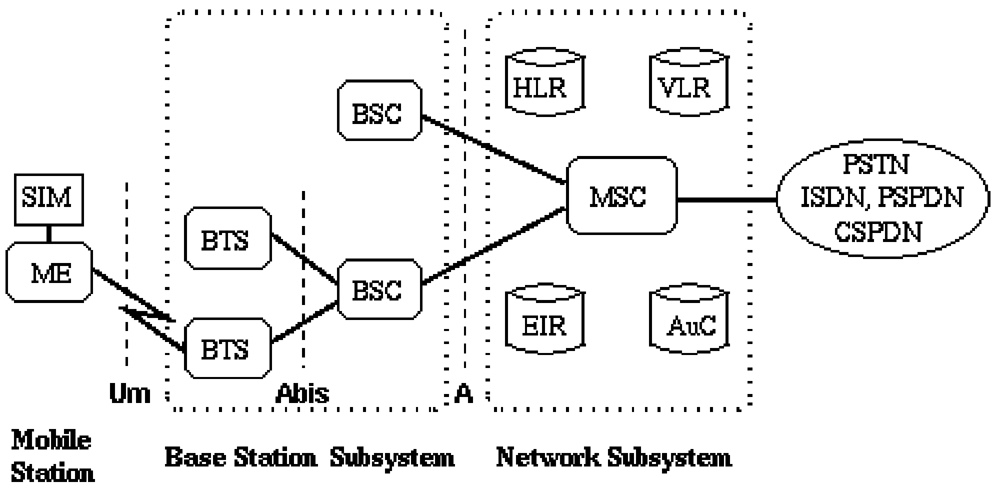
In reality, because of gyro drifting effects and the accumulation effects in acceleration-distance integration, this localization implementation scheme demands extraordinary efforts in manufacturing as well as usage calibration to make it work reliably. The result is that such a system is very expensive and suitable only in special applications such as submarine navigation, international ballistic missiles navigation systems.

However, the MEMS technology dramatic advancement provides a great potential to build micro-gyroscopes and even micro inertial navigation systems. There are many research projects going on in this area. As an example, the DARPA (US Government Department of Defense Advanced Research Project Agency) MEMS INS (Micro-Electromechanical Sensor Inertial Navigation System) program aims to improve the silicon-based, inertial sensors (gyros and accelerometers developed in the MEMS technology program), and integrate them with navigation software into a low-power, small, lightweight, tactical grade inertial navigation system. (http://www.darpa.mil/spo/programs/memsins.htm). It doesn't seem too bold to predict that, with the help of some calibration system (such as, a DGPS enabled system) the MEMS navigation/localization technology will be the dominant technology for all location and proximity information service just like the role of GPS localization today.

5.1.3. GSM wireless communication

In 1982, the Conference of European Posts and Telegraphs (CEPT) formed a study group called "Groupe Special Mobile" (GSM) to study and develop a pan-European standard public land mobile system. The GSM Phase I specifications were published in 1990, and the term GSM has been redefined as “Global System for Mobile communications” — the trademark of the system. The GSM specifications have been soon adopted beyond European countries and GSM now is the largest system for mobile communications in the world. The United States and Canada version of GSM is called PCS1900 (Personal Communications Systems).

The specified GSM is a fully digital cellular system allowing both speech and data services and allowing roaming across networks and countries. The GSM network can be divided into three main parts: (1) the mobile stations (MS) that are usually handheld terminals being carried by the subscribers; (2) the base station subsystems (BSS) that controls the radio link with those mobile stations; (3) the network and mobile service switching subsystem that performs the switching of calls and handles mobility management operations. The covering area of an operator network is divided into many cells, with the size of each cell being decided by the power of the BTS (base transceiver station) in the base station subsystems. The GSM architecture is shown as: [49]



SIM Subscriber Identity Module
ME Mobile Equipment
BTS Base Transceiver Station
BSC Base Station Controller
HLR Home Location Register
VLR Visitor Location Register
MSC Mobile services Switching Center
EIR Equipment Identity Register
AuC Authentication Center

There are a few technology alternatives for locating mobile phones based on overlay triangulation technologies: from the timing or angle of signal transmission, electromagnetic signal energy dissipation, or the Cell of Origin information. They are briefly described in the followings paragraphs.

The **COO (Cell Of Origin) method** uses the mobile network base station (BTS) cell area as the indicator of the caller location. It can be easily inferred that the COO positioning accuracy depends on the size of the cell: in urban areas with the deployment of Pico-cell sites it can be at the scale of 100 meters. The COO method is the only technology that is widely deployed in wireless network today to meet the Federal mandate[7] for geo-locating cellular E911 calls in USA.

[7] The Federal Communications Commission (FCC) of the USA adopted a ruling in June 1996 (Docket no. 94-102) that requires all mobile network operators to provide location information on all calls to “911”, the emergency services. The FCC mandated that by the October 1, 2001, all wireless 911 calls must be pinpointed within125 meters, 67% of the time. On December 24 1998, the FCC amended its ruling to allow terminal based solutions as well as network based ones (CC Docket No. 94-102, Waivers for Handset-Based Approaches).

The **E-OTD (Enhanced Observed Time Difference) method** places reference beacons at multiple sites geographically dispersed in a wide area (overlaid on the cellular network) and measures the time difference of the arrival signal (synchronized signal among reference beacons) at the MS handsets. The differences in time stamps are then combined to produce intersecting hyperbolic lines: with at least three base station measurements, the location can be estimated. The system needs to build the reference beacon infrastructure as well as E-OTD enabled software in the handsets, which means that this service cannot be provided to the existing customers. The E-OTD method offers a greater positioning accuracy than COO method, between 50 and 125 meters, but have a slower speed of response, typically ~ 5 seconds.

The **TOA (Time Of Arrival) method** is similar to the E-OTD method but measures the time of arrival of signal from a mobile device to three BTS's. To measure the time differences the GSM network BTS's need to be synchronized (usually through GPS time information or atomic clocks) and the investment required for an operator may be very high. The system response is far slower, typically ~ 10 seconds, but the very good thing is that there is no need to modify the existing handsets to provide them with this localization services.

The **AOA (Angle Of Arrival) method** uses a complex 4~12 small aperture directional antenna array at each of several cell site locations to measure the direction of the handset, thus to find out the intersection of the directional projection lines. The **Signal Attenuation method** estimates a cellular phone's location using the principles of signal attenuation as the mobile phone moves away from a base station. These two methods are less used, as their reliability is relatively low.

5.1.4. IrDA: active-badge, smart-badge

Light, whether it is infrared light, visible light, or ultraviolet light, physically is a kind of electromagnetic wave phenomenon. Our human vision system, our eyes, can see light whose wavelength ranges from ~0.3 μm to ~0.7 μm. The "infrared" light usually refers to the light whose wavelength is greater but close to 0.7 μm. As all objects we daily deal with are several factors larger in dimension, the infrared electromagnetic waveform refraction and diffraction efforts are so weak that they can be ignored in almost all applications. When it is used in communication, this factor characterizes the communication channel in that it works only when the light transmitter and receiver are in a line-of-sight — otherwise even if a connection via light reflection passages does exist, it is usually very unreliable.

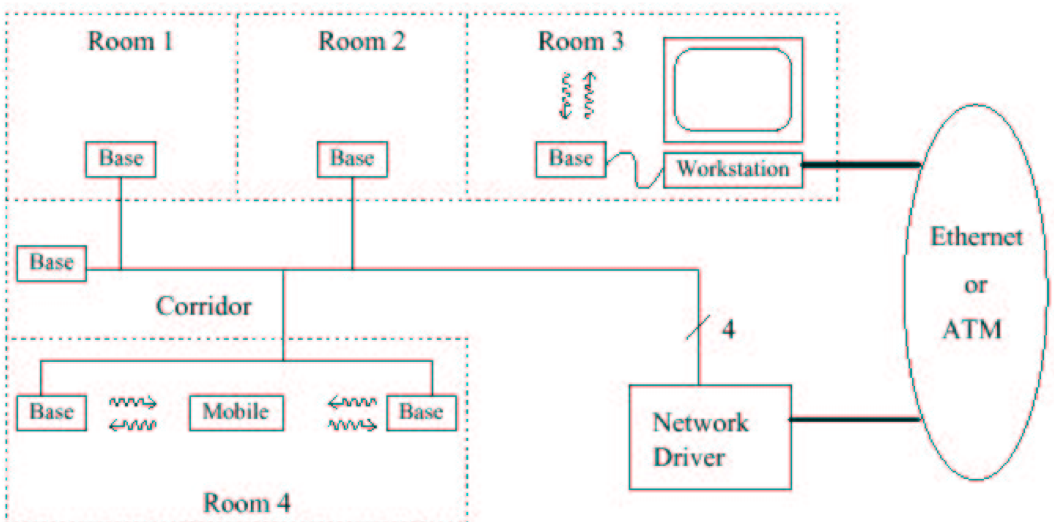
An infrared communication connection includes an infrared LED, a photodiode, and a few IC's in a one-way information data flow. The LED transmits data as an optical signal. The photodiode receives data from the optical signal, and the IC's control the translation from data to a signal and back to data. The IC also controls the signal strength at different transmission distances.

The Infrared Data Association (IrDA) specifies three infrared communication standards: IrDA-Data, IrDA-Control, and a new emerging standard called AIr. They all regulate specifications and protocols regarding cordless connection using infrared light. The IrDA protocols specify that an IrDA compatible communication connection should be able to perform continuous operation from contact (0 distance) up to at least 1 meter away. For most IrDA products, their actual operation range can usually reach about 2 meters far.

Once an infrared communication connection is established, with the location knowledge of one transceiver, the proximity of the other transceiver can be surely inferred; this mechanism provides the basis of many so-far developed infrared location systems. The "Active Badge"

People and Objects Tracking System was the first context-ware or ubiquitous computing related system developed at the Olivetti Research Laboratory, which is widely referred in the context-ware computing research area (http://www.uk.research.at.com/ab.html). It was actually primarily intended for use in locating people for the particular purpose of more effectively directing incoming calls to the telephone extension nearest to the being called person's location.

The active badge system is comprised of a number of active badges (transmitters) and active badge sensors (receivers). The active badge itself is a small device that can be clipped to one's jacket pocket or belt. It emits an intermittent beacon (in the form of an infra-red transmission) approximately every 10 seconds identifying its owner by a unique number. The sensors receive such signals and forward the sightings on to a server. The topology of the fixed network of the wireless receivers (sensors) is the basis for the location system, as the position of a mobile wireless transmitter (the active badge) is determined by the identity of the receiver sensor in the network. A dedicated location server builds and maintains a database of location information gathered from the sensor sources. The system is illustrated in the following figure. [28]



As infrared signals cannot penetrate the borders of a room, the location cells will typically be the size of a room, although in large spaces such as open plan offices there could be more than one sensors per room[8]. The room size cells are then naturally mapped into units of location in the location server, which are immediately accessible to users, i.e. room numbers.

A simple query of the location database allows a user to ascertain where a particular person is. More importantly, for context-aware computing it also allows the users to find out where they are and additionally who and what (any inanimate object such as a printer can also be tagged with badges[9] just as effectively as a person) they are with.

[8] The range of infrared transmission system can reach around 13 meter without much difficulty, but increasing the density of receivers allows the transmitter power to be reduced and the granularity of location to be refined.

[9] The badges for equipment objects are tuned to declare their existence every five minutes to extend their battery life to several years as compared to about 1-year battery life for people tracking badges. Special support for equipment badges is available whereby a simple device status reading can be additionally transmitted.

5.1.5. Smart Card and iButton (Java Ring)

A smart card is typically a credit-card-sized plastic card with a small-embedded computer microchip which can store and process electronic data. The embedded chip provides significantly more memory than conventional magnetic-stripe cards, currently about 100 times more storage capacities.

Smart cards can be classified into two categories: memory cards and processor cards. A memory card may only store data and requires an outside processor to access and work with its data; thus, it provides only minimum security. A processor card has its own microprocessor embedded, complete with its own operating system, and can process and store data on its own. A processor card processes and verifies data, such as a PIN code or biometric data, internally and only issues a valid authentication certificate, thus making it almost impossible to tamper with the sensitive data flow. A processor card that carries an encryption coprocessor for additional security is also called an encryption card.

The physical size of a smart card as well as the software data interface is defined by the international standard ISO 7810. This standard also defines physical characteristics of the card, including temperature range and flexibility, physical position of electrical contacts. Data communication between the card and terminal takes place through one single serial line, and the data transfer has also been standardized by ISO standards.

The communication between a smart card and a terminal (usually at a card application service station such as POS, telephony, transportation, banking, and healthcare transactions etc.) is always triggered by the terminal. The process of data communication is as follows: (1) the terminal sends a command to the smart card; (2) the smart card microprocessor process the command; and (3) the smart card sends out an answer which may already contain data, an acknowledgement, or an error code relating to a command that may or may not have to be executed.

JavaCard is a smart card that is capable of running Java byte codes. iButton is a variation of smart card product from Dallas Semiconductor Corporation (http://www.ibutton.com/ibuttons/). It has three different types: memory iButton, Java-powered cryptographic iButton, and thermochron iButton. The third one, thermochron iButton, is a little further variation from smart card in that it integrates a thermometer, a clock/calendar, a thermal history log, and 512 bytes of additional memory to store shipping manifest.

Physically an iButton is a 16 mm computer chip armored in a stainless steel can. Because of this unique, durable package, and because it can “steel” power from the data line with the included lithium cell battery only for maintaining data in volatile RAM memory when not connected to a LAN, the iButtons can be used in very harsh application environments, such as used in every-day wearable jewelry fashion as the Java Ring products.

Because of their durability and cheap costs, the Java Ring fashioned smart cards are suitable for context-aware computing applications — to announce the existence (location proximity) of a person or an object in order to trigger events. However, as it needs the user to physically touch the iButton Ring against the LAN terminal to initiate an event, it is usually considered as an “intrusive” location method: arguably, a user will often feel troublesome or forget to do it.

5.1.6. RFID, Near-field Radio

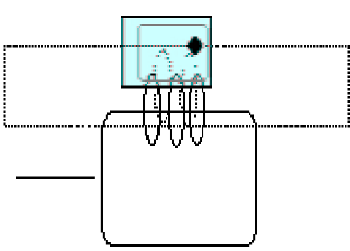
Radio Frequency Identification (IFID) technology differs from infrared connection technology essentially in that the frequency it uses is much lower – so low that its waveform diffraction and refraction effects can no longer be ignored but to the opposite, it plays such important role that it completely changes the implementation models. Here as the electromagnetic wavelength is comparable to or larger than our daily-life objects, the communication connection no longer depends on line-of-sight signal transmission. Given an RFID design and with controlled power, it can also be used for people or objects' location detection purposes just like IrDA technology, and their principles are the same.

A basic RFID system consist of three components: (1) an antenna or coil, (2) a transceiver (with decoder), and (3) a transponder (commonly called an RF tag) that is electronically programmed with unique information. Often the antenna is packaged with the transceiver and decoder to become a reader (a.k.a. interrogator). The reader emits radio waves in ranges of anywhere from one inch to 100 feet or more, depending upon its power output and the radio frequency used. When an RFID tag is within the electromagnetic zone, it detects and responds to the reader's activation signal according to the encoded information in its integrated circuits (silicon chip). The reader picks up the tag's responses, decodes the data and decides whether the signal is a repeat transmission, if it is a repeat transmission it may then instruct the transponder to cease transmitting otherwise the tag's response is forward to a network server. This method of possible contention resolving is known as the "Command Response Protocol", or sometimes referred to as "Hands Down Polling". An alternative, more secure, but slower tag polling technique is called "Hands Up Polling" which involves the interrogator looking for tags with specific identities, and interrogating them in turn.

As RFID can be used in non-contact and non line-of-sight applications in harsh operation conditions such as through snow, fog, paint, crusted grime etc., it is evolving very fast recently. Many RFID systems are very smartly designed and implemented, they already give great impact in many applications. Take Texas Instruments' "Tag-it" RFID as an example: the tags come in two size: 1.8×1.8 in. and 1.8×3.0 in. at 15-mil – thin enough to run through a standard printer — with a chip and antenna incorporated (the tags cost between $0.25 to $1.00 depending on volume); the corresponding readers operate at ~13.56 MHz and the read distance is as away as 1.5 feet. — Systems like this would surely be good candidates for context-aware computing applications.

Basically the frequency difference but in another view, two methods distinguish and categorize RFID systems: one based on close proximity electromagnetic or inductive coupling and the other based upon propagating electromagnetic waves.

5.1.6.1. RFID of close electromagnetic coupling



Zimmerman [29] has demonstrated a fringing-field technology termed PAN (Personal Area Network). Data rates of up to 50 kbps have been demonstrated with an elegant, low power circuit implementation. However, PAN requires contact electrodes and the data rate is limited to the poor high frequency characteristics of the human body.

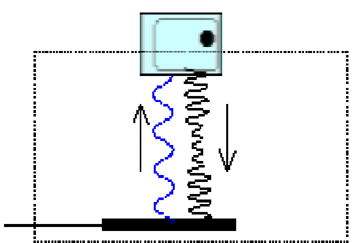
To overcome this limitation, the Hewlett-Packard Laboratory's "Foot-Bridge" [25] instead uses a near-field radio system, which does not require explicit contact to an electrode. Foot-Bridge is not claimed as a competitor to PAN; but instead it is a complimentary technology where physical contact is not required and capable of relatively high data rates. Foot-Bridge and PAN can equally co-exist in a wearable computer.

The precise definition of the “near-field” is somewhat complex and is a function of both frequency and antennae geometry. A common definition uses the distance r, $0<r<\lambda/2\pi$, from the transmitting antenna. Within the near-field region, energy transfer between a transmitter and receiver is via the electric or magnetic field component. A transmitter/receiver system using loop antennae will utilize the magnetic field component in much the same way as the coupling between the primary and secondary of a transformer, so allowing information (e.g. an identifier) to be passed from the tag to a detector.

The main advantage of near-field coupling stems from the restriction of operating range to a fraction of the carrier wavelength (6 m at 8 MHz). This arrangement virtually eliminates the problems of reflections and standing waves that disrupt tagging systems based on infrared or conventional radio [24].

The first prototype near-field radio communication, Pinger — developed in Hewlett-Packard for “situated computing” (i.e. context-aware computing) — can be used to identify people, places and objects. Its magnetically coupled radio link has communication bit-rate of 10 kbps with a range of 3 meters. The next prototype “Foot-Bridge” employs the same principles but with shorter range – up to 1 foot. [25]. The further limited range of transceiver enhances the communication bandwidth to 300 kbps ~ 1 Mbps and enables finer proximity detection. It is designed to be used as the radio transceiver (the radiating/receive antennae is a single turn loop of approximately 8 ~ 10 cm in diameter) worn on the on the user’s wrist or shirt cuff.

5.1.6.2. RFID of electromagnetic wave propagation coupling



RFID systems generally can be categorized according to their frequency range as low (100~500 kHz), intermediate (10~15 MHz), and high (850~950 MHz, and 2.4~5.8 GHz). Within each frequency range there are specific laws and regulations that govern the use of the frequency, including its power levels, interference as well as frequency tolerances. Higher frequency corresponds to short electromagnetic wavelength, and at the higher end correspondingly the systems more directionality. For the high frequency RFID systems, the coupling between the tags and the transceivers are through electromagnetic wave propagation. Although the level of available power is the primary determinant of the communication range, the manner and efficiency in which that power is deployed and the environmental conditions and structures also have big influences. Generally speaking, systems of electromagnetic wave propagation coupling are designed to work at greater distance than the systems of near-field radio RFID systems.

5.1.6.2.1. Bluetooth technology: (http://www.bluetooth.com/)

Bluetooth is a personal area wireless networking standard, which is a low-power, short-range, wireless technology designed for local area voice and data communications. The system operates in the 2.4 GHz ISM (Industrial, Scientific, and Medical) band, providing license-free operation in the United States, Europe and Japan. With over 1000 members of the Bluetooth Special Interest Group (SIG), development of version 1.0 of the standard was announced in July 1999.

The nominal link range is 10 cm to 10 m, but links can be extended to more than 100 m by increasing the transmit power. In general, the Bluetooth specification has two power levels defined; a lower power level that covers the shorter personal area within a room, and a higher power level that can cover a medium range, such as within a home. Software controls and identity coding built into each microchip ensure that only those units preset by their owners can communicate.

The Bluetooth technology uses Frequency Hop (FH) spread spectrum, which divides the frequency band into a number of hop channels, and during a connection, radio transceivers hop from one channel to another in a pseudo-random fashion. The Bluetooth wireless technology supports both point-to-point and point-to-multipoint connections. It supports up to 8 devices in a “piconet” — with the current specification, up to seven “slave” devices can be set to communicate with a “master” radio in one device. The connection is omni-directional, which does not require line-of-sight clearance but the transmission can go through walls and briefcases.

As the Bluetooth technology supports both isochronous and asynchronous services, it can be easily integrated into TCP/IP networking. Several of these ‘piconets’ can be established and linked together in ad hoc ‘scatter-nets’ to allow communication among continually flexible configurations. All devices in the same piconet have priority synchronization, but other devices can be set to enter at any time. The topology can best be described as a flexible, multiple piconet structure. Further more, the Bluetooth protocol stack contains a service discovery protocol (DSP) that enables the retrieval of information that can be used to configure the stack to support several end-user applications.

Till now, I have not seen any reports about incorporating Bluetooth technology into context-aware research yet, however, from the technology availability versus application demands point of view, I would expect it come up soon.

5.1.6.2.2. PinPoint Corp’s 3D-iD indoor positioning system:

(http://www.pinpointco.com/technology/technology.htm)

The PinPoint Corp’s 3D-iD® indoor positioning tracking system is an another RFID implementation example — a little more beyond just RFID: there are several antennae attached to a Cell Controller and the time delays between the initial signal and the retransmitted signal for each antenna are calculated, with this information the more accurate location information is calculated from triangular geometry. The system can read tag signals from distances of up to 200 feet to within ±10 feet resolution.

The 3D-iD® system is claimed as being designed around PinPoint’s proprietary L3RF (Long Range, Long Life, Low Cost Radio Frequency) technology.

It consists of five major components:

(1). **Tags**: tags are attached to stationary equipment, mobile assets or inventory, or even used as personnel badges. At a constant duty cycle, the Tag transponds (trans-mitter + res-ponder) and reflects (or "transflects") a low-power, 2.4GHz radio signal that is transmitted by the system's Antennas. The emitted signal, which has a unique Tag Serial Number encoded in it, is a 5.8GHz radio signal. The 3D-iD® Tag architecture is built upon an open standard that enables third parties and OEMs to create custom Tag designs and Tag packages for specific applications.

(2). **Cell Controller Network (with Antennas)**: The 3D-iD® Cell Controllers continuously track hundreds of L3RF Tags in real-time (as quickly as every 0.5 seconds a new location can be calculated). The network of Cell Controllers can attach to existing LAN via TCP/IP. Each Cell Controller can have up to sixteen (16) Antennas attached and can cover an area as much as 5 acres (200,000 sq. ft.).

(3). **Cell Controller Firmware**: The Cell Controller Firmware is embedded within each Cell Controller. Using the Linux Operating System, the system collects Tag location data in real-time and publishes the data to a PinPoint Server.

(4). **ViewPoint Administrator Server Software**: The server software utilizes the Publish / Subscribe mechanism and proprietary NT services to retrieve data from the Cell Controller, manipulate it, and prepare it for use by any end-user applications. The software allows an administrator to configure all communications, maintain the Tag database, configure alerts, and determine in which database to store the location information.

(5). **End-user Software (ViewPoint Administrator):** The ViewPoint Administration software provides tools to view and monitor tagged items, such as assets, inventory, and personnel in real-time.

Although the system is claimed to have been successfully used in several applications, it is said that this technology is so mature — it was easily interfered with other telecommunication signals and cannot reliably work when being tested in the Motorola Labs.

5.1.7. Radar detector, laser rangefinder, and smart-floor

5.1.7.1. Miniature radar detector

A radar system sends out high-frequency electromagnetic wave pulses and detects the reflection of the waveform, measuring the intensity change and the elapsed flight time to detect objects location and motion. A Doppler radar uses the Doppler effects in physics to provide higher resolution measurement, that when a moving object reflects the radio waveform the reflected radio waveform is modulated by the motion of the object. For field applications, the Doppler radar motion sensors are widely used to measure the speed of moving vehicle by police. By carefully controlling the signal power to make sure that it is safe to human health, Doppler radar motion detection sensors can also be used to sense user's location and motion as used in MIT Immersive Environments Projects. [61]

The detailed specifications about how the MIT Media Lab used microwave Doppler radar motion sensors are not included in the publications. An intuitive guess would come to an estimation of not so high resolution: as with the significant radio side-lobes and not so sophisticated signal processing factors being considered.

5.1.7.2. Laser rangefinder distance measurement

A laser rangefinder emits a laser light pulse and detects the light spot (laser reflection). The speed of light is already known, by measuring the time that it takes a laser pulse to travel to the target and back with a precision, crystal-controlled time base, the traveled distance then can be calculated. To increase accuracy, many rangefinders use very short laser pulses and do multiple measurements, utilizing a least squares method in calculating the range.

Laser rangefinder using low intensive safe laser is a very good distance and thus location measurement method. It can achieve a very high accuracy with a relative low cost. However, a drawback of using a laser rangefinder in context-aware computing applications is that the measurement is relatively slow — it cannot be used to measure large rapid moving objects.

5.1.7.3. Smart-floor weight measurement

People's body-weight population is reasonably distributed, and different people often have their own walking styles in terms of using footfall forces. This information can be gathered by measuring the applied force on the floor and, it can be used as supplement information to identify users and track their activities in the facilitated room. The implementation can vary: the Magic Carpet in MIT Media Lab's Immersive Environment project uses a grid of piezoelectric wires hidden under a foot carpet [61], whereas Georgia Institute of Technology's GUV center uses load cell under floor tiles [62] to measure the applied forces. Other possible implementation can be grids of deformation sensitive optical fibers and networks of vibration sensors attached to the underside of the flooring.

5.1.8. Ultrasonic and other multisensor location system

Stereovision is one of the most commonly used object recognition and measurement method, whereas the object location estimation accuracy depends on cameras' resolution, deployment as well as the objects' own characteristics (which will affect the point matching calculation as discussed in section 3.2.2). There are many similar or disparate location sensing modalities being used besides those described thus far, below are two examples using ultrasonic measurement as part of their key sensing method.

5.1.8.1. AT&T Labs, Cambridge "Active-bat" 3D ultrasonic location system

The Active-bat 3-D ultrasonic location system is developed in AT&T Labs, Cambridge. The system is for indoor environments based on the principle of trilateration — finding position by measurement of distances. The system consists of two parts: transmitters that are attached to the being tracked objects, and receivers matrix that are amounted in the room ceilings. The system is described as following. (http://www.uk.research.att.com/bat/)

A small, wireless transmitter is attached to every object point that is to be located. The transmitter consists of a microprocessor, a 418MHz radio transceiver, a Xilinx FPGA and a hemispherical array of five ultrasonic transducers. Each prototype mobile device has a unique 16-bit address.

A matrix of receiver elements is mounted on the ceiling of the room to be instrumented. Receivers are placed in an array, 1.2m apart (the prototype system has 16 receivers in a four-by-four square grid). Each receiver has an ultrasonic detector, whose output is passed through an amplifier, rectifier and smoothing filter before being digitized at 20KHz by an ADC. The ADC is controlled by a Xilinx FPGA, which can monitor the digitized signal levels. Receivers also have a serial network interface, through which they are individually addressable, and are connected in a daisy-chain to a controlling PC.

Every 200ms, a radio message consisting of a preamble and 16-bit address is transmitted in the 418MHz band by a controller also connected to the PC. The PC dictates which address is sent in each message. The radio signals are picked up by the transceiver on each mobile device and decoded by the on-board FPGA. The single addressed device then drives its transducers for 50ms at 40KHz, and an ultrasonic pulse is broadcast in a roughly hemispherical pattern around the top of the unit.

The FPGAs on each receiver then monitor the digitized signals from the ultrasonic detector for 20ms, calculating the moment at which the received signals peak for the first time. The short width of the ultrasonic pulse ensures that receivers detect a sharp signal peak. The controlling PC then polls the receivers on the network, retrieving from them the time interval between the reset signal and detection of the first signal peak (if any signal was detected).

The distances from the transmitter to each receiver then can be calculated by the measured time delay. The Active Bat location tracking system is a good choice for context-aware computing applications; it is reported that when averaged with 10 samples 95% measured points can fall within 8 cm accuracy range.

5.1.8.2. Motion tracking systems in VR & VA applications

Position and Orientation Tracking is also a very import part in Virtual and Augmented Reality (VR, AR). With many different graphics applications that depend on motion tracking technology, a wide range of interesting motion-tracking solution have been invented. Compared with general context-aware location tracking requirements, the motion-tracking systems in VR and AR usually are required to be more accurate and to work at a smaller range with which tethering cables are acceptable. However, some VR and AR motion tracking systems have improved so much that they are close to fulfill some context-aware applications' requirements. Among those systems, the InterSense's (http://isense.com) IS-900 wide area object tracking system is the most impressive one.

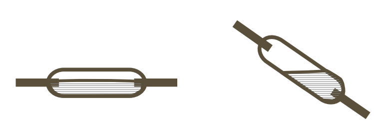
Each InterSense IS-900 system is comprised of a SensorFusion base unit, a series of pre-assembled ultrasonic SoniStrips that can cover up to 900 m^2 (LAT configuration only), and multiple omni-directional tracked stations. At each omni-directional tracked station, the inertial sensors simultaneously measure 9 physical properties: angular rates, linear accelerations, and magnetic field components along all 3 axes. Micro vibrating elements are used to measure angular rates and linear accelerations. The angular rate signals are integrated in a direct manner to obtain orientation. The linear acceleration signals are double integrated to track changes in the position. This method, however, leads to an unacceptable rate of positional drift and must be corrected frequently by some external source. Hence, the system uses an acoustic ultrasonic time-of-flight ranging system to prevent position and orientation drift. The accelerations are also fed into an error estimator to help cancel pitch and roll drift. The yaw drift is corrected by acoustic range measurements simultaneously with the position drift.

Using its proprietary SensorFusion software to combine the latest advances in inertial and ultrasonic tracking, the IS-900 delivers tracking resolution in the millimeter range for position, and below 0.1° for orientation. The system specifications are among the best for people location tracking system, the main hindrance for the IS-900 is perhaps this expensive price.

5.2. Orientation Sensing Implementation

For vertical orientation detection, in the simple case where the being detected object is not in obvious vibration and jittering movement status, two simple tilt switches are enough to roughly measure whether the object is in the up-straight pose. As an example, the Smart Badge system uses four tilt switches as its vertical orientation detection sensors.

Mercury switch is a simple type of tilt switch. It consists of a small (usually glass) bulb with a contact at each end and a ball of mercury in the middle. When the switch is level, the mercury touches both contacts and allows current to flow through a circuit. When the switch is tilted, no current can flow. [26]



Mercury switch, Closed — Mercury switch, Open

Mercury switches are easily available and cheap, but they only make binary measurement. To make finer vertical orientation measurement, array of mercury switches need to be used in a parallel direction with each mercury switch open/close at specific different tilt angle.

Horizontal orientation was traditionally measured with a floating ball compass, which was made of magnetic material and could easily align itself to the earth magnetic field. More modern version compasses use the hall-effects to sense the earth magnetic field, and their output can be in analog or digital fashion. As compasses can be easily affected by nearby ferric material, some sensors such as those used in vehicle guidance applications have built into it a electronic compensation circuit and some give out a warning signal when they detect existence of such influence.

Digital compass can be easily made as small as a dime, and they are commercially available with quite cheap prices (components are about a few dollars, and a system usually costs about $40~50). Their accuracy typically is about 5°.

5.3. Contexts of location, proximity, and orientation information

All the previously discussed location, proximity and orientation sensing technologies, by themselves, only provide indexes or entrance to the more content-rich property contexts, — although they include lots of information and constitute tremendous challenges for performance enhancement or functional improvements. When we conclude that location, proximity, and orientation information consists a very important aspect of contexts, what we mean is not just the "latitude, longitude, and height above sea level" parameters, but also its associated properties such as name and supposed functions etc.

As the localization information can be at different granularities and be described at different level of abstraction or ambiguity, when referring "close to somewhere" it may imply that the "close to" granularity also associate with the "somewhere" granularity (at this specific point, in a cubic, in a room, this building, street block, "downtown", city, etc). Therefore, the seemly simple context description question actually is another challenge — if we want to take a systematic approach to describe the already sensed low-level location information so that our context-aware system can easily adapt to different applications. Here, some kind of sensor fusion or information abstraction is needed, as well as some kind of standards are needed to describe the context information.

6. User Mental Activity Sensing

All human thoughts and emotion are the result of the human nervous system activities in the form of tiny electrical flows in brain. Psychophysiological monitoring has been used for decades to try to reveal the relationship between physical and emotional states. Electroencephalography deals with sensing the electrical activity of the brain — by means of electrodes attached to the surface of the skull — and studing the patterns on the electroencephalogram (EEG). With the joining force of MRI (Magnetic Resonance Imaging) technology advancement, people can now associate human mental activity with some specific area of brain cortex activities and recognize some brain wave patterns[10]. The biggest limitation of MRI and EEG technologies is that they are not easy to implement in field sensing, as context-aware computing applications usually require.

Although today's technology is still very far away from that it can "read the user's mind", there do exist many interesting physiological sensing measurements that are tightly related to the

[10] For example, the alpha wave is an indicative of a relaxed state, the theta wave (usually found in children) is thought by some to reflect creative activity in adults.

user's health conditions as well as to his/her emotional status to give valuable cues of users' mental activity contexts.

6.1. Mood versus physiological responses

A human being's brain is the source of his/her thoughts, moods, and emotions. The brain is the hub of one's nervous system; all of a person's movements, sensations, thoughts, and emotions are products of his/her nervous system. The human nervous system has three main parts, the central nervous system (i.e., brain and spinal cord), the peripheral nervous system (which carries all the messages sent between the central nervous system and the rest of the body), and the autonomic nervous system.

The autonomic nervous system is a special part of the peripheral nervous system that regulates the automatic bodily processes such as breathing and digestion, without conscious control by the brain, to maintain a stable internal environment. The autonomic nervous system has two parts, the sympathetic system and the parasympathetic system. The sympathetic system responds to the body's needs during increased activity and in emergencies — speeding up the heartbeat, sending additional blood to the muscles, and enlarging the pupils of the eyes to use all available light. The parasympathetic system, in general, opposes the actions of the sympathetic system — slowing down the heartbeat, diverting blood from the muscles to the stomach and intestines, and contracting the pupils of the eyes. The balance of activity between the two systems is controlled by the central nervous system.

The theories involved in psychophysiology date back to the early part of this century. In the past 60 years, psychophysiological research has examined the relationship between emotional states and physiological states. Within the past 28 years, attempts have been made to affect these emotional states, as well as states of consciousness, through a process called biofeedback. [66] Biofeedback is the process in which a subject receives information about his biological state[11]. The application of specific biofeedback techniques and procedures began only 25 years ago. The use of electromyograph (EMG) and peripheral skin temperature are the principle modalities used today in clinical practice.

Mittelmann and Wolff (1943) defined emotional stress as anxiety, anger, embarrassment, humiliation, joy with anxiety, depression with hostility, guilt, fear of abandonment, and conflict over the use of the hands for aggressive and sexual purposes. It was clear that EMG measurements provided substantial physiological information regarding the activity of the central nervous system. However, there was a need to monitor the autonomic nervous system. Traditionally, the galvanic skin response (GSR) had been used to monitor the sympathetic part of the autonomic nervous system. It has been found that the measurement of peripheral skin temperature could provide a more reliable measure of sympathetic-autonomic nervous system activity.

6.2. Mood sensing with unobtrusive measurements

For the physiological sensors to be used in context-aware computing applications, besides their safety requirements as compulsorily enforced by FDA (US Food and Drugs Administration)

---

[11] Usually a subject is not aware of his physiological functions, especially those controlled by the autonomic nervous system, such as heart rate and peripheral vasoconstriction. Biofeedback creates an external loop by which a subject can monitor one or more of his physiological states.

regulations, they ought to be lightweight, portable, comfortable to wear, and robust to changes in user position.

6.2.1. Heart Rate Variability (HRV)

Blood carries oxygen and food to all the human body, with each heartbeat, the life-giving blood is pumped out throughout the body to power the body. The heartbeat rate is controlled by the autonomic nervous system, being increased or decreased depending on the body's needs. The heart pumps slowly and relatively small amounts is provided to the body while a person sleeps, but, the heart rate can be quickly speeded up and enormous oxygen output is provided when he/she excises, becomes frightened, or needs to fight or run. The speedup of heart rate is the most obvious indicator to a person's extreme excitement or stress mood responses.

Heart rate can be measured easily. There are various kinds of small handy heart rate products commercially available, many of which are in fashions of wristband watch, chest belt, and finger pad etc. with conventional electrode-contact measurement and telemetric data reporting capabilities, their prices are in the range of $30 ~ $200. As an example, Polar Heart Rate Monitor series (http://www.polar.fi/productfinder/) has several models for customers to choose from, all models incorporate electrodes encased in a lightweight chest transmitter that continuously picks up the heart's electrical impulses, and then wirelessly transmits this information to a corresponding wrist receiver. (Their reference prices: M22 Monitor $150; M71ti Monitor $210; Transmitter $40; Transmitter-coded $60).

6.2.2. Blood volume Pulse (BVP)

Another way of hart rate measurement is to detect the blood flow rate in the finger or earlobe of a person with a photoelectric sensor. Called as photoplethysmography, the process applies a light source to the skin of human body extremities and measures the light reflection: at each contraction of the heart beat, blood is forced through the peripheral vessels, producing engorgement of the vessels under the light thus modifying the amount of light to the photo-sensor.

Heartbeat rate and blood volume flow are just two intimately related aspects of the body responses — vasomotor activity (activity which controls the size of the blood vessels) is also controlled by the sympathetic nervous system, the BVP measurements can display changes in sympathetic arousal. An increase in the BVP amplitude indicates decreased sympathetic arousal and greater blood flow to the fingertips. [64]

6.2.3. Respiration monitoring

Emotional excitement and physical activity are reported to lead to faster and deeper respiration. Peaceful rest and relaxation lead to slower and shallower respiration. A state of stress would therefore be indicated by frequent respiration; however, sudden stressors such as a startle tend to cause momentary cession of respiration.

Human lungs are enclosed in the thorax, or chest. The rib cage gives shape to the thorax and the muscular diaphragm is a dome-shaped partition at the base of the rib cage. During inspiration or breathing in, the ribs are expanded by the action of muscles that surround them and the diaphragm contracts to increase the volume of the thorax. Relaxation of the diaphragm allows the opposing set of muscles to return the chest to its natural position, forcing the air from the lungs during expiration (exhalation, or breathing out).

Respiration can be measured by detecting the amount of expansion of the chest cavity as a person breathes in and out. The sensor would usually consist mainly of a large velcro belt which extends around the chest cavity and a small elastic which stretches as the subject's chest cavity expands. The amount of stretch in the elastic is measured and the result is usually represented as a voltage change to be recorded. Along with other biofeedback equipment such as EEG and ECG measurement instruments, respiration-monitoring sensors are commercially available to different situations.

6.2.4. Galvanic skin response (GSR) and electromyogram (EMG)

As mentioned before, a human being's emotional fluctuation is always associated with autonomic nervous system activity, which will cause a change in the skin's conductivity. The overall degree of arousal of the hemispheres, and indeed the whole brain, can be shown by the readings of the GSR psychometer, which does not differentiate between the hemispheres, or between cortical and primitive brain responses. [65] Higher arousal (such as occurs with increased involvement) will almost instantaneously (0.2 - 0.5 sec) cause a fall in skin resistance; reduced arousal (such as occurs with withdrawal) will cause a rise in skin resistance.

Galvanic skin response (GSR) is actually a measurement of electrical resistances to the passage of a very small electric current between two electrodes in contact with the skin, across any two points on the body. GSR meters have been used as polygraph or lie-detector for decades. When a person is reminded of certain past events, or when a change of mood is induced in him, the needle in the GSR meter would jump erratically — the degree of jump is in proportion to the strength of unconscious reaction. In skilled hands, the meter could be used to locate a particular mental content, the nature of that content, the location of that content in space and time, and the amount of force contained within it.

However, one cannot simply associate the GSR reading directly with some mental depressed or joyful status. It has long been known in biofeedback research, that meditation and relaxation procedures cause a rise in skin resistance. It has therefore been assumed that high and low skin resistance correlate directly with relaxation and stress respectively, and that a high resistance indicates a pleasant relaxed state of mind, whereas low resistance indicates tension. But the reverse is also true in a psychotherapy session. When repressed material is coming to the surface (e.g. material associated with guilt or pain), the skin resistance rises and the client experiences feelings of tension; thus in a therapy session, high skin resistance indicates tension, and not relaxation as in meditation. Then, when the repressed material reaches the surface and the negative emotion discharges, there is usually a sudden large drop in skin resistance and the client experiences relief. This demonstrates a correlation between low skin resistance and relaxation of tension, which is in contradiction to the pattern of research findings in meditation.

The electromyographic sensors measure the electromyogram (EMG) — the electrical activity produced by the muscles when they are being contracted in a human being's voluntary nervous system. EMG readings can also be used to measure the relax extent of body.

6.2.5. Body temperature

As the problem with GSR measurement is that, it is subject to many artifacts, for example, a deep breath or tightening a muscle group would decrease the electrical resistance of the skin. Body temperature measurement can be a good alternative to monitor the subject's autonomic nervous system activities. The changes take place more slowly than those of the GSR, and the changes are not subject to the annoying artifacts of the GSR.

Past research has reveals the tight relationship between a human subject's emotional disturbances with his body temperature variation. First, there is a clear link between emotional tension and finger temperature. Second, there is a correlation between the degree of the fall in finger temperature and the intensity of the subject's anxiety. Third, the extent to which a subject is aware or unaware of an emotional conflict or sense of tension will not prevent the lowering of the finger temperature. This result suggests that skin temperature could be considered a more "objective" index of a subject's level of stress or anxiety although his subjective sense of the emotional feelings may not reflect this anxiety. Fourth, a subject can — through belief in his sense of well-being, security, or superiority — override his tension. If the subject can relax and objectify his emotional reactions, he can maintain his finger temperature at a high level. This latter finding suggests that subjects can have some voluntary control over their peripheral skin temperature by way of their mental state.

A person's finger temperature or body temperature can be easily measured and there are many commercial thermometer products for various application situations.

6.3. Affective contexts

A human being's mood change (emotion, arousal, and attention etc.) is linked with his/her sympathetic nervous system activation, which in turn changes the level of sweat in the eccrine sweat gland, among other responses such as heart rate and respiration etc. change. As mentioned before, though there exist strong relationships between a human being's physiological measurements and his/her emotional activities, but to detect a person's mood is a very difficult task because the human mood activity is extremely colorful and intrinsic and the mapping is extremely complicated.

There is no definitive model of emotions to describe affective contexts. Psychologists have been debating for years how to define them. The pattern recognition problem consists of sorting observed data into a set of states (classes or categories), which correspond to several distinct (but possibly overlapping, or "fuzzy") emotional states. Which tools are most suitable to accomplish this depends on the nature of the signals observed.

7. Vicinity Environment Properties

Outside environment is another very important part of a user's context as people's comfortable feelings are largely determined by their environment parameters. Beside the sound and visionary environment aspects we discussed in the previous sections, environment aspects include: temperature, humidity, barometry, oxygen concentration, smell.

7.1. Temperature

In physics, the term *temperature* is defined as "the property that determines whether an object will be in thermal equilibrium with other objects, two objects in thermal equilibrium with each other are at the same temperature". Temperature is expressed in one of three measuring scales: Celsius, Fahrenheit, and Kelvin; and it is measured by thermometer sensors — traditionally a scale-marked tubes, within it a liquid height indicating the temperature readings.

There are many temperature sensor products in various forms. The following is just a list of forms that are thought to be among the temperature sensor candidates for context-aware computing applications:

- **Thermocouple**s are based on the Seebeck effect that occurs in conductors that experience a temperature gradient along their length. They are pairs of dissimilar metal alloy wires joined at least at one end, which generate a net thermoelectric voltage between the two ends according to the size of the temperature difference between the ends, the relative Seebeck coefficient of the wire pair and the uniformity of the wire's relative Seebeck coefficient. Thermocouples are among the easiest temperature sensors to use and obtain and are widely used in science and industry. (http://www.temperatures.com/tcs.html)
- **Thermistor**s are tiny bits of inexpensive semiconductor materials with highly temperature sensitive electrical resistance. Thermistors abound inside many devices as well as in specialty temperature sensing probes. Some of those new-fangled medical thermometers that get stuck in one's mouth by a nurse with an electronic display in her other hand are based on thermistors. They typically work over a relatively small temperature range and can be very accurate and precise within that range. (http://www.temperatures.com/thermistors.html)
- **RTD's (Resistance Temperature Detectors)** are based on the positive temperature coefficient of electrical resistance. They are wire wound and thin film devices that work on the physical principle of the temperature coefficient of electrical resistance of metals. They are nearly linear over a wide range of temperatures and can be made small enough to have response times of a fraction of a second. They require an electrical current to produce a voltage drop across the sensor that can be then measured by a calibrated read-out device. RTDs are among the most precise temperature sensors commercially used. (http://www.temperatures.com/rtds.html)
- **Radiation Thermometers (RT's)** are non-contact temperature sensors based on Planck's Law of the thermal emission of electromagnetic radiation. The RT sensors measure temperature through sensing the amount of thermal electromagnetic radiation, and it can take many kinds of optical infrastructures in the forms of spot measurement, line-measurement, or area-measurement as infrared cameras. They can further be classified according to their thermo-radiation-spectrum sensitivity wavebands, their portable/mounting manner.
  The RT sensors may be referred to by other names as: "Optical Pyrometers", "Radiation Pyrometers", "Total Radiation Pyrometers", "Automatic Infrared Thermometers", "Continuous radiation Thermometers", "Line-Scanners", "Thermal Imaging Radiometers", "Infraducers", "Infracouples", "Fiber-optic Thermometers", "Gold Cup Pyrometers", "Surface Pyrometers", "Ratio Pyrometers", "Two-Color Pyrometers", "Infra-Snakes", etc. (http://www.temperatures.com/rts.html)
- **Fiber Optic Thermometers**. Many so-called "fiber optic thermometers" devices utilize fiber optics to aid in measuring temperature, though most of them are actually just slight variations of radiation thermometers. The fiber optic thermometers use optic fibers to conduct thermo-radiation energy from the "free-end" to a measuring system that collects the desired radiation and processes it into a temperature value.

7.2. Humidity

Humidity is regarding the moisture content in the atmosphere. The atmosphere always contains some moisture in the form of water vapor whose maximum amount — the amount of vapor that will saturate the air — increases with a rise in temperature. Moisture terms and units all fall under the area of psychrometrics, the study of water vapor concentration in air as a function of temperature and pressure.

Humidity is specified in several different ways: (1) The weight of water vapor contained in a volume of air is known as the absolute humidity and is expressed in grams of water vapor per cubic meter. (2) The ratio of the weight of water vapor to the total weight of air (including the water vapor) is known as the specific humidity — expressed in grams of water in kilograms of air. (3) The ratio of the weight of water vapor to the weight of dry air (with the water vapor removed) is known as the mixing ratio — expressed in grams of water vapor per kilogram of air. (4) Relative humidity is the ratio between the actual vapor content of the air and the vapor content of air at the same temperature saturated with water vapor.

When the atmosphere is saturated with water, the level of discomfort is high because the evaporation of perspiration, with its attendant cooling effect, is impossible. The temperature-humidity index, also called the discomfort index, expresses in numerical values the relationship between comfort or discomfort and temperature and humidity. It provides an apparent temperature, or how hot the air feels. For example, an air temperature of 38ºC (100°F) and a relative humidity of 60 percent produce an apparent very hot temperature of 54°C (129°F). When the index is 20°C (68°F), most people are comfortable; an index of 25°C (77°F) means that half are uncomfortable.

Humidity sensors are often referred to as hygrometers. There are mainly three categories of portable humidity sensors. Extension hygrometers use some material's (such as hair, plant fiber, or nylon etc.) special property that they will noticeably contract in length when absorb more moisture. Absorption hygrometers use the phenomenon that some material will easily gain or lose weight due to moisture absorption. Whereas electronic hygrometers are based upon the fact that some material's impedance or capacitance value is affected by the environment humidity conditions. [71]

Of all the three hygrometer categories, capacitive relative humidity sensors dominate both atmospheric and process measurements. Capacitive relative humidity sensors are the only types of full-range RH measuring devices capable of operating accurately down to 0% RH. Because of their low temperature effect, they are often used over wide temperature ranges without active temperature compensation.

8. People Interaction and Social Relationship

People identification and their relationship/interaction are the most important yet most difficult contexts for computers to understand and exploit. To explore this labyrinth, the first step is to ask "who you (the user) are and, who are the people out there"; then, "what you are doing and, what those people are doing"; and finally perhaps to address the more complicated "why and what is interesting" questions. This is more an intelligently comprehensive inference rather than a sensing or perception task, at current computational artificial intelligent research stages, it is too preliminary to provide a panorama situation understanding, yet, there are still a few things very important that the computers can do with the help of some kind of biometric sensors.

8.1. Single user identification, biometrics

An accurate automatic personal identification is critical to a very wide range of applications in various domains of modern society such as welfare benefit disbursement, financial security, sensitive information access control etc.. Traditional major automatic personal identification approaches are either "token-based" – showing some object (a badge, a driver's license, a credit card, etc.) to prove an identity with the access privilege, or “knowledge-based” – showing some knowledge (password, a PIN number, etc.) to prove an identity with the access privilege.

To improve identification reliability and defeat imposters, modern biometrics technology identifies an individual based on his/her physiological or behavioral characteristics. Biometrics deals with two kinds of identification tasks: (1) verification: authenticating a claimed identity – am I the person who I claim as? and, (2) recognition: determining the identity of a given person from a database of people known to the system – who am I?

A biometric system is essentially a pattern recognition system, which makes a personal identification by establishing the authenticity of a specific physiological or behavioral characteristic possessed by the user. [38] Logically, it can be divided into two modules: “enrollment module” and “identification module”. The enrollment module is responsible for enrolling individuals in the biometric system, during which the biometric characteristic of an individual is first scanned by the biometric sensors and, in most cases, some further features are extracted and kept as a template in a database. The identification module is responsible for identifying individuals at the point of access, during which the individual is scanned by biometric sensors and the newly extracted characteristic or features are compared with the preset templates.

There are many types of biometric technologies including: signature recognition, voiceprint recognition (speaker identification), fingerprint image sensing (optical, thermal, ac-/de-capacitive), hand-geometry measurement and palm scanning, face recognition, retina identification, and iris recognition etc.. The most important factor that affects biometrics’ wide application is the cost of their infrastructures, which overall drops dramatically during the last a few years. It seems that there will be a big boom of biometrics — if the public’s privacy concerns related to it can be properly addressed.

8.1.1. Signature recognition

Signature is currently the most commonly used authorization and identification method, but unfortunately the least reliable one. As a static signature (a written signature) can easily be simulated by forgers, newer signature biometrics incorporates “dynamic information” as well to improve reliability such as the stroke speed, pen pressure against the paper, and the angle the pen is used to generate the signature.

The now quite commonly used signature recognition biometric sensors (mostly in commercial retail industry, as in department stores) often consist of a pad that contains a resistive grid or a 2-D array of ultrasonic sensors. LCI Computer Group’s (http://www.lcigroup.com/) Smartpen goes further incorporating a minicomputer and a small radio transmitter. It looks like, and functions as, a regular ball point pen, but its special sensors can detect how a person writes his/her signature with the force applied to the paper, the acceleration of the pen, and the angle at which the pen is held.

8.1.2. Fingerprint recognition

A fingerprint is the pattern of ridges and furrows the surface of a fingertip. Its formation is determined during the fetal period: each fingerprint is unique and its pattern doesn’t change over one’s lifetime. Fingerprints are one of the most understood and studied biometrics and may be expected to lead the biometric application in future with their implementation now becoming increasingly more affordable.

In middle 1998, there were already four IC manufacturers (SGS-Thomason, Veridicom, Harris Semiconductor Corp., and Thomson-CSF) embodied unusual fingerprint-sensing technologies in new chips, and dozens of companies announced fingerprint-sensing units. [37] Their sensing implementation includes optical, thermal image, as well as dc-capacitive and ac-capacitive technologies.

8.1.3. Voiceprints, speaker identification

As described in the sound context section, speaker verification/identification can be used to make sure whether the user is who claims. Though people's sound may influenced by various factors such mood, stress, fatigue, and how much time has passed since being awake, voiceprint experts insist that enough characteristics of one's voiceprint will remain constant under all circumstances that voiceprint can reliably verify the speaker's identity.

8.1.4. Hand vein, hand geometry, footstep profile feature recognition

Hand vein patterns can be easily captured with an infrared camera. Hand vein patterns are unique to each individual and it is very difficult to change the formation of the hand vein pattern of in individual, thus a hand vein based biometric system has the potential to achieve a reasonable accuracy. The drawback for hand vein pattern biometrics is that if can be obliterated due to medical conditions such as obesity and aging.

Hand geometry measurement can include its shape, lengths and widths of the fingers etc.. Its biometric measurement can be made without much technical difficulty, but just like hand vein biometric system, hand geometry information may not remain invariant over the lifespan of a person — especially during childhood.

A person's walking feature, the footstep profile, can also reveal his/her identity – though the recognition is at rough resolution. This feature is usually suggested as a supplement biometric measurement.

8.1.5. Face recognition, facial thermogram

Human beings primarily use face recognition plus figure features to recognize people. The facial recognition has been a active research topic in computer science for many years and there are many interesting reports about their performance improvements. Roughly, the research can be classified as either geometrical modeling (e.g. decomposition of face into a number of canonical faces) or image modeling (treating face simply as some color or black/white image change pattern). It seems that image modeling with neural network algorithm now does a better job in face recognition, though the people identification rate on the whole is still quite low and it is too much subject to non-intrinsic condition variations such as his shaving states.

A facial thermogram is a facial image taken by an infrared camera. It is believed that a face thermogram based recognition is superior to a face recognition using regular CCD cameras because it is not vulnerable to disguises, and even plastic surgery which does not reroute the flow of blood through the veins cannot change the formation of the face thermogram. However, thermogram depends heavily on a number of factors such as emotion and the body temperature of the subjects etc, and it has not been demonstrated as sufficiently discriminative as regular optical images for people identification purpose.

8.1.6. Retinal pattern, iris pattern recognition

The pattern formed by veins beneath the retinal surface in an eye is stable and unique. The image can be captured by projecting a low-intensity beam of visual or infrared light into the eye and using a camera (arranged similar to a retinascope) to take the picture of the illuminated retina. In order that a fixed portion of the retinal vasculature is used for identification, the subject is required to close gaze into an eye-piece and focus on a predetermined spot in visual field. Retina scan is currently perceived to be the most secure biometric technique, as a large number of it are installed in several highly secure environments (e.g., prisons). [38]

Iris is the annular region of an eye bounded by pupil and sclera (white of the eye). Everyone's iris is unique and a person's left and right irises are not very even similar. The visual texture of iris stabilizes very early (first 2 years) in life. Like the retina recognition (but more readily imaged than retina), iris based identification perhaps offers the highest accuracy. A great effort is made to drop the cost and enhance convenience in use, as an example, IrisScan Inc.'s authentication system can scan an individual's iris even if the person is wearing glasses.

8.2. Interacting people detection

The people sensing implementation should blend naturally with the habitual action of the people being identified to provide reliable information. There is not a single sensor available that can un-intrusively and reliably detect people's proximity and identity.

Active Badge was one of the earliest successful systems for sensing the presence, identity and even location in a room within a building. The downside of it is that the users must wear their badges all the time (keeping them in pockets doesn't work as they use infrared communication channel to announce the presence), only the pre-registered people with badge ID can be detected, and on the other side some users may not willingly accept it as it is hard to dissimulate.

Other sensor systems have their own limitations and disadvantages. For example, passive RFID (Radio Frequency Identification) technique requires every person bear a RF transmitter tag whose unique ID is pre-registered just like active badge system, and cannot reliably work when more than two tags are simultaneously present in the field of the antenna. Smart floor requires only one person at a time is walking on the measuring floor tile. [23]

Human speech recognition and face recognition are perhaps the two most promising technologies for people activity and intention detection.

Speech recognition is sometimes regarded as a somewhat intrusive method though, because it requires the person deliberately speak something. As discussed before, at current stage the recognition capability is quite limited and its performance is not very reliable yet.

People image recognition for activity and intention recognition, mostly facial recognition, is still under intensively research. The image understanding technology is not very practical yet as it now suffers from camera perspective, lighting condition and occlusion variation difficulties. At current research stage, a facial recognition system usually requires a user to stand still in front of the camera at a preset position for a picture to be taken.

8.3. Social Relationship

Social relationship is very high-level context information far from being just sensed by some simple sensors, though occasionally some reports are shown in popular science, that some psychologists acclaim that something as being subjective as a couple' intimacy relationship can be measured or predicted by some electronic instruments.

9. Context Extraction and Beyond

9.1. Intelligent sensing

Sensing is the core and the first step of “intelligence”. Context can range from simple preliminary information (such as temperature reading) to very abstract concept (such as “she feels a little hot with the noise music”). Therefore, the extraction of various contexts from sensors processing is no longer a one-to-one simple mapping, but an intrinsic computation with artificial intelligence heavily involved.

9.1.1. From sensors to contexts – information abstraction

The increased availability of various sensing technologies will make it more viable to collect a variety of information from a vast network of sensing resources, however, before we can make a good use of the information, a great effort is needed to define a consistent methodology to properly extract, organize and maintain the context information. Due to the intrinsic complexity of *general* contexts, until now most context-ware applications have been built in an ad hoc manner, i.e. they are heavily influenced by the underlying technology used to acquire the context, we still lack a consistent way to describe as well as use contexts.

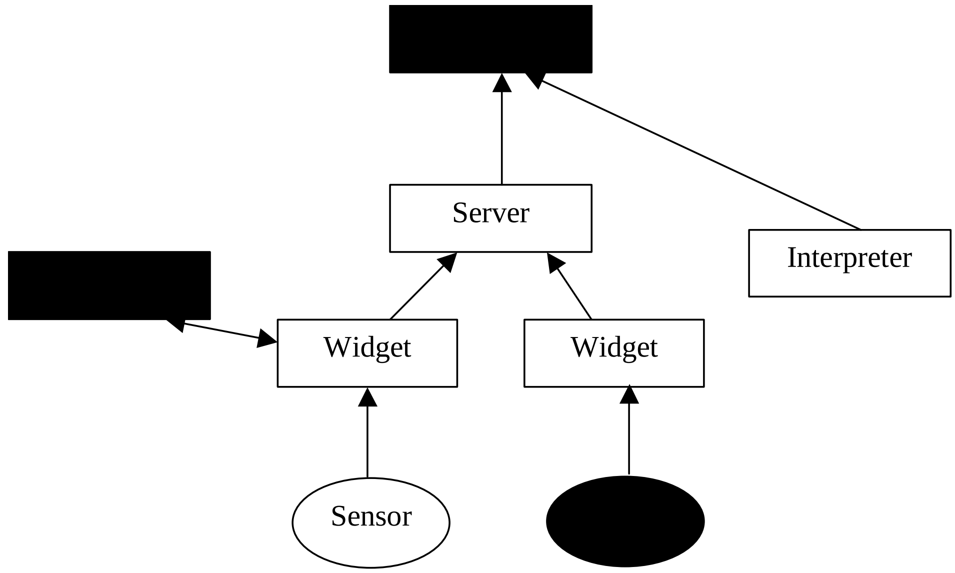
9.1.1.1. Contexts insulated from sensors

The first step towards a systematic and reusable contexts handling methodology is to insulate sensors’ implementation details. There are several system address this research direction: the metaDESK project [19] used sensor proxies; the Situated Computing Service [20] had a single server for both context acquisition from sensors and context abstraction; the AROMA [21] system’s architecture used the concepts of sensor abstraction and interpretation to provide awareness of activities between remote sites.

Based upon their previous CyberDesk [22] system (a modular but non-distributed architecture for separating context acquisition, abstraction and notification) research, Dey et al [18] further suggested the following requirements to support a general mechanism for building context-ware applications:

- Allow applications to access context information from distributed machines in the same way they access user input information from the local machine
- Support execution on different platforms and the use of different programming languages
- Support for the interpretation of context information
- Support for the aggregation of context information
- Support independence and persistence of context widgets
- Support the storing of context history

Illustrated in the diagram bellow, the further proposed architecture that fulfills the requirements consists of three main types of objects: context widgets, context servers and context interpreters. Each type of the objects (maybe written in different languages) is autonomous: they are instantiated independently of each other and execute in their own threads in a single computing device or multiple devices distributed in a network.



Context Widgets receive information from sensors and provide contexts, making the sensing implementation transparent to the above objects. Context widgets are defined by their attributes and callbacks. Attributes are pieces of context that it makes available to other components via polling or subscribing. Callbacks represent the types of events that the widget can use to notify subscribing components. Context widgets support both the polling and notification mechanism to allow components to retrieve current context information. They also allow components to retrieve historical context information.

Context Severs implement the aggregation abstraction, subscribing to every widget of interest and acting as proxies. They are used to collect all the contexts about a particular entity, such as a person, for example, so that an application needs not to subscribe every context widget that could provide information about the person of interest. The Sever class is sub-classed from the Widget class, inheriting all the methods and properties of widgets. It can be thought of, then, as a compound widget. Just like widgets, context severs have attributes and callbacks, they can be subscribed and polled, and their history can be retrieved.

Context Interpreters are responsible for implementing the interpretation abstraction, for example, converting a room location into a building name or taking the location, identity and sound information to determine whether a meeting is occurring. By separating the interpretation abstraction from applications, it allows the reuse of interpreters by multiple applications as an interpreter can be called by widgets, servers, applications and even other interpreters.

Though such an architecture will hide the details of the context sensors thus allowing the underlying sensors to be replaced without affecting the applications, as it was stated [18], it was not meant to claim a complete solution. Their further research regarding data processing and context acquisition will address supporting sensor uncertainties, physically distributed components' sensor fusion, and automatic sensor/context resource discovery.

9.1.1.2. Sensor resource and context automatic discovery

As discussed before, typically sensors are physically distributed and need to communicate their data, further more, typically raw sensory data are subject to complicated signal processing and artificial intelligence techniques in order to provide context information to applications. Urnes et al [27] believe that these sensor distribution and sensor discovery difficulties can be tackled with Jini dynamic discovery protocols. They declared their success of this approach by claiming having built several context-aware applications (smart map, a communication and coordination system for families, and a smart kitchen) using Jini as a support infrastructure.

Though the generality of this kind of various approaches may soon become a hot topic in context-aware computing research field, it would be helpful to explore more in this direction. I don't have the details of the applications' implementation to fully understand and appreciate the novelty, the following are largely from their report [27]. The approach explanation begins with a brief introduction of Jini dynamic discovery protocols.

9.1.1.2.1. The Jini Dynamic Discovery Protocol:

Dynamic discovery protocols are the key to the convergence of consumer electronics and computer applications. Now a wide range of new networking technologies can connect all kinds of electronic devices, but only dynamic discovery protocols allow such devices to be discovered, to offer services, and to interoperate with each other and applications – all without requiring manual installation and configuration. Jini is a new dynamic discovery protocol from Sun Microsystems.

Realized as an extension to Java's mechanism for remote method invocation (RMI), Jini offers a simple and elegant API for both service providers and application developers. Code mobility is central to Jini. Each Jini service – realized as a hardware device, a software program, or a mixture of both – offers its interface as a serialized Java object.

Known as a proxy, a serialized object is a representation of an object that can be downloaded and instantiated on any device with a Java virtual machine (JVM). Activation of a proxy's methods typically results in the remote invocation of the corresponding service methods[12].

To allow services to be offered and used, most dynamic discovery protocols provide a service registry where service providers register themselves to signal availability to service consumers. Jini implements the registry as a special Jini service, the lookup service. Service registration simply involves entering an appropriate proxy in the lookup service.

The lookup service always contains its own proxy as an entry. Both service providers and consumers must obtain this proxy prior to interacting with the lookup service. A simple multicast protocol is used to locate the lookup service on the local area network and fetch the lookup service proxy. The lookup service proxy is used to register new services, to query the service database, and to subscribe to various kinds of events, among other things.

Handling dynamic issues (the entering, leaving, and partial failure of service providers, applications, and lookup services etc.) is not easy; Jini uses the concepts of "leasing, distributed events, and transactions etc." and implementation API's to help developers tackle spontaneous changes in the Jini architecture. Jini also consists of a programming model aimed at simplifying many of the tasks required to offer and use Jini services.

9.1.1.2.2. Sensors as Jini Services:

Sensors may be viewed as electronic devices offering context services. Many of the problems related to distribution of sensors, heterogeneous sensor hardware, and lack of interoperability between sensors and applications are not unique to context-aware applications, they also appear

---

[12] Jini assumes that JVM's be available on all devices. If a device does not possess a JVM, then necessary information must be mirrored on a Java-enabled machine elsewhere (using some proprietary protocol to maintain consistency between device and mirror).

in other device networks such as home networks. Dynamic discovery protocols were designed to solve such problems. It would therefore likely be desirable to register sensors as Jini services, hence relieving developers from dealing with the said problems.

Many sensors generate data at varying rates making it cumbersome for applications to detect context changes through sensor polling. Jini's distributed event system allows sensors instead to notify applications when certain pre-defined changes occur. Sometimes, a piece of contextual information may require combining data from a large number of sensors. In such cases, it may be appropriate to collect the sensory data in a central database and register the database, rather than the sensors, in the Jini lookup service.

According to the report, the first attempts are successful and they are now investigating dynamic aspects more closely, building Jini-based sensor networks on top of the Bluetooth protocol for wireless communication. The conclusion is that dynamic discovery protocols can greatly simplify the problems related to the sensor distribution and contexts' dynamic aspects, although incorporating sensors into a dynamic discovery protocol such as Jini is not trivial.

9.2. Enabling technologies

Sensing and sensor technology is an applied science and technology; its advancement is based upon the advancement of its each embedded science and technology branch or component. There are many promising research activities going on, which may have huge impacts on the sensing and sensor's technology capability or performance enhancement. The following are just small samples that are thought to have the biggest contribution.

9.2.1. MEMS technology

Things behave substantially differently in the micro domain. Forces related to volume, like weight and inertia, tend to decrease in significance; forces related to surface area, such as friction and electrostatics, tend to become large; and, forces like surface tension that depend upon an edge become enormous[13]. Interestingly, once was considered a flaw in semiconductor devices — a "released layer" or a loose piece of circuit material in the micro-space above the chip surface — now become the key to build MEMS (Micro-Electro-Mechanical System) sensors, after people gradually understand and can control the interaction between the "loose piece" and its surround analog environment. This opens a door of building cheap and accurate tiny sensors in various applications.

Perhaps the first mass-produced micro device was the inkjet printer head in the mid 1970's. Recently, the MEMS technology develops very fast, with the support from DARPA (US Government Department of Defense Advanced Research Project Agency) Program, there are a large number of research projects (http://www.darpa.mil/MTO/MEMS/Projects/index.html) going on in this area. There are many very impressive progresses in these research projects; the most successful two are perhaps the MEMS accelerometer and MEMS pressure sensors.

Today, many MEMS accelerometer devices are batch fabricated and have reached wide acceptance in automobile industry including in air bag and chassis control, in side-impact detection, and in antilock braking systems; many MEMS pressure sensors are already used for automobile engine control such as Motorola's MPXA4100A in manifold absolute pressure (MAP) measurement and MPX4115A in barometric absolute pressure (BAP) measurement.

---

[13] It takes awhile to get one's micro intuition sorted out: an ant carrying many times its weight or a water bug walking on the surface of a pond are just two manifestations of this different micro world.

The opportunity of MEMS technology to be used in context-aware application is huge as it not only makes sensor's much smaller but also makes those nearly impossible measurement possible and smart. Some examples of this possibility can easily imagined in biometric or spying applications.

9.2.2. Biochips, Digital DNA

Biochip is a new technology consists of a cross-sectional technology combining molecular biology with micro-system technology. The Argonne National Laboratory and the Russian Academy of Science's Englehardt Institute of Microbiology, Pakard Instrument Company, and Motorola are working together to develop biochip technology. Motorola is developing manufacturing processes to mass-produce biochips. [35]

Biochips are actually carriers that come in the shape of quadrangular class slides. On the surface of each glass slide, it contains up to 10,000 tiny gel pads serving as mini test tubes, and a large number of bio-molecules (short strands of DNA — the unique set blueprints that determine the building blocks of every living spices) are attached onto this high-density gel pads in a defined micro-arrangement.

The information in DNA is encoded in long sequences of four molecular units, or bases – adenine (A), cytosine (C), guanine (G) and thymine (T). The precise pairing of A on one strand with T on another strand and G with C, allows DNA to form its "double helix". By fixing only one strand of the double helix to each gel pad, the chip employs the natural tendency of each DNA base to pair with its complementary base.

As an example, in the tuberculosis (TB) strain identification tests, when tests begin, a sample of unknown single strands of TB DNA will be spread on a chip; — based on ink-jet printing technology, the Parkard's BioChip Arrayer delivers tiny droplets of the assay fluid to the thousands of individual micro-gels on the chip's surface quickly and accurately. The assay fluid naturally pairs up with single strands of known TB DNA already in the gels – a direct match will identify drug resistant TB strains. Then the BioChip Imager, a laser optic imaging device, analyzes the thousands of biological reactions on the biochip simultaneously and displays the results in a user-friendly manner. [34]

By changing the DNA samples in the gels, this technique can be used to quickly and efficiently diagnose an unlimited range of other diseases, find differences between individual human immune systems, and monitor the environment. This technology advancement makes the before almost impossible sensing and measurement feasible, one can easily imagine that its application may soon go beyond the human user health monitoring — with some interesting discovery about some kinds of relationship between human DNA genes coding and personal characters, between body fluid composition and mood change etc.

9.3. Sensor Technology Envision

According to Paul Saffo [2], microprocessors' processing revolution indicated by PC popularization in 1980s reshaped information landscape; cheap lasers (disc player, CD-ROM, long-distance optical telephone lines) symbolized by Internet network access made 1990s "access decade"; and the 2000s will be the "sensor/effector decade" — where ubiquitous devices will not only observe things, they will also manipulate them.

The Institute of the Future's (http://www.iftf.org/) ten-year forecast series mentioned the following list as presenting some of the most important research and promising opportunities:

- Micro-machines [2]
- Nano-technology (10+ years) + microelectronic technology [1]
- Embossing circuit patterns onto silicon with arrays of scanning tunneling electron microscopes (e.g. work by Cornell University, IBM Research)
- Resonant tunneling transistors built from nano-scale silicon, integrated into conventionally fabricated circuits (e.g. work by Texas Instruments, MIT Lincoln Labs) [1]
- Use of conductive polymer "buckytubes" instead of silicon interconnection on a chip (by Purdue University, Delft University, and Yale University) [1]
- Micro-power Impulse Radar — "Personal Radar" (e.g. work by Lawrence Livermore National Laboratory)
- Piezo-materials: surface-mount sensors as well as effectors [2]
- VLSI Video, cheap video CCD camera with everything packed in a chip [2]

Sensors as the core and first step towards ubiquitous "smart stuff" or "smart material" that actively senses and responds to its surrounding environment in future, their capability improvement is based on individual sensing technology research fields' breakthrough as well as the global intelligent computation/network advancement. Thought some technology such as the nanotechnology[14] breakthrough may fundamentally change the landscape of our technology and society, an equally important aspect resides in the cheaper and ubiquitous computation and connection, which makes "ubiquitous sensing" possible — even with today's sensing technology capabilities.

[14] Some researchers use the word "nanotechnology" to refer to high resolution technology, however, more often it refers to the technology that enables manipulating materials at the nano-scale or molecular-scale level, with the features of constructional material blocks' self-assembly and/or positionally controlled reactions in manufacturing, thus to make most products lighter, stronger, smarter, cheaper, cleaner and more precise.

10. References

[1]. Ten More Years of Torrid InfoTech Growth, *Tim Oren*, *Institute for the future, Library, Technology: 1998 Ten-year forecast*. (http://www.iftf.org/html/iftflibrary/technology.html)

[2]. Sensors: the Next Wave of InfoTech Innovation, *Paul Saffo*, *Institute for the future: Library: Technology: 1997 Ten-year forecast*. (http://www.iftf.org/html/iftflibrary/technology.html)

[3]. Beyond Silicon Valley: Other Technologies and Trends to Watch, *Stephanie Schachte*, *Institute for the future: Library: Technology: 1997 Ten-year forecast*. (http://www.iftf.org/html/iftflibrary/technology.html)

[4]. Health and Health Care Forecast, Executive Summary, Medical and Information Technologies, *Institute for the future: Health and Health Care 2010, Chapter 1*.

[5]. Anind K. Dey and Gregory D. Abowd, "Towards a Better Understanding of Context and Context-Awareness", ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-22.pdf

[6]. Lecture Notes in Computer Science 1707: Handheld and Ubiquitous Computing (Proceedings of the First International Symposium HUC'99, Karlsruhe, Germany, September 1999), edited by G. Goos, J. Hartmanis and J. van Leeuwen, Springer-Verlag Berlin Heidelberg 1999, ISBN 3-540-66550-1

[7]. Perceptual Intelligence, *Alex Pentland*, HUC'99[5] (page 74~88)

[8]. Advanced Interaction in Context, *Albrecht Schmidt, Kofi Asante Aidoo, Antti Takaluoma (from Nokia) et al*, HUC'99[5] (page 89-101)

[9]. A Tutorial on MicroMechanics and MEMS, http://home.earthlink.net/~trimmerw/mems/tutorials.html

[10]. Nanotechnology, (http://www.zyvex.com/nano/)

[11]. Polar Heart Rate Monitors, http://www.polarusa.com/

[12]. Affective Computing: Sensing Human Affect, http://www.media.mit.edu/affect/AC_research/sensing.html

[13]. Microphone Topology, http://www.tangible-technology.com/microphones/tech1.html

[14]. Sensors and Analyzer Handbook, Harry N. Norton, Prentice-Hall, Inc., 1982, ISBN: 0-13-806760-0

[15]. Microphones, http://hyperphysics.phy-astr.gsu.edu/hbase/audio/mic.html

[16]. Professional Audio Page, http://www.us-epanorama.net/audiopro.html

[17]. Rawesak Tanawongsuwan, Alexander Stoytchev, and Irfan Essa, "Robust Tracking of People by a Mobile Robotic Agent", College of Computing, GVU Center, Georgia Institute of Technology

[18]. Anind K. Key, Daniel Salber, Gregory D. Abowd, and Masayasu Futakawa, "An Architecture To Support Context-Aware Applications", GVU Technical Report GIT-GVU-99-23. June 1999, ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-23.pdf

[19]. Ullmer, B. and Ishii, H., "The metaDESK: Models and prototypes for tangible user interfaces", In Proceedings of UIST'97 (Oct. 14–17, Banff, Alberta, Canada), ACM/SIGCHI, N.Y., 1997, pp. 223-232.

[20]. Hull, R., Neaves, P., and Bedford-Roberts, J. “Towards situated computing”, In Proceedings of International Symposium on Wearable Computers (October, Boston, MA), IEEE, 1997, pp. 146–153.

[21]. Pederson, E.R. and Sokoler, T., “AROMA: Abstract representation of presence supporting mutual awareness”, In Proceedings of CHI'97 (March 22–27, Atlanta GA), ACM/SIGCHI, N.Y., 1997, pp. 51–58.

[22]. Dey, A.K., Abowd, G.D. and Wood, A., “CyberDesk: A framework for providing self-integrating context-aware services”, in “Knowledge-Based Systems” 11, 1999, pp. 3-13.

[23]. Daniel Salber, Anind K. Dey, Rob J. Orr, and Gregory D. Abowd, “Designing for Ubiquitous Computing: A Case Study in Context Sensing”, GVU Technical Report GIT-GVU 99-29

[24]. Richard Hull, Philip Neaves, James Bedford-Roberts, “Towards Situated Computing”, Hewlett-Packard Labs Technical Report, April 1997, http://www.hpl.hp.com/techreports/97/HPL-97-66.pdf

[25]. Philip Neaves, James Bedford-Roberts, “Dynamic Connection of Wearable Computers to Companion Devices Using Near-Field Radio”, Hewlett-Packard Laboratories, Bristol, UK, Technical Report HPL-98-66, April, 1998, http://www.hpl.hp.com/techreports/98/HPL-98-66.pdf

[26]. Cory Kidd, “Sensors, Overview of sensor technology”, http://triton.cc.gatech.edu/ubicomp/111

[27]. Tore Urnes, Arne S. Hatlen, Pal S. Malm, and Oystein Myhre, “Building Distributed Context-Aware Applications”, http://pi.nta.no/users/tore/chi.html

[28]. Andy Harter, Andy Hopper, “A Distributed Location System for the Active Office”, IEEE Network, Vol. 8, No. 1, January 1994, or AT&T Technical Report 94.1, ftp.uk.research.att.com:/pub/docs/att/tr.94.1.pdf

[29]. Thomas G. Zimmerman, “Personal are networks: Near-field intra-body communication”, IBM Systems Journal, 35 (3):609-617, 1996, http://www.research.ibm.com/journal/sj/mit/sectione/zimmerman.pdf

[30]. Michael Gorlick, (gorlick@aero.org), “Location Awareness is Where Items At”, The Aerospace Corporation, TWIST’99, University of California, Irvine, August 20, 1999.

[31]. Judith Markowitz, “Visits and Interviews”, http://www.jmarkowitz.com/road.htm

[32]. William S. Meisel, “Can Speech Recognition Give Telephones a New Face” in BCR April 1999 Voice 2000, http://www.bcr.com/voice2000/v2n4/p18.htm

[33]. Judith Markowitz, “Motorola’s Vision of the Future”, Speech Technology Magzine, October/November 1998, http://www.speechtechmag.com/st21/judith21.htm and http://www.jmarkowitz.com/road.htm

[34]. Linda Wang, “Fending Off Drug-resistant TB Strains with Biochip Technology”, http://www.anl.gov/OPA/logos17-2/tb1.htm

[35]. logos New Biochip Products, update: “Biochip Agreement Products Introduced by Packard”, http://www.anl.gov/OPA/logos17-2/bioprod.htm

[36]. Ralph C. Merkle, "It's a Small, Small, Small, Small World", MIT Technology Review, February/March 1997, http://www.zyvex.com/nanotech/MITtechRvwSmlWrld/article.html

[37]. Dan Strassberg, "Biometrics: You Are Your Password", http://www.biometricaccess.com/biometrics.htm

[38]. Anil Jain, Lin Hong, and Sharath Pankanti, "Biometrics: Promising Frontiers for Emerging Identification Market", http://www.cse.msu.edu/cgi-user/web/tech/document?ID=436

[39]. Henry Rowley, Shumeet Baluja, and Takeo Kanade, "Human Face Detection in Visual Scenes"

[40]. "Meade Pictor XT-Series CCD Autoguider/Imagers, Section 12: CCD Tutorial", http://www.meade.com/manuals/pictor/pictor12.html

[41]. Gary W. Elko, "A Simple Adaptive First-Order Differential Microphone", Acoustic and Speech Research Dept., Bell Labs, Lucent Technology, http://www.darpa.mil/MTO/sono/presentations/lucentelko.pdf

[42]. Vasc in Robotics Institute, Carnegie Mellon University, "Modeling by Videotaping", http://www.vasc.ri.cmu.edu/~mbv/

[43]. Joao Costeira and Takeo Kanade, "A Multi-body Factorization Method for Motion Analysis", http://www.ri.cmu.edu/pubs/pub_2097.html

[44]. Charles M. LaPierre**, "**Personal Navigation System for Visually Impaired, Chapter 5: GPS, DGPS, and Backup Systems", http://www.jammys.net/lapierre/masters/chapter5.html

[45]. Richard May, Stu Turner, Kevim Audleman, and Ioannis Kassabalides, "Wide Area Tracking for Computer Supported Collaborative Work", http://www.hitl.washington.edu/people/tfurness/courses/inde543/reports/1doc/

[46]. Terry (Cef) Pearson, "How a Gyroscope Works", http://www.accs.net/users/cefpearson/gyro.htm

[47]. Ivan Chang, "Ultrasonic, Inertial and Hybrid Tracking Technologies", http://www.cs.nps.navy.mil/people/faculty/capps/4473/projects/chang2/Full.htm

[48]. Svein Yngvar Willassen, "A method for implementing Mobile Station Location in GSM", http://www.willassen.no/msl/diplom.html

[49]. John Scourias, "Overview of the Global System for Mobile Communications", http://ccnga.uwaterloo.ca/~jscouria/GSM/gsmreport.html

[50]. Javier Gozalvez Sempere, "An Overview of the GSM System", http://www.comms.eee.strath.ac.uk/~gozalvez/gsm/gsm.html

[51]. Mobile Lifestreams, "Mobile Positioning: An Introduction to Mobile Position", http://www.gsmworld.com/technology/mobposit.zip, also in, http://www.mobilePositioning.com

[52]. SiberLink web-page: http://www.siberlink.com/siber_niche_smart.html

[53]. Ellen Walsh, "Smart Card Technology: Overview", http://163.18.14.55/datapro/37216-1.htm

[54]. Rinaldo Di Giorgio, "Interview: Smart Card Guru Answers Questions", http://www.javaworld.com/javaworld/jw-12-1997/jw-12-javadev.interview_p.html

[55]. Rinaldo Di Giorgio, “Smart Cards: A Primer”, http://www.javaworld.com/javaworld/jw-12-1997/jw-12-javadev_p.html

[56]. iButton Overview, http://www.ibutton.com/ibuttons/index.html

[57]. Rinaldo Di Giorgio, “iButtons: The First Ready-to-buy 2.0 Java Card API devices”, http://www.javaworld.com/javaworld/jw-04-1998/jw-04-ibuttons_p.html

[58]. Stephen M. Curry, “An Introduction to the Java Ring”, http://www.javaworld.com/javaworld/jw-04-1998/jw-04-javadev_p.html

[59]. Steven Meloan, “Inside the Java Ring Event”, http://java.sun.com/features/1998/07/ring-project.html

[60]. Dan Awtrey, “Transmitting Data and Power Over a One-Wire Bus”, Dallas Semiconductor, http://www.sensorsmag.com/articles/0297/onewire/main.shtml

[61]. Joseph Paradiso, Graig Abler, Kai-yuh Hsiao, and Matthew Reynolds, “The Magic Carpet: Physical Sensing for Immersive Environments”, MIT Media Laboratory, http://www.acm.org/sigchi/chi97/proceedings/short-demo/jp.htm

[62]. Robert J. Orr, and Gregory D. Abowd, “The Smart Floor: A Mechanism for Natural User Identification and Tracking”, Technical report, ftp://ftp.cc.gatech.edu/pub/gvu/tr/2000/00-02.pdf

[63]. Stelian Persa, and Pieter Jonker, “On Positioning for Augmented Reality Systems”, Proceedings of the First International Symposium HUC’99, Karlsruhe, Germany, September 1999

[64]. MIT Media Labs, “Research on Sensing Human Affect”, http://www.media.mit.edu/affect/AC_research/sensing

[65]. Peter Shepherd, “The Galvanic Skin Response (GSR) Meter”, http://www.trans4mind.com/clarity/gsr.html

[66]. Donald E. O’Hair, “Biofeedback: Review, History and Application”, http://www.users.cts.com/crash/d/deohair/psychoph.html

[67]. Thought Technology’s References and Bibliography, http://www.thoughttechnology.com/refer.htm

[68]. EEG Biofeedback and Neurofeedback, http://www.futurehealth.org/nfcntral.htm

[69]. Brooke’s Sensors Page, http://www.pacificsites.com/~brooke/Sensors.shtml

[70]. Sensor Consulting Research, Inc (http://www.sensors-research.com/index.html), “Resources for Sensors, Measurements, Instrumentation and related Sciences”, http://www.sensors-research.com/links.htm

[71]. Mark Brownawell, “Interchangeability: Key to Choosing RH Sensors”, article appeared in the April 1995 issue of Sensors Magazine, http://www.protimeter-usa.com/Support/zinter.htm