ActionSpaces: Device Independent Places of Thought, Memory and Evolution

Rudolf Melcher, Martin Hitz, and Gerhard Leitner

University of Klagenfurt, Universitätsstraße 65-67,
9020 Klagenfurt, Austria
{rudolf.melcher,martin.hitz,gerd.leitner}@uni-klu.ac.at

Abstract. We propose an inherently three-dimensional interaction paradigm which allows individuals to manage their personal digital artifact collections (PAC) regardless of the specific devices and means they are using. The core of our solution is to provide unified access to all user artifacts normally spread across several repositories and devices. Not till then individuals may foster and evolve persistent multi-hierarchical artifact structures (PAS) fitting their cognitive needs. PAS subsets can be arranged and meaningfully related to virtual habitats or even mapped to physical contexts and environments they are frequenting to solve their tasks.

Keywords: 3DUI, interaction paradigm, semantic desktop metaphor, ubiquitous computing, distributed computing, distributed cognition, mixed realities, concept maps, virtual file system, post-WIMP, post-desktop, digital artifacts, information space.

1 Introduction

In our work we emphasize the need for a paradigm change as a direct consequence of complex personal device infrastructures. We therefore try to determine the minimum characteristics facilitating device-independent interaction in the face of virtual artifacts (i.e. digital entities like files, folders, repositories, emails, contacts, appointments, web pages, database-views, services views, rendered objects, etc.) Virtual artifact management is a difficult cognitive task on its own at least with regard to artifact classification. But now we have to discuss additionally where multi-device usage will finally lead to and how its potentials may be leveraged. In our opinion techniques subsumed as *cloud computing* are not sufficient. We recognize a demand for user-centric, device-independent, and non-hierarchical structures carrying the artifacts of individuals persistently while being ubiquitously accessible. Our approach is to determine an adequate set of concepts and methods for so-called workspace-level integration [1]. The driving vision and assumption is that in the course of time ubiquitous augmented realities (UARs) related to user's cognitive panoramas will succeed.

In terms of system architecture we propose two new layers reorganizing the artifact management between hardware and operating systems on the one hand and applications on the other hand. For all that, the feature sets and capabilities specific to distinct appliances shall remain unlimited.

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 355–364, 2009.

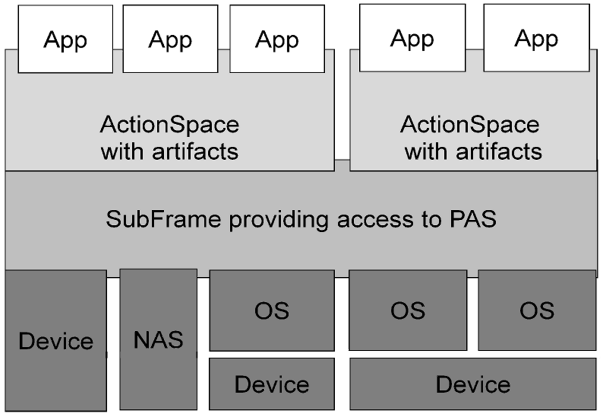

Fig. 1. Four layer approach to facilitate device-independent interaction

First, we outline the middleware layer called *SubFrame* which masks the specific file handling mechanisms and unifies the artifact references across an individual's heterogeneous device infrastructure. Persistent user-specific structures for virtual artifacts (PAS) are accessed, manipulated, and stored on this basis.

Second, subsets of PAS are embedded and presented in 3D work spaces called *ActionSpaces*. We define *ActionSpaces* as session-persistent geometrical areas populated by virtual artifacts according to an individual's concept map or mental map. An *ActionSpace* may be rendered at arbitrary points on Milgram's reality-virtuality continuum [2]. Accordingly, it is an extension and virtualization of workplaces and habitats normally experienced in the physical world with a view to tool and artifact arrangements. *ActionSpaces* are more flexible than virtual desktops. Current desktop implementations are bound to screen characteristics and represent only a small fraction of possible manifestations along Milgram's continuum. In this sense, *ActionSpaces* manifold virtual desktop scenarios on common PC-systems as well.

There seems to be a firm conviction in the HCI community that 3DUI technology is only of use in complex and expensive visualization scenarios and probably will never penetrate common workplaces, with the exception of gaming. In opposition to that we are convinced, that inherently three-dimensional work-spaces represent an important step towards ubiquitous augmented realities (UARs) [3]. We found two intriguing arguments for this. On the one hand 3DUIs are the most comprehending virtual hull to arrange and render all kinds of virtual artifacts such as text, image, video, sound, animated 3D graphics, etc. and to express relations between them. On the other hand virtual three-dimensional spaces can be mapped to physical locations without the slightest effort, provided appropriate paradigms are applied. This, in turn, is an important foundation for pervasive computing scenarios and location based services.

With the two layers proposed we follow a top-down approach in order to develop a device-independent thought-pattern for interaction.

2 Related Work

The research assignment to specify a device-independent and user-centric interaction paradigm is related to many distinct research areas partially summarized here. This multi-disciplinary interdependence, the need to develop a basic understanding of the heterogeneous issues related and the difficulty to specify formal models makes it difficult to prove the assumptions, potentials and consequences of our approach in this early stage. The potentials and consequences have to be discussed separately and investigated issue by issue in view of every distinct area after a complete prototypical framework is available. For now we want to provoke an interdisciplinary discussion about its need.

Cognitive Models and Semantic Desktops

As Winograd and Flores stated "we recognize that in designing tools we are designing ways of beeing!"[4] There is a lot of empirical evidence for the defects of the desktop metaphor discussed closely in [1]. Regarding possible workspace-level implementations for non-hierarchical artifact structures we can build on the semantic desktop paradigm defined by Sauermann et al.: *"A Semantic Desktop is a device in which an individual stores all her digital information like documents, multimedia and messages. These are interpreted as Semantic Web resources, each is identified by a Uniform Resource Identifier (URI) and all data is accessible and queryable as RDF graph. Resources from the web can be stored and authored content can be shared with others. Ontologies allow the user to express personal mental models and form the semantic glue interconnecting information and systems. Applications respect this and store, read and communicate via ontologies and Semantic Web protocols. The Semantic Desktop is an enlarged supplement to the user's memory."* [5] However, by some means or other these semantic desktop solutions remain application-, platform-, or device-dependent. These dependences are predetermined breaking points for the whole paradigm. The user-centrism aimed is not consistently portable and the visualization of ontologies remains problematic.

3DUI, MR, UARs

Kirsh argues that „how we manage the spatial arrangement of items around us is not an afterthought: it is an integral part of the way we think, plan, and behave."[6] Gregory Newby argues that in exosomatic memory systems, the information spaces (geometry) of systems will be consistent with the cognitive spaces of their human users. [7] Following this discussion we conclude that 3D user interfaces (3DUIs) summarized by Bowman et al. [8] and AR interfaces discussed by Haller et al. [9] can fit these requirements. Issues related to the tracking and rendering performance on common user platforms were striking problems in the last decades. Now, the major hindrances are caused by application, platform and device dependences. As Sandor and Klinker conclude "Increasingly, an overarching approach towards building what we call ubiquitous augmented reality (UAR) user interfaces that include all of the just mentioned concepts will be required" [3]. At the end of the day "user interfaces for ubiquitous augmented reality incorporate a wide variety of concepts such as multi-modal, multi-user, multi-device aspects and new input/output devices." [10]

3 SubFrame

Our ultimate goal is to establish true device-independence while supporting consistent artifact management across the whole range of user devices and appliances. We therefore propose a middleware layer called *SubFrame* which we introduced in [11]. We assume the existence of such a layer for the discussion of *ActionSpaces*. For sake of completeness, we will briefly address it here.

The *SubFrame* holds and provides a complete set of references to the user's artifacts. It is, however, neither responsible for nor involved in the rendering of artifacts. The terms "personal information cloud" and "cloud computing" immediately coming to mind are very popular but fuzzy. While clearly related to our work we need to distinguish from them as they are disputed with regard to service orientation and we discuss user-centered artifact management here. Throughout this paper we use the term *personal artifact collection* (PAC) being more precise regarding artifact sets and *personal artifact structures* (PAS) being more precise regarding their inter-relationships. The whole concept is based on the following set of key assumptions:

Table 1. Key assumptions of device-independence

0	Devices are networked (temporarily)
1	PACs don't belong to devices.
2	PACs don't belong to applications or services.
3	PACs are private and non-substitutable.
4	PACs are partially shared.
5	PAS are unique in their characteristics.
6	PAS are ever-changing and evolving.

In [11] we suggest employing a *personal proxy server* (PPS) for the purpose of sub framing. The PPS holds a central instance of the user's PAS. Every single user appliance has to request the artifacts through the PPS. Even – and that is very important – local artifacts have to be requested through the PPS (loopback) to enable classification and temporal tracing. Despite evident redundancies this is maintainable, as artifacts are handled by reference and not by containment. In a second step this approach may also be used to implement transparent backup mechanisms.

The consequence of workspace-level integration yields the ground rule not to foster proprietary solutions. For prototyping we use Squid proxy servers and we investigate how the RDF format may be used in conjunction to express and annotate non-hierarchical personal artifact structures.

At the end of the day this layer implements the five functional requirements for personal information management defined by Ravasio et al. [12], which are: unified handling, multiple classification, bi-directional links, tracing of documents' temporal evolution, and transparent physical location of a particular piece of information. Finally, PAC and PAS together represent an instance of the personal information cloud, a kind of exo-somatic memory [7] at our disposal.

4 ActionSpaces

Now let us assume individuals have unobstructed access to every single digital artifact in every collection they own or share with others. We have to identify the fundamental building blocks needed to make them (re-)presentable and manageable on every device, every interface, in every virtual environment we can think of.

On typical WIMP (windows, icons, menus, and pointer) platforms application windows are used to render one or even more documents simultaneously. File manager applications of all kinds and proprietary file dialogs are playing an outstanding role on these platforms. They provide hierarchical artifact access and classification means. But this powerful technique cannot be implemented consistently and efficiently across the heterogeneous device infrastructure and Milgram's reality-virtuality continuum [2]. Even on the historical target platforms for desktop computing we face a lot of problems in the view of cognitive load, efficiency and consistency [1]. Thus we have to find an alternative better suitable for our needs and consistently implementable.

We propose to rely on the principle of abstraction since it is common sense that a monolithic "universal interface" is neither feasible nor eligible in terms of specification, implementation, and usage. We quest for a user-centric thought pattern, where all possible interactive features and interface implementations fit in. This thought pattern will allow for smooth transitions with minimal cognitive loads across the user's device infrastructure.

In the year 2000 Dachselt suggested a metaphorical approach called action spaces to structure three-dimensional user interfaces. He argues, that "new applications will be built in the near future, where the focus does not lie on navigation through more or less realistic worlds, but rather on 3D objects as documents in an interactive three-dimensional user interface."[13] He defines action spaces as task-oriented scenes of actions or more precisely as virtual 3D spaces with interface controls serving an associated task. Noteworthy he points out that action spaces do not have to be rooms in a geometric sense and "there is a need for the integration of these spaces in a more general visual application framework, in a geometric and metaphorical structure". [13]

Based on Dachselt's work we reformulate and generalize *ActionSpaces* as session-persistent mixed-reality areas (geometry and location) populated by virtual artifacts (PAC) which are rendered according to the individual's mental models (PAS). According to this definition *ActionSpaces* are a kind of generalized display and interaction areas. Traditional screen spaces, now part of a more generic scheme, represent only a small fraction of possible instances.

4.1 Device Independent Places

Rendered or not, an *ActionSpace* has a geometry or cubic expansion respectively. The geometric bounding box may be based on the underlying hardware (e.g. screen size), it may be specified explicitly (e.g. fish-tank VR) or its dimension may be implicitly based on a real-world subspace (e.g. desk, wall, room). For now we define the bounding box dimensions of *ActionSpaces* as:

$$AS = x \cdot y \cdot z \quad with \quad x > 0, y > 0, z \geq 0. \tag{1}$$

The point of origin and the orientation may be bound to an absolute or relative location, i.e. a point in physical/virtual space or to a physical/virtual object (e.g. by using GPS coordinates or other tracking techniques). There are several possible bindings listed in Table 2. In accordance with Feiner et al. [14] we have to carefully

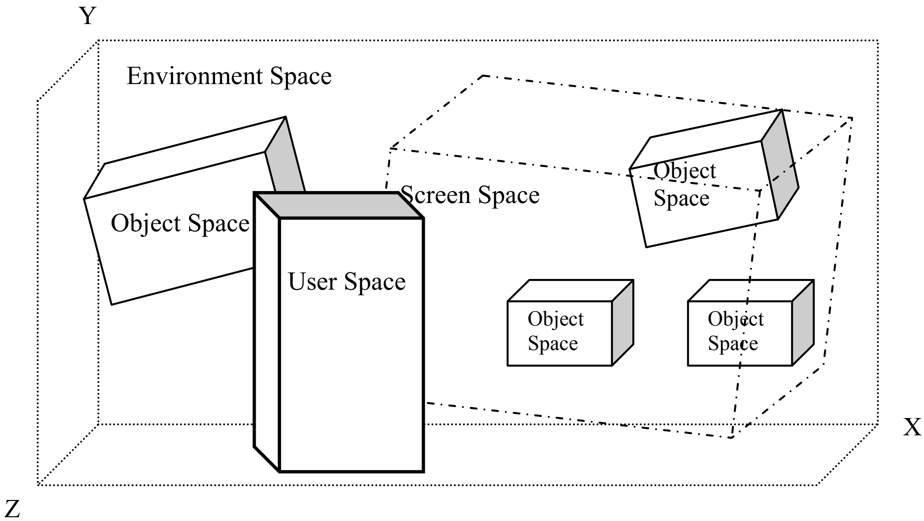

Fig. 2. Types of reference-spaces to distinguish carefully. Cubes represent the local coordinate systems involved.

distinguish environment spaces, object spaces, screen spaces, and user spaces as depicted in Figure 1.

Following Bowman et al. [15], we should establish comprehensible mappings between the different spaces in terms of user interaction and interplay between real and virtual. Therefore we discuss the characteristics of *ActionSpaces* and possible mappings across the heterogeneous device infrastructure.

Table 2. Binding examples between ActionSpaces and (physical) reference space

Space Binding	Physical Link	Parameters used in ActionSpaces
environment	GPS, Fiduciary Marker, …	position and orientation
object	screen, prop, …	object dimensions, orientation
screen	cell phone, monitor, …	screen dimensions
user	individual	perspective, distance,
dynamic	3DMouse, …	position, orientation and scale
unbound	---	determined programmatically

ActionSpaces themselves may be rendered explicitly as a composition of virtual objects (i.e. all kinds of semi-opaque virtual environments) or not (i.e. as transparent augmented realities). The examples in Table 3 give a rough impression of interface paradigms possibly consulted to render *ActionSpaces*. The attribution of entries is ambiguous because the paradigms available today were not designed with the suggested classification in mind.

In any mapping-case *ActionSpaces* serve as reference coordinate systems for the positioning and location of artifacts which always have to be rendered visually as they

Table 3. Rendering examples and paradigms for ActionSpaces

Space Binding	(Semi-)opaque	Transparent
environment	Virtual Room (Cube VR)	augmented workbench (AR), Handheld AR
object	World in Miniature (WiM), Control Panel (Virtual Prop)	Augmented engine manual (AR), Marker-based AR, Tangible Computing
screen	Desktop, Today Screen	TV-Inserts, RT Video-Overlays
user	Head-Up-Display (HUD) (head-tracked VR)	AR HUD (head-tracked AR)
dynamic	Fish-tank VR, Dome VR	Pervasive Annotation Layers,
unbound	Animation, Film	RT Video

would be not accessible otherwise. Together, an *ActionSpace* and the artifacts populating it are specified in an XML based file format. Hence, *ActionSpaces* are artifacts themselves and an ActionSpace may contain other *ActionSpaces*. This is an important aspect of our concept allowing complex (i.e. non-hierarchical) relationships between artifacts and spaces and advanced information and artifact exchange between individuals. The two later show great promise for future collaboration scenarios, but they cannot be discussed here for now.

By design an *ActionSpace* may be accessed on demand everywhere and at any time (universal accessibility). We fulfill this requirement by requesting it with its URI from a *personal proxy server* (PPS). For instance, on a desktop machine the process may appear similar to the request of a free-form HTML file in a full-screen browser.

The appropriate type of space binding according to Table 3 is subject to the user's location, device and platform (interface paradigm). Even time may be used as a parameter to support pervasive scenarios. The possibilities are manifold. Until the potentials and consequences can be studied in detail, we suggest an obvious design rationale, provided that the actual platform supports it: An *ActionSpace* shall be mapped to the environment-space when its position has to be fixed persistently. It shall be mapped to an object-space when it has to be moved frequently or transported. Screen-space mappings are used for currently prevalent location-independent applications and for remote scenarios, e.g. to access virtual artifacts located on my home-office shelf. User-space mappings capturing the user's attentiveness are used to present artifacts available for interaction, activity guidance, and communication, e.g. note taking.

We now will briefly look into *ActionSpaces* and discuss the abstract characteristics and features supporting activities like thinking, memorizing and evolving.

4.2 Thought, Memory, and Evolution

Thinking involves the activity of classification. [16] Because spatial classification works well for individuals we need to provide proper means for this task.

Depending on the device or platform on-hand artifacts should be rendered directly in full-detail, or represented intermediately as 3D objects, 2D icons, and text entries. Each artifact has a well-defined position, orientation and scale relating to the

ActionSpace it is member of. To support the act of classification all three parameters are separately adjustable. Wherever possible, direct interaction shall be used to facilitate this parameterization. Automatic grouping, ordering and alignment functions shall be provided for more complex activities to be conveniently performed. The actual values shall be session-persistent, i.e. they are saved together with the artifact references in the XML file specifying the *ActionSpace*.

As already mentioned, *ActionSpaces* may contain other *ActionSpaces* in addition to artifacts since *ActionSpaces* are artifacts themselves. The possibilities cannot be exposed in detail here, but related to cognitive activities this feature yields expressions like to “change a topic”, “step through a process”, “go into detail”, “get an overview”, and many more. That way we have plenty of means to browse multi-hierarchical structures in relation to physical and virtual contexts now.

By design an artifact is always created inside an *ActionSpace* or imported from another one. Interaction techniques akin to drag-and-drop (pointer) and linguistic interaction (speech recognition, command line interface) are used for these tasks so that position, orientation, and scale are implicitly defined at first. Complex manipulation of “opened” artifacts is – aside from some advanced platforms and paradigms – still the prevalent domain of applications. The application layer is not within the scope of this work.

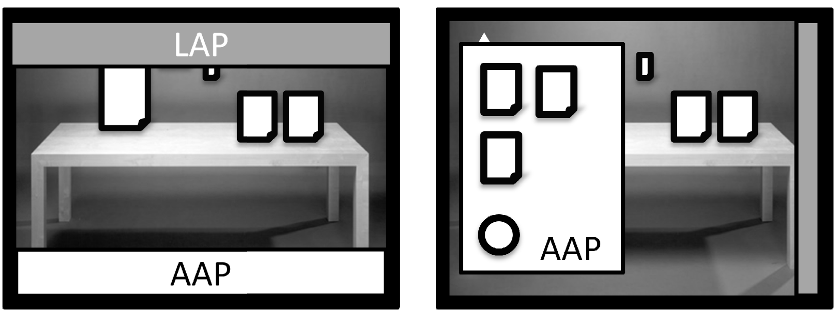

Fig. 3. At least three ActionSpaces are active with two of them bound to screen space (LAP and AAP) and one of them bound to the physical environment space as actual workspace

As depicted in Figure 3 there are at least three *ActionSpaces* active, even if not visible, at any time. Two of them are special instances and are always bound to the screen space or the user space respectively. Their position and size depends on user preferences. The first one represents the device-bound local artifact pool (LAP) and makes its content accessible. It may for example contain several image files recently shot by the user and not yet integrated (i.e. not classified) in the PAS. LAPs are the counterpart to the prevalent file managers. The second one represents the ambient artifact pool (AAP), i.e. the environmental resources accessible in a given context. It makes their artifacts available (e.g. thumb drives content) and allows sending artifacts to them (e.g. printer queue). In advanced mixed reality scenarios the “real device” and its artifacts may be directly addressed at their physical locations, by dragging virtual text documents and dropping them onto the printer with gestures. In most paradigms an iconographic representation is still essential to make AAPs handling convenient.

Structural persistence is guaranteed across the device infrastructure. Geometric arrangements related to gestalt principles are kept persistent or they are emulated to support individual recognition. For instance, if an *ActionSpace* bound to an environment space is accessed remotely, the geometric proportions and distances remain valid. But they will probably be scaled to establish comprehensible viewpoints and perspectives. Even if the artifacts contained in such an *ActionSpace* are listed in pure textual form, the original positions will be preserved. Hence, the content arrangements in an *ActionSpace* may be temporarily reorganized using non-persistent automatic modes or they may be manually and persistently rearranged by the user.

The *personal proxy server* (PPS) allows the implementation of transparent backup strategies and encrypted artifact pools. It will be of great importance for individuals that with our approach every single artifact is accessible, and in terms of device-dependence nothing gets lost accidentally. In fact we work towards lifelong solutions for artifact management and thus lifelong exosomatic memories. For instance, the damage of one's laptop yields no loss of artifacts since all *ActionSpaces* and PAS arrangements are still available from the PPS. Transparent backups and versioning make up for a potential loss of local copies. If an artifact is not findable via navigating the PAS it may be still found with common search techniques applied on the PPS. That way and not surprising search is still complementing the structural access.

5 Conclusions

Every tool/device/appliance has its own strength and weaknesses. For that reason users want to interact with more than one. The availability of devices in different contexts (e.g. tasks, times and locations) is another reason. In the quest for device independence we are not well advised to seek for the intersection (i.e. greatest common divisor) of feature sets. With the widely discussed problems of current interaction paradigms and the steadily growing diversity of electronic appliances in mind we try to determine a minimal set of interaction means unifying artifact handling in a device-independent manner. That way the cognitive load for users can be minimized and the potentials of heterogeneous infrastructures may be leveraged.

Both layers proposed – the *SubFrame* and the *ActionSpaces* – challenge the hierarchical file systems across the user's device infrastructure and the dated desktop metaphor on PC devices. With both layers consequently realized, we will see new possibilities in artifact handling and sharing which will meet the cognitive conditions and needs of individuals. The concrete representation of artifacts, interrelations, and *ActionSpaces* depends on the individual's needs and flavors. Therefore it could not be specified and, in view of design style guides, it never will. Following the argumentation of Ravasio and Tscherter, every single *ActionSpace* "should be a place where users, and only users, are able to engrave personal preferences and tastes." [1]

Acknowledgments. The authors would like to thank Prof. R. Mittermeir at the University of Klagenfurt for his constant input and challenging discussions and our student Bonifaz Kaufmann for the brave development of prototypical solutions.

References

1. Kaptelinin, V., Czerwinski, M.: Beyond the Desktop Metaphor. MIT Press, Cambridge (2007)
2. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: A class of displays on the reality-virtuality continuum. In: SPIE, Telemanipulator and Telepresence Technologies, vol. 2351, pp. 282–292 (1994)
3. Sandor, C., Klinker, G.: A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. Personal Ubiquitous Comput. 9(3), 169–185 (2005)
4. Winograd, T., Flores, F. (eds.): Understanding computers and cognition. Ablex Publishing Corp., Norwood (1985)
5. Sauermann, L., Bernardi, A., Dengel, A.: Overview and outlook on the semantic desktop. In: Proceedings of the 1st Workshop on The Semantic Desktop at the ISWC 2005 Conference (2005)
6. Kirsh, D.: The intelligent use of space. Artificial Intelligence 73(1-2), 31–68 (1995)
7. Newby, G.B.: Cognitive space and information space. J. Am. Soc. Inf. Sci. Technol. 52(12), 1026–1048 (2001)
8. Bowman, D.A., Kruijff, E., LaViola, J.J., Poupyrev, I.: 3D User Interfaces: Theory and Practice. Pearson Education, Inc., Boston (2005)
9. Haller, M., Mark Billinghurst, B.T. (eds.): Emerging Technologies of Augmented Reality: Interfaces and Design. Idea Group Publishing, USA (2007)
10. Hilliges, O., Sandor, C., Klinker, G.: Interactive prototyping for ubiquitous augmented reality user interfaces. In: IUI 2006: Proceedings of the 11th international conference on Intelligent user interfaces, pp. 285–287. ACM, New York (2006)
11. Melcher, R.: Device-independent handling of personal artifact collections. Submitted to Interact 2009 (2009)
12. Ravasio, P., Vukelja, L., Rivera, G., Norrie, M.C.: Project infospace: From information managing to information representation. In: HumanComputerInteraction Interact 2003, 8092, Swiss Federal Institute of Technology, Zurich, Switzerland, pp. 864–867. IOS Press, Amsterdam (2003)
13. Dachselt, R.: Action spaces - a metaphorical concept to support navigation and interaction in 3d interfaces. In: Proceedings of 'Usability Centred Design and Evaluation of Virtual 3D Environments, April 13-14, 2000, Shaker Verlag, Aachen (2000)
14. Feiner, S., MacIntyre, B., Haupt, M., Solomon, E.: Windows on the world: 2d windows for 3d augmented reality. In: UIST 1993: Proceedings of the 6th annual ACM symposium on User interface software and technology, pp. 145–155. ACM, New York (1993)
15. Bowman, D.A., North, C., Chen, J., Polys, N.F., Pyla, P.S., Yilmaz, U.: Information-rich virtual environments: theory, tools, and research agenda. In: VRST 2003: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 81–90. ACM Press, New York (2003)
16. Bowker, G.C., Star, S.L.: Sorting Things Out: Classification and Its Consequences (Inside Technology). The MIT Press, Cambridge (1999)