This is a preprint of the paper to be presented at the 1st Immersive Learning Research Network Conference (iLRN), Prague, Czech Republic, July 13–14, 2015. The final version will be published by IOS Press.

An Overview of Capturing Live Experience with Virtual and Augmented Reality

Mikhail FOMINYKH [a,b], Fridolin WILD [c] Carl SMITH [d], Victor ALVAREZ [e], and Mikhail MOROZOV [b]

[a] *Europlan UK Ltd, London, UK*
[b] *Volga State University of Technology, Multimedia Systems Lab, Yoshkar-Ola, Russia*
[c] *The Open University, Knowledge Media Institute, Milton Keynes, UK*
[d] *Ravensbourne, Learning Technology Research Centre, London, UK*
[e] *Murdoch University, Chancellery Learning and Teaching, Perth, Australia*

Abstract. In this paper, we review the use of virtual and augmented reality technologies for capture and externalization of tacit knowledge from complex activity in knowledge-intensive professions. We focus on technologies for converting experience hidden in activity with the aim to boost industry competitiveness, innovation, and facilitate learning on the job. As such types of knowledge and experience are difficult to capture and represent in traditional media, we explore emerging technology along two lines of investigation. First, we look at applications of virtual reality to then, second, focus on using sensors, augmented reality, and wearable technologies. We discuss existing and future applications of experience capturing with virtual and augmented reality technologies. This review provides a comprehensive overview for those interested in recording virtual, real, and augmented activities, methods for delivering the recorded data, and extracting knowledge.

Keywords. Virtual Reality, Augmented Reality, Wearables, Recording, Tacit Knowledge, Professional Development.

1. Introduction

Competence is a potential for action, and human activity is the process leading to competence development. To convert experience to knowledge and enable learning, the information relevant for a master level of performance (aka 'theory'), traditionally, is separated from the immediate experience of competent action (aka 'practice'). In the past key roles in this conversion have been held by skilled teachers using direct instruction as well as by authors producing media such as textbooks or instructional films, to name but a few. With technical advances and with growing access to wearables, however, new opportunities arise for capturing, sharing, and re-enactment that do not rely on strict separation of knowledge from its application, also giving more room to new roles in knowledge production beyond the specialist curator.

Capturing activity at the right level of abstraction, while still retaining all relevant detail is a complex task, not only from the technological point of view, as a significant part of the actual experience has to be subducted in the process. Conventional approaches such as video recording or screen capturing provide only limited points of view, significantly reducing the wealth of information available from direct experience as well as its immediacy and authenticity. These technologies map activities in the real world to flattened representations with limited affordances. Learning needs more than that: its feedback loop character requires an iterative cycle of invention, observation, reflection and action/revision [1]. The first two stages are concerned with how contextualized information is measured, captured, stored and relayed to the individual, whereas the last two create a pathway to judgment and decision making [2].Virtual Reality (VR) and Augmented Reality (AR) technologies provide perspectives that may help overcome present reductionist limitations by implementing more open loops.

VR simulates spaces, objects, humans, and activities that often reproduce a precise image of the reality [3]. AR increasingly accommodates innovative devices and technologies, while at the same time leveraging our understanding of human-computer symbiosis by enhancing sensory feedback and enabling interaction with augmented information based on actions in the physical world [4].

The objective of the work presented in this paper is to survey the state of the art and assess the potential for VR, AR, and wearable technologies to capture experience and deliver recorded data directly to the user, looking at whether and how well re-enactment of a recording allows re-living of experience. In order to fulfil this objective, we describe work conducted within research projects and commercial platforms, identifying issues and challenges and characterize the approaches taken so far to overcome them. Prior reviews of the state of the art in using VR and AR do not focus in depth on experience capturing but provide insight into their application in general, falling short of inspecting facilities and approaches required for recording. This review consolidates existing work from many sources and relates them in a framework for capturing experiences as an overview of the field.

The rest of this contribution is organized as follows. In Section 2, we provide a theoretical background of activity theory, tacit knowledge, and workplace learning. In addition, we introduce a characterization framework to structure the work. In section 3, we present an overview of the most representative examples of capturing *virtual* activities, while Section 4 presents a review of AR applications for capturing *real-life* activities and extracting knowledge from them. Section 5 then condenses findings in an overview and discusses the state of the field. Finally, in Section 6, we describe future challenges that require further work and research.

2. Experience Capturing and Knowledge Conversion

From an epistemological perspective, knowledge resides in and is accessible from community memory repositories. Knowledge appears in many forms. Tangible knowledge may be stored as written instructions or in databases. Intangible knowledge appears as activity, practice, relation between participants, and in their shared experience. The former type of knowledge is known as explicit and the latter as tacit [5]. In education and professional development, the two types of knowledge are mutually dependent and constituting [6, 7]. Tacit knowledge has been recognized to be important for industry competitiveness, innovation, and workplace learning [7].

Tacit knowledge can be converted to explicit knowledge through narratives – in addition to iterative training that aim to create embodied experience through activity. Capturing tacit knowledge from activity facilitates sharing and transfer, giving rise to new applications such as the accurate preservation of activity (also registering involved methods and procedures), retrospective analysis and reflection on action, and scripted re-enactment in related trainings [8].

Examples for rich experiences include safety training simulators, forensic recreation of crime scenes, task sharing from expert to novice at the workplace, or training of operation or maintenance of advanced machinery. Executing such activity typically requires specialist knowledge in a particular domain and a high level of skill to achieve a master level of performance. It is evident that capturing 'rich' experiences from such 'fluid' containers poses a significant challenge, not only to traditional repositories. Still, the knowledge they carry is essential for many high-skill professions [9]. Drawing upon the work in activity theory [10], we may see activity as a primary source for the development of knowledge from experience. As knowledge containers, activities can be characterized by their narratives, the collaboration, actors and objects involved, their governing rules, etc.

With respect to Virtual and Augmented Reality applications, there are several dimensions along which systems and approaches differ significantly from each other. We identify six such key dimensions: learning affordances, perspectives & viewpoints, levels of abstraction, editing facilities, social scope, and sensors & senses (Fig 1).

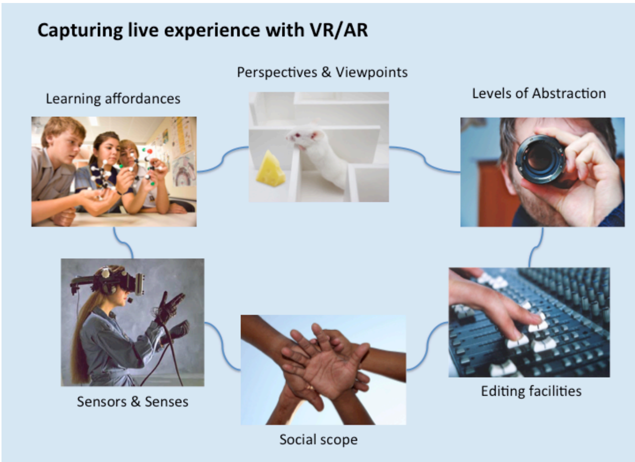

Figure 1. Dimensions for capturing VR/AR experiences

Learning affordances thereby refer to "the perceived possibilities for both thinking and doing", derived from interaction with tools or content [11]. When reviewing the state of the art in Section 3 and 4, we will highlight the difference between systems and approaches in what they afford, for example, whether users utilize them for assessment, reflection, rehearsal, guidance, regulation, or planning.

With *perspectives and viewpoints* we look closely at how much flexibility is foreseen for revisiting the recorded experience: is there support for free replay or of fixed sequences? Can roles be changed? What about visual aspects such as viewing angles or support for cut-out views, overlays, or explosion charts?

This is related to, but independent of the supported *levels of abstraction* of the recording, turning attention to the segue from unfiltered, sensor raw-data oriented recording towards more abstract representations such as activity descriptions or annotations. While any recording renders tacit knowledge from the observed activity to explicit knowledge available in the data captured, the degree to which the essence of it becomes salient differs from case to case.

Furthermore, we will look at what *editing facilities* are provided to support conversion between such different levels of abstraction. This relates to such questions such whether it is possible to remove, add, or modify an object or person in a recording.

Under *social scope*, we investigate whether systems and approaches are primarily collaborative or single-person (or both). Moreover, we look at how sharing takes place. Thereby, collaborative activity and practice sharing are of special interest, as they play a key role in activating explication of otherwise internal thoughts and routines.

Last, but not least, we analyze, what *sensors* are deployed and which *senses* monitored. This relates to the level of reality inspected, whether observation and capturing stays merely audio-visual – or goes beyond that by, for example, capturing well-being data or machine diagnostics. Re-immersing into captured experiences using a virtual, augmented, or mixed reality environment can trigger deep learning, especially when complemented with information that otherwise would not be directly observable.

3. Capturing Virtual Experiences

The realism of VR-based systems improved significantly with advances in computer graphics. Modern desktop-free human-computer interfaces increase value and transferability of virtual experiences, especially since human cognition is closely coupled to sensory-motoric experiences [12].

The potential of VR for capturing activities and extracting value from observing from multiple points of view and perspectives was acknowledged as early as in the late 90s. MASSIVE-3 supported a mechanism called 'temporal links', which overlaid recordings of prior and present activities [13]. The CAVE Research Network had an application called Vmail which supported recording of an avatar's gestures and audio together with the surrounding environment [14].

Another example is Asynchronous Virtual Classroom (AVC) aimed at solving the problem of time-sharing in distance learning. AVC allowed a group of students to watch a video image of a certain lecture, while software agents were playing additional participants [15]. The system is designed to provide collaborative experience, as the agents acted based on the scripts of previously recorded real-live student activities (questions, answers and annotations), which defined the level of abstraction. The learning affordances of AVC are awareness of discussions in previous iterations of the same lecture and therefore – social learning.

The N*Vector project developed a technology for overcoming time-zone differences using three approaches to support annotations for collaboration in VR. These approaches included: VR-annotator – an annotation tool that allows attaching 3D VR recordings to objects; VR-mail – an email system built to work entirely in VR; VR-vcr – a streaming recorder to record all transactions that occur in a collaborative session [16]. VR-vcr allows experiencing a virtual recording from multiple perspectives and points of view. Captured experience is composed of state updates of 3D objects and tracker data for the users. VR-mail was designed to work in a CAVE environment and supported gesture tracking and sound capture. VR-annotator supports text and sound notes attached to 3D objects – another level of abstraction.

More recently, an Event Recorder feature was implemented (though then not developed further) within the Project Wonderland (aka. Open Wonderland). The Event

Recorder implements capturing and playback of ‘events’ caused by activities of users and agents in such a way that during playback a user is able to view the activities that those events caused. 'Events' are recorded into an external form that can then be replayed to all the users in the space. A specialized plugin records the states and transitions of objects that cannot be represented as events (e.g., audio conversations).

vAcademia is a functional VR system that supports virtual recording [17]. This feature allows everything to be captured in a given virtual location including the state of objects and avatars, media contents, and text and voice communication. In addition, any recorded session can be attended by a group of users. A new group can work within a recorded session (with their avatars) and record it again making multi layered documentation capturing the evolution of an activity. A set of editing tools is provided such as cutting out parts on the timeline and modifying sound streams of users. vAcademia supports gesture tracking with Kinect and displays the environment on desktop or HMD (Oculus Rift). The learning affordances of vAcademia include collaborative review and analysis of captured collaborative experience.

4. Capturing Augmented Reality Experiences

The characteristics of AR [18] are strongly related with the senses of presence, immediacy, and immersion. The nature of AR provides a return to embodiment [19] that conditions the way users interact with their surrounding physical context. Analysts predict that we are on the verge of ubiquitously adopting AR to enhance our perception and help us see, hear, and feel in new and enriched ways [20]. Mobile technology is moving us from ‘information communication’ to ‘experience communication’[21].

The first technologies that were used to capture a physical contextualized experience enabled the use of AR tracking thanks to the sensors available in handheld devices. Motion capture is a way of effectively storing and re-enacting experience enabling stylistic analysis [22]. A more complete perception of the context implies being able to sense not only the contents of that environment but also to establish a perceptually correct connection between those real-world objects and digital content. This is currently achieved using sensor-based, vision-based and hybrid tracking [23].

‘Perceptual technologies’ [24] are capable of capturing users’ mental and physiologic states using bio-signals and physiological phenomena. In The Mind-Mirror [25], a combination of EEG and AR enables a 3D visualization of the subjects own brain activity. The Transformative Technology Lab (http://transtechlab.org) uses AR in biofeedback applications to capture and transform processes in minds and bodies of their subjects that can then be experienced externally. In some cases, this results in going beyond first-hand experience to social or collective experience. HearNow (http://www.biofluent.com/hearnow) is intended to focus attention for behavior change by converting brain and heart activity into ‘immersive soundscapes’ with the goal of increasing self-awareness and consciousness.

The ability within one field of view, to be both in the world and to see yourself in it [26] is an important tool for capturing live experience with AR. Field of view technologies (FOV), such as the FlyVIZ 360 headset, transform the real time visual system of users by giving them 360-degree vision compressed to fit into a 180 degree visor. It takes 15 min for the brain to adjust before this new way of seeing is ‘accepted as normal’ [27]. The AR-brain-machine interface (AR-BMI) enables a third-person view of the real environment by tele-operating an agent robot using brain signals [28].

AR is being adopted to create new sense modalities, 'perception beyond human senses' or 'sixth senses'. AR provides new perspectives and viewpoints, allowing a color blind person to capture and convert his experience of color (http://cyborgism.wix.com/cyborg). FeelSpace enables magnetic field sensing using a belt that produces vibro-tactile notifications (http://feelspace.cogsci.uni-osnabrueck.de).

Wearable sensing technologies allow a much closer association with the user and a higher degree of freedom. The wearable sees, hears and perceives the user's physical state. Wearable health-monitoring systems have drawn the attention of medical researchers and the industry. AR headsets are used to assist amputee victims with re-training the brain to recognize and integrate a fully working limb into its mind map.

The potential impact on industry competitiveness, innovation, and workplace learning is still emerging alongside the potential use cases. The concept of Interactive Augmented Prototyping combines virtual and physical prototypes for recording and recollection of design review meetings [29]. Novel editing tools support creating high-quality video tutorials for physical tasks using AR recording demonstrations [30]. Tangible AR interfaces enable intuitive manipulation and interaction with physical objects, and are expected to find use in product design, manufacturing, assembly, industrial processes and workplace instruction.

5. Summary

Using examples from our overview, we now attempt to summarize the features of VR and AR technologies which define the characteristics of the framework (Table 1).

Table 1. State of the art analysis against the dimensions of the experience capturing framework

VR	AR
Learning affordances	
The learning affordances related to captured experience are reflection, analysis and assessment. These are thoroughly supported in packages such as vAcademia and can be recorded repeatedly resulting in both an original experience and the review experience in a single recording. Pre-recorded experience can be used as background for life actions.	AR has compelling educational and pedagogical implications, providing a medium for understanding concepts and phenomena by integrating physical and digital worlds. Learning affordances include developing skills in context, tangible manipulation and exploration, improved immersion, ubiquitous and situated learning, and facilitation of social and collaborative learning tasks.
Perspectives & Viewpoints	
Most of the existing VR systems allow generation and use of multiple perspectives and viewpoints. The recorded experience can be observed from any point of view in the virtual space. An additional role of a reviewer is common, except for systems such as AVC which support only one role in both live and recorded experience. Most of the VR systems supply free reply functionality, allowing any additional actions during re-enactment. AVC stands out as having a fixed sequence of actions (timeline), while previously captured experience overlays over the live experience.	The combination of FOV technologies, see-through displays, third-person view and ambient-awareness are radically evolving this characteristic of the framework. AR enables wearable context awareness, and a number of self-awareness / self-observation AR systems are being developed (e.g. The Mind-Mirror). Some striking results of this characteristic are enabling perceptual awareness both from inside and outside the anatomy of the human body (e.g., AR-BMI) and the ability for us to create entirely new senses (feelSpace).
Levels of Abstraction	
Various levels of abstraction are used for capturing experience in VR systems. Sensor data is captured in some VR systems (e.g., N*Vector and vAcademia) to be re-enacted in virtual simulation. Such simulation	In AR, tangible objects act as triggers representing concepts and their interrelationships. This aspect opens a broad range of interpretations within the context of a reality-virtuality continuum. AR has a

can copy the reality – close to raw data. It can also be different – high level of abstraction. For example, AVC represents experiences as activity scripts. Most systems support text, sound or graphical annotations (e.g., AVC, N*Vector/VR-annotator and vAcademia).	powerful potential of creating new body - brain maps associations (Cyborg Foundation), extend sensing capabilities (feelSpace), and reshape and transform the perception of the surrounding physical context and our own body (HearNow).
Editing facilities	
Most VR systems provide control over the re-enactment (e.g., fast-forwarding). However, facilities for editing captured VR experience are not common. vAcademia supports cutting out parts of the recorded experience on the timeline, performing new activities with the addition of scripted objects while re-enacting, and removing/modifying sound streams of the recorded users.	Although still in its early stages, AR is a potentially powerful paradigm for annotating and editing the environment as well as our mental and physical states. Democut supports creating effective tutorials for physical tasks using AR recordings and enabling audio and video analysis, automatic organization, user annotations and video editing effects.
Social Scope	
In the past, VR systems allowed single-user re-enactment of experience (N*Vector, VR-mail and CAVE Research Network, Vmail). MASSIVE-3, AVC, vAcademia, and Wonderland allow collaborative observation of re-enactments by the same or other users. Observation of pre-recorded experiences can create an impression that these experiences are live during the re-enactment – a pseudo-collaborative experience.	AR tangible and enhanced experience represents a step forward into shared experience and collective communication. The third-person view encourages the ability to go beyond first-hand experience to collective identity. The Transformative Technology Lab uses biofeedback to create contextual embodiment and collective consciousness. These features enable previously unknown levels of collective insight and social collaboration.

6. Conclusion and Outlook

Although the study of capturing live experience with VR and AR technologies is still in its infancy, the advances and innovations over recent years direct us towards new possibilities for practical applications. Creating and applying a characterization framework, we identified the key properties of past and existing experience capturing systems. Features that implement experience capturing and re-enactment are summarized in section five against the dimensions of the framework.

Features for some of the dimensions are yet to be implemented. In VR-based experience capturing, sensing can advance beyond motion tracking and voice capturing. As the level of abstraction can be controlled more easily than in AR, this criterion can be experimented upon. In AR-based experience capturing, the sensing component has advanced recently but not all the senses can be captured and re-experienced (e.g., haptics). The level of abstraction can be developed up to and including holographic representation. The perspectives and viewpoints lack peripheral vision.

Both VR and AR technologies are used in a variety of professions. Repetition of certain procedures in a safe (VR-based) environment and practicing under supervision of (AR-based) assistance are well explored and used techniques. Experience capturing and re-enactment allows additional techniques such as self-observation, review, and analysis of experience with some flexibility along the dimensions of our framework.

References

[1] J. D. Sterman, "Learning in and about complex systems," *System Dynamics Review,* **10**, (1994), 291–330.
[2] T. Goetz. Harnessing the power of feedback loops. *Wired magazine* 2011.

[3] J. Steuer, "Defining Virtual Reality: Dimensions Determining Telepresence," *Journal of Communication,* **42**, (1992), 73–93.

[4] B. Furht, *Handbook of Augmented Reality*, Springer, New York, NY, 2011.

[5] M. Polanyi, *The Tacit Dimension*, Peter Smith, Glouchester, MA, 1966.

[6] I. Nonaka and H. Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, New York, 1995.

[7] A. Lam, "Tacit Knowledge, Organizational Learning and Societal Institutions: An Integrated Framework," *Organization Studies,* **21**, (2000), 487–513.

[8] M. Fominykh, E. Prasolova-Førland, L. M. Hokstad, and M. Morozov, "Repositories of Community Memory as Visualized Activities in 3D Virtual Worlds," in *Hawaii International Conference on System Sciences*, 2014, 678–687.

[9] T. Haldin–Herrgard, "Difficulties in diffusion of tacit knowledge in organizations," *Journal of Intellectual Capital,* **1**, (2000), 357–365.

[10] Y. Engeström, "Activity theory and individual and social transformation," in *Perspectives on Activity Theory*, Cambridge University Press, Cambridge, UK, 1999, 19–38.

[11] K. Pata, "Revising the Framework of Knowledge Ecologies: How Activity Patterns Define Learning Spaces," in *Educational Social Software for Context-Aware Learning: Collaborative Methods and Human Interaction*, IGI Global, Hershey, PA, USA, 2010, 241–267.

[12] L. W. Barsalou, "Grounded cognition," *Annual Review of Psychology,* **59**, (2008), 617–645.

[13] C. Greenhalgh, M. Flintham, J. Purbrick, and S. Benford, "Applications of Temporal Links: Recording and Replaying Virtual Environments," in *Virtual Reality*, 2002, 101–108.

[14] J. Leigh, M. D. Ali, S. Bailey, A. Banerjee, P. Banerjee, K. Curry, J. Curtis, F. Dech, B. Dodds, I. Foster, S. Fraser, K. Ganeshan, D. Glen, R. Grossman, Y. Heil, J. Hicks, A. D. Hudson, T. Imai, M. A. Khan, A. Kapoor, R. V. Kenyon, K. Park, B. Parod, P. J. Rajlich, M. Rasmussen, M. Rawlings, D. Robertson, S. Thongrong, R. J. Stein, S. Tuecke, H. Wallach, H. Y. Wong, and G. Wheless, "A Review of Tele-Immersive Applications in the CAVE Research Network," in *Virtual Reality*, 1999, 180–187.

[15] K. Matsuura, H. Ogata, and Y. Yano, "Agent-based Asynchronous Virtual Classroom," in *Computers in Education*, 1999, 133–140.

[16] T. Imai, Z. Qiu, S. Behara, S. Tachi, T. Aoyama, A. Johnson, and J. Leigh, "Overcoming Time-Zone Differences and Time Management Problems with Tele-Immersion," in *Internet Society Conference*, 2000.

[17] M. Morozov, A. Gerasimov, and M. Fominykh, "vAcademia – Educational Virtual World with 3D Recording," in *Cyberworlds*, 2012, 199–206.

[18] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *Computer Graphics and Applications,* **21**, (2001), 34–47.

[19] K. P. Fishkin, "A taxonomy for and analysis of tangible interfaces," *Personal Ubiquitous Comput.,* **8**, (2004), 347–358.

[20] D. W. F. van Krevelen and R. Poelman, "A Survey of Augmented Reality Technologies, Applications and Limitations," *The International Journal of Virtual Reality,* **9**, (2010), 1–20.

[21] A. D. Cheok, "Multi modal sensory human communication in the Internet society," in *Mixed and Augmented Reality*, 2011, 1–1.

[22] D. Vlasic, R. Adelsberger, G. Vannucci, J. Barnwell, M. Gross, W. Matusik, and J. Popović, "Practical motion capture in everyday surroundings," *ACM Transactions on Graphics,* **26**, (2007), 35.

[23] Z. Feng, H. B. L. Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR," in *Mixed and Augmented Reality*, 2008, 193–202.

[24] R. Stiefelhagen, "Perceptual Technologies: Analyzing the Who, What, Where of Human Interaction," in *Computers in the Human Interaction Loop*, Springer, 2009, 9–10.

[25] J. Mercier-Ganady, F. Lotte, E. Loup-escande, M. Marchal, and A. Lecuyer, "The Mind-Mirror: See your brain in action in your head using EEG and augmented reality," in *Virtual Reality*, 2014, 33–38.

[26] E. J. Gibson and E. J. Gibson, "The perceived self," in *Ontogenesis of the perceived self*, Cambridge University Press, Cambridge, UK, 1994, 25–42.

[27] J. Ardouin, A. Lécuyer, M. Marchal, C. Riant, and E. Marchand, "FlyVIZ: a novel display device to provide humans with 360° vision by coupling catadioptric camera with hmd," in *18th symposium on Virtual reality software and technology*, 2012, 41–44.

[28] K. Kansaku, N. Hata, and K. Takano, "My thoughts through a robot's eyes: An augmented reality-brain–machine interface," *Neuroscience Research,* **66**, (2010), 219–222.

[29] J. Verlinden, I. Horváth, and T.-J. Nam, "Recording augmented reality experiences to capture design reviews," *International Journal on Interactive Design and Manufacturing,* **3**, (2009), 189–200.

[30] P.-Y. Chi, J. Liu, J. Linder, M. Dontcheva, W. Li, and B. Hartmann, "DemoCut: generating concise instructional videos for physical demonstrations," in *26th annual symposium on User interface software and technology*, 2013, 141–150.