

学 位 論 文 題 名

Social Avatar Language for
Natural Avatar-Based Human Communication

(アバタ利用による自然なヒューマンコミュニケーションを目指した
社会的要因を考慮した化身話の研究)

学位論文内容の要旨

A modern society requires harmonic, interdependent, or cross-cultural human relationships among people regardless of age, culture, language, or race. A friendly, face-to-face, easy-to-understand, and real-time human communication is at the root of leading to a pleasant life journey. That is why avatar-based human communication has intrigued researchers and developers from academia and industry for several decades.

Communication by avatars is a very interesting topic that is both socially relevant and scientifically oriented. It provides us a huge virtual space where we can either simulate our daily life in reality or stretch our imagination. Its many advantages, especially its interaction, attract researchers from fields like education, military, entertainment, etc. The ideal natural avatar-based communication requires multidisciplinary efforts, including computer science, artificial intelligence, and social science. Recently, many exciting results have been reported in the areas of human figure modeling, rendering, animation, natural language processing, speech recognition and synthesis, cognitive science, psychology and linguistics, etc. However, current systems remain difficult in natural and convincing interaction. The major challenge comes from describing and representing social context and personality for avatars.

Some researchers are working toward personalized or emotional avatars. For example, EMOTE (Expressive MOTion Engine) offers a methodological tool for creating personalized agents, which provides a method of synthesizing movements by controlling a relatively small number of parameters, and enables a higher level of understanding communicative gestures. BEAT (Behavior Expression Animation Toolkit) allows animators to insert their preferred personalities, motion characteristics, or other particular features in the final animation. However, most

existing systems do not provide an effective and easy-to-use tool for addressing the *social context* for avatars that is especially useful for cross-cultural communication. Moreover, most studies to date focus on either the body animation or the face animation. Integrating and especially synchronizing the body and face pose challenge in this field. While hand gestures have received much attention, the function of arm gestures and body postures has been mostly neglected. Furthermore, the existing commercial animation tools for avatars usually rely on pre-authored audio, but natural language real-time instruction for avatars, especially multiple natural language instruction, is rarely studied. These challenges create new opportunities and motivate us to seek a new approach.

An avatar is a graphical representation of a real person in virtual environment. As real humans, we have different positions and play different roles in society. Our behaviors are influenced by social context such as culture, profession, religion, to mention a few. It is understandable that an American probably displays different nonverbal behaviors from a Japanese. But as stated above, the social context is not considered in most current avatars. Thus, we propose a new concept, Social Avatar Language (SAL), to describe the nonverbal behaviors for avatars with regard to social background. And we name this new type of avatars that can conduct the social behaviors Social Avatar (SA). That is, SAL represents social behaviors of an avatar through animations of facial expressions and body gestures in cyberspace.

We originally propose to use ontology technology for addressing the social context for avatars. We define here ontology as the specifications for describing and representing our proposed SAL. This ontology is used to match and bind data to 3D (dimensional) avatar presentations including facial expressions, lip movements, body postures and gestures. We propose two new ontologies, Avatar Modeling Ontology (AMO) and Social Context Ontology (SCO). While the AMO is used for the anatomic constraints for animation, the SCO is used to associate the social context with an avatar. It is important to understand that our SA is at first an avatar, and we focus on the influence of the social context on the behaviors and the appearance of the avatar. What we are especially interested in is how to associate the social context with an avatar and how to extend an SA by more social aspects. SAL can reveal the social background of avatars with multiple natural language processing function in the mean time. To achieve this, a set of operations and mapping is required, which includes semantic extraction from the input text, and translation between natural languages and nonverbal languages. For animation synthesis, besides seamless body gestures, we also create rich facial expressions with appropriate eye, eyebrow, and mouth movements, for representing human emotions (such as happiness, sadness, or anger). The current version of this system can deal with nonverbal behaviors for Japanese, Chinese, and American cultures to some extent.

学位論文審査の要旨

主 査 教 授 青 木 由 直

副 査 教 授 北 島 秀 夫

副 査 教 授 荒 木 健 治

学 位 論 文 題 名

Social Avatar Language for Natural Avatar-Based Human Communication

(アバタ利用による自然なヒューマンコミュニケーションを目指した
社会的要因を考慮した化身話の研究)

インターネットの普及によって国境を越えたコミュニケーションが拡大している。諸事情で共通言語が利用できない場合、コンピュータの画面の前で身振り手振り、表情等でコミュニケーションを行う事が予想され、言葉が通じない状況では健常者もコミュニケーション障害者の立場になる。そこで、聴覚、発声の障害克服のため利用している「手話」を用いた国際コミュニケーションを含んで、新しく「化身話(Avatar Language)」という概念での非言語コミュニケーション方式を提案した。これは手話が既に社会に受け入れられていて、それゆえの制限の下にあり、本研究での手話の意味は健常者が用いる手話であるとの区別も明らかにする点からも、化身話の概念を提案しこれを用いている。この概念に基づき、具体的システム構築を行い、方式の妥当性を論じている。

第1章は研究の背景と論文全体の構成について述べている。第2章は本研究に関連する従来の技法を引用しながら、本研究で用いているコンピュータグラフィックス(CG)技法による「化身(Avatar)」の表現法や、心理学や言語学も参考にしながら化身を動かして意志や感情を伝える非言語コミュニケーションの基本的な検討を行っており、シミュレーション実験で提案方式の有効性について知見を得ている。

第3章は、社会的要因を考慮した化身の利用について検討を行っている。第4章では、化身という明示的な実態化を行う“オントロジー(ontology)”モデルを提唱して、非言語コミュニケーションにおける有効性を確かめている。

第5章では、化身の動作や表情で表現される化身話とテキストによる表現の対応関係について述べている。基本的にはある言語でのテキストによる記述に対応する化身動作を作っていくことで、化身話の表現としている。

第6章では知的通信方式での化身話コミュニケーションについて検討している。知的通信では、化身話のアニメーションをそのまま伝送するのではなく、CGアニメーションを

作り出すパラメータ等を伝送し、このパラメータの受信後にパラメータを用いてCGによるアニメーションを再構成する方式であり、伝送速度が低い通信路での化身話コミュニケーションの実現のための基礎データを得ている。

第7章では化身話の応用として仮想教室を想定して検討を行っている。教材としては風力や信号処理に関するものを取り上げている。第8章は結論とこれからの課題について述べている。

本論文の要点は以下のようにまとめることができる。

(1) 健常者が異言語の壁を越えてのコミュニケーションを行うため、健常者の手話ともみなせる化身話による非言語コミュニケーションが有効な方法である事を明らかにした。

(2) 国籍を明示して、国によつての社会的特性を化身に吸収させることが国際コミュニケーションでは有効である。化身の国籍と仮定すると、その国の習慣による身振り、手振りを行わせ、その国で理解され易い動作で意志を伝える方式を検討した。それはその国で用いられている手話を参考にする事でもあり、手話はその国の社会的要因に基づいており、化身もそのようなものとして利用すべきであるとの考えを実験的にも確かめている。これは、社会的な約束事を化身の取る動作に反映させる社会文脈オントロジー(Social Context Ontology: SCO)という本論文で提唱している新しい概念につながっている。この概念について、具体的システムを構築して新知見を得ている。

(3) 単語に対応した動作のアニメーションを予め制作しデータベース化しておいて、ある言語での単語の並べ方に従って、このデータベースからの個々のアニメーションをつなぎ合わせて文章に対応した化身話文章を表示する方法を提案し、実験的システムを構築して検討を行っている。この方式は知的通信方式にも適合した方式であり、化身話コミュニケーションに有効である点を確かめた。

これを要するに、著者は言語の壁を越える目的での化身話コミュニケーションという新しい非言語コミュニケーションの分野において新知見を得ており、情報メディア工学に貢献するところ大なるものがある。

よって著者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。