

論文の内容の要旨

論文題目 **AUTOMATIC MULTIMODAL BEHAVIOR GENERATION FOR EMBODIED VIRTUAL CHARACTERS BASED ON TEXT**

(仮想キャラクターのためのテキストからの自動マルチモーダル行動生成)

氏名 **ブライトフス ウェーナー**

Virtual agents represent a powerful human-computer interface, as they can embody behavior that a human may identify with, this ability may encourage users to engage in a more natural and immersive interaction and establish bonds with them. Facilitating multiple modalities like speech, facial expressions, body postures and gestures they can relay information on different channels. Combining these modalities and creating human-like conversational behavior like gaze and gestures is a challenging and tedious task for human animators. The increasing popularity and use of virtual characters in more and more applications, such as computer games, online chat, and virtual worlds pushes the need for an automatic approach. Since text is often the easiest available information for the conversations of different virtual characters and foremost simplest way of creating dialogues for human authors, it forms a natural source for input. Thus, there have been some attempts to generate non-verbal behavior for embodied agents automatically based on text. The drawback of most current systems and tools, however, is that they consider only one agent, or only suggest behaviors, such that the animator still has to select appropriate ones by him- or herself.

This thesis focuses on the automatic generation of non-verbal behavior such as gestures and gaze using only linguistic and contextual information retrieved from the input text. It presents a method to transform text into agent behavior enriched by eye gaze, movement and conversational gesture behavior. A salient feature of our approach is that we generate the behavior not only for the speaker agent but also for the listener agent that might use backchannel behavior in response to the speaker agent. Employing two presenter agents holding a dialogue is advantageous, since watching (or interacting with) a single agent can easily become boring and it also puts pressure on users, as they are the only audience. Furthermore, two agents support richer types of interactions and “social relationships” between the interlocutors. The system presented in this thesis can be used to provide natural gestures for both types and since all behaviors are generated automatically, there is no extra effort the user would have to contribute to increase the naturalness of the characters behavior and so provides a convenient method to have multimodal conversations in virtual environments.

The agents’ gaze behavior is informed by theories of human face-to-face gaze behavior, that looked into multiparty gazing patterns and also gaze behavior in situations where only two human hold a dialogue. The theories we used are based on empirical user studies yielding information that enabled us to design the rules for our gaze generation method.

Gestures are generated in three steps , for each step a software module was implemented. First the analysis of linguistic and contextual information of the input text like part-of-speech information, lemmas and syntax information. Facilitating this we can split the text in to rheme/theme, objects and action parts which form the basic structure for our behavior generation . Second is the adding of the appropriate gesture. In this step we use different methods for different classes of gestures such as an algorithm for multimodal referring expressions and a gesture dictionary for iconic, metaphoric and emblematic gestures. The final steps adopts the gaze and gesture behavior of each character and produces a graph-based behavior representation. The third module uses this representation and produces a playable script which then can be used to animate the actual characters in different virtual settings.

Our gesture generation system has been evaluated in several user studies providing us with result enabling us to further improve and enhance our approach. Many of these studies showed that the behavior in conversations between virtual characters produced by our system was conceived as human-like, natural and easy to follow.