

Association for Information Systems AIS Electronic Library (AISeL)

AMCIS 2012 Proceedings

Proceedings

Understanding Avatar Sentiments Using Verbal and Non-Verbal Cues

Yulei Zhang

Northern Arizona University, Flagstaff, AZ, United States., yulei.zhang@nau.edu

Yan Dang

Northern Arizona University, Flagstaff, AZ, United States., yan.dang@nau.edu

Susan Brown

University of Arizona, Tucson, AZ, United States., suebrown@email.arizona.edu

Hsinchun Chen

University of Arizona, Tucson, AZ, United States., hchen@eller.arizona.edu

Follow this and additional works at: <http://aisel.aisnet.org/amcis2012>

Recommended Citation

Zhang, Yulei; Dang, Yan; Brown, Susan; and Chen, Hsinchun, "Understanding Avatar Sentiments Using Verbal and Non-Verbal Cues" (2012). *AMCIS 2012 Proceedings*. 7.

<http://aisel.aisnet.org/amcis2012/proceedings/VirtualCommunities/7>

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2012 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Understanding Avatar Sentiments Using Verbal and Non-Verbal Cues

Yulei Zhang

Computer Information Systems
The W. A. Franke College of Business
Flagstaff, Arizona, 86011
yulei.zhang@nau.edu

Yan Dang

Computer Information Systems
The W. A. Franke College of Business
Flagstaff, Arizona, 86011
yan.dang@nau.edu

Susan A. Brown

Department of Management Information Systems
Eller College of Management
University of Arizona
Tucson, Arizona, 85721
suebrown@eller.arizona.edu

Hsinchun Chen

Department of Management Information Systems
Eller College of Management
University of Arizona
Tucson, Arizona, 85721
hchen@eller.arizona.edu

ABSTRACT

With the increased popularity of virtual worlds, hundreds of thousands of people from different physical locations can join virtual worlds. In this computer-based simulated 3D environment, avatars can both interact with each other and the environment. This new type of world has important implications for business, education, and society at large. In order to fully use the benefits of virtual worlds, it is important to know how the residents (i.e., avatars) behave, such as how they express sentiments. This research in progress seeks to study avatar sentiments in virtual worlds to examine whether and how sentiments are conveyed by avatars. Both verbal and non-verbal cues will be utilized in the sentiment analysis. To conduct the study, an advanced data collection method is leveraged to obtain various types of avatar data from a large number of real virtual world residents in Second Life in an effective and efficient way.

Keywords

Avatar Sentiments, Second Life, Verbal Cues, Non-Verbal Cues, Avatar Data Collection.

INTRODUCTION

A virtual world is an electronic environment that can simulate the world where people physically live. People can interact with each other, virtual objects, and the environment by controlling their digital representations, called avatars (Bainbridge 2007). Virtual worlds have been gaining more and more attention in recent years. It is expected that virtual worlds will grow in societal importance and will influence various aspects of people's lives such as education, collaboration, business and research (Messinger et al. 2009). In addition, a virtual world serves as a new place to do business in new ways and has been bringing a lot of business opportunities. The emerging virtual world market could reach billions of dollars in the coming years (Hendaoui et al. 2008). For example, the money supply of the most popular virtual world, Second Life, grew to US \$29.3 million in the third quarter of 2011, 8.5 percent of increase over 2010; the web merchandise sales volume grew to L\$1.183 billion, or approximately US \$4.7 million (the average exchange rate was 1US\$=251L\$) in the third quarter of 2011 (The Second Life Economy in Q3 2011; <http://community.secondlife.com/t5/Featured-News/The-Second-Life-Economy-in-Q3-2011/ba-p/1166705>).

To make the best use of this new medium, previous research has compared it with the physical world and examined its adoption in various areas. For example, some studies have compared the efficacy of e-learning in virtual worlds with real-world classroom learning (Franceschi et al. 2009; Hassell et al. 2009; Lester and King 2009). Holzwarth et al. (2006) found that using avatar sales agent led to more satisfaction with the retailer, a more positive attitude toward the product, and a greater intention to purchase. Other studies have examined the social interaction patterns between male and female avatars (Yee et al. 2007). However, few studies have empirically examined avatar sentiments, an important aspect of avatar behavior and expression.

In this research in progress, we address the following question:

RQ1. Do avatars express sentiments in virtual worlds?

RQ2. How can we leverage both verbal and non-verbal cues to conduct avatar sentiment analysis in virtual worlds?

RQ3. Are there any sentiment differences across avatars in different categories (such as male avatars vs. female avatars)?

We aim to examine avatar sentiments using both verbal and non-verbal cues. Verbal cues are chat messages used by avatars to communicate with each other. Non-verbal cues include avatar actions and facial expressions that may also be leveraged by avatars to express sentiments. To conduct text-based sentiment analysis, previous literature has summarized related techniques into two categories: machine learning and semantic orientation approaches (Liu 2007). A machine learning approach identifies sentiments by leveraging a set of language features and training a classification model. A semantic orientation approach determines sentiments by checking against a pre-defined sentiment lexicon.

LITERATURE REVIEW

Importance and Applications of Virtual World Research

In addition to the world where people physically live, another type of world – virtual world – has recently been playing an important role in various areas (Chen 2009). Reports have shown that millions of people in different places over the world spend an average of 22 hours per week in virtual worlds to interact with others (Yee and Bailenson 2007). Currently, one of the most popular social interaction based 3D virtual worlds is Second Life (<http://secondlife.com>) that was developed by Linden Lab and publicly launched in 2003. Users of Second Life can interact with each other, the virtual objects, and the highly immersive environment via controlling their avatars. The total number of users of Second Life has grown dramatically, from 2.2 million in December 2006 to 8.3 million in August 2007 (Messinger et al. 2009). As reported by Second Life, there was an average of 1.046 million monthly unique logged-in users in the third quarter of 2011 (<http://community.secondlife.com/t5/Featured-News/The-Second-Life-Economy-in-Q3-2011/ba-p/1166705>).

The advances and rapid development of Web-based technologies have changed the way the Internet serves its users – from only allowing them to passively acquiring information to enabling them to actively participate and contribute to the Internet (O'Reilly 2005). A virtual world is a place where users can experience it via their avatars and actively interact with others by using both verbal and non-verbal cues. The highly interactive and immersive nature of virtual worlds could be of interest and importance to researchers in various areas including business, social science, communication, politics, and education (Chen 2009; Hendaoui et al. 2008; Messinger et al. 2009). For example, virtual worlds enable new models for marketing; business transactions on virtual objects and goods may influence users' purchase behavior in the physical world and could play an important role on consumer engagement with a brand (Hemp 2006). The vivid 3D representation of virtual worlds makes them an ideal platform for e-learning as well (Franceschi et al. 2009; Hassell et al. 2009).

In addition, virtual worlds can offer many potential benefits to researchers. For example, it can be difficult to recruit a large number of subjects from different physical locations. However, virtual worlds can make this possible and much easier since there is almost no cost to travel from one region to another in a virtual world – users can click to switch regions (Bainbridge 2007). It is also possible to study the social and ethical issues (e.g., comparing alternative governmental regulations, and studying people's perceptions towards the ones who have serious transmitted diseases by manipulating the appearance and behavior of avatars) that cannot be performed in the physical world (Bainbridge 2007; Gordon et al. 2009).

Previous Empirical Research on Avatar Behavior

Previous research has conducted empirical studies on virtual worlds to examine avatar behavior. For example, when comparing male and female avatars, Yee et al. (2007) found that male avatars tended to keep larger interpersonal distances with other male avatars compared to the interpersonal distances between two female avatars. They also found that male avatars tended to have less eye contact with other male avatars than the frequency of eye contact between two female avatars (Yee et al. 2007). Harris et al. (2009) conducted a lab experiment using 80 subjects and found that, in general, avatar activity declined with increased use experience, suggesting that experienced avatars tended to conduct fewer activities.

Previous research has also examined how avatars' relationships are affected by different environmental factors in multiplayer online game based virtual worlds such as EverQuest II (Shim et al. 2011). For example, Keegan et al. (2010) found significant differences in the trade networks of gold farmers and non-gold farmers in EverQuest II by measuring social network metrics such as in-degree, out-degree, and centrality. They found that gold farmers had significantly lower in-degrees and out-degrees compared with non-gold farmers, indicating that gold farmers typically tended to conduct

transactions with only a few trusted characters. Kawale et al. (2009) studied the problem of player churn in EverQuest II by analyzing player engagement and network statistics such as the number of neighbors and number of churning neighbors, and significant differences were observed between networks of general players and churning players.

Other research has examined the relationship between avatars and their users. For example, previous research has found that virtual world users would conform to expected behaviors associated with their avatars' virtual identities and behave accordingly (Yee et al. 2011); Suh et al. (2011) found that the more closely an avatar resembled its owner in terms of appearance, the more likely the owner would have positive attitudes toward the avatar.

Sentiment Analysis

Sentiment analysis is the automatic detection of opinions from free text (Abbasi et al. 2008; Dang et al. 2010). The computer-based detection and analysis of sentiments has been a growing interest among researchers in recent years (Pang and Lee 2008). Previous sentiment analysis studies include determining whether a text is objective or subjective (Pang and Lee 2008), or whether a subjective text contains positive or negative sentiments (Wiebe et al. 2001; Wiebe et al. 2004).

Two approaches in sentiment analysis have been widely used: the machine learning approach and the semantic orientation approach (Liu 2007; Pang and Lee 2008). The machine learning approach treats sentiment analysis as a topic-based text classification problem. Any text classification algorithm can be employed. For example, Pang et al. (2002) compared the performances of Naïve Bayes, Support Vector Machine (SVM), and Maximum Entropy using movie reviews and found that Naïve Bayes and SVM performed well in determining if a piece of review was negative or positive.

The semantic orientation approach performs classification based on positive and negative sentiment words and phrases contained in each evaluation text and no prior training is required in order to mine the data (Pang and Lee 2008). This approach often relies on external knowledge resources beyond raw data and thus is knowledge-rich (Zhou and Chaovalit 2008). It generally proceeds in three steps: (1) extracting words or phrases that express semantic orientations; (2) determining the polarities of the extracted words or phrases; and (3) computing the polarity of the text by aggregating the polarities of individual words or phrases in the text (Zhou and Chaovalit 2008).

Research Gaps

It is widely believed that virtual worlds have brought more opportunities in various areas. They've expanded the ways in which communication, business, education, etc. can be conducted. Users use their avatars to experience a virtual world and interact with others and the virtual environment. Previous research has examined different aspects of avatar behavior such as interpersonal distance and eye contact patterns (Yee et al. 2007), interaction network differences (Shim et al. 2011), and the similarity between avatar appearance and the user appearance (Suh et al. 2011). However, few studies have been seen to examine avatar sentiments.

Sentiment analysis has gained a lot of attention in recent years. This may be because of the emerging of new communication platforms with the advent of Web 2.0 such as Web forums, blogs, social networking sites and most recently virtual worlds. Understanding sentiments expressed in these new media can be of importance to business, education, government, etc. Different from other Web 2.0 media, conducting sentiment analysis on virtual worlds could be more challenging, since avatars can express their sentiments not only via verbal cues (mostly by using instant chatting messages) but also non-verbal cues (such as actions and facial expressions). However, little research has been done on avatar sentiment analysis. Thus, this research in progress seeks to conduct avatar sentiment analysis using both verbal and non-verbal cues. We also plan to examine sentiment differences across avatars in different categories (such as male avatars vs. female avatars).

RESEARCH METHOD

Research Site

The study site we plan to use is Second Life (<http://secondlife.com>) that is currently the most popular social interaction based 3D virtual world. To access Second Life, users need to create their own avatars that are the digital representations of themselves. Avatars can then login to Second Life and interact with other avatars and virtual objects.

Data Collection Methods

To collect avatar behavioral data from Second Life, most previous research has utilized self-report questionnaires from recruited subjects. The major limitations of this method include the unreliability of self-reported measures and the relatively small sample size that can be collected (Yee and Bailenson 2008). To overcome the drawbacks, Yee and Bailenson (2008)

have developed an advanced method to semi-automatically collect avatar-related data from Second Life based on the Linden Scripting Language (LSL). LSL is a state-event-driven scripting language allowing Second Life users to control their avatars' behavior. Leveraging this language, Yee and his colleagues have developed the semi-automatic, bot-based approach to collect avatar behavior data.

Although quite advanced, their approach still has some limitations due to the restrictions of LSL. For example, a bot can collect only a relatively small number of records from avatars that are close to it. Thus, human intervention is needed to manually move the bot close to an avatar in order to collect its behavioral data. In addition, since bots are created as objects (visible or invisible) attached to avatars, when the avatars change their outfits, the attached bots will be automatically detached. Furthermore, the method can collect limited types of avatar data. It can collect an avatar's current action and physical position but cannot collect avatar profile information (such as gender and age) or the verbal information (i.e., chatting messages between two or more avatars).

To address these limitations, we developed an improved bot-based approach that can effectively and efficiently obtain more types of avatar data from Second Life (Zhang et al. 2010). Instead of using LSL, we leverage LibOpenMetaverse (<http://openmetaverse.org/>), a software library that can be used as a third-party application to communicate with the Second Life server, to develop the data collection method. LibOpenMetaverse enables the creation of avatar-like bots and allows each bot to conduct a much larger-scale data collection with no human intervention. Different libraries in LibOpenMetaverse can be used to collect various types of avatar data including verbal and non-verbal cues.

We will leverage both machine learning and semantic orientation approaches to conduct sentiment analysis. Evaluation will then be performed to examine the analysis results. Further exploration will also be conducted to try to identify different patterns of sentiments expressed by avatars with different identities (such as in different genders).

Research Status

Research Phase 1

The goal of phase one is to conduct a large-scale avatar data collection using the improved bot-based approach. Different components have been developed in order to collect different types of avatar data, including verbal cues, non-verbal cues, and avatar profile data. One component takes charge of collecting verbal cues, that are chat messages between avatars. The chat messages that are open to the public can be collected. For chat messages that are set as private, Second Life does not allow the collection of them because of privacy issues. Another component deals with collecting non-verbal cues such as avatar actions and facial expressions. LibOpenMetaverse provides some predefined categories for capturing avatar actions and facial expressions. Examples related to positive sentiments include belly laughing, blow kissing, kissing, laughing, smiling, with a big tooth smile, winking, and happy expression on face; examples of negative sentiments are afraid, angry, bored, frowning, sad and worried expressions on face, crying, and showing dislike. The last component is to collect avatar profile data so that sentiment differences can be compared among avatars with different characteristics such as different genders. This phase includes data collection, data preprocessing and data cleaning. We plan to have a dataset ready for the next phase by April.

Research Phase 2

The goal of phase two is to conduct the sentiment analysis and evaluation. Once phase one is completed, sentiment analysis will be performed. For non-verbal data, we will look at the predefined actions provided by LibOpenMetaverse. For verbal data, we plan to use both machine learning and semantic orientation approaches to conduct the analysis. Language features will be extracted from the data collection and be leveraged in building sentiment analysis models. A sentiment-based lexicon, such as SentiWordNet (Esuli and Sebastiani 2006) will be leveraged to determine sentiments of the extracted language features. Evaluation will be conducted to examine the performance of the analysis models. Based on avatar profile data, further exploration can be conducted to study different sentiment patterns among avatars with different characteristics.

CONCLUSION

In this paper, we describe a research in progress that aims to examine avatar sentiments using both verbal and non-verbal cues. As a new, innovative social interaction platform, virtual worlds have provided opportunities for various areas and are gaining increased attention among researchers, businesses, educators, and the government. Understanding the behavior of virtual world residents such as their sentiments can help us make the best use of the new medium.

Specifically, the contribution of this study is that it will be among the first to empirically examine avatar sentiments in Second Life. The analysis results can help expand our understanding of avatar sentiments in virtual worlds. In addition,

understanding how avatars express sentiments can inform our understanding of sentiment expressions in the physical world. We can look at how sentiments are expressed individually based on different avatar characteristics (e.g. gender) in virtual world, and compare the patterns obtained from virtual worlds with face-to-face communications. This study will be of importance and interest to researchers and practitioners who would like to know more about virtual worlds and their impacts and plan to leverage the new medium in various areas.

REFERENCES

1. Abbasi, A., Chen, H., and Salem, A. (2008) Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums, *ACM Transactions on Information Systems*, 26, 3, 1-34.
2. Bainbridge, W.S. (2007) The Scientific Research Potential of Virtual Worlds, *Science*, 317, 472-476.
3. Chen, H. (2009) AI, E-Government, and Politics 2.0, *IEEE Intelligent Systems*, 24, 5, 64-67.
4. Dang, Y., Zhang, Y., and Chen, H. (2010) A Lexicon Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews, *IEEE Intelligent Systems*, 25, 4, 46-53.
5. Esuli, A., and Sebastiani, F. (2006) Sentiwordnet: A Publicly Available Lexical Resource for Opinion Mining, *LREC-06, the 5th Conference on Language Resources and Evaluation*, Genova, Italy, 417-422.
6. Franceschi, K., Lee, R.M., Zanakis, S.H., and Hinds, D. (2009) Engaging Group E-Learning in Virtual Worlds, *Journal of Management Information Systems*, 26, 1, 73-100.
7. Gordon, R., Björklund, N.K., Smith, R.J., and Blyden, E.R. (2009) Halting Hiv/Aids with Avatars and Havatars: A Virtual World Approach to Modelling Epidemics, *BMC Public Health*, 9, S13, 1-6.
8. Harris, H., Bailenson, J.N., Nielsen, A., and Yee, N. (2009) The Evolution of Social Behavior over Time in Second Life, *Presence: Teleoperators and Virtual Environments*, 18, 6, 434-448.
9. Hassell, M., Goyal, S., Limayem, M., and Boughzala, I. (2009) Being There: An Empirical Look at Learning Outcomes in 3d Virtual Worlds, *15th Americas Conference on Information Systems (AMCIS)*, San Francisco, California, 1- 11, Paper 733.
10. Hemp, P. (2006) Avatar-Based Marketing, *Harvard Business Review*, 84, 6, 48-56.
11. Hendaoui, A., Limayem, M., and Thompson, C.W. (2008) 3D Social Virtual Worlds: Research Issues and Challenges, *IEEE Internet Computing*, 12, 1, 88-92.
12. Holzwarth, M., Janiszewski, C., and Neumann, M.M. (2006) The Influence of Avatars on Online Consumer Shopping Behavior, *Journal of Marketing*, 70, 4, 19-36.
13. Kawale, J., Pal, A., and Srivastava, J. (2009) Churn Prediction in Mmorpgs: A Social Influence Based Approach, *Proceedings of the IEEE International conference on Computational Sciences and Engineering: IEEE Computer Society*, 423-428.
14. Keegan, B., Ahmad, M., Srivastava, J., Williams, D., and Contractor, N. (2010) Dark Gold: Statistical Properties of Clandestine Networks in Massively Multiplayer Online Games, *Proceedings of IEEE 2nd international conference Social Computing: IEEE Computer Society*, 201-208.
15. Lester, P.M., and King, C.M. (2009) Analog vs. Digital Instruction and Learning: Teaching within First and Second Life Environments, *Journal of Computer-Mediated Communication*, 14, 457-483.
16. Liu, B. (2007) *Web Data Mining*. Berlin Heidelberg: Springer.
17. Messinger, P.R., Stroulia, E., Lyons, K., Bone, M., Niu, R.H., Smirnov, K., and Perelgut, S. (2009) Virtual Worlds - Past, Present, and Future: New Directions in Social Computing, *Decision Support Systems*, 47, 3, 204-228.
18. O'Reilly, T. (2005) What is Web 2.0? Design Patterns and Business Models for the Next Generation of Software, <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
19. Pang, B., and Lee, L. (2008) Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, 1, 1-2, 1-135.
20. Pang, B., Lee, L., and Vaithyanathan, S. (2002) Thumbs Up? Sentiment Classification Using Machine Learning Techniques, *the ACL-02 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Philadelphia: ACL Press, 79-86.
21. Shim, K.J., Pathak, N., Ahmad, M.A., DeLong, C., Borbora, Z., Mahapatra, A., and Srivastava, J. (2011) Analyzing Human Behavior from Multiplayer Online Game Logs: A Knowledge Discovery Approach, *IEEE Intelligent Systems*, 26, 1, 85-89.
22. Suh, K.-S., Kim, H., and Suh, E.K. (2011) What If Your Avatar Looks Like You? Dual-Congruity Perspectives for Avatar Use, *MIS Quarterly*, 35, 3 September, 711-729.
23. Wiebe, J., Wilson, T., and Bell, M. (2001) Identifying Collocations for Recognizing Opinions, *Proceedings of the ACL/EACL Workshop on Collocation*, Toulouse, France: ACL Press.

24. Wiebe, J., Wilson, T., Bruce, R., Bell, M., and Martin, M. (2004) Learning Subjective Language, *Computational Linguistics*, 30, 3, 277-308.
25. Yee, N., and Bailenson, J.N. (2007) The Proteus Effect: The Effect of Transformed Self-Representation on Behavior, *Human Communication Research*, 33, 271-290.
26. Yee, N., and Bailenson, J.N. (2008) A Method for Longitudinal Behavioral Data Collection in Second Life, *Presence: Teleoperators and Virtual Environments*, 17, 6, 594-596.
27. Yee, N., Bailenson, J.N., Urbanek, M., Chang, F., and Merget, D. (2007) The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments, *Journal of Cyberpsychology and Behavior*, 10, 115-121.
28. Yee, N., Ducheneaut, N., Yao, M., and Nelson, L. (2011) Do Men Heal More When in Drag? Conflicting Identity Cues between User and Avatar, *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (ACM CHI 2011)*, Vancouver, BC, Canada: ACM, 772-776.
29. Zhang, Y., Yu, X., Dang, Y., and Chen, H. (2010) An Integrated Framework for Collecting and Analyzing Avatar Data from Virtual World: An Exploratory Study in Second Life, *IEEE Intelligent Systems*, 25, 6, 17-23.
30. Zhou, L., and Chaovalit, P. (2008) Ontology-Supported Polarity Mining, *Journal of the American Society for Information Science and Technology (JASIST)*, 59, 1, 98-110.