3.2 Developing Culture-Adaptive Competency through Experiences with Expressive Avatars

Developing Culture-Adaptive Competency through Experiences with Expressive Avatars

Daniel S. Silverglate
DanS@vcom3d.com

Edward M. Sims, PhD
EdS@Vcom3D.com

Vcom3D, Inc.

Gerald Glover, PhD
jgloverocd@yahoo.com
Hawaii Pacific University

Harris Friedman, PhD
harrisfriedman@hotmail.com
University of Florida

Abstract. Modern Warfighters often find themselves in a variety of non-combat roles such as negotiator, peacekeeper, reconstruction, and disaster relief. They are expected to perform these roles within a culture alien to their own. Each individual they encounter brings their own set of values to the interaction that must be understood and reconciled. To navigate the human terrain of these complex interactions, the Warfighter must not only consider the specifics of the target culture, but also identify the stakeholders, recognize the influencing cultural dimensions, and adapt to the situation to achieve the best possible outcome. Vcom3D is using game-based scenarios to develop culturally adaptive competency. The avatars that represent the stakeholders must be able to portray culturally accurate behavior, display complex emotion, and communicate through verbal and non-verbal cues. This paper will discuss the use of emerging game technologies to better simulate human behavior in cross-cultural dilemmas.

Nomenclature: culture, adaptive, values, cultural values dimensions, dilemmas, virtual humans, non-verbal communications

1.0 INTRODUCTION

This paper addresses a technology and methodology for the development and delivery of serious games for instruction in cross-cultural adaptability. The Cross-cultural Adaptivity Trainer (CCAT) system has been developed over the course of several projects, including cultural training for the US Navy in the Pacific, the Royal Australian Air Force (RAAF), the Air Force Negotiation Center of Excellence (NCE), and the Joint Warfare Center (JWFC).

Whereas the present paper describes the technology we have developed to create and deliver Cross-cultural Adaptability Training (CCAT), a companion paper by the same authors [1] focuses on the underlying theory and pedagogical design.

2.0 CULTURALLY ADAPTIVE COMPETENCY

Contemporary Warfighters are expected to conduct operations and successful missions in many areas of the world. The global community includes a wide variety of socio-cultural contexts, each one with unique challenges for cross-cultural communication and adaptive competency. It is impossible to learn everything about all cultures in the world, or even everything about one culture. Yet we can observe that some people are able to effectively operate in a wide variety

of socio-cultural contexts, even when they do not have prior knowledge or experience of the local culture.

This ability to adapt to a variety of socio-cultural contexts is primarily learned, and is based on the development of adaptive predispositions for behaving effectively in a global community. Specifically, being culturally adaptive is based on predispositions in which a person is able to: 1) recognize and respect cultural differences; 2) make sense of signals and cues in a culturally diverse situation; 3) create synergy from diversity and resolve cultural dilemmas; and 4) view the socio-cultural encounter from a holistic perspective.

As shown in Figure 1, culturally adaptive competency requires that a person be able to recognize, respect, resolve, and reconcile cultural differences in socio-cultural encounters, even in socio-cultural contexts in which the person has no prior knowledge nor experience.

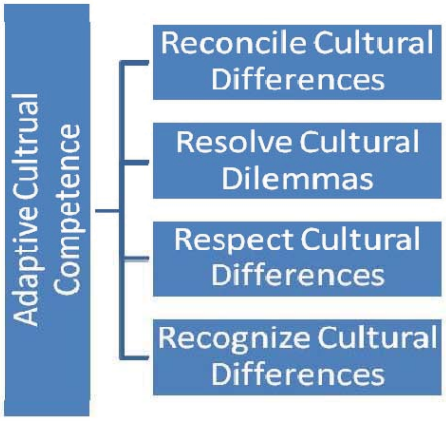

Figure 1. Adaptive Cultural Competence

A framework for understanding and making sense of cultural encounters is the underpinning for developing CCAT. Culture has many levels of understanding and abstraction, including behaviors, beliefs, and values. Values are ways that cultures provide meaning to human behavior and beliefs, and values represent what is important to people. Human behaviors and beliefs can be seen as falling within the extremes of opposite positions on values dimensions. These values dimensions are not stereotypes, but represent dynamic tensions between opposites, such as individual versus collective orientations.

Cross-cultural dilemmas occur when actors hold seemingly opposing values related to a situation, issue, negotiation, or other aspect of a cross-cultural encounter. Cross-cultural dilemmas can be understood and effectively navigated when actors understand their own cultural values in relation to the values of other actors, and know how to resolve or reconcile important differences. Resolve refers to obtaining a workable compromise, whereas reconcile refers to obtaining a higher-order solution that exceeds a compromise and involves creating an innovative solution to a dilemma. Adapting cross-culturally is related to being able to appropriately resolve and reconcile cultural dilemmas through adaptive decision making. We use the concepts of cultural values dimensions, cultural dilemmas, and adaptive decision making to build the CCAT approach, which provides learners with opportunities to experience and respond to cultural dilemmas through recognizing and respecting different value dimensions and resolving or reconciling them through adaptive decision making.

3.0 OVERVIEW OF CCAT

CCAT focuses on developing the learners' cross-cultural adaptive predispositions by providing the learner with a working knowledge of culture through an easily useable framework for making culture operational, and opportunities to engage with scenarios and simulated experiences for applying this knowledge in a variety of cross-cultural encounters.

While video recordings of live role-playing scenarios could be an option for CCAT, we are developing the training using interactive

3D serious game technology. This approach provides several advantages:

- Scenarios and simulations can be created and edited rapidly without access to live actors.
- The technology supports naturalistic interactions, in which computer-generated characters take initiative, in contrast to branching video.
- Advances in virtual human technology support the inclusion of subtle nonverbal cues.

CCAT is currently under development and will be available from a Web site for universal access, as well as on mobile tablets and smartphones. Training will be delivered by avatar-based scenarios and simulations involving cultural dilemmas through which actors holding divergent cultural value orientations will need to navigate. Learners will view and respond to these in virtual cross-cultural encounters. In order to provide the variation and repetition required to support recognition of cultural values from context [2] ,we envision the development of 35 scenarios and 5 simulations to provide cross-cultural experiences in a globally-appropriate collection of cultural encounters. In a relatively brief time period, learners are expected to gain cultural knowledge and experience typically available only to persons who have traveled the world extensively for years.

4.0 TECHNOLOGY REQUIREMENTS

Experiential training for cross-cultural adaptability poses new requirements for serious games. Specifically, the learner needs to be exposed to simulated situations in which they must recognize subtle variations in behavior that are indicative of higher-level differences in cultural beliefs and values. The learner must then apply culture-general knowledge and metacognitive skills to form a mental model of the different cultural values influencing the situation.

We have identified the following technical requirements as key to the successful development of CCAT training:

- Computer-generated characters must be highly expressive, with nonverbal behaviors indicative of emotional response and focus of attention.
- The actors' behaviors must be highly dynamic, showing subtle variations in response to learners' actions.
- The scenarios and simulations should use authentic, relevant storylines to engage the student and keep them motivated.
- The technology should be capable of being adapted to emerging gaming platforms, including mobile tablets, smartphones and virtual worlds.

5.0 EXPRESSIVE AVATARS

Zbylut and Metcalf [3], in an ARI study, have indicated that skill at interpreting nonverbal behavior is one of the most frequent and important communication skills that advisors reported using while deployed. Effective training of culture-adaptive competency requires computer-generated characters capable of reproducing complex nonverbal behaviors such as body gestures, facial expressions, and eye movements.

Body gestures may be either universally recognizable body language or deliberate gestures unique to some culture. We have created a system to allow cultural experts to build their own gestures utilizing Inverse Kinematics (IK). As opposed to Forward Kinematics (FK), where the 3D character is posed by direct manipulation of each joint, Inverse Kinematics allows the hands or fingers to be placed relative to a set of focal sites on the body [4, 5]. A complex set of algorithms then determine the joint angles necessary to achieve this result. The algorithms take into consideration the stresses and limitations of each joint so that pose emulates proper human motion, but

can still be adjusted to ensure the gesture conveys the intended meaning. Since gestures created in this manner are target based, they can be applied to characters with different body types.

As reported by Paul Ekman [6], the perception of emotion through facial expressions is universal throughout all cultures. However, the rules for displaying emotion vary significantly according to the values of a culture [7]. For computer-generated characters to be perceived as having emotions, it is important to apply the fundamental aspects of how facial expressions are formed and their relationship to emotions. Ekman and his colleagues developed the Facial Animation Coding System (FACS) [8] as a mechanism to quantify the movement of facial muscles and relate them to the expressions they form.

According to FACS, facial expressions are the result of the activation of individual muscles or muscle groups beneath the skin. Similarly, the movement of the mesh of a seamless 3D model of the face is affected by the articulation of elements known as bones. Each 3D bone affects some set of vertices in the mesh by some percentage called the weight. Each vertex of the face mesh may be influenced by more than one bone but the cumulative weights should total 100%.

We have developed a new bone-based rig for our 3D characters' faces. Each of 39 "bones" represents one of the FACS action units and by extension the muscle or muscle group of the human face. The bones can be animated to simulate the movement of their real world muscle counterparts. Figure 2 shows the process of creating a facial expression for an avatar. On the left is an image of the control rig which allows the animator to move control points (blue circles) to manipulate the movement of groups of muscles that usually work in concert. The middle image shows the bone-driven, muscle system of anatomically accurate muscles. These bones influence the movement of the mesh to display the intended expression as shown in the image on the right.

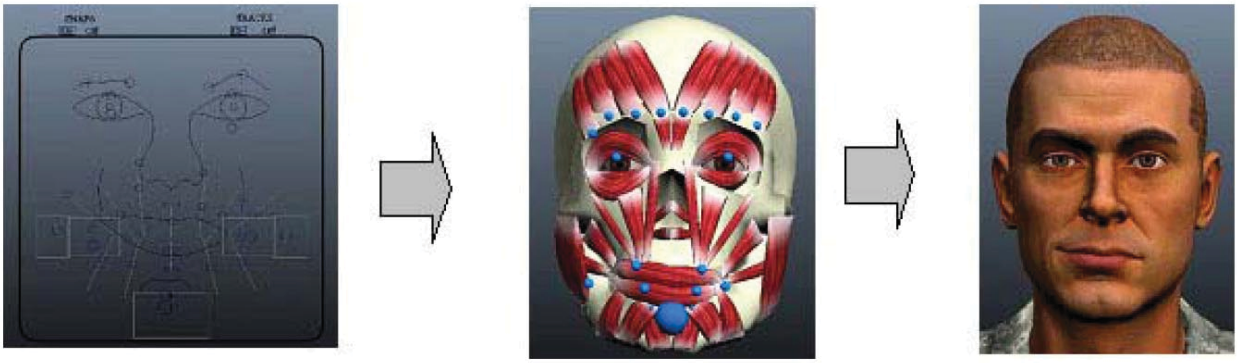

Figure 2. Muscle-based facial rig

The FACS system can also be used to validate that the avatar's expression conveys the same emotion as a human. First, we generate the FACS score for a human subject or reference photo which demonstrates the target expression. Independently, we generate a FACS score for the avatar's expression and then compare the results.

Another element essential to facial expressions is the inclusion of believable lip-synching. Traditional methods of lip-synching are often expensive, time consuming, and may still produce results that are distracting for the observer. We have developed a system to automatically generate a correctly timed lip-synch animation for an audio file of recorded speech based on the corresponding phoneme sequence. Phonemes are the smallest unit of unique speech sounds. Each phoneme has a visual counterpart or mouth shape called a viseme. Using the same process for creating a facial expression, we have created a validated library of visemes. The resulting animation properly simulates human behavior and can be blended with the underlying facial expressions.

6.0 DYNAMIC AND ENGAGING CONTENT

CCAT's instructional content takes the form of scenarios and simulations that embody cultural dilemmas that have been reported by Warfighters and other stakeholders with experience in conducting such missions as Humanitarian Aid / Disaster Relief (HA/DR), Counter-Insurgency (COIN), and Provincial Transition Teams (PTTs). Virtual mentors from each of the services, as shown in Figure 3, introduce compelling stories which embed these dilemmas.

Figure 3. CCAT Virtual Mentors

To meet the requirement for dynamic and engaging conent, CCAT uses Vcom3D's Plug-and-Play Cultural Avatar (PnPCA) system. As shown in Figure 4, PnPCA integrates the physically expressive virtual human role players with a cognitive/affective model of the characters' behavior, as influenced by cultural values. As described in [9], we have used Soar Technology's Cultural Cognitive Architecture (CCA) to provide characters that not only react to learner decisions, but also pursue their own agendas. The CCA combines Hierarchical Value Maps (HVMs), cultural schema, and values-based appraisals to simulate characters whose behaviors are consistent with the culture being modeled.

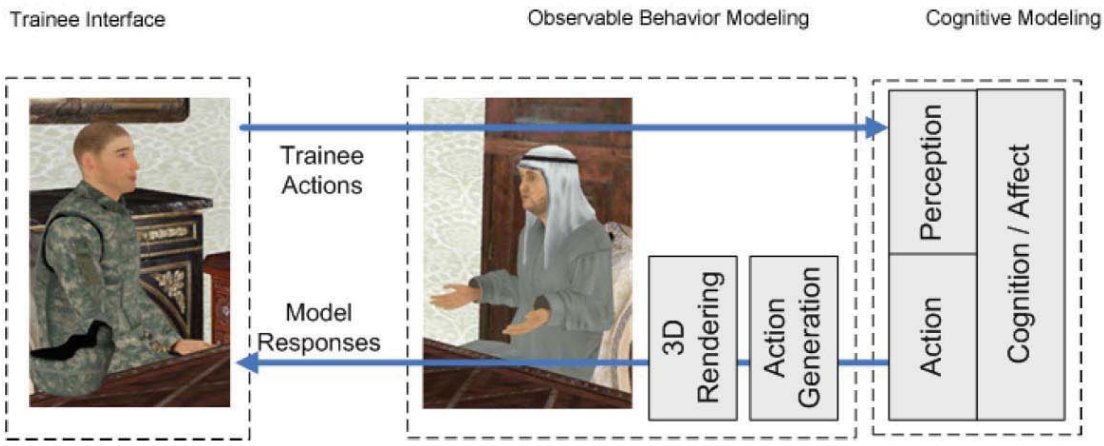

Figure 4. Integration of Physical and Cognitive Model into a Training Environment

7.0 DEVELOPMENT AND DEPLOYMENT

The CCAT process for developing culturally adaptive competency training scenarios begins with a storyboard, which includes the setting, characters, dialog, camera shots, special behavior descriptions, and control flow. Whether the content is to be deployed on the Web, PC, or a mobile platform such as iPhone or Android, the content creation process is nearly the same.

The Content Creation Pipeline, as shown in Figure 5, begins by recording the dialog as performed by native speakers. An animation script author imports each audio file along with the transcript into Vcommunicator Studio which automatically creates the lip-synch animation for the expressive avatar. Vcommunicator Studio is a tool we developed to expedite the character animation process. It allows high-level users, such as subject matter experts or instructional system designers, to directly create the behavior animation sequences. Users work with an intuitive, timeline-based interface to composite body gestures, facial expressions, head and eye tracking, with the lip-synched audio. Studio provides libraries of body gestures which can be filtered by cultural region. The gestures may be layered on top of one another to combine for example the lower body of one gesture with the upper body of another. New upper body gestures can be created with Gesture Builder, imported into Studio, and composited with existing gestures. The user aligns the resulting behaviors to the dialog to create a completed animation sequence for this character.

The animation is then exported for Autodesk's 3ds Max modeling and animation software where it is automatically applied to a reference character. 3ds Max is also used to create other 3D assets such as the environments, vehicles, props, and any special animations.

As shown in the Game Development Pipeline of Figure 5, the 3D assets are prepared for a game development tool. Once such tool is called Unity and is capable of deploying to a wide range of platforms including Android, iPhone, Windows, Mac, and web browsers. Unity includes an integrated development environment (IDE) where 3D assets are connected and controlled through code and an optimized real-time 3D renderer that uses standard graphics libraries.

Within the IDE, the cultural scenario's game logic is implemented. The game logic will control the flow of the scenario, the character movement, and the camera to

form a cohesive user experience. It will also handle user interaction and trigger the behavior animations for each character.

The final step is to deploy the integrated scenario for the target Operating System.

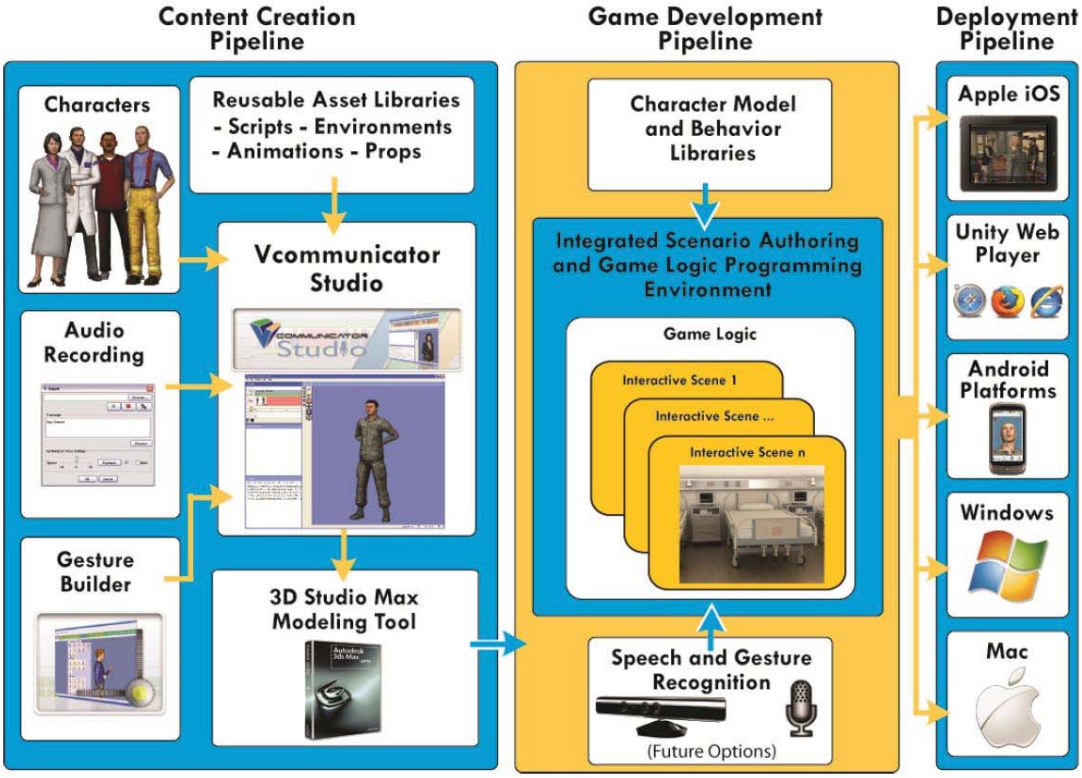

Figure 5. Development and Deployment Pipeline

8.0 EXAMPLE SCENARIO

Utilizing these expressive avatars, we use the described development process to create interactive scenarios depicting authentic cultural dilemmas. The player/learner assumes a defined role with a specific mission objective. Figure 6 shows an example scenario between the player, who assumes the role of a soldier, an Afghan Army officer, who is serving as his interpreter, and a local Afghan Elder. When the player reaches the cultural dilemma with the other stakeholders in the scenario, they are prompted to choose between multiple courses of action. The other stakeholders must be able to respond realistically to enable the student's task of identifying the cultural values dimensions involved in this dilemma.

Figure 6. Example Scenario

The player must also discern the particular values orientation of each of the other stakeholders. Figure 7 shows the cultural values dimension of Ascription vs.

Achievement and the assignment of the stakeholders' values orientation. In the style of modern "reality show" story-telling techniques, the stakeholder's reaction may be augmented with an "after the fact" interview shown immediately after the reaction. Finally the player is given the opportunity to reflect upon their decision and even observe the result of alternate choices.

Figure 7. Mapping Cultural Values

After working through a number of scenarios in which they learn to recognize and respect cultural differences, the learner then has the opportunity to practice these skills in a mission simulation. These simulations require the learner to discern the cultural values of the virtual characters and to use this understanding to successfully complete missions by resolving or reconciling differences.

9.0 CONCLUSION

In this paper we have described the design considerations that we have applied to develop the CCAT system for cultural adaptivity training. This system combines authoring tools and a multi-platform delivery system for embedding expressive computer-generated characters within a serious games scenario. The system facilitates several training objectives not possible with other methods of learning including:

- Learning thorough variable mission outcomes
- Recognition of non-verbal cues
- Inference of cultural elements through observation
- Recovery from maladaptive behavior
- Time sensitivity - performance of actions in a timely manner and in an appropriate sequence

10.0 REFERENCES

[1] Glover, G., Friedman, H., Sims, E., & Silverglate, D. (2011). Developing Culturally Adaptive Competencies: A Dilemmas Based Approach.

[2] Klein, G.A. (1989). Recognition-primed Decisions. In W.B. Rouse (Ed.), Advances in man-machine systems research (Vol. 5, pp. 47-92). Greenwich, CT: JAI Press.

[3] Zbylut, M., Metcalf, K., et al. (2009). The Human Dimension of Advising: An Analysis of Interpersonal, Linguistic, Cultural, and Advisory Aspects of the Advisor Role. US Army Research Institute for the Behavioral and Social Sciences, Technical Report 1248. June 2009.

[4] Sims, E., and Pike, W. (2004). Reusable, Lifelike Virtual Humans for Mentoring and Role-Playing. In 2004 Proceedings of Interservice/Industry Training Simulation & Education.

[5] Sims, E. (2005). Simulating Believable, Context-aware and Culture-specific Human Behaviors. in 2005 Interservice/Industry Training Simulation & Education Conference. Orlando, FL.

[6] Ekman, P. (1971). Universals and cultural differences in facial expressions of emotions. in Nebraska symposium on motivation. Lincoln, NB: University of Nebraska Press.

[7] Matsumoto, D., S. Takeuchi, S. Andayani, N. Kouznetsova, and D. Krupp, The contribution of individualism-collectivism to cross-national differences in display rules. Asian Journal of Social Psychology, 1998. 1: p. 147-165.

[8] Ekman, P. and Friesen, W. (1976). Measuring Facial Movement. Journal of Nonverbal Behavior, 1976 - Springer

[9] Sims, E., & Taylor, G. (2009). Modeling believable virtual humans for interpersonal communication. Proceedings of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), 2009, Orlando, FL.

11.0 ACKNOWLEDGEMENTS

The CCAT training concept described in this paper was developed as a collaboration between Vcom3D and our co-author consultants, Drs. Jerry Glover and Harris Friedman.

This CCAT authoring and delivery system is an extension of the Plug-and-Play Cultural Avatar system, based in part on work performed under US Government contract W91WAW-08-C-0052, monitored by the Army Research Institute and US Government contract W91CRB-09-C-0004, administered by the Combating Terrorism Technical Support Office (CTTSO). Opinions expressed in this paper are those of the authors and do not necessarily represent the views of these sponsoring organizations.