An Intelligent Multimodal Mixed Reality Real-Time Strategy Game

Sascha Link Berit Barkschat Chris Zimmerer Martin Fischbach Dennis Wiebusch
Jean-Luc Lugrin Marc Erich Latoschik*

Human-Computer Interaction
University of Würzburg

Abstract

This paper presents a mixed reality tabletop role-playing game with a novel combination of interaction styles and gameplay mechanics. Our contribution extends previous approaches by abandoning the traditional turn-based gameplay in favor of simultaneous real-time interaction. The increased cognitive and physical load during the simultaneous control of multiple game characters is counteracted by two features: First, certain game characters are equipped with AI-driven capabilities to become semi-autonomous virtual agents. Second, (groups of) these agents can be instructed by high-level commands via a multimodal—speech and gesture—interface.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—;

1 Introduction

Digital interactive surfaces augment traditional physical setups (i.e., pen & paper) with virtual elements and touch-based interactions to create a Mixed Reality (MR). Such environments overcome the static nature of traditional multi-user tabletop games [7]. Still, they preserve fundamental characteristics of these games, i.e., haptic interaction using physical game elements (cards and playing pieces), a co-located space that maintains mutual accessible frames of reference, and the rich variety and subtleties of interpersonal communication signals for social interactions [6].

We introduce a novel combination of interaction styles and gameplay mechanics for MR-based tabletop role-playing games (RPGs). To the best of our knowledge it is the first approach which 1) abandons the traditional turn-based style in favor of an increased simultaneous real-time interaction 2) based on intelligent virtual agents 3) controlled via a multimodal—speech and gesture—interface. Such interfaces provide promising interaction alternatives given the dynamic nature of many applications in MR as well as in Virtual and Augmented Reality (VR and AR).

2 Related Work

Previous work already explored the benefits of MR-based systems in the context of turn-based tabletop games [1, 7, 10]. Our current work follows Scott et al., who identified important requirements for systems to facilitate co-located collaborative work [9]. Our new system especially allows flexible user arrangements (req. 7 in [9]) and simultaneous user interactions (req. 8 in [9]) in a real-time strategy game. However the new interactive gameplay increases the cognitive load of participants. It has been shown that users tend to shift to multimodal communication as load increases with task difficulty and communicative complexity [8]. Therefore, we replaced subsidiary tangible characters by intelligent virtual agents that can be controlled via multimodal commands.

*email: marc.latoschik@uni-wuerzburg.de

Figure 1: Demonstration at the Mobile Media Day 2015.

3 Proof of Concept

Game Design: The basic game design adopts rules and concepts from classical tabletop RPGs and relies on tangibles i.e., cards and playing pieces. In contrast, the traditional turn-based gameplay is replaced by continuous and simultaneous interactions similar to (purely virtual) real-time strategy games to create a much more dynamic game experience. Around the table, up to four players, three *heroes* and one *evil game master*, take control of their available game characters. Heroes are able to move their characters simultaneously, but manually, around a battlefield and perform desired actions using the tangible interface elements. As a result, idle times for players are drastically shortened. Although balancing is not in focus of the current prototype, cooldowns for actions are introduced to avoid the gameplay being utterly dependent on dexterity.

In the storyline, heroes fight the game master. Hence, the game master is facing a superior number of opponents simultaneously attacking, while she has to command her army of minions in a very short time. In order to decrease her cognitive and physical load, the minions of the game master are developed as semi-autonomous intelligent virtual agents and not controlled manually but by multimodal speech and gesture commands. Auditive and visual feedback support the game experience. Overall, we have optimized the available interaction space for all players to further reduce cognitive demands. Game rules are evaluated automatically by the system and appropriate visual feedback is given depending on the current action. Additionally, virtual augmented tangible panels are available to further customize the user interface and available screen estate depending on the game context (see Figure 1).

Gameplay: The following section showcases one typical interaction sequence: In order to attack a minion, a player has to get in range and move his hero's playing piece. A virtual circle appears indicating the maximum movement radius (see Figure 2 a)) once the playing piece is lifted. The player places his playing piece inside the circle and is now within attack range. The minion recognizes the hero and strikes back. To support her minion, the game master sends reinforcements by selecting a group of minions by speech

Figure 2: Interaction techniques: a) Simultaneous tangible interaction b) Command virtual agents via multimodal utterances.

and ordering them to move while pointing at the target (see Figure 2 b)). Another hero rushes to his companion's side and both players collaborate with simultaneous attacks against the approaching minions. The game master tells all melee minions to attack the second player using the hero's name. Other minions are commanded to burn down the village of the heroes. Eventually minions autonomously search, ignite and use torches to execute the command.

System Design: The application is implemented with Simulator X, an open-source software platform for intelligent real-time interactive systems targeting AR, VR, and MR [5]. Samsung's SUR40 multi-touch table serves as input and output device, capable of tracking multiple touch inputs and fiducial markers. Speech input is processed by an external machine using the Kinect SDK.

The state of the application is monitored by a game logic module based on a finite-state automaton. A planning module that incorporates artificial intelligence methods controls the virtual agents' actions. The default behavior comprises approaching, attacking heroes in sight, and withdrawing at low health. Multimodal utterances can be used to assign tasks to the virtual agents (e.g. moving to specific locations). Therefore speech and touch inputs are passed to the multimodal processing framework miPro [4] via a TUIO and VRPN server. The multimodal grammar is defined by a revised version of the augmented transition network (ATN) from [3]. A command consists of three parts: *Selection*, *Action* and *Target* which each is described in respective Sub-ATNs. Figure 3 conceptually depicts the structure of (a) the abstract command and (b) a specific *Target* Sub-ATN facilitating multimodal fusion. The planning module interprets recognized commands by checking preconditions and eventually deriving basic actions which can be executed by virtual agents.

Preliminary User Feedback: As a first informal evaluation, the system has been showcased at two public exhibitions. The received feedback is encouraging. Our prototype not only attracted non- or casual tabletop gamers sharing a affinity towards new media (Mobile Media Day '15), but also dedicated tabletop game fans (RPC Köln '15). Both user groups enjoyed the real-time aspect and easily engaged and interacted with the system. A user study will be conducted to confirm these promising results.

4 Conclusion

In this paper we presented a novel concept for mixed reality tabletop games. The proposed approach abandons the traditional turn-based style by incorporating simultaneous real-time interactions. To reduce cognitive load for the users caused by the increased game speed, intelligent virtual agents have been developed which can be commanded using a multimodal—speech and gesture—interface. The received feedback by first informal evaluations is promising. Future work includes the formal evaluation of the user experience and performance, in particular the influence of multimodal interactions on perceived cognitive load and overall game experience. In addition, we want to explore possibilities to extend multimodal interactions for all players.

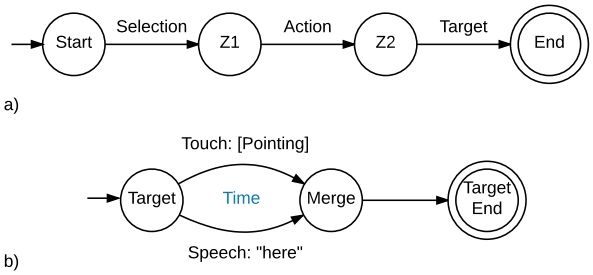

Figure 3: Excerpt of the implemented ATN: a) Abstract command structure b) *Target* Sub-Atn structure.

Acknowledgements

The authors wish to thank Dimitri Reisler and Pegasus Spiele for supporting the XRoads project [2].

References

[1] S. Bakker, D. Vorstenbosch, E. van den Hoven, G. Hollemans, and T. Bergman. Weathergods: tangible interaction in a digital tabletop game. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 151–152. ACM, 2007.

[2] M. Fischbach. XRoads. `http://hci.uni-wuerzburg.de/projects/xroads.html`, 2016.

[3] M. E. Latoschik. Designing Transition Networks for Multimodal VR-Interactions Using a Markup Language. In *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces ICMI'02*, pages 411–416. IEEE, 2002.

[4] M. E. Latoschik and M. Fischbach. Engineering variance: Software techniques for scalable, customizable, and reusable multimodal processing. In *Proceedings of the HCI International Conference 2014*. Springer, 2014.

[5] M. E. Latoschik and H. Tramberend. Simulator X: A scalable and concurrent architecture for intelligent realtime interactive systems. In *Virtual Reality Conference (VR)*, pages 171–174. IEEE, 2011.

[6] J. Leitner, M. Haller, K. Yun, W. Woo, M. Sugimoto, and M. Inami. Incretable, a mixed reality tabletop game experience. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*, pages 9–16. ACM, 2008.

[7] C. Magerkurth, M. Memisoglu, T. Engelke, and N. Streitz. Towards the next generation of tabletop gaming experiences. In *Proceedings of Graphics interface 2004*, pages 73–80. Canadian Human-Computer Communications Society, 2004.

[8] S. Oviatt, R. Coulston, and R. Lunsford. When do we interact multimodally?: Cognitive load and multimodal communication patterns. In *Proceedings of the 6th International Conference on Multimodal Interfaces*, ICMI '04, pages 129–136. ACM, 2004.

[9] S. Scott, K. Grant, and R. Mandryk. System guidelines for co-located, collaborative work on a tabletop display. In K. Kuutti, E. Karsten, G. Fitzpatrick, P. Dourish, and K. Schmidt, editors, *ECSCW 2003*, pages 159–178. Springer Netherlands, 2003.

[10] C. Zimmerer, M. Fischbach, and M. E. Latoschik. Fusion of mixed reality tabletop and location-based applications for pervasive games. In *Proceedings of the 2014 ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 2014.