

Analog Personal Information – When you just can't get what you know is available

Helge Hüttenrauch, Botond Pakucs
School of Computer Science and Communication,
Royal Institute of Technology (KTH)
Stockholm, Sweden
hehu@nada.kth.se, botte@speech.kth.se

Abstract

This paper investigates the remote access to analogue information through a Universal Speech Access server by directing a service robot to the site of the needed information. In a proof-of-concept study the transmitting back of an image of the previously electronically inaccessible information is tested to investigate whether analogue personal information resources can be made available to mobile users.

Key words: Personal Information Management, Universal Access, Mobility, Robotics

1. Introduction

In daily life we sometimes need to access analogue information. Consider for instance the situation when you realize that you have to work over to finish a report. Question: Can you put in this extra effort today or was there not something else you had to attend to? You remember that there was a *rescheduled* football match from your son's team this week and that you signed up for transportation. Uncertainty is growing on your part – and what's even worse: You just know that your Personal Information Management (PIM) on your mobile device does not have the answer. The *regular* schedule is there, both for your professional time-keeping as well as the jokingly called Family Information Management (FIM) or "kids-business". What's *not digitized in your PIM* however is the *scribbled down note* on the calendar residing next to the fridge, or for that part, most of the analogue information at home or on your job. In short: You have to access a crucial piece of information and you *know exactly where the needed information is without being able to get to it*. In theory the above situation should never occur due to the powerful communication devices we have available today – but personal experience and empirical evidence on how people organize and manage different analogue information [Nässla & Carr (2003); Swan & Taylor, (2004)] tell us differently. In combining previous work on mobile devices as remote control to service robots [Hüttenrauch & Norman (2001)] and the possibilities of accessing information through a Universal Speech Interface [Pakucs (2004)] we present our concept to explain the relevance, conceptual ideas, and our attempt in conversing technologies of mobile communication, speech dialogue, and robotics. Our system is intended to show how media breaks in retrieval and handling of personal information can be bridged by controlling a service robot in a remote location via speech and receiving in return, information from this robot, e.g. on the mobile device.

2. Background

2.1 Personal Information Management and Analogue Information

Managing the complexity of "just daily life" can be a challenge. Many different information and media types need to be handled including electronic and analog messaging. Keeping

schedules and staying on top of what feels like parallel and time-competing tracks of activities is sought to be managed with electronic Personal Digital Assistants (PDA), by PIM applications, or in analog form of filoFax[®] managers. Features and functions of such PIM applications are numerous, but often a core of a calendar, an activity-journal, an address book, to-do-lists, and notes can be found. A flora of handling reminders using different analog artifacts or (mobile) electronic devices are additionally used – all of these sources of information compete for attention and require active administration not only for professionals on the job, but also in managing private life.

From PIM-research different issues are well established: Information needs to be worked with actively, i.e. *finding*, *keeping*, and *managing* [Jones (2005a)] of information are central themes. Furthermore have researchers looked at the aspects of communication, coordination, and co-operation [Poltrick, Grudin (1999)] as part of the field of Computer Supported Co-operative Work (CSCW). Another problem addressed is the handling of multiple, parallel work tasks and the interruption and time-allocation/task-switching between these tasks [Hudson et al. (2002)].

While pointing towards interesting issues an important challenge is missed out upon. Many people are at different physical places during one day and as a consequence do they only have subparts of their information context [Kirsh (1995); Dourish (2004)] and information available to them: When one leaves the home, information attached, e.g., to the fridge door is not available any longer, despite the fact that the *whereabouts* of this information is well known. In this paper a novel approach combining work on mobile devices as remote control to service robots and the possibilities of accessing information through a Universal Speech Interface [Pakucs (2004)] will be introduced: When a user is in a remote location, she can pick up her phone and call an interactive speech server. This server then takes up contact with a robotic system in the home. The robot can be instructed to physically go to places and check upon information available in the home (e.g., notes, calendars etc.). Smart phones can thus be used to retrieve analogue information from a remote location.

The remaining paper is structured as follows: The current PIM situation is rephrased according to our conceptual idea in the next section. Then a short introduction to service robotics and universal speech interfaces is given to motivate our efforts in drafting this concept and introduce the used technologies. Finally, a proof-of-concept scenario experimentation and a discussion concludes this paper.

2.1 A different PIM Perspective

The border between work and private life activities is increasingly blurred out. This development can also be seen as consequences of the development towards ubiquitous and pervasive computing environments, where more computing power is becoming accessible from a multitude of devices and appliances as envisioned by [Weiser (1991)]. However, the claim is made that this development also requires a strengthening of PIM access solutions that are not yet available.

Some questions might be helpful to introduce the idea of universal access to PIM functions: What does it mean to have universal access to PIM? Should every device and software tool have a (more or less well synchronized) copy of our personal information or is a unique single data haven required? Who “owns” the/my personal information? Who should pick the tools

for handling and using personal information, e.g., on the workplace? In the following some guidelines from [Banavar *et al.* (2000)], are given for consideration. The authors claim that

- a device is a portal into an application/data space, *not* a repository of custom software managed by the user,
- an application provides means by which a user performs a task and *not* a piece of software that is written to exploit a device's capabilities, and finally
- the computing environment is the user's information enhanced physical surroundings, *not* a virtual space that exist to store and run software.

In other words, personal information should be stored in at least one or possibly in a few locations in a way that enables different devices and software tools to gain access to these repositories. These central repositories should furthermore provide support for a Personal Unification Taxonomy (PUT) and different strategies to use and manage personal information [Jones *et al.* (2004)]. Accordingly, ***PIM data should not be kidnapped by any application or device.***

Users may for example want to access and use analog data even from remote locations without being forced¹ to first go manually through the process of transferring this information to their currently used PIM or digital device. A mobile and embodied service robot² may consequently just provide the means to bridge the analog with the digital world. For instance can a personal robot be used in the home environment. It can have multiple (service) tasks including manipulation and information management. With these features a service or personal robot can therefore be seen as a local agent that informs, alarms and (re-)acts upon user requests even if (s)he is in a remote location ("sensing and acting over distance") [Goldberg (2000)].

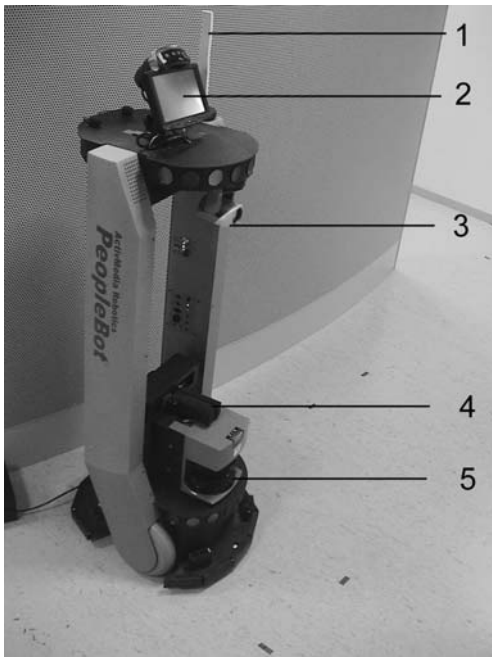


Figure 1: ActivMedia PeopleBot at the Royal Institute of Technology.

2.2 Service Robots – A brief Introduction

Service robots or socially interactive robots [Fong *et al.* (2003)] are different from, e.g., industrial robots. They are intended to work in the close proximity of users where, e.g., touching is not only a possibility, but could be one of the accepted modes of operation [Breazeal (2004)]. The environment of operation will be the private home or other domestic settings.

The robotic research platform employed in this study is a commercially available ActivMedia PeopleBot³ (see Figure 1). It is a wheel-based platform with a wireless LAN (1), a touch-screen (2), a pan-tilt-zoom camera (3), a gripper (4), and a laser-range-finder (5). Additionally system components used, but neither shown nor discussed further are for example the system's navigation components⁴.

Interesting for our scenario are robotic systems that work in *socially interactive* or *semi-autonomous* fashion. Mission (or utility) function(s) can be speci-

¹ In contrary to what is proposed by [Bell (2001)]

² Software agents are explicitly excluded from our discussion; we are aware that service robots might communicate *with* or *through* software agents with users or other hard- and software components.

³ www.mobilerobots.com

⁴ We gratefully acknowledge the support of Elin A. Topp, Centre for Autonomous Systems (CAS);

fied in (at least) two ways: Either pre-programmed task can be invoked or the system's features can be directly controlled. This semi-autonomous state of the command and control can be achieved by multiple, multimodal interfaces: For our scenario the commanding of the robot from a mobile device is focused upon. In summary are service robots an interesting addition to interactive artefacts as they combine *perceptual*, *computational*, *communicative*, and *manipulative* abilities.

2.3 Universal Speech Access

Speech based interaction is a promising solution for mobile and ubiquitous computing environments. For people without hearing impairments, speech is appealing because of the naturalness of the modality. Speech is not only a highly efficient and flexible way of human-to-human communication [Bass (2001)], but humans also have a lifelong training in verbal communication. Speech dialogue systems do not require hands to operate and are "always" available, unless the social setting or the physical environment rules out the usage of speech. Accordingly, speech is especially desired in "hands & eyes occupied" situations, e.g., while driving.

In spite of the advances in speech technology during the last decades are speech based interfaces still considered unreliable. Mobile applications with good usability characteristics are especially rare. An application-centered design and speech processing functionality embedded in, and performed by the appliances themselves are some of the main reasons behind these shortcomings. In mobile environments with diverse and concurrent speech interfaces, the usability and HCI related issues are expected to create new challenges. Interface consistency, multi-domain approach, support for user control and service coordination are some of the challenges which have to be considered when designing and developing mobile speech interfaces [Pakucs (2004)]. A human-centered approach [*ibid.*] has been proposed for handling of the above challenges. The central idea in the proposed solution is to know as much as possible about the individual user, such as speech and behavior patterns, history of previous interactions, preferences etc. Every user is expected to operate a single, highly individualized universal speech interface for accessing a multitude of services and appliances. Accordingly, a high level of personalization and adaptation to the context of use are some of the major characteristics and advantages of the proposed solution. This solution is believed to support PIM related activities.

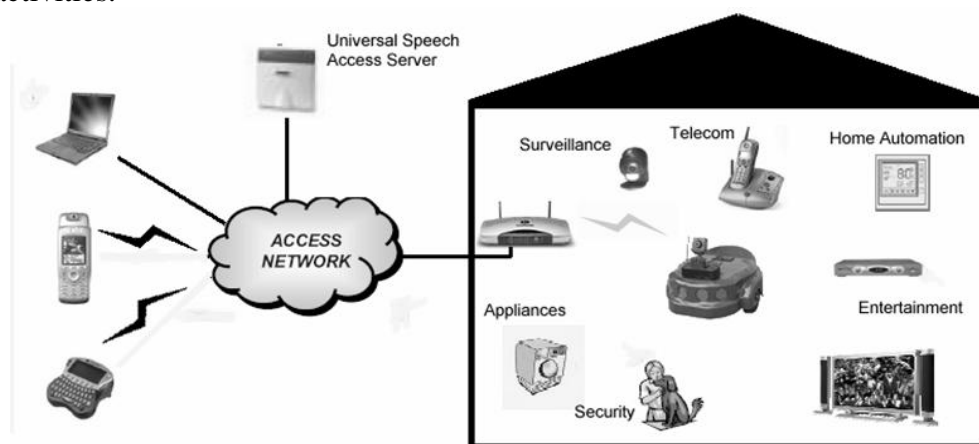


Figure 2: Universal Speech Access to the Networked Home Automation System

3. Proof-of-Concept

3.1 System overview

For our proof-of-concept we wanted to allow users to “call in” to the robot (i.e., through a universal speech server) that connects to the robot in a remote location. The robot will be able to communicate via this speech-server, perform tasks like looking up analog information sources (see figure 4 below) and transmit back to the user this information either by visual, textural, or spoken representation/feedback. If available it is also attainable that the robot is able to provide or control other appliances, systems, or services in the home (figure 2). The robot could thus be a mobile artefact that has the ability to be part of the home- entertainment, information-, health-monitoring-, communication-, security-, heating-/ irrigation, or white-ware appliances world as sketched by researchers in the field of ubiquitous / persuasive computing, home-automation and/or “ambient intelligent” systems.

3.2 Basic Functions of a Remote-Information Robot Manager

The following architecture and features were identified for a service robot that can be called-in upon via a universal speech interface to “look up” analog information. We focus here only on the basic set-up; an extended feature-set allowing for example to connect to other home-automation is deferred. The “Basic set-up” will be implemented by data exchange between the robot and the speech server based on a XML-protocol that defines the entities and current values. Communication between the mobile device and the universal speech server is assumed to exist whether provided for by a service provider (e.g., via GPRS) or in private hand (WLAN, Bluetooth or similar networking techniques). The universal speech access server will in turn take contact with the home-appliance networked devices like the robot through broadband TCP/IP-based networking. The robot needs to be able to determine its current position and how to get to a new instructed position in the operation area of the home while avoiding obstacles. Certain pre-determined places where information content can be assumed to be available (e.g., location of a shared calendar, the fridge door) should be known to the robot up front. During a mission the robot should permanently update its current position in order to communicate locomotion progress back to the user (“going from hallway into kitchen”). The basic functions of the remote-information robot manager will enable the robot to

- move to different <predefined> places,
- focus the camera on objects,
- take pictures [record video sequences],
- transmit the material back to a user.

3.3 Speech based Access to Personal Information: Calendar as prototype

The AthosCal, a speech based calendar prototype, was developed in the framework of the DUMAS project [Jokinen and Gambäck (2004)]. The research goal of the project was to investigate adaptive multilingual interaction, study the generic features of the Athos framework, and extend the AthosMail functionality with access to calendar data through speech access.

Support for the iCalendar standard [Dawson and Stenerson (1998)] was developed to allow the voice-controlled calendar to be used as a complement to desktop-based calendars. A tool-kit has been implemented for parsing iCalendar files into the back-end database and for generating iCalendar files from the database. The generated files can be accessed by any calendar application with support for the iCalendar standard.

The VoiceXML-based dialogue descriptions are generated by using a template-based solution. Initially, for handling the dialogue descriptions, the proprietary SpeechWeb platform⁵, has been used. For accessing the voice controlled calendar, users only need to make a telephone call to the Speech-Web platform. Handling the dialogue can be achieved with other generic speech platforms such as the Athos platform or the SesAME [Pakucs, (2003)] generic dialogue manager.

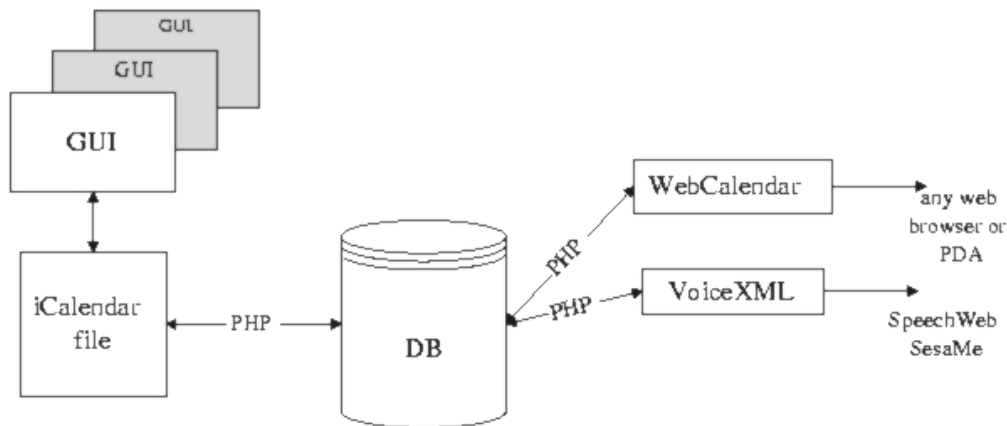


Figure 3: The architecture of the voice controlled calendar.

The WebCalendar GUI has been adapted to small screens allowing for example PDA-based terminals GUI access. By employing a voice over IP based telephony solution and wireless LAN-network a stand-alone PDA-based solution has been enabled. The AthsoCal prototype was evaluated in a small user study [Pakucs & Huhta (2004)]. One of the major conclusions was that the speech based access to PIM data is a feasible and efficient solution for mobile users.

3.4 Speech I/O-based, remote Service Robot control

The remote, speech-based access to a service robot is implemented as a modification to the above described voice-controlled calendar (see figure 3). From a system design perspective the service robot application retrieves information (through the robot) that is seen as “just another” data-type made available through a speech-based request.

In our implementation scenario a user “calls-in” to the robot-speech-server with a mobile phone/PDA and by speech retrieves the robot’s current status, including mission-status, location, and/or battery-voltage. The robot can then be send to different places and there take photos with its on-board-camera. These images will be transmitted back to a shared (robot/speech-server) repository in order to make this image available to the incoming caller, i.e. the user of a mobile phone or PDA.

3.5 Experimentation

An implementation of our system has been tried to test our idea, to collect real robot mission data for system testing, and to study the overall feasibility of our idea. The PeopleBot robot was send on three missions in an office environment. The first mission was to go to the analogue sign-in board (figure 4, left) to check the current status of members of the laboratory. The second mission was to take an image of the local blackboard in the kitchen for a concert series announced on a poster (figure 4, middle). Finally, the analogue calendar at the shared

⁵ h20208.www2.hp.com/opencall/products/media/index.jsp

meeting room entrance was checked with help of the robot (figure 4, right) to see whether the room was available for a booking.



Figure 4: Robot images of sign-in board (left), concert schedule (middle), and meeting room calendar (right)

The robot could go to the intended places and “report back” with the images seen in figure 4⁶. From this initial experimentation a couple of issues for future refinement became evident: The robot’s field of vision might require manual user adjustment (i.e., the exact pan-tilt-zoom) of the camera once the robot has reached a target location. As the mission length in time duration varies, a prompting of the user once the robot has reached its intended goal-position seems advisable. Another feature that we did not initially plan for was the *post-processing* of the images that is likely required to optimize for small-screen viewing of the images. An automatic differentiation according to connected client device abilities needs to be investigated.

4. Discussion

In this paper we put forward ideas that couple the speech based access to PIM data with the Teleoperation of sending a robot to a only locally available, analog information source, and transfer this information back to its user/owner. From literature we found evidence that these information sources exist and that they are important. Furthermore, the type of information focused upon is generally not being transferred to other PIM formats and applications. As a consequence do users know very well *that* the information exists, *where* it is to be found, but nevertheless are unable to access it, once they left the location of the information. We have therewith identified the need to allow users to get a view on this data over distance through a universal speech access server and a mobile robot that is directed to access and transfer this information. On a more abstract level this remote access through the robot can be understood as a system helping to unify the access of PIM records as the speech server is contributing to a “virtual unification”: If the demand arises, the information can be accessed independent upon placement and modality.

The service robot that we used to conceptualize the idea of a remotely accessible source of personal, but analog information, can be seen as part of a ubiquitous/pervasive computing environment for home automation. Such systems include, but are not limited to monitoring and controlling appliances, lightning-, heating-, entertainment systems etc.

Our proposed system is not *technically* advanced. The novelty of our proposal is instead an innovative mix of advanced, but already available technological solutions. The concrete

⁶ Names have been blurred manually for privacy.

practical realization of the overall system is straightforward as an initial proof-of-concept and limited in its (clear defined) utility. However, we can currently only speculate about the future of such information services and remote monitoring, control, and handling of information and systems. Different application metaphors and/or marketing niches might come into existence, e.g. as “PC-home of the future”, a “Personalized Butler”, “MyBender”, a “Home Control”, or as an add-on to a future game-console offering.

References

- [1] Bass, L. (2001). Interaction Technologies: Beyond the Desktop. *User Interfaces for All*. Ed: Stephanidis, Constantine. Lawrence Erlbaum Associates, London, pp. 81-95.
- [2] Banavar, G.; Beck, J, Gluzberg, E., Munson, J., Susman, J., Zukowski, D. (2000). Challenges: An Application Model for Pervasive Computing. *Proceedings of Mobicom 2000*.
- [3] Bell, G. (2001) A personal digital store. *Communications of the ACM*, 44(1): p. 86-91.
- [4] Breazeal, C. (2004). Social interaction in HRI: the robot view. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, May 2004, Vol. 34, No. 2, pp. 181-186.
- [5] Dawson, F. and Stenerson, D. (1998). Internet Calendaring and Scheduling Core Object Specification (iCalendar). RFC2445, Lotus, Microsoft, November. URL: www.faqs.org/rfcs/rfc2445.html.
- [6] Dourish, P. (2004). What We Talk About When We Talk About Context. *Personal and Ubiquitous Computing*, 8(1), 19-30.
- [7] Fong, T., Nourbakhsh, I. and K. Dautenhahn (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, Vol. 42 (3-4), pp. 143-166.
- [8] Goldberg, K. (ed.) (2000). *The Robot in the Garden: Telerobotics and Telepistemology in the Age of the Internet*. Cambridge, MA: The MIT Press: 366 pp
- [9] Hudson J. M., Christensen J., Kellogg W. A. and Erickson T. (2002). I'd be overwhelmed, but it's just one more thing to do: Availability and interruption in research management. *Human Factors in Computing Systems: Proceedings of CHI'02*. New York: ACM Press, pp. 97-104.
- [10] Hüttenrauch H. and Norman, M. (2001), PocketCERO – mobile interfaces for service robots, *Proceedings of Mobile HCI 2001, IHM-HCI 2001 Lille, France, Sep. 2001*.
- [11] Jokinen, K. & Gambäck, B. (2004). DUMAS - Adaptation and Robust Information Processing for Mobile Speech Interfaces. *Proceedings of The 1st Baltic Conference "Human Language Technologies - The Baltic Perspective*, pp. 115–120, Riga, Latvia.
- [12] Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, Vol. 9 (3), URL: firstmonday.org/issues/issue9_3/jones/index.html (2005-06-03).
- [13] Jones, W. (2005a). Personal Information Management (PIM) 2005. Lecture Notes. PIM-Course at the Royal Institute of Technology, Stockholm, Sweden.
- [14] Kirsh, D. (1995). The Intelligent Use of Space. *Artificial Intelligence*, Vol. 73 (1-2), pp. 31-68.
- [15] Nässla, H. & Carr, D. (2003). Investigating Intra-Family Communication Using Photo Diaries, *Proc. of HCI International 2003*, pp. 984-988.
- [16] Pakucs, B. (2003). Towards Dynamic Multi-Domain Dialogue Processing. *Proceedings of the Eurospeech'03*, Vol. 1, pages 741–744.
- [17] Pakucs, B. (2004). Butler: A Universal Speech Interface for Mobile Environments. *Lecture Notes in Computer Science 3160*. Springer. pp. 399-403.
- [18] Pakucs B. & Huhta, S. (2004). Developing Speech Interfaces for Frequent Users: the DUMAS Calendar Prototype. *Proc. of COLING 2004 Workshop on Robust and Adaptive Information Processing for Mobile Speech Interfaces*. pp. 65-68.
- [19] Poltrock S. and Grudin, J. (1999). CSCW, groupware and workflow: experiences, state of art, and future trends. *Proceedings of CHI '99: abstracts on Human factors in computing systems*, pp. 120-121. ACM Press.
- [20] Swan, L., & Taylor, A. S. (2004, April 24-29). An ethnography of lists in mothers' work. *Conference on Human Factors and Computing systems*, CHI Fringe 2004, URL: research.microsoft.com/~ast/home.aspx, 2005-05-30
- [21] Weiser, M. (1991). The Computer for the Twenty-First Century, *Scientific American*, pp. 94-10.