Exigent: An Automatic Avatar Generation System

Dominic Kao, D. Fox Harrell
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA
{dkao,fox.harrell}@mit.edu

ABSTRACT

Avatars are pervasive in video games and virtual worlds. The automatic generation of these avatar promises to reduce player effort and provide system-defined mappings between "real" (physical) player characteristics and virtual identities. We present an avatar generation system called Exigent; given a photograph of a human face, Exigent creates an avatar. Exigent leverages two recent computer vision packages and extends them with a custom facial analysis and generation sub-module. The generated avatars incorporate the user's gender, face shape, hair color, facial expression, eye shape, and more[1]. This system is different from most avatar creators in its consideration of finer-grained details such as curvature of the face, distance between eyes, size of features, etc.

Categories and Subject Descriptors

K.8.0 [**Personal Computing**]: General – Games

Keywords

Automatic avatar creation, avatar generation

1. INTRODUCTION

Virtual identities, in the likeness of a specific user or not, is common across games, social networks, and learning platforms. Facial recognition is used to generate a Nintendo Mii from a photo of the user's face on the 3DS [5]. Games such as Bloodmasque [6] and FIFA [1] also let players use their face photos in game avatars. Avatars can be used in a wide array of settings (e.g., images on forums, characters in games, etc.). The modularity of the images made by our system gives it a broad range of applications.

Mapping between an input photo to a virtual avatar can be understood as a "semiotic morphism" [3, 4]. A morphism is a structure-preserving mapping from one set of signs to another (e.g., choosing a file name, choosing a suitable icon, translating a LaTeX document to html [2]). The design of our own system aims to make its morphism flexible, such that representing certain aspects of a user's "real" characteristics can be altered, enhanced, or discarded altogether. Here, we present Exigent, a system for automatically generating avatars from an input image of a human face.

[1]Video demo: https://youtu.be/rvu7J_Xor5Y

Proceedings of the 10th International Conference on the Foundations of Digital Games (FDG 2015), June 22-25, 2015, Pacific Grove, CA, USA.
ISBN 978-0-9913982-4-9.

2. PRIOR WORK

Facial recognition has been used to generate avatars [1, 5, 6]. However, many of these avatar creators require photos or video of the user from multiple angles [1, 7]. The ones that do not are constrained in their ability to detect finer-grained details accurately; this is even more onerous a task when considering less than ideal conditions (e.g., poor lighting, etc.). We overcome many of these limitations by leveraging recent computer vision advances [8, 9] that outperform commercial systems, and augment this with our own system that provides additional analysis, classification, and generation.

3. SYSTEM DESIGN

The aim of the system design was to make components that are adaptable. For instance, should one wish to extend the system to use a new set of assets (e.g., modify the art style), or add procedures (e.g., a "sepia" tint), it is simple to do so.

The core components of Exigent, forming the computational basis for generating likeness can be seen in Figure 1. The *Landmark Localization* procedure finds points of interest on the human face, such as the eyes, nose, etc. [9]. Here we also find the degree of facial curvature in the face. The *Attribute Extraction* procedure determines high-level attributes such as gender, and the presence of facial hair and accessories [8].

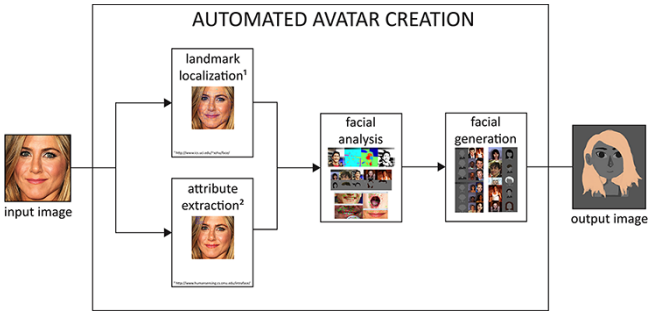

Figure 1: System overview.

The *Facial Analysis* and *Facial Generation* sub-module performs many analyses. This includes Procrustes analysis to

find a suitable face shape, ascertaining hair color, and generating output. The final avatar is a 3,000x3,000 pixel image.

We calculate the size of individual facial features (e.g., eyes, nose, mouth, etc.) that affect corresponding output features. Positions also matter; for instance, the relative position of eyes are identified as wide-set, proportioned, or close-set. We also capture some facial expressions (e.g., smiling, surprised, etc.). See Figures 2 and 3 for examples[2].

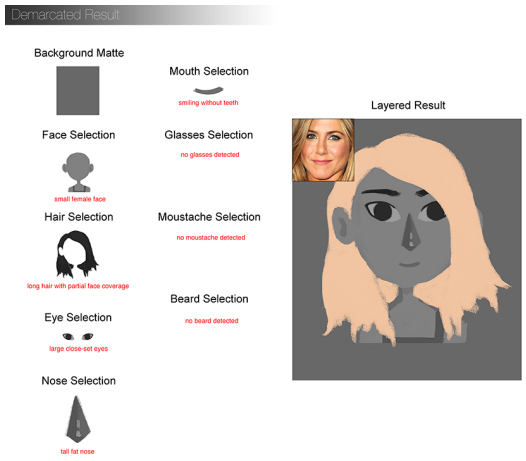

Figure 2: Layering process.

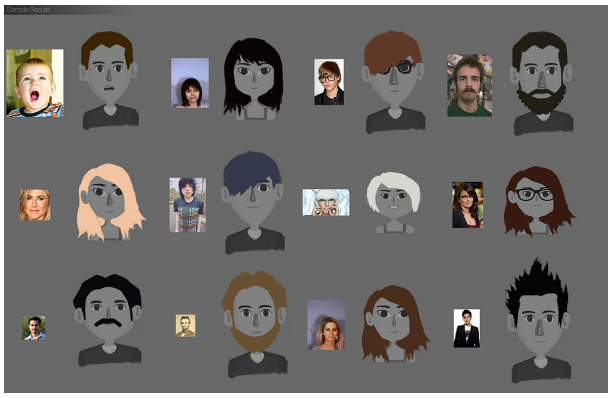

Figure 3: Sample results.

4. EVALUATION

We conducted an evaluation using Amazon's Mechanical Turk (MT). We completed an assessment as to whether the avatars created by Exigent bore resemblance to the original face. We were interested in whether our semiotic morphism preserves resemblance of users. We chose to compare avatars automatically generated by our system to ones generated at random (with the correct gender).

There were a total of 13 participants, restricted to Amazon "Master" workers. Using a total of eight images, these were processed once using our system and once with assets and parameters chosen at random. Participants were asked to evaluate the original images vis-à-vis the avatars. We then asked participants a series of Likert-scale questions (e.g., the avatar's eyes bear resemblance to those in the human face).

[2]In these examples, hair style was *not* picked automatically.

5. RESULTS

The scores collected are based on a 5-point Likert scale (1: *Strongly Disagree* to 5: *Strongly Agree*). See Table 1.

Criterion	Exigent Avatar	Random Avatar
Overall Resemblance	3.3	2.4
Eyes Resemblance	2.9	2.7
Nose Resemblance	3.4	3.1
Hair Resemblance	3.1	2.4
Face Shape Resemblance	3.7	2.4
Mouth Resemblance	3.4	2.4
Hair Color Resemblance	3.8	1.6

Table 1: Mean scores averaged over all test images.

Users were asked a comparative question asking if the system-generated avatar bore a higher resemblance than the randomly generated one. The mean comparative resemblance score was 4.2, suggesting that on average, participants agree that the system generated avatar bears higher resemblance.

6. CONCLUSION

We have presented the Exigent system, which creates avatars given input in the form of an image of a human face. The system has been designed as a series of extensible, modular components. Therefore, alterations to an individual component is painless. We have also done a user evaluation that demonstrates our system is creating a semblance of likeness.

Additionally, we have prototyped a functional web service with which users can upload their photos for processing. In roughly 20-30 seconds time, the server returns Exigent's avatar. We hope that systems like Exigent testify to the potential of automatic avatar generation; reducing player effort, and forging inventive approaches to mapping the photoreal onto the virtual.

Future work may involve generation of avatars in a variety of art-styles or with specific types of systematic alterations: e.g., preserving some features and transforming others (such as fashion or gender), caricature, generating NPCs related to the users' social group, and more.

7. REFERENCES

[1] Electronic Arts. Game Face, 2015.

[2] J. Goguen. Semiotic morphisms. 1997.

[3] J. Goguen. An Introduction to Algebraic Semiotics, with Application to User Interface Design. *Computation for metaphors, analogy, and agents*, 1999.

[4] D. F. Harrell. *Phantasmal Media: An Approach to Imagination, Computation, and Expression.* The MIT Press, 2013.

[5] Nintendo. Nintendo 3DS, 2011.

[6] Square Enix. Bloodmasque, 2013.

[7] Villa-Uriol et al. Automatic creation of three-dimensional avatars. In *Electronic Imaging (EI'03) -Videometrics VII*, 2003.

[8] X. Xiong et al. Supervised Descent Method and its Applications to Face Alignment. *CVPR*, 2013.

[9] X. Zhu et al. Face Detection, Pose Estimation, and Landmark Localization in the Wild. *CVPR*, 2012.