FaceMaker—A Procedural Face Generator to Foster Character Design Research

Valentin Schwind, Katrin Wolf and Niels Henze

Abstract Understanding the effect of facial features on human's perception and emotion is widely studied in different disciplines. In video games, this is especially important to improve the design of virtual characters and to understand their creation process. Virtual characters are widely used in games, virtual therapies, movie productions, and as avatars in e-commerce or in e-education. Studying the design of virtual characters is challenging as it requires to have tools at hand that enables the creation of virtual characters. Therefore, we developed a system that enables researchers to study the design process of virtual faces as well as the perception of such faces. We developed a 3D model of the Caucasian average face and implemented design parameters that can be manipulated to change the face appearance. We integrate the face creation system into a web application, which allows us to conduct studies in the large. The application has been validated through a cluster analysis of procedurally generated faces from 569 participants which created 1730 faces.

1 Introduction

Virtual characters are commonly used in games, virtual therapies, movie productions, and as avatars in e-commerce or in e-education solutions. Due to the recent success of virtual reality and head-mounted displays, the importance of virtual characters will likely further increase over the following years. The success and the acceptance of an avatar in a movie, a game, or a therapy highly depend on how the

V. Schwind (✉) · N. Henze
University of Stuttgart, Stuttgart, Germany
e-mail: valentin.schwind@vis.uni-stuttgart.de

N. Henze
e-mail: niels.henze@vis.uni-stuttgart.de

K. Wolf
HAW Hamburg, Hamburg, Germany
e-mail: katrin.wolf@acm.org

O. Korn and N. Lee (eds.), *Game Dynamics*,
DOI 10.1007/978-3-319-53088-8_6

character is perceived. In particular, the general appearance, the facial features, and the character's emotions influence the reaction of the player, customer, audience, user, or patient. Despite a body of work from different disciplines, the effect of a character's appearance is not completely understood.

Furthermore, self-customization and individualization of personal avatars are commonly used mechanisms to strengthen the connection between users and their virtually created characters. Game developers know that this connection intensifies the immersion in virtual worlds. Today, avatar creation systems are major aspects, especially at the beginning, of current role-play games (RPGs). A lot of recent game series or titles (Fallout, The Elder Scrolls, World of Warcraft, Mass Effect, The Sims) make use of individual customizations of virtual avatars to increase the emotional relationship between players and the game.

Avatar generators are also used in a non-gaming context. Virtual chat applications such as in AltspaceVR, Twinity, or at IMVU.com offer customized appearances and styles of virtual and help users to express themselves and to communicate their personal attitudes. With virtual clothing e-shopping assistants and virtual fitting rooms, such as clothing techniques presented by FITLE.com, users can scan their own bodies using their smartphone and customize their virtual self with garments before buying them.

However, little is known about how users create a virtual appearance of procedurally generated faces. To determine preferred characteristics of virtual faces, we conducted a user study in the large, based on an avatar generation system, which we called faceMaker. These results were already presented in a foregoing study (Schwind et al. 2015). In this article, we explain the system in detail and present the results of a further validation study of analysis using cluster algorithms. This demonstrates how the system can be used to extract user data and draw conclusions. We show how the system works and how it can be extended through further research. Thus, our contribution delivers a new kind of open-source tool at hand for game developers and researchers to understand how people create and perceive their own virtual creations.

Our work aims to provide a research tool that enables researchers of various VR domains, such as games, education, and therapy, to investigate effects of facial features and the look and appearance of a face on the perception of the virtual character, such as their influence on UX and emotions or even on the effect on the acceptance of an application. Similar to state-of-the-art role-play games (RPGs), our tool enables users to personalize virtual faces through providing a rich variety of parameters and options. We created faceMaker, a browser-based and real-time face generation application to provide the research community with a tool to investigate virtual faces in laboratory studies as well as in the large. We enable individual customizations of virtual faces to conduct reliable studies using a unified and complete unbiased avatar face model. Moreover, our web-based tool can be used as research apparatus in participatory studies where participants use our tool and researchers analyze the face design and measured results.

In this paper, we present main features of that system and explain how it can help to discover new findings in the perception of avatars and consequently in human–avatar interaction. The main contributions of our systems are (1) an open-source avatar creation system based on the average face, (2) a web-based user engaging a research-in-the-large tool, (3) the complete recording system of user activities and facial preferences, and (4) a back-end for data exploration and aggregating results.

2 Related Work

The connection between the appearance of a virtual character and ways of its customization through users is a research topic affecting many disciplines while being underexplored. For example, it has been found that physical attractiveness positively influences the opinion about other properties of a person (Dion et al. 1972; Eagly et al. 1991). Moreover, it is widely accepted that childish facial features, such as big eyes, trigger sympathy and let us find a character attractive (Langlois et al. 1991). In contrast, artificial characters, as robots, especially those who try to achieve human likeness are perceived as uncanny (Mori et al. 2012). These aspects can be transferred to virtual characters and often depend on factors as aesthetics and cultural context (Schwind 2015). However, it is currently unknown how customization and individualization of characters influence these aspects.

In human–computer interaction research and psychology, recent work is especially interested in the use of character creation systems to learn how virtual characters are perceived and how they are designed. Chung et al. (2007), for example, suggest that the process of avatar creation leads to a stronger sensation of their cyber-self-presence and psychological closeness to their customized character. This is supported by a study of Bessière et al. (2007), which indicates that a player's self-customized character in World of Warcraft leads to more favorable attributes than their own self-rated attributes, especially for people with a lower psychological well-being.

Avatar generation systems can also help to examine and identify differences between certain groups of players: Heeter et al. (2008), for example, observed teens in designing games over three years and found that females use a very high level of avatar customizations, while males rather use predefined characters. A study by Rice et al. (2013) indicates similar interests in customization of an avatar between different age groups. Ducheneaut et al. (2009) compared avatar creation systems for three different virtual environments. Their study shows that users emphasize the capability to change body shapes, hairstyles, and hair colors.

Studying virtual characters and especially their design requires character creation tools. Developing usable character creation tools can be challenging and requires significant effort. Various demands on face model and on the system, like visual quality, adaptability, extensibility, and a neutral base model, impede the fast development of reliable avatar creation systems.

Apostolakis and Daras (2013) developed the reverie avatar authoring tool RAAT with the aim to provide a research tool that allows individual customizations of whole-body characters. The online tool is based on a JavaScript library that helps developers and researchers to address avatar feature requirements. One of the application's features is a server-side module, which allows users to automatically incorporate mesh and texture of their own face on an avatar. RAAT supports interactive whole-body customizations like clothes, hairstyles, and changing face textures but no parametrized morph blendings of a neutral face model like the average face.

3 Parametric Average Face

3.1 Model Requirements

One aim of our system is to offer a solution that enables researchers to understand how people design characters and how the designed characters look like. For reliable investigations, the initial state of an avatar face should meet the following requirements: (1) The face should contain a minimum of characteristics which could be preferred or rejected by participants. (2) The model should have a balanced and uniform distance from typical facial proportions for procedural changes. (3) The start face must originate from the surveyed population (in our case Caucasoids). (4) The start face should not include any additional or biasing content.

To fulfill the four requirements, we developed a model of the Caucasian average face which is widely used in anthropomorphism and attractiveness research, e.g., (Langlois et al. 1994; Rikowski and Grammer 1999). Previous work primarily relied on image compositions of photographs. However, composing multiple images of human faces removes skin details and leads to unrealistic facial symmetry. The psychological research found that especially these two properties lead to higher attractiveness (Grammer and Thornhill 1994). To enable researchers to determine when people try to add (or to remove) realistic features, we decided to restore these properties and to introduce the skin details parameter—a realistic skin texture with asymmetrical and irregular skin details. As the average human face has any individual characteristics, researchers are now able to implement additional parameters for their experiment and minimize unwanted effects.

3.2 3D Face Model

The base face of our system is a 3D model of the Caucasian average face aiming for a neutral appearance as starting point for individual parameter configuration of facial features. Based on these parameters, corresponding 3D models can be derived. Thus, the generating system enables morphings of the model as well as blendings of texture maps to customize the appearance of the face.

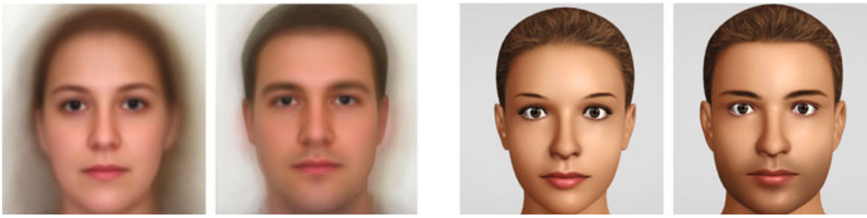

Fig. 1 *Both images left* Image compositions of the female and male average face, *both images right* average 3D face models

To address gender-related differences between females and males, we developed develop two separate average faces. The result based on image compositions including a large number of frontal images of neutral human faces. Images of 117 males and 151 Caucasian people from 18 to 40 years were retrieved from the online Face Database of the Parking Aging Mind Laboratory[1] (PAL) Database and from 3d.sk.[2] Children and older adults were not taken into account for the construction of the face model. This should be considered in further revisions or branches of the system. We used the automatic morphing method of PsychoMorph[3] developed by Tiddeman et al. (2001) to compose both average faces of adults (see both images on the left of Fig. 1).

The average faces were used as input for the PhotoFit feature of FaceGen[4] to create a first 3D model of the face. However, irregular polygon sizes, small artifacts, triangles that avoid subdivision smoothing, as well as the low image resolution of the calculated texture, were not useful for our purposes. Therefore, both meshes were retopologized, subdivided, and retextured using Autodesk Maya™ 2014 and Mudbox™ 2014 by two experienced CGI artists. The final results are shown in both renderings at the right in Fig. 1.

Physical attractiveness based on facial symmetry and golden ratio is considered as a result of averaging faces (Grammer and Thornhill 1994; Langlois et al. 1994). In order to determine whether the generated average faces met the assumption of facial symmetry and golden ratio, we applied Stephen Marquardt's φ-Mask as suggested by Prokopakis et al. (2013). The mask was developed to determine physical attractiveness and to determine deviations of facial symmetry. Thus, we assume that the φ-Mask can be applied to the generated 3D average faces. Figure 2 shows how Marquardt's φ-Mask is applicable to the female as well as to the male 3D average face.

---

[1]Parking Aging Mind Laboratory Database: http://agingmind.utdallas.edu/facedb.

[2]3d.sk Database: https://www.3d.sk.

[3]PsychoMorph Software: http://users.aber.ac.uk/bpt/jpsychomorph.

[4]FaceGen Software: http://www.facegen.com.

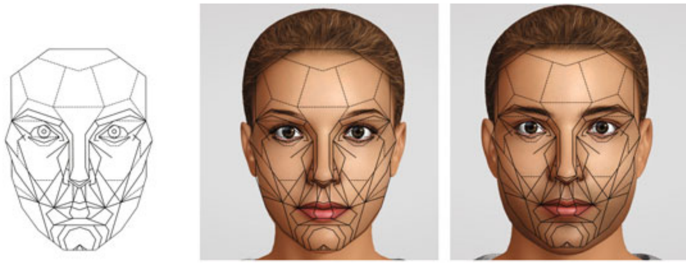

Fig. 2 Facial graph of Marquardt's golden ratio and symmetric φ-mask applied on both average faces

3.3 *Parametrized Morphings*

The female, as well as the male average face, can be continuously morphed using the face gender parameter. Additionally, according to our research questions investigated in our previous work (Schwind et al. 2015), we introduce three common character parameters: skin brightness, face style, and hair color. Face details were introduced to counteract the resulted skin smoothness of the average face.

To get an impression of prevalently used facial customization techniques, we examined the avatar customization systems of 9 commercial RPG: Mass Effect III, The Elder Scrolls: Oblivion/Skyrim, The Sims II/III, World of Warcraft, Destiny, Dragon Age I/II. Similar parameters (e.g., nose length and size) were aggregated, and inappropriate features (e.g., tattoos, scars, elf ears) were not taken into account. Due to their complexity and lack of parameterization, additional hair styles were not developed yet. Because of the limited space of a user interface and because all the parameters should be treated and distributed equally, we chose a maximum of 32 parameters. Table 1 shows all implemented parameters, their impact from left (−100%) to right (+100%) and the default value (*). Every parameter corresponds to a certain target morphing model (m) or texture blending (t). All morphing targets were modeled by hand. Model changes were only performed by vertex transformations in areas of the corresponding facial region (Fig. 3).

Morphings (also known as blend shapes or vertex displacements), as well as texture blendings, are used to change the appearance of the average face and to allow multiple parameterized customizations at the same time. In order to structure all parameters meaningfully, we developed a classification scheme that groups adjustments in eight facial domains (c.f. Fig. 3): common face parameters, eyes, eyebrows, nose, outer face, jaw and cheeks, mouth and lips, and makeup. Interface groups and infographics of the final application were separated according to this scheme. Currently, the system does not support facial domains or blendings. However, asymmetric surface details can be added using the skin detail texture blending, which includes no symmetries between the left and right half of the average face.

Table 1 All facial parameter scales

Parameter	–	Values	+	Type
Face gender	Female	Androgynous*	Male	tm
Face style	Realistic*		Cartoon	m
Face details	None	Half*	Full	t
Skin color	Black	Average*	White	t
Hair color	Black brunette	Med. Blonde*	Red bright blonde	t
Eyes color	Black brown*	Amber blue	Lt. Blue Green	t
Eyes shape	Droopy down	Round oval*	Almond up asian	m
Eyes opening	Narrow	Average*	Wide	m
Eyes size	Small	Average*	Big	m
Eyes height	Up	Average*	Down	m
Eyes distance	Narrow	Average*	Wide	m
Eyes orbit	Bulgy	Average*	Cavernous	m
Eyes rotation	In	Average*	Out	m
Eyebrows color	Black brunette	Med. Blonde*	Red bright blonde	t
Eyebrows shape	Pointed straight	Average*	Round hooked	m
Eyebrows strength	Thin	Average*	Thick	t
Nose shape	Snub	Average*	Hooked	m
Nose length	Short	Average*	Long	m
Nose width	Thin	Average*	Thick	m
Nose bridge	Thin	Average*	Thick	m
Nose cartilage	Round	Average*	Flat	m
Forehead size	Down	Average*	Up	m
Ear size	Small	Average*	Big	m
Throat size	Thin	Average*	Thick	m
Jaw shape	Triangle	Average*	Squared	m
Jaw length	Long	Average*	Short	m
Chin shape	Pointed	Average*	Cleft	m
Cheeks shape	Full	Average*	Scraggy	m
Lips volume	Thin	Average*	Full	m
Lips size ratio	Upper lip	Average*	Lower lip	m
Mouth shape	Down	Average*	Up	m
Mouth width	Wide	Average*	Narrow	m
Mouth height	Up	Average*	Down	m
Mouth depth	Backwards	Average*	Forwards	m
Makeup eyes shadow none*	None*		Full	t
Makeup lipstick	None*		Full	t
Makeup rouge	None*		Full	t

*default value, *t* texture blending, *m* mesh morphing

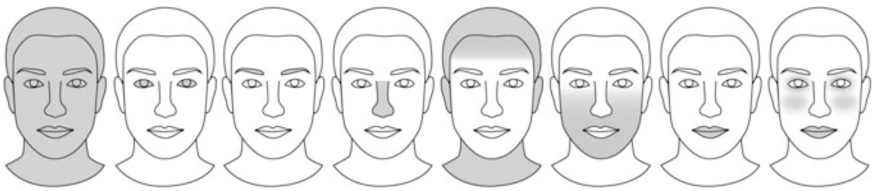

Fig. 3 Area classification used for vertex selection and infographics of the final application

4 FaceMaker Implementation

A system that allows manipulating the human average face model should meet the following criteria: (1) Reach a large number of participants under similar conditions as when players usually build avatars (at home). (2) A contemporary and interactive rendering engine, which is able to blend all morphings. (3) An easy to understand user interface, which could be randomized for every session to avoid sequence effects. (4) Provide support that helps to complete certain objectives and questionnaires.

In order to meet all requirements, we developed faceMaker. The browser application was developed to control the facial changes of the average face using parametric values. Our application design prevents the calculation of mean characteristics. First, multiple features on a single scale allow no reliable assumptions about a new means. And second, participants are not forced to change any values. However, to understand the concepts of different faces and to ensure that participants are able to create them, the 6 following objectives were introduced as design tasks: (A) A personal arbitrarily avatar face; (B) an uncanny, repulsive face; the stereotypical, positive-related face of an attractive (C) heroine; and (D) hero; the stereotypical face of a (E) female villain; and (F) male one. Therefore, we assume that faces of the category A, C, and D rather evoke positive feelings and that faces of B, E, and F are rather related to negative associations.

4.1 *Online Face Generator*

The application was developed with HTML, JavaScript, and jQuery. MySQL and PHP were used for client–server communication. For the implementation, WebGL is used as a rendering engine.[5] Hardware acceleration of more than 3 morphing targets simultaneously is currently not supported by Three.js. Therefore, we implemented a CPU-based software algorithm for multiple blendings of morph targets. Three directional lights and a slight ambient light in neutral white are used

[5]ThreeJS, WebGL Engine: http://threejs.org.

for lighting. The key light casts shadow maps. The orbiting camera can rotate within 180° in front of the face by clicking the left mouse button. The face consists of the head model, eyes, and eyelashes. No animations were added to the face. A gray t-shirt was added for a neutral transition from décolleté. The application runs in full screen of a browser window. The background is dark gray. Anti-aliasing is enabled.

For statistical and usability reasons, we decide that all parameters should be controlled with linear scales starting from a neutral point which corresponds to the human average face. This was realized using sliders but led to the disadvantages that participants sometimes will not change the default value. This can cause bias to the center, which is not intended in some studies or results in the need for a larger sample to find significant differences. Nevertheless, we decide that participants should not be forced to do something they do not want to change or should fix randomized parameter values they do not know.

All facial parameters can be controlled using horizontal sliders. The pixel width of each slider is 200. To the left and to the right, each tenth segment is highlighted with a stepped marker. No cursor snaps are used. Each parameter and all group boxes are labeled with the name of the facial parameter or domain. Additional states (e.g., colors) between left and right are labeled using tool tips below the slider. Each parameter can be set back to the default value with a button at the right. The majority of parameter sliders is gray. Only color changing sliders are equipped with color scales according to their parameter range. For every new user, all sliders within a domain group as well as the group box itself are randomly distributed on the left or right column of the browser. A help icon for each group box opens a help description, which informs the user about the corresponding changes.

4.2 *Requirements and Compatibility*

The system currently supports the browsers Firefox, Chrome, and Opera and offers multi-language support. An internet connection is only required while loading the application and submitting a face. The system needs a PC or mobile device with hardware-based 3D graphic acceleration to run smoothly. We recommend a graphics card with the performance of a NVIDIA GeForce GT 650M or higher.

4.3 *Measurements, Back-End, and Extensions*

The application is able to detect a couple of measurements: processing times, facial changes, resets, self-assessments, and facial parameters. All measurements are saved in cookies and are stored in the server's database after submitting a face. To filter for demographics, tasks, to view faces, and to save aggregated samples of certain face groups, we developed an application back-end. Data from here can be

Fig. 4 Screenshots of the faceMaker back-end for evaluation purposes and icon generator/model downloader tool

used to view the faces participants create and to see which parameters they prefer. For example, the back-end can be used to see faces created by different genders, age groups, or country. The tool is also able to aggregate faces in a certain degree of perceived realism, attractiveness, or likeability. Graphical bars emphasize averaged or summed-up values behind each number.

Our system also includes an avatar image creation tool for saving the generated images to motivate people to use faceMaker. The small editor opens access to all faces in the database. Users can now change the background of the image and render each face from any perspective in a certain image size (optimized for Skype, Twitter, Gravatar accounts). The image can be shared via Facebook, Twitter, Google Plus, and LinkedIn. The tool also enables users to save the faces as OBJ and to download skin, hair, and eyes textures. Our tool can easily be combined with standard web surveys, which allows gathering additional information. Such combination of the user-designed faces and questionnaire inputs provided by faceMaker opens up a good platform for investigating virtual characters and for better understanding how their appearance is influencing the perception of an avatar at the first place, and the resulting effects on the acceptance of applications in a second place (Fig. 4).

4.4 Delimitation from Avatar Generators in Games

Our avatar generation system differs in many respects from avatar generation systems in games. For example, our system only supports human-like characters. Body decorations (e.g., tattoos, jewelry), aliens, or fantastic creatures (elves, orcs, etc.) are not supported. Thereby, our system differs from avatar generation system in commercial game series like Mass Effect, The Elder Scrolls, or Fallout. Another difference is animacy during the creation process. Avatar creation systems in games as in The Sims present their character in an interactive style with animated postures while changing parameters. Due to limited resources in a browser-based system, we only apply render updates to the render engine when a user triggers changes. Due to the lack of animation of the avatar face, which has to be rendered all the time, our system does not continuously consume system resources. Thus, faceMaker runs on devices in battery mode over a long period of time.

Another difference between avatar generation systems and faceMaker is the structure of the graphical user interface. To avoid biases, the parameter sliders of faceMaker are randomly arranged, which is not the case in games. Here, we found the most important settings (gender, race, etc.) at more accessible places in the GUI than, for example, finer adjustments of the mouth. Another distinguishing feature between GUIs in faceMaker and game is that faceMaker only use faces not the rest of the body, hair styles, or clothes.

5 Cluster Analysis and Statistics

To validate faceMaker, we conducted two studies: In a foregoing study, we validated the practical use of faceMaker in "the wild" through determining the preferred characteristics of stereotypical avatars. The results (14) show that it is possible to recruit a larger number of participants with faceMaker and to draw conclusions about the user's concepts and preferences of virtual avatars. We compared arbitrary faces with other categories. The results of this study finally showed that people rather prefer to create attractive female heroic faces than other faces.

As a second analysis, which we present below, we conducted a clustering analysis to understand which kind of faces was created without considering the objectives directly. We assume that our objectives deliver a higher contrast between different kinds of faces and relieve clustering without using any tasks for participants.

5.1 Procedure

The application procedure includes five steps: (1) A session starts with demographics including gender, age, origin, consummation of games and movies. Furthermore, we asked them to accept the terms of use. (2) The application starts and a pop-up window appears where a participant receives the instructions. After confirming the instructions, a participant could use the 37 sliders to change the average face. (3) Before submitting, a participant had to fulfill four assessments, which were not considered in the cluster analysis. (4) The application reloads after submitting and goes back to step 2. (5) After submitting 6 faces, a participant could view and download the generated faces. Cookies ensured that each task could only be performed once and repeated after the sixth trial. Cookies were deleted after 7 days. To avoid sequence effects, we used a balanced Latin square design starting with the face type which has the fewest participants.

5.2 Participants

We collected the results of 569 participants (313 males, 247 females, 9 n.a.) who created 1730 faces. The mean age was 30.58 (SD = 11.86). The participants were mainly recruited via mailing lists, social networks, and advertisements. 127 (22%) participants pointed out that the play games daily, 108 (19%) more than once a week, 68 (12%) once a week, 40 (7%) once a month, 125 (22%) play infrequently, 101 (18%) never. A total of 192 (14%) participants pointed out that they watch movies every day, 192 (34%) more than once a week, 171 (30%) once a week, 53 (9%) once a month, 62 (11%) watch infrequently movies, and 10 (2%) never. The median browser resolution on a participant's display was 1708 $\times$ 925. A total of 291 (51%) participants used Firefox, and 278 (49%) used Chrome as a browser.

5.3 Clustering and Multi-dimensional Scaling

To group and visualize the procedural faces systemically, cluster analysis was employed. Among different clustering algorithms, we applied the expectation–maximization (EM) algorithm for several reasons: EM does not require a predefined number of cluster (as kmeans), and the method iteratively searches for the maximum likelihood and was adopted in the previous work, e.g., for face detection as used by Rujirakul et al. (2014). It is known, that different cluster algorithms often produce inconsistent results. Different distance measurements can be used to validate the consistency of an algorithm. We decided to use the Euclidean distance metric because other metrics such as Ward's method tend to establish equal cluster sizes. To understand how the clusters are related to each other, we developed a spatial map to visualize the results. The distance matrices were used to conduct a multi-dimensional scaling (MDS). This approach cannot provide the accuracy of a face classification or the complete clustering; however, vividly illustrates the similarities or differences of faces. EM clustering was conducted in Weka, distance metrics, and multi-dimensional scaling was computed using R cmdscale.[6]

5.4 Cluster Analysis and Statistical Results

Based on the training set only given by the created facial parameters, EM clustering delivers 6 nodes: Cluster 0: 282 (17%), Cluster 1: 362 (21%), Cluster 2: 169 (10%), Cluster 3: 209 (12%), Cluster 4: 229 (13%), Cluster 5: 457 (27%). The 6 objectives that people created were verified using a cluster assignment. Cluster 0 was assigned

[6]cmdscale—A R-library for multi-dimensional scaling: Retrieved June 2016 from https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cmdscale.html.

to F (male villain face), Cluster 1 to C (attractive heroine face), Cluster 2 to A (arbitrary face), Cluster 3 to B (repulsive face), Cluster 4 to E (female villain face), and Cluster 5 to D (male hero–villain). We would like to note that the results of the EM algorithm and the assignment procedure are incidentally identical to the number of objectives. The parameter results of the cluster analysis were added to the plotted diagram of the multi-dimensional scaling (see Fig. 5). 55.0% of the faces were incorrectly clustered instances. The class attribution table (not illustrated) reveals, that Cluster 1, for example, shares 257 instances with objective A and C, which could be explained by the tendency of participants to create faces in the arbitrary task that are similar like in the female hero task.

The plotted MDS map (Fig. 5) reveals a dense filament including two main consolidations at the top and clusters of faces that people prefer to create. Parametric values of the cluster centers were used to render faces and were placed on the MDS map. Two main clusters are connected in the main filament structure. Cluster 1, 2, and 4 as well as the average faces of objective C and E are on the female “side” on the map. Cluster 0 and 5 as well as faces of objective D and F are on the male “side” of the map. Since all parametric changes were weighted equally, the arrangement and the results of this cluster analysis reveal that the sum of parametric changes is made according to gender and appeal. The arbitrary face (A), which participants created without any restrictions, is very close to the female cluster of faces. We also see that the cluster center of female villains (4) is much closer to female heroines (1) than the male villain center (0) from the male heroes one (5). The cluster center of the repulsive face is outside from the cluster centers of male or female faces.

The MDS map of the cluster analysis shows that facial properties of the created faces depend on gender and appeal. Male faces tend to be placed at positive x-values; female faces are rather placed at negative ones. Appealing faces (heroes) tend to have positive y-values, and not appealing (repulsive faces) have negative y-values. Using this approach, we are now able to derive stereotypical faces and avatars from procedural ones. For example, to find very female facial properties, we can look at samples very close to the cluster centers 1 and 2. In contrast, very repulsive faces are outside of the main cluster of females and males. The repulsive face samples show a very distributed pattern. This can be explained by the fact that repulsive faces have no certain features or patterns. However, they differently deviate from the human ideal and human average proportions, which people try do not violate while creating appealing faces.

Besides the results of the cluster analysis, we looked deeper into the statistics of the 37 parameter scales. Figure 6 shows how often people interact with these parameters. The bar charts show the average values of parameter changes, resets, and views (instead of using the orbit camera). Face gender, hair color, eye shape, and skin color are changed very often in contrast to ear size, rouge, lip stick, and mouth height. The resets indicate which parameters were often reset to the human average. The most resetted parameters are related to the eyes. That were in particular eye depth, eye distance, and eye rotation. The parameter views in Fig. 6 show which parameters were often visited by the mouse cursor. Hovering a slider

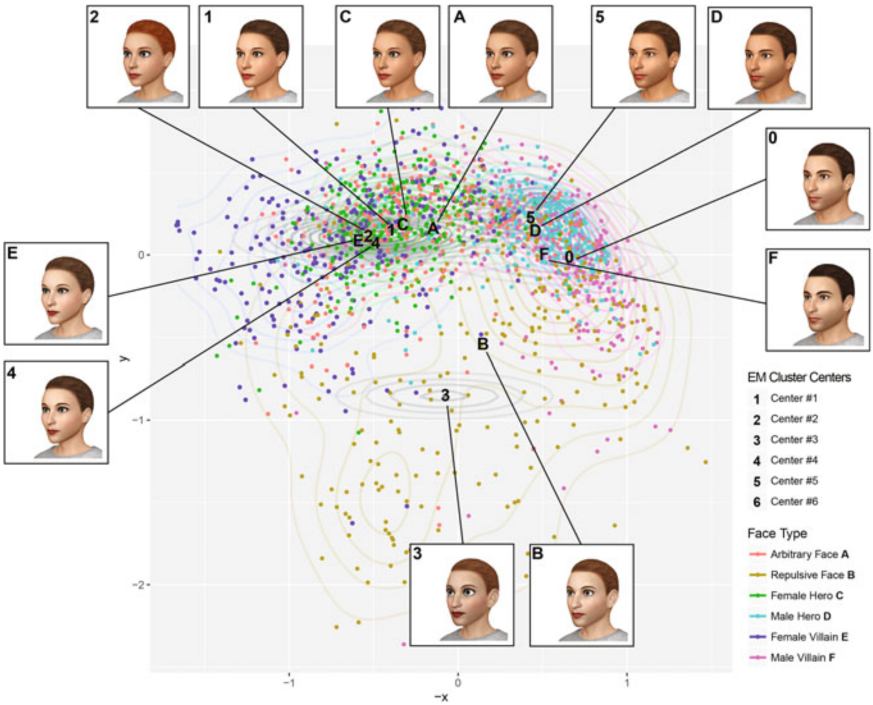

Fig. 5 Plot of the MDS analysis including 1730 faces from 569 participants and location of the average faces (*A*–*F*) resulted from the users' objectives and the 6 cluster centers (*1*–*6*) delivered by expectation–maximization (EM)

zooms or rotates the camera perspective to a certain region. The diagram shows on which facial regions users seem to be interested in while creating avatars.

6 Discussion

In this study, we present the results of a cluster analysis based on 1730 procedural faces created by 569 participants which used the online avatar generator faceMaker. To validate faceMaker, we used EM clustering. We found the same amount of cluster nodes as given objectives (arbitrary face, repulsive face, female face, male face, female villain, male villain). Through cluster alignment, we determined which cluster generally corresponds to the stereotypical average face. Plotting the results of a MDS analysis reveals how the faces are related and which face types rather deviate or correspond to the human average and human ideal. The map reveals two main distributions of male and female faces and shows how participants create faces according to their concept of stereotypes. Class attribution reveals that Cluster 1

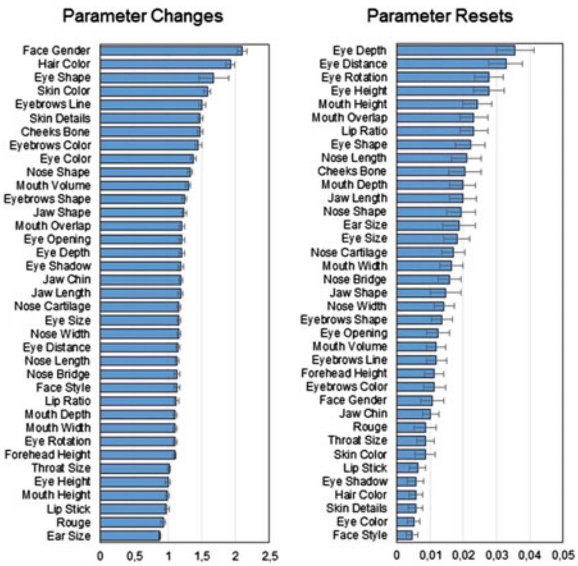

Fig. 6 Bar charts of parameter changes, resets, and views. *Error bars* show standard error (SE)

shares 257 instances with the arbitrary and female hero face. If having a choice, our participants are more willing to create female faces instead of male faces. Positively associated faces (female and male heroes) get smooth skin, realistic and attractive proportions, and natural average skin color. Villains get bright skin, distinctive cheek bones, and exaggerated jaws. Female villains receive strong lip stick, male villains get strong eyebrows. Repulsive faces get features that strongly deviate from the human norm, which is divided into two main distributions of male and female faces. They were exaggerated with unnatural violations against the human average. Thus, using these procedural types of faces, we are now able to look into preferred or not preferred stereotypical concepts of avatar faces and can deduce which facial properties are rather preferred or not preferred.

The avatar and face creation system faceMaker[7] enables researchers to conduct valid online studies reaching a large number of participants. In the cluster analysis, we revealed groups of faces and their relation to each other. We found how procedurally generated faces could be used to identify stereotypical faces and which parameters they possess. Game developers can now deduce which avatar faces are rather preferred or not preferred by users. Knowledge about preferred or not preferred facial characteristics of avatars gained by faceMaker can be considered in design decisions about the appearance of game characters and future avatar generators. The usage of parameters provided in faceMaker allows optimizing such avatar generators. For example, we showed that gender, hair color, and eye shape

[7]FaceMaker Online Application: http://facemaker.uvrg.org.

are the most changed parameters. This knowledge could be considered in the development and optimization of user interfaces of game character generators in RPGs.

The application is available on GitHub[8] and can be downloaded and used under the general public license (GNU) v2. It opens up new branches of customized avatar-related research instead of presenting predefined images. The program can be used as experimental apparatus in laboratories as well as web-based application in the large. This is now also possible for other researchers who investigate the human perception or attitude toward virtual avatars. For investigations of differences in other cultures or different ages, it is necessary to use average faces from different ethnic groups like Asians or Africans faces.

Researchers in the fields of human–computer interaction, games, social sciences, psychological sciences, medicine, and other disciplines are now able to use faceMaker for their own purposes. The application is available for free disposal and can profit through new findings by investigating users and their generated avatars instead of analyzing predefined face models only. Game developers can use our system to conduct prestudies to investigate which facial properties main characters or stereotypes in their game should have. They can also use the system to optimize location-based changes or questions about their main characters or user interfaces. Furthermore, we suggest future research about how users create procedural faces they already know (as celebrities), look like themselves, or investigate gender or cultural differences in the face creation process.

Acknowledgements Special thanks go to Katharina Leicht, Verena Dengler, Julia Sandrock, and Solveigh Jäger. We thank the German Research Foundation (DFG) for financial support within project C04 of SFB/Transregio 161. This work was also supported by the cooperative graduate program "Digital Media" of the University of Stuttgart, University of Tübingen, and the Stuttgart Media University.

References

Apostolakis, K. C., & Daras, P. (2013). RAAT - The reverie avatar authoring tool. In 2013 18th International Conference on Digital Signal Processing (DSP) (pp. 1–6). IEEE. 10.1109/ICDSP.2013.6622788

Bessière, K., Seay, A. F., & Kiesler, S. (2007). The Ideal Elf: Identity Exploration in World of Warcraft. CyberPsychology & Behavior, 10(4), 530–535. 10.1089/cpb.2007.9994

Chung, D., DeBuys, B., & Nam, C. (2007). Influence of Avatar Creation on Attitude, Empathy, Presence, and Para-Social Interaction. Human-Computer Interaction. Interaction Design and Usability, 4550, 711–720. 10.1007/978-3-540-73105-4

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. Journal of Personality and Social Psychology, 24(3), 285–290. 10.1037/h0033731

Ducheneaut, N., Yee, N., Wadley, G., Alto, P., & Alto, P. (2009). Body and Mind : A Study of Avatar Personalization in Three Virtual Worlds, 1151–1160.

[8]faceMaker at GitHub: https://github.com/valentin-schwind/FaceMaker.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but … A meta-analytic review of research on the physical attractiveness stereotype. Psychological Bulletin, 110(1), 109–128. 10.1037/0033-2909.110.1.109

Grammer, K., & Thornhill, R. (1994). Human (Homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness. Journal of Comparative Psychology, 108(3), 233–242. 10.1037/0735-7036.108.3.233

Heeter, C., Egidio, R., Mishra, P., Winn, B., & Winn, J. (2008). Alien Games: Do Girls Prefer Games Designed by Girls? Games and Culture, 4(1), 74–100. 10.1177/1555412008325481

Langlois, J. H., Ritter, J. M., Roggman, L. a., & Vaughn, L. S. (1991). Facial diversity and infant preferences for attractive faces. Developmental Psychology, 27(1), 79–84. 10.1037/0012-1649.27.1.79

Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? Psychological Science, 5(4), 8.

Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. IEEE Robotics & Automation Magazine, 19(2), 98–100. 10.1109/MRA.2012.2192811

Prokopakis, E. P., Vlastos, I. M., Picavet, V., Nolst, G., Thomas, R., Cingi, C., & Hellings, P. W. (2013). The golden ratio in facial symmetry. Rhinology, 51(4), 18–21. 10.4193/Rhino12.111

Rice, M., Koh, R., Lui, Q., He, Q., Wan, M., Yeo, V., … Tan, W. P. (2013). Comparing avatar game representation preferences across three age groups. In CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13 (p. 1161). inproceedings, New York, New York, USA: ACM Press. 10.1145/2468356.2468564

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry and attractiveness. Proceedings. Biological Sciences/ The Royal Society, 266(November 1998), 869–874. 10.1098/rspb.1999.0717

Rujirakul, K., So-In, C., & Arnonkijpanich, B. (2014). PEM-PCA: A parallel expectation-maximization PCA face recognition architecture. The Scientific World Journal, 2014. 10.1155/2014/468176

Schwind, V. (2015). Historical, Cultural, and Aesthetic Aspects of the Uncanny Valley. In Collective Agency and Cooperation in Natural and Artificial Systems (pp. 81–107). Cham: Springer International Publishing. 10.1007/978-3-319-15515-9_5

Schwind, V., Wolf, K., Henze, N., & Korn, O. (2015). Determining the Characteristics of Preferred Virtual Faces Using an Avatar Generator. In Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15 (pp. 221–230). New York, New York, USA: ACM Press. 10.1145/2793107.2793116

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. IEEE Computer Graphics and Applications, 21(October), 42–50. 10.1109/38.946630

Author Biographies

Valentin Schwind is Ph.D. student in the field of Human–Computer Interaction (HCI) under the supervision of Prof. Dr. Albrecht Schmidt at the University of Stuttgart. After completing his diploma in media computer science at the Stuttgart Media University (HdM), Schwind worked as freelancer and consultant for visual effects and game companies in Germany, Austria, England, and the USA. Later, he was research assistant and lecturer at the Stuttgart Media University and involved in research projects investigating virtual faces and avatars.

In 2015, he became researcher at the Institute for Visualization and Interactive Systems (VIS) of the University of Stuttgart focusing on the Uncanny Valley in Human–Computer Interaction and deploying metrics in the visual perception of virtual characters. He is now doctoral researcher of the collaborative research center in the SFB-TRR 161 investigating quantitative methods for visual computing. He published in renowned conferences and journals, received the best paper award at the conference Mensch & Computer 2015, and has been reviewer for leading conferences and journals.

Schwind teaches in the field of computer graphics and game development. His interests range from gaming, interaction, human perception, and cognition. In particular, he is interested in human–avatar interaction, virtual reality, and visual computing. He developed and supervised the creation of games, computer animations, and mobile applications including procedurally generated content.

Katrin Wolf is professor for Media Informatics at Hamburg University of Applied Science where she is teaching in the Media Systems program. Before she was professor for Media Informatics at BTK—University of Art and Design in Berlin. She was a postdoctoral researcher in the Human–Computer Interaction Group in the Institute for Visualization and Interactive Systems and the SimTech Cluster for Simulation Technology at the University of Stuttgart.

Wolf was a researcher and doctoral student at the Telekom Innovation Laboratories at the Technical University of Berlin, where she investigated ergonomics in mobile HCI. Her studies at the Berlin University of the Arts, where she received two German diplomas, provide her with a background in media, interface, and interaction design as well as with communication science.

Particularly, she is interested in human–computer interaction, interaction design, and using parametrically or procedurally generated content in games. Wolf published peer-reviewed journals and conferences. She organized scientific workshops, served as program committee member, and has been reviewer for leading conferences in her field. Wolf lectures interaction design as well as human–computer interaction in practical and theoretical courses for several years.

Niels Henze is assistant professor for socio-cognitive systems in the Institute for Visualization and Interactive Systems and the SimTech Cluster for Simulation Technology at the University of Stuttgart. Beforehand, he was postdoctoral researcher in the Human–Computer Interaction group at the University of Stuttgart. After receiving a diploma in computer science in 2006, he worked for European research projects at the OFFIS Institute for Information Technology. In 2008, he became a researcher and doctoral student in the Media Informatics and Multimedia Systems group at the University of Oldenburg. He worked for international projects and was responsible for teaching and tutoring. Henze finished his Ph.D. in 2012 with his thesis about Camera-based Mobile Interaction with Physical Objects under the supervision of Prof. Dr. Susanne Boll.

Henze's research interests are mobile–human–computer interaction and pervasive computing. Particularly, he is interested in large-scale studies using mobile application stores as a research tool, attention and smart notification management, as well as multimodal interfaces. Niels published in scientific journals and competitive conferences. He received awards from CHI, MobileHCI, and Mensch & Computer. He organized several scientific workshops, served as a guest editor for the International Journal on Mobile Human–Computer Interaction, and has been a reviewer for the leading conferences and journals in his field. Henze lectures human–computer interaction for several years. He developed and supervised the development of mobile applications to conduct large-scale studies that have been installed more than a million times. He is interested in using procedural generated content in games for conducting user studies in the large.