A Generic UPnP Architecture for Ambient Intelligence Meeting Rooms and a Control Point allowing for integrated 2D and 3D Interaction

Ali A. Nazari Shirehjini

Fraunhofer-IGD
Departments "Interactive Multimedia Appliances"
Fraunhoferstr. 5, 64283 Darmstadt, Germany
ali.nazari@igd.fraunhofer.de

Abstract

In this paper we present a generic UPnP Presentation Architecture for AmI meeting rooms. It allows the development of applications based on standardized access mechanisms. This architecture introduces besides standard lighting devices also a UPnP design for complex projection settings, analog audio-video devices, shutter blinds and media repositories. Using this architecture, AmI developers benefit form UPnP device discovery as well as standardized access to devices and media repositories. This allows the development of interaction solutions working in various AmI environments that may be composed by different device infrastructures. Based on this architecture, we present the PECo system, a novel Control Point which provides integrated and intuitive access to the user's surrounding and media repositories [1] allowing to control and manage intelligent environments. PECo uses an automatically created 3D visualization of the environment. Entering a room, PECo discovers the infrastructure and available devices and builds the integrated user interface. The 3D visualization creates a logical link between physical devices and their virtual representation on the user's PDA. By doing so, the user can easily identify a device within his environment based on its position, orientation and form. There he can access the identified devices through the 3D interface and manipulate them directly within the scene. For example he can click on a 3D object to turn on a light. The 3D interface allows the user to access the infrastructure without demanding knowledge about specific device names, IP-numbers, URLs etc.

1. Personal Access to AmI Meeting Rooms

Ambient Intelligence (AmI) is the vision of a world where we are surrounded by a huge amount of intelligent and small devices, which are seamlessly embedded in the user's environment [2, 3, 4]. They pro-actively support the user by performing everyday activities. Considering intelligent multimedia environments such as modern conference rooms, today there already exist countless devices including ambient audio-visual rendering equipment, complex lighting infrastructure, intelligent HVAC equipment, intelligent media storage devices, as well as personal appliances such as notebooks, tablet PCs, smart pens, PDAs, MP3-players or smart phones. In such a world, major challenges are providing the user ambient interaction means to exploit all those capabilities without interrupting user activity flows. One of key questions to provide the mentioned interaction systems is how to allow integrated access to a) physical environment and b) multimedia documents. Considering again a typical presentation scenario, an intuitive interaction system should allow the user for example to present his slides in various conference rooms, each of them possibly with different infrastructure and capabilities. Another fact is that users mostly store their personal media across their personal devices. The user should be able – in the mentioned example – to present his media on available display devices, according to his situation and personal preferences. Adjusting complex light settings, starting video conferencing or just a video clip are other usual activities performed in such multimedia environments. Even supporting such a simple presentation scenario requires device transparent access to distributed media as well as infrastructure transparent access to devices. Beyond these, the user should be able to orient himself within the complex, foreign environment. He should be able to find and intuitively select and control a display among many even without having knowledge about specific technical information such as IP-numbers or unique device etc. Last but not least, he should be able to link his personal media with environmental (foreign) devices, such as displays or audio equipment.

1.1 Integrated Media and Environment Access

With the PECo system, we provide the user a novel interaction metaphor allowing access to both intelligent environments and personal media [1, 5]. PECo uses a 3D visualization, which makes a direct link between physical devices and their virtual representation. Figure 1 shows a complex meeting scenario and the PECo UI for that room. The automatically created and updated 3D visualization helps the user to easily orient him within a room without requiring knowledge about technical details of the available infrastructure in his current (foreign!) environment. There he can identify a device (e.g. the camera as shown in Figure 1) based on its position, orientation and form. Changes to the environment, new devices or re-positioned devices are identified and implicate an update of the UI. Henceforth, the user can access identified devices through the 3D interface and directly manipulate them. For example, he can just click on a 3D object to turn on a light.

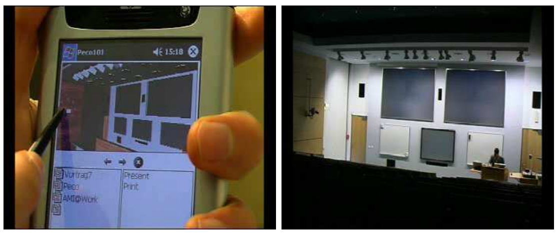

Figure 1: A user interacting with an intelligent conference room using PECo

The second important feature of PECo is its Personal Media Management (PMM) functionality. It provides the user a single point of access to his distributed personal media. It discovers available media and provides them on the users PDA regardless of the physical storage location of those documents.

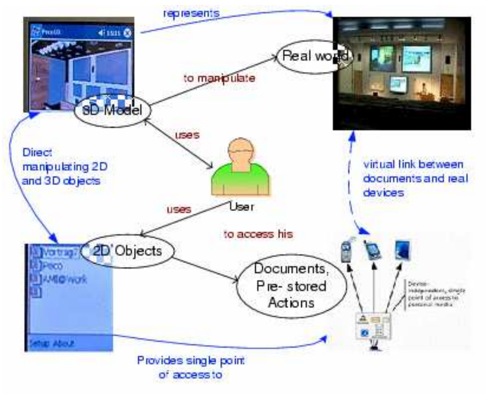

Figure 2: The idea of 3D-based interaction with intelligent environments

1.2 PECo Interaction Metaphors

PECo uses a 2D interface for document access and a 3D interface for room and device interaction (see Fig. 2). The user is able to transfer objects from his 2D world to his 3D world (Fig. 1). By doing so, PECo integrates the virtual media repository as well as the physical environment of the user into a unified personal environment. This enables the user for example to move a PowerPoint document - which is stored on his notebook – to the beamer by just one drag&drop operation. For both 3D and 2D interaction, PECo provides unified metaphors (pointing, selecting, drag&drop, clicking) so the user can handle beamers and lights in the same manner he also interacts with his files and directories. By doing so, we extend the well-known metaphor of direct manipulation to the physical world.

More detailed description of PECo user interface and automatic 3D creation is provided in [6, 7, 5].

1.3 Requirements on Underlying Middleware

The realization of discussed a) Personal Media Management component, b) the dynamically created 3D visualization of foreign environments and their devices, as well as the physical access to different environments requires (see Figure 3):

- Device Discovery: entering an environment the 3D scene of the raw room (without devices) is downloaded from a known database. Next, available devices should be discovered and their corresponding 3D objects inserted within the 3D scene at the right position and orientation [1]
- Media Management: available media on several devices should be accessible by the user
- Standardized Device Access: devices of various environments should be accessible by a standard way, allowing to control foreign environments

In following sections we propose the usage of the UPnP standard to meet above requirements. Beside UPnP, PECo also requires position and orientation information, which we gather by geometry management system, using a combination of the Ubisense [?] system for positioning persons. We developed a passive RFID based system [8] to gather precise orientation and position information for mobile multimedia devices. Stationary devices such as fixed back projection systems provide their location information by their self. For this purpose, we extend the UPnP device profile. For each (UPnP) discovered device we look up its position and orientation and disclose devices located in other rooms of the same sub-network (UPnP discovery delivers every UPnP device of a subnet). We extend each UPnP device by an UPnP service (optional; not part of the standard) providing dynamic location information for that device. However, in this paper we neither focus on the location-aware device discovery nor on mentioned positioning services. Moreover, we describe a generic design of an UPnP presentation architecture, since most of meeting room devices are not directly standardized by the UPnP Forum. We deployed this design architecture to all real life meeting and conference rooms at Fraunhofer-IGD (Darmstadt, Germany). This architecture serves as a basis for systems, targeting AmI meeting rooms, such as the PECo system.

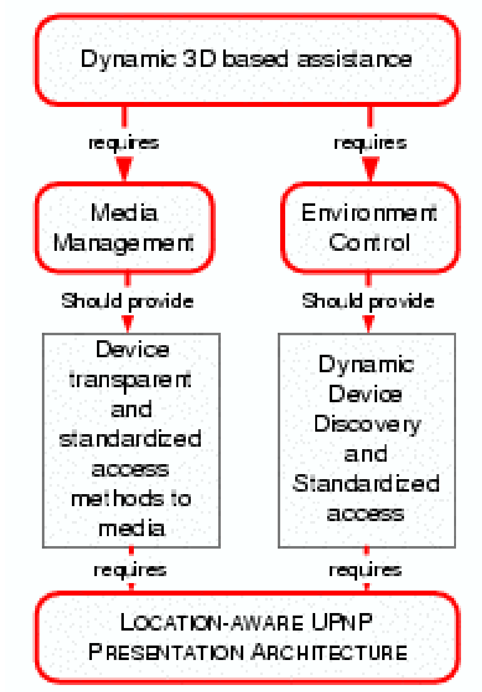

Figure 3: 3D-based dynamic assistance requires UPnP

2. Designing a UPnP Presentation Architecture Allowing Standardized Access to Meeting Rooms

2.1 Pre-UPnP Era

Such Rooms are composed of the physical space (walls, windows, doors, etc.), embedded stationary and mobile devices (e.g. lights, TV or beamers) as well as the user's personal appliances such as his notebook or smart phone. There are several actors, e.g. lighting devices, visualization devices, video conferencing equipment, HVAC infrastructure and shutters. Some devices already come equipped with standard access and control mechanisms, such as the EIB (European Installation Bus). Others provide proprietary access mechanisms - often based on RS232. To control EIB devices, we use of-the-shelf Instabus connectors (such as of the GIRA company) and the open source software EIB Home Server [9]. A person has a set of connections for video, network and power, where he can connect his laptop, and one microphone input. In parallel to PECo, there is a conference room control panel allowing access to the room functions. The available devices and commands are listed in Table 1. An analog audio-video bus interconnects the room devices. There are several sources connected to this bus and several renderer devices that can show any of the inputs of the bus. Considered renderer devices are back projection systems, mobile beamers, video-telephones and audio mixers.

Device	UPnP	Actions
Light	Dimmable Light	on/off, brightness up/down
Blind	Solar Protection Blind	up/down/stop
Projector	AV Media Renderer with embedded Server	on/off, source selection, shutter open/close
VCR	AV Media Renderer with embedded Server	play, stop, pause, rec, ffwd, rew, eject
Visualizer	AV Media Renderer	on/off, light on/off, focus more/less, zoom more/less, paper size A4/A5/A6
Camera	Security	Camera PIP on/off/change, move up/down/left/right, zoom more/less
Video-Phone	AV Media Renderer with embedded Server	Volume up/down/mute, source selection
Mixer	Two AV Media Renderers, one with an embedded Media Server	mic. volume up/down/mute, line in volume up/down/mute

Table 1: Devices available in the room and functions associated with them

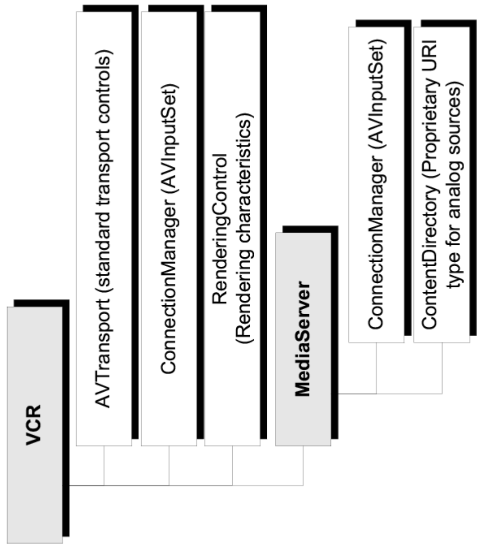

Figure 4: UPnP Design for a VCR

2.2 Challenges by applying UPnP for Meeting Rooms

Considering the above-described environment architecture, which represents todays networked (non-UPnP) multimedia workspaces, major challenges by providing an UPnP architecture for those environments are:

- Absence of standardized DCP for some devices, making impossible a complete standard UPnP Meeting Room
- Absence of bidirectional communication. This is a problem when we interact using a low-level state modifier (like a physical switch) because the UPnP system never receives a notification of the change
- Some controls like dimming in the lamps, are numerical in UPnP. Today, most dimmable (non-UPnP) lights only allow the increase or decrease fixed steps with no reference value for the system
- UPnP AV architecture differs from classical AV, especially the way the media is processed and sended to the devices

2.3 Design of UPnP Devices

Current AV projection devices present a complex behavior. The roles of Media Renderer and Media server are not so well defined when we are dealing with classical AV devices such as large visualization walls, CAVE systems or even usual projectors. To deal with this problem, Intel has proposed its AV Architecture [10][11] to design more complex AV devices. According to them, classical AV devices could be represented through a hierarchy of embedded devices. The current state of the development including UPnP Devices and implementations can be downloaded at www.igd.fraunhofer.de/ anazari/UPnP.

2.3.1 UPnP VCR and Analog AV Device Inputs

Following this idea, we represent an analog output (see Figure 4), or a logical group of them, as an AV Media Renderer. Analog inputs are represented as entries in an embedded Media Server. This Media Server is the child of the top most parent device in the hierarchy that can render content sourced from the analog input.

2.3.2 UPnP Projection Devices

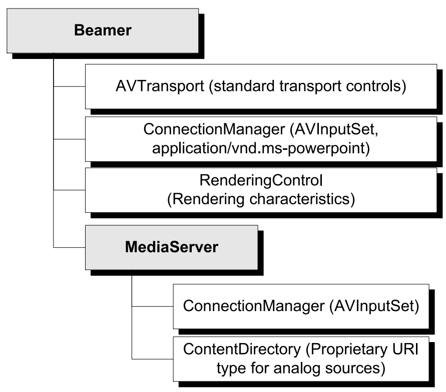

Figure 5: UPnP Design for a Projector

To represent projection systems such as display walls (e.g. GraphXMASTER CX60-100U) or mobile back projection systems (e.g. Rear Projection Smart Board 3000i) we will follow the Intel AV Architecture recommendations, but have to do some modification. As one of important extensions to the basic UPnP AV Architecture, media rendering agents are embedded into the projectors. They are located in a presenter PC directly bounded to the projector. So a beamer can visualize/auralize all content types supported by its embedded agents (currently PPT, MPEG, AVI, MP3). By adding rather agents, more media types (e.g. live video, 3D streams, PDF files, MindMaps) could be rendered. These embedded media rendering agent access desired files and render them directly to the displays.

2.4 Embedded and External Media Servers

Media to be rendered are provided by Media Server. In the current presentation architecture there exist two different Media Servers: one in the computer embedded to the projection system that offers the local stored data and another external one that offers externally stored data. For rendering analog sources (VCR, video telephone) which are not available by Media Servers as a real file, the extended UPnP design of a projector allows to (physically) select those analog devices as analog projector input sources. These sources are represented as CDS entries in the UPnP Media Server that is embedded in the beamer. Moreover, the user also is able to interact directly with those devices, for example in order to play a DVD or to move a room camera.

The usage of Media Servers provide means to a) integrate Media of several distributed sources, b) to realize everywhere access, c) to transport media to rendering devices, as well as d) to provide a standardized interface for higher level applications such as the PECo interaction appliance.

2.5 The contribution of the presented UPnP architecture for AmI applications

As a result of the current presentation architecture in which we logically aggregate and include rendering and media providing agents into a projector (Figure 5), AV devices become more advanced devices than simple AV Renderers or Media Servers. The current design is a scalable and flexible multimedia platform that can be scaled to support new media types and arbitrary compositions of projection systems, e.g. multi-display environments (e.g. as in Figure 1). Deploying the described UPnP presentation architecture and providing UPnP profiles and services for devices listed in Table 1 enables higher level applications (e.g. the PECo system) to use a) device discovery services, b) storage independent retrieval, access and rendering of multimedia documents by using standardized mechanisms, c) standardized access to infrastructure and devices allowing to interact with various environments built up by different physical infrastructure, d) future compatibility when new UPnP devices would be added to Meeting Room.

3. PECo: an intuitive Control Point

The overall system architecture of PECo (Figure 3) comproses the environment (this is the room), a Personal Media Management agent, an Environment Controller agent. The Interaction Appliance - a PDA visualizing the novel 3D-based user interface - serves as a control point for the whole system including media servers. The User Interface has already been described in [1, 6, 5] and presented at CeBIT 2004 and 2005. It is also has been demonstrated at PERVASIVE05, Munich.

3.1 The Communication Infrastructure

Three different networks provide the Communication Infrastructure underlying the described UPnP architecture. The communication between software components is based on WLAN or LAN. The communication between all audio-visual equipment is based on a proprietary analog system (optical communication). This analog A/V bus system interconnects several notebooks and PC VGA outputs, beamers, cameras and audio equipment. The third channel includes connections between the Environment Controller agent and the actors and devices. These connections are established by RS232 and EIB.

3.2 Environment Controller Agent

The Environment Controller provides services for accessing the user's physical environment [1]. Using the UPnP mechanisms (described before) it performs the detection and integration of devices available in a room (location-aware device discovery). Furthermore, it provides an environment model, which comprises device profiles and corresponding access services, a geometry model allowing for location-aware device discovery and gathering position information, as well as 3D model of the environment that contains discovered devices.

3.3 Personal Media Manager Agent

The Personal Media Manager is a device-independent single point-of-access, providing a unified facility to organize and access the different media repositories, which the user has at his disposal. It provides a 2D User Interface to accesses and use (see Figure 5) UPnP Media Servers.

4. The PECo Core System

(Figure 6 shows the software architecture of the PECo Interaction Appliance (the control point) and its most important entities. Currently, the PECo Interaction Appliance is based on a set of C++ classes running on a PocketPC device. All 2D controls are realized by standard MFC controls. This architecture is responsible for creating the user interface, to process UI events and to access devices in order to perform actions.

Device	UPnP	Actions
Projector	AV Media Renderer with embedded Server	Renders PPT and MP3 files and a proprietary content type that represent the different analog inputs. These CDS entries are offered by the embedded Media Server, that can offer other types (PPT, MP3) too
VCR	AV Media Renderer with embedded Server	Renders a proprietary content type data that represent the different analog inputs. These CDS entries are offered by the embedded Media Server. This allows us to record from different sources
Visualizer	AV Media Renderer	Has no content type to be reproduced, as we consider it as a live stream, so we never control with UPnP the content rendered.
Camera	Security Camera	Security Camera device offers us a set of standardized actions to control the camera. For the projector or VCR, it is a live stream so they have only need to select this source. Video-Phone AV Media Renderer with embedded Server Media Renderer represents the input of the video telephone; that is the stream to be sent. This renderer has a proprietary content type which represents the different analog sources. The Media Server offers the different possible analog inputs as CDS entries.
Mixer	Two AV Media Renderers, one with an embedded Media Server	It contains two different Media renderers that control independently each input (Line-In and microphone). There is an embedded Media Server that controls the different sources for Line-In. We consider the microphone as a static renderer with no content type

Table 2: Enhanced UPnP Devices for AmI Meeting Room

PECo Core is responsible for launching the control point system and for processing user interface events. After the control point has been started, PECo Core retrieves devices available in the current environment. The list of available devices is managed by PECo EnvManager which is to be understood as

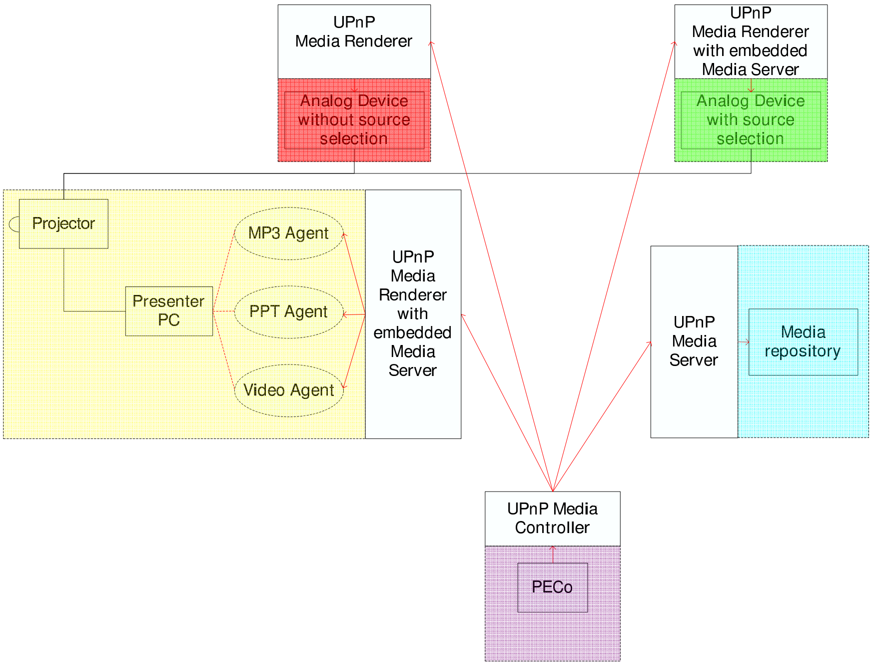

Figure 6: UPnP Architecture for AmI Meeting Room

a client stub for the Environment Manager Agent described in section 3.2.

After retrieving available devices, a Device Factory is used to create an instance for each Device. Created devices are inserted in PECo DeviceRegistry for look-up purposes

A UI Command is encapsulated by a specific class which contains a reference to the current Environment, to the Device an Action could be performed on and the Action itself as well as optional parameters.

PECo CommManager provides high-level communication means for various network and bus systems. Devices can use CommManager to get a channel regardless what kind of physical network (GPRD, Bluetooth, WiFi) currently is available.

PECo EnvManager manages meta-data describing the room. It contains and manages Devices available in an environment. It also provides access to 3D representations for the room as well as for each device.

Devices manage their own capability and property profiles and provide means to perform Actions supported by them. Each Action object provides an execute() method. This method performs the action physically on the device it is assigned to. For example, Action ON (a class derived from Action) for the Device Beamer sends the corresponding UPnP SOAP event to the desired UPnP Service of the device. When PECo Core receives a User Interface Event to process, it retrieves the corresponding device instance from the lookup table. Then it leaves the Device object to execute the required command on the physical device. This is done by calling the execute() command on the desired Action of the selected Device.

5. Conclusion

5.1 Contributions to the presented work

In this paper, we argued that in order to support the user in intelligent multimedia workspaces intuitive interaction means for Media Management and Environment Control are required. We proposed an automatically created 3D-based user interface to meet the required intuitiveness. The synchronized 3D view of the user's real physical environment provides him an orientation within the physical space and allows him to manage even foreign complex environments without requiring technical knowledge about underlying infrastructure. We introduced an integrated interaction metaphor allowing access both to devices and media. We discussed that the provision of the 3D interface as well as the access to foreign environments require location-aware device discovery [1] and standardized control mechanisms. We also discussed requirements for the provision of single-point of access to personal media. For meeting identified requirements a generic UPnP Presentation architecture has been introduces. As an important benefit, the presented architectures provides full UPnP mechanisms (device discovery; access to device profiles, states and services; media access) thus serving as a development basic for further AmI research (e.g. reactive agent systems) and deploying third party interaction solutions. As shown in Table 2 and Figure 6, we have presented the following generic UPnP architecture for meeting rooms with extensions to existing AV Architecture:

5.2 Related Works

InfoPoint [12] is a hand-held device that allows appliances to work together over a network. It provides a unified interface that gives different types of appliances drag-and-drop-like behavior for the transfer of data. Moreover, it can transfer data from/to non-appliances such as pieces of paper. As a result, InfoPoint allows appliances to work together, in the real world environment, in terms of data transfer. However, the focus of InfoPoint is not to provide the user with intuitive facilities for direct-manipulation of his physical environment, as it is the major goal of PECo. It can not control appliances, such as turning a light on or putting a jalousie down. Nevertheless, both systems, PECo and InfoPoint, allow the transfer of media between appliances, e.g. to drag and drop a presentation from a smart phone to a beamer. However, in order to achieve this, InfoPoint requires special hardware and a 2D marker attached to each device that should be accessible through it. InfoPoint also does not support the user to intuitively identify which devices of his environment are accessible. In contrast, PECo is based on of-the-shelf appliances and allows to easily identify accessible devices by providing its novel UI approach. This specially is helpful when the user is interacting in a complex environment for the first time. The Phillips ProntoPro Next Generation [13] provides a dynamic, digital user interface for directly accessing devices through RF or InfraRed. It allows customizing the ProntoPro NG by adding, deleting, and labelling new buttons or menus for new devices. ProntoPro NG can control virtually any infrared (IR) device. It comes with pre-installed remote control codes for more than 800 brands. However, also the Phillips ProntoPro NG focuses on smart home environments and uses a WIMP based interaction mataphor to allow device access. For controlling a device, it requires explicite knowledge about device names. For example, in order to be able to turn on one of three TVs available in a room, the user still needs to know the ID which represents the specific TV within ProntoPro NG. The iPronto system comes with a special handheld and is not available for arbitrary handheld devices such a PocketPC 2003. In contrast, PECo runs on every PocketPC device.

5.3 Discussing the Interaction Approach

Allowing the user to easily access virtual and physical objects surrounding him is the goal of PECo. There exist several solutions for providing access to physical environment and to the user's personal media: The Personal Universal Controller (PUC) [14], Universal Information Appliance (UIA) [15] and The Information Furnace [?] are examples . . . However, most of them have drawbacks for varying reasons. In contrast, PECo solves common problems of existing systems by:

1. Intuitive access. PECo allows to intuitively access complex infrastructures of the physical world based on the 3D-visualization of the user's environment. The 3D visualization makes a direct link between physical devices and their virtual representation within the 3D model. By doing so, the user is able to access devices based on their position, orientation within the room and their look. He does not need to have knowledge about special device names, IP numbers or device paths.
2. Unified access. The user interacts with his physical world in the same manner he also interacts with his documents and virtual objects (via mouse clicks and drag&drop). By doing so, we are extending the metaphor of direct manipulation - which is known from the desktop world - to the real world.
3. Media Management. PECo provides uniformed access to the personal media space by means of the Personal Media Manager Agent.

Using PECo the user can just drag a presentation from his media repositories and drop it to a physical display of his real world by just one interaction step (drag&drop). without interrupting the flow of his main activities.

6. References

[1] A. A. N. Shirehjini, A novel interaction metaphor for personal environment control: Direct manipulation of physical environment based on 3d visualization, in: Computers & Graphics, Special Issue on Pervasive Computing and Ambient Intelligence, Vol. 28, Elsevier Science, 2004, pp. 667–675.

[2] E. Aarts, R. Harwig, M. Schuurmans, Ambient intelligence, McGraw-Hill, Inc., New York, NY, USA, 2002, pp. 235–250.

[3] J. L. Encarnação, T. Kirste, Ambient intelligence: Towards smart appliance ensembles, in: From Integrated Publication and Informations Systems to Virtual Information and Knowledge Environments, Springer, 2005. http://www.informatik.uni-rostock.de/mmis/paper.pdf

[4] T. Kirste, Smart environments and self-organizing appliance ensembles, in: In Aarts E, Encarnação J L (eds): True Visions, Springer, 2005.

[5] A. A. N. Shirehjini, Peco: 3d-based interaction with a upnp meeting room, OCG, Austria, 2005.

[6] A. A. N. Shirehjini, M. Hellenschmidt, T. Kirste, An integrated user interface providing unified access to intelligent environments and personal media, in: Proceedings of the 2nd European Union symposium on Ambient intelligence, ACM Press, 2004, pp. 65–68.

[7] T. Kirste, Ambient intelligence and appliance ensembles: Architectural considerations, in: Conference on Human Factors in Computing Systems (CHI), Netherlands, 2004.

[8] A. A. N. Shirehjini, Ein umgebungsmodel für situierte interaktion in ami meetingsräumen, report Number: 05i003-FIGD (April 2005).

[9] EIB, Eib home server (2004). http://sourceforge.net/projects/eibcontrol/

[10] M. Renderer, Designing a upnp av media renderer., white Paper 5-20-2003 (2004).

[11] M. Server, Designing a upnp av media server., white Paper 6-25-2002 (2004).

[12] J. R. Naohiko Kohtake, Y. Anzai, Infopoint: A device that provides a uniform user interface to allow appliances to work together over a network, Personal and Ubiquitous Computing 5 (4) (2001) 264–274. http://citeseer.nj.nec.com/542270.html

[13] iProntoPro, Prontopro next generation - home theater control panel (2004). http://www.remotecontrol.philips.com

[14] J. Nichols, B. A. Myers, M. Higgins, J. Hughes, T. K. Harris, R. Rosenfeld, M. Pignol, Generating remote control interfaces for complex appliances, in: Proceedings of the 15th annual ACM symposium on User interface software and technology, ACM Press, 2002, pp. 161–170.

[15] K. F. Eustice, T. J. Lehman, A. Morales, M. C. Munson, S. Edlund, M. Guillen, A universal information appliance, IBM Systems Journal 38 (4) (1999) 575–601. http://www.research.ibm.com/journal/sj/384/eustice.pdf