PARAMETERIZED GENERATION OF AVATAR FACE DATASET

Justin N Oursler, Mathew Price, and Roman V. Yampolskiy
Computer Science and Engineering
JB Speed Hall, 106
University of Louisville
Louisville, Kentucky, USA
jnours01@louisville.edu, mahatah@gmail.com, roman.yampolskiy@louisville.edu

KEYWORDS
Artimetrics, Avatar, Dataset, Face, Second Life.

ABSTRACT

Virtual communities such as *Second Life* and *Entropia Universe* are quickly becoming the next frontier of cybercrime. With GDPs of virtual economies of persistent worlds approaching billions of dollars it is necessary to develop tools for protection of virtual environments similar to those utilized to secure real world's infrastructure, such as biometrics security systems. The first step in the development of such security methodologies is wide availability of datasets necessary for training and testing of those biologically inspired systems. In this paper we present a manual and automated approach to generation of parameterized datasets containing facial images of avatars representing typical entities in the virtual worlds. Our work on making such standardized datasets makes it possible to develop novel security systems which function just like biometric recognition systems, but can be applied to recognition and verification of non-biological entities.

INTRODUCTION

Virtual economies of persistent worlds have Gross Domestic Product (GDP)s approaching billions of dollars, far exceeding GDPs of some real countries. Virtual crime results in real financial damages which will only worsen if measures are not taken to secure authentication and communication in virtual reality worlds. In the virtual world increasingly populated by non-biological characters there are just no existing techniques for identity verification of intelligent entities other then self-identification.

Artimetrics, which is defined as the science of recognition and verification of intelligent software agents, domestic and industrial robots and other non-biological entities aims to address this problem. This future oriented sub-field of security has broad applications (Yampolskiy, 2008). Artificially Intelligent programs are quickly becoming a part of our everyday life. Virtual assistants, shopping bots, and smart search engines, to give just some examples, are used daily by millions of people. As such programs become closer in their abilities and intelligence to human beings, the need will arise to recognize and verify the identity of such entities just like it is necessary to authenticate the identity of people. With respect to artificial agents, such reasons include but are not limited to (Yampolskiy, 2008):

- Preventing malicious intelligent software from obtaining access to information or system resources and granting it to authorized agents and by doing so improving security of virtual communities, social networks, and country's cyber-infrastructure especially vulnerable in the post 9/11 world.
- Finding out which agent has performed a given task in case a number of possible alternatives exist, either for demanding responsibility or assigning reward.
- Securing interaction between different pieces of intelligent software or between a human being and an instance of intelligent software/robot.
- Determining who has the authorship rights to the results of computation and creative output produced by an AI entity.
- Identifying semi-autonomous software tools used by hackers to attack systems and networks.
- Making it possible for scientists in fields as diverse as biology, communications, and e-business to securely work with intelligent assistants, and robots.

PREVIOUS WORK

Artificially generated datasets are a common approach in development of biometric security systems. Once a system for producing simulated data is developed it is fast and cheap to produce large quantities of high quality biometric data without any privacy or security concerns to worry about

(Yampolskiy, 2008). Such data can be used for testing newly developed biometric systems, benchmarking well developed security systems, testing scalability of authentication systems or for certification of commercially available packages. Production of synthetic biometrics allows researchers to better understand individuality of biometric patterns and allows parametric sensitivities of algorithms to be investigated in greater detail (Orlans, 2004; Ma, 2005; Makthal, 2005).

Biometric data comes in four different formats: images, feature measurements, matching scores and decision data. Each type can be potentially simulated for experimental purposes (Ma, 2005).

- **Images** This is the best researched type of synthetic biometric represented by the raw image as it comes out from a digital scanner or a camera and depicts a two-dimensional view of the body part/structure in question (Ma, et al. 2005). Fingerprint, face and iris images are examples of the most frequently produced artificial data (Cappelli, 2000; Cui, 2004; Ayers, 2004; Cappelli, 2002; Sumi, 2006; Makthal, 2005).
- **Matching Scores** Similarity functions attempt to measure closeness of the input feature vector to the one stored in a template. A decision on user verification or recognition is produced by the system based on such a measurement. Generation of matching scores may be valuable for testing a model for the actual distribution of matching scores for a particular biometric (Wein, 2005; Ma, 2005).
- **Decision Data** This is the highest level of biometric data which can be generated. All biometric systems produce a binary decision in terms of authenticating the user or not. Synthetic decision data can be used to estimate error rates produced by a biometric system (Schuckers, 2004).
- **Feature Measurements** Extraction of statistically valuable information from the raw images produces feature measurements such as the number of minutiae points in a fingerprint or the inter-eye distance in face recognition.

Currently well researched synthetic biometric data generation approaches exist for many physical biometrics such as fingerprint (Cappelli, 2000; Ayers, 2004; Cappelli, 2002), face (Sumi, 2006), and iris (Cui, 2004; Makthal, 2005). Additionally, ability to generate many others comes from other fields such as computer graphics for face (DeCarlo, 2008; Gao, 2000) and iris (Lefohn, 2003), speech synthesis for speech (Taylor, 1999), document image analysis for signatures (Oliveira, 1997), and speech processing for lips (Reveret, 1998).

Virtual World Selection

For the purposes of avatar facial data generation, various virtual worlds and avatar creation software were considered based on the needs of this project. These included:

1) Mutable attributes to avatar facial features
2) Ability to view avatar from multiple angles
3) Selection of contrasting facial features in generating new avatars
4) General ease of use and versatility

While many virtual worlds such as *Entropia Universe* and avatar creation software such as *Poser* were considered, the Massively Multiplayer Online Role Playing Game (MMORPG) *Second Life* (Second Life, 2009) was found to best fit the above criteria. Second Life, a three-dimensional, virtual world where the Internet community creates avatars to interact with one another in real-time, provided the largest, quickest, easiest to use, and most versatile set of avatars and creation materials of all software considered.

Second Life's benefits were many in consideration for the creation of a dataset, but four attributes stood out among all others. First, dozens of physical facial attributes, such as length and right-to-left symmetry, could be adjusted to create a truly unique avatar. Secondly, Second Life implemented an avatar randomizer which can create a new, unique avatar with the press of a button. Thirdly, camera pan, tilt, and zoom are controllable by the user which makes it possible to gather several angles from the same avatar. Finally, Second Life allows the user to manipulate environmental elements through the use of an in-game scripting language, adding versatility to the collection of data.

DEVELOPED METHODOLOGY

Manual Approach Each unique avatar face was randomly generated and saved via the Gadwin PrintScreen (Screen, 2009) application which allowed us to quickly capture the screen and save it in the desired directory as a certain file type. The following steps were taken to capture the profile of each avatar:

1) Open the Appearance Menu and randomized each category of the avatar's physical appearance, ending with the eyes. By ending

with the eyes, the focus would automatically go to the avatar's face, allowing consistency in the dataset.

2) Adjust the view of the screen, putting the avatar in the center of the screen, capture the screen, and save the image in the dataset directory with the appropriate filename.
3) Rotate the camera to the left where the avatar is still facing forward but we can still capture a left-handed angle in between this view and the center view. Then, capture as before.
4) Rotate the camera in between the far left view and the center and then capture.
5) Rotate the camera to the far right, again, so the avatar is still facing forward but where a capture can still be performed between this view and the center view. Then, capture the screen.
6) Rotate the camera between the center view and far right view and capture the screen.
7) Focus the camera on the center again, but then rotate the camera to a view below the avatar, with the avatar still facing forward. Then, capture as before.
8) Rotate the camera straight through the center to a view above the avatar with the avatar still facing forward. Then, capture the screen.

These eight steps generated one image set for each avatar. Problems with this approach mainly consist of possible human error and irregularities in the collected data. Though a possible problem, this could also be viewed as a benefit since a true application of our research will likely work with noisy data. Giving the future algorithms for avatar face detection and recognition a random selection of views makes them better equipped for real-world applications. Also, manually producing avatar facial database is a very time consuming task, and so is not a feasible approach for larger size datasets.

The dataset generated by the manual approach consists of one hundred different avatars with seven pictures from different angles, totaling seven hundred images. The images were captured as Tagged Image File Format (TIFF) (TIFF, 2009) images at a resolution of 1280 X 1240, resulting in each image being 3,843 KB in size (see Figure 1-4). The TIFF format was used because it is a flexible format that allows for customization in its tags (image information) and it is uncompressed, maintaining the quality of the image. The first fifty avatars (three-hundred fifty images) consist of randomly generated male avatars and the last half of the images is of female avatars. An angle from the a) front, b) far left, c) mid left, d) far right, e) mid right, f) bottom, and g) top, is captured for each avatar. The images are named in a consistent format; stating the program, character, and angle. For example, the image SL-051b.tif, is a capture using *Second Life* (SL) of the 51st character (051) from the far left (b) that happens to be female because it is in the last half of the dataset.

Automated Approach For a second approach, we designed and implemented a scripting technique to automate the above process. Using the programming language AutoIt (AutoIt, 2009) as well as a scripting language native to *Second Life*, better known as Linden Scripting Language (LSL) (LSL, 2009), a successful generation of random avatars was achieved. The following is a walkthrough of this process for the creation of one hundred randomly generated avatars:

1) Using the scripting language AutoIt, it was possible to simulate key presses and mouse control in a Windows environment. During the first run of the AutoIt script, simulated keyboard commands are used to circle the *Second Life* camera around the avatar such that the front of the avatar's face is exposed.

2) The script is paused and requests the user to center the avatar's face with the horizon using the movement control. This is only needed on the first run and constitutes the last interaction with the user.

3) The AutoIt script then activates the LSL script by clicking on a button attached to the avatar's hub.

4) The LSL script locks the *Second Life* camera's position and rotation as well as controls from the game's automated functions (such as camera changing on clicking).

5) The AutoIt script then takes a screen shot of the avatar using the *Second Life* tool "screen shot". The script then labels avatar "Avatar 'x' face 'y'", where x corresponds to the number of avatar created (1 - N) and y corresponds to the screen shot for that avatar (1 – 10).

6) The script then zooms into the avatar's face before taking another screen shot and using the same labeling system as in step 5.

7) The AutoIt script then rotates the camera at eight specific angles (upper left, center left, lower left, upper center, lower center, upper right, center right, and lower right) taking screen shots at each.

8) The script then selects "edit", then "appearance", bringing up the avatar editing tool. From here the script randomizes a body for a new avatar. Body

height, torso length, and leg height all must be set to 50% in order to preserve the camera angle, which is done automatically by the script.

9) The AutoIt script then clicks on the body parts sub menu items “skin”, “hair”, and “eyes” randomizing each of them as they are entered. The save “all button” is pressed, saving the avatar to begin the screen shot process again.

10) The script zooms away from the avatar before taking the new avatar's center body screen shot.

After step 10, the AutoIt script restarts at step 7 until one hundred avatars are created and one thousand images have been taken. A sample segment of Autoit source code responsible for GUI interaction is given below:

```
Func snapshot ($picture)    dim $picture
mouseClick("Left", 440, 756, 1)   ;snapshot button
sleep(2000)
mouseClick("Left", 102, 296, 1)   ;save button
sleep(3000)
send("{DOWN}{ENTER}")
findname($picture)                    EndFunc
```

The datasets generated by the scripted approach consists of avatars with ten pictures from different angles. The images captured are in the Portable Network Graphics (PNG) (PNG, 2009) format at a resolution of 1024 X 768 resulting in each image being between 110KB and 450KB in size. One upper body picture is taken as well as nine facial pictures, all differing in angles. These angles include the top, center, and bottom of the left, center and right side of each avatar's face. The images are named in a consistent format; stating the program, gender, avatar number, and angle. For example, the image “SecondLife Male Avatar 4 gesture 5.png” refers to the image of an avatar that looks like a male character, the fourth in the dataset, and the fifth picture taken in this avatar's set of 10 (see Figures 5-8). The gender of the avatar is dependent upon the user’s selection at the beginning of the process.

CONCLUSIONS AND FUTURE RESEARCH

Now that we have a completed dataset of avatar faces, our next step is to utilize the dataset in different security related experiments. We will first experiment with avatar face detection and then face recognition. Face detection is the process of determining whether or not there is a face in the image and precisely locating it. The experiments involving face detection will consist of utilizing and improving existing face detection algorithms currently used for human face detection in biometric applications. After we are successfully able to detect faces we will begin experimentation with face recognition algorithms. Face recognition aims to identify if a specific face is included in the dataset of previously enrolled faces. Again, experiments will consist of applying existing biometric algorithms and improving their design until we can successfully authenticate avatars via face recognition. First, our algorithms will only be tested with still photos taken from in-game screenshots. This allows us to keep the problem simple for testing purposes. Our ultimate goal is to be able to utilize an algorithm that can detect and recognize an avatar in a three-dimensional environment like in the virtual world *Second Life*.

Being able to identify software agents not just by certain specific codes (ID numbers, serial keys, etc.), but also visually adds extra security to the authentication process. If a person, through an avatar, is wanting to exchange money for goods and services in an application such as *Second Life*, the program can authenticate that avatar by not only their factual information but also by the way they are expected to look, which provides a second level of authentication or alternatively allows for decentralized (serverless) virtual government with vibrant economy. The security in our real world is ever changing and the security in our virtual world must do the same in order to keep up with the real crimes committed via wires and waves. No longer do we only need to worry about human faces and identities but non-human entities with real abilities to cause harm.

BIOGRAPHIES

Justin N. Oursler is currently studying Computer Engineering and Computer Science at the J.B. Speed School of Engineering at the University of Louisville. Justin is an active member of the Sigma Chi fraternity and representative of the Speed School Student Council.

Matthew E. Price holds an undergraduate degree in Justice Administration from the University of Louisville. He is currently pursuing an MS degree in Computer Science and Engineering from the University of Louisville. Mr. Price's interests include computer forensics and biometric systems.

Roman V. Yampolskiy holds a PhD degree from the department of computer science and engineering at the University at Buffalo. He is currently an assistant professor at the University of Louisville. Dr. Yampolskiy’s main areas of interest are computer security and biometrics. Dr. Yampolskiy is an author of over 40 publications including multiple books.

Fig. 1. Manually collected image SL-002a.

Fig. 2 Manually collected image SL-024c.

Fig. 3. Manually collected image SL-054e.

Fig. 4. Manually collected image SL-095g.

Fig. 5. Automatically collected image Second Life Male Avatar11 gesture3_001.

Fig. 6. Automatically collected image Second Life Male Avatar11 gesture5_001.

Fig. 7. Automatically collected image Second Life Male Avatar22 gesture2_001.

Fig. 8. Automatically collected image Second Life Male Avatar22 look7_001.

REFERENCES

Ayers, T., R. Federkeil, et al. (2004). Modeling Fingerprints: Components for the Task of Synthesis. Proc. BT2004 Int'l Workshop on Biometric Technologies, AB, Canada.

Cappelli, R., D. Maio, et al. (August 2002). Synthetic Fingerprint-Database Generation. in proceedings 16th International Conference on Pattern Recognition (ICPR2002), Quebec City, Canada.

Cappelli, R., D. Maio, et al (2000). Synthetic Fingerprint-Image Generation. 15th International Conference on Pattern Recognition (ICPR'00).

Cui, J., Y. Wang, et al. (23-26 August 2004). An Iris Image Synthesis Method based on PCA and Super-resolution. IEEE International Conference on Pattern Recognition, United Kingdom.

DeCarlo, D., D. Metaxas, et al. (1998) An Anthropometric Face Model using Variational Techniques. SIGGRAPH '98.

Gao, W., J. Yan, et al. (September 2000). Individual 3d face synthesis based on orthogonal photos and speech-driven facial animation. In IEEE International Conference on Image Processing (ICIP), Vancouver, Canada.

Jermyn, I., A. Mayer, et al. (August 23-36, 1999). The Design and Analysis of Graphical Passwords Proceedings of the 8th USENIX Security Symposium, Washington, D.C.

Lefohn, A., R. Caruso, et al. (November 2003). An Ocularist's Approach to Human Iris Synthesis. IEEE Computer Graphics and Applications.

Ma, Y., M. Schuckers, et al. (October 2005). Guidelines for Appropriate Use of Simulated Data for Bio-Authentication Research. 4th IEEE Workshop on Automatic Identification Advanced Technologies (AUTO ID), Buffalo, NY.

Makthal, S. and A. Ross (September 2005). Synthesis of Iris Images using Markov Random Fields. in Proc. of 13th European Signal Processing Conference (EUSIPCO), Antalya, Turkey.

Oliveira, C. d., C. A. A. Kaestner, et al. (November 1997). Generation of Signatures by Deformations BSDIA'97, Curitiba,Brazil.

Orlans, N. M., D. J. Buettner, et al. (2004). *A Survey of Synthetic Biometrics: Capabilities and Benefits*. International Conference on Artificial Intelligence (IC-AI'04), CSREA Press.

Reveret, L. and C. Benoit (Dec. 4-6, 1998). A New 3D Lip Model for Analysis and Synthesis of Lip Motion in Speech Production. Proc. of the Second ESCA Workshop on Audio-Visual Speech Processing, AVSP'98, Terrigal, Australia.

Schuckers, M. E., A. Hawley, et al. (2004). A comparison of statistical methods for evaluating matching performance of a biometric identification device- a preliminary report. Proceedings of SPIE Conference on Biometric Technology for Human Identification, Orlando, USA.

Sumi, K., C. Liu, et al. (Jan. 2006). Study on Synthetic Face Database for Performance Evaluation. International Conference on Biometric Authentication (ICBA2006).

Taylor, P. and A. Black (1999). Speech synthesis by phonological structure matching. In *Eurospeech99*, Budapest, Hungary.

Thorpe, J. and P. v. Oorschot (December 6-10, 2004). Towards Secure Design Choices For Implementing Graphical Passwords. 20th Annual Computer Security Applications Conference, Tucson, Arizona

Wein, L. M. and M. Baveja (May 2005). Using Fingerprint Image Quality to Improve the Identification Performance of the U.S. Visitor and Immigrant Status Indicator Technology. In Proceedings of the National Academy of Sciences.

Yampolskiy, R. V. and V. Govindaraju (2008). Behavioral Biometrics for Verification and Recognition of Malicious Software Agents. Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VII. Orlando, Florida.

Yampolskiy, R. V. and V. Govindaraju (2008). Generation of Artificial Biometric Data Enhanced with Spatial-Temporal and Environmental Information. Biometric Technology for Human Identification V. Orlando, Florida.

Second Life | What is Second Life? [cited 2009 March 01]; Available from: http://secondlife.com/whatis/.

Screen Capture and Print Screen software. [cited 2009 March 01]; Available from: http://www.gadwin.com/printscreen/.

AutoIt Script Home Page. [cited 2009 March 01]; Available from: www.autoitscript.com/autoit3/index.shtml.

LSL Wiki: HomePage. [cited 2009 March 01]; Available from: http://www.lslwiki.net/lslwiki/wakka.php?wakka=HomePage.

TIFF. [cited 2009 March 01]; Available from: http://partners.adobe.com/public/developer/tiff/index.html.

PNG (Portable Network Graphics) Home Page. [cited 2009 March 01]; Available from: http://www.libpng.org/pub/png/.