EPiC Series in Engineering

Volume 1, 2018, Pages 11–20

ReVo 2017: Laval Virtual ReVolution 2017
"Transhumanism++"

"Real Baby - Real Family": Holdable tangible baby VR

Rex Hsieh, Yuya Mochizuk, Takaya Asano, Marika Higashida, Daiki Agatsuma, Tatsuya Nishikizawa, and Akihiko Shirai

Kanagawa Institute of Technology, Atsugi City, Kanagawa, Japan
shirai@ic.kanagawa-it.ac.jp

Abstract

"Real Baby - Real Family" is a Virtual Reality baby nursery simulator that aims at conveying the parental and family love to the audiences. It has since been exhibited at Laval Virtual 2017, Anime Expo 2017, and SIGGRAPH 2017. This article will attempt to discuss the feedback and reviews gathered during the international exhibitions and shed light on the future development of this project.

1 Introduction

Babies are the fruits of parental love. Such "conjugal love" and "love in family ties" can be used as a synonym for a child, or a baby. This project, named "Real Baby - Real Family"[5], is a virtual reality baby nursery simulator that allows friends, homosexual couples, and heterosexual couples to experience the joy of raising a child. It is our hope to invoke deep discussions about family ties and love with this VR project. In this paper, the team reports the motivation and concept of this project in its initial phases during the 24th International Collegiate Virtual Reality Contest (IVRC 2016) and the different iterations during the many international exhibitions such as Laval Virtual 2017, Anime Expo 2017, and SIGGRAPH 2017. Reactions and feedback from the attendees as well as future possibilities of this baby nursery simulator are discussed in detail.

2 Related Works

Below are some of the previous researches that also uses baby as their subject matter. NeuroBaby [8],Infanoid [2], Kismet [3], Cog [7], Babybot [4], YOTARO [6]. Amongst these researches, YOTARO can be considered as a work that manifests a realistic baby. YOTARO is a baby robot that has a runny nose and many different emotions. However, while YOTARO can be considered as a precedent which represents a realistic VR baby robot, it's face is that of an illustration and not a real human being, making it look as if the baby is not related to the players and is instead a fictional character. "Real Baby - Real Family" generates a baby from facial images of the heterosexual or homosexual couples playing it thus those experiencing this project get a baby looking just like the themselves.

A. Shirai, L. Chretien, A.-S. Clayer, S. Richir and S. Hasegawa (eds.), ReVo 2017 (EPiC Series in Engineering, vol. 1), pp. 11–20

3 Design

"Real Baby - Real Family" is characterized by a virtual family enabling the players to communicate with a baby through haptic, visual, and audio feedback while wearing a head-mounted display (HMD). According to our hypothesis, creating a believable baby looking like our players will make people more emotionally invested with the experience. Thus we created a baby generated from players' photographs. The most important part of our project; however, is that we constructed a physical baby that can interact with the audience. Projects such as YOTARO features a lot of interactivity but none allows the player to interact with the baby through haptic means. This means one cannot hug the baby which is a fundamental interaction in raising a baby. On the other hand, projects such as Laerdal's high performance infant medical training simulator "SimBaby™" and "SimNewB®"[1] focuses on treatment and is too expensive for the general public. The biggest difference between baby simulator in the medical and VR world is those developed for medical purpose aims at patient treatment while VR baby simulator aims at creating a believable experience that touches the hearts of the players. "Real Baby - Real Family" is unique for allowing players to (1) hug and interact with the baby physically, (2) creating a baby looking just like the players, (3) contains audio components.

3.1 Holdable baby device

In order to hug a baby doll without experiencing discomfort while wearing the HMD, there is a need for high precision position tracking of the baby doll. However, this is hindered by image recognition ability of a one eyed camera therefore slowing down the position tracking speed. Our proposed method uses multipoint image sensor implemented in the controller of HTC Vive to enable high precision position tracking. This allows players to hug the baby without getting visually disoriented (fig.1).

Figure 1: HTC Vive and Baby mock-up

3.2 Visual Face Generator

Below are three characteristics of our baby face generator: (1) Generating a face fitting the players' skin color, (2) Creating the baby's face from multiple photographs, (3) Reversing the age of the players to fit that of the baby.

3.2.1 Average Baby Face Generation

It determines the skin color of the baby by averaging the skin tone of the two photographs. Our system also morphs the colors and shapes of the faces using OpenCV. Lastly it collects 16 baby face images using Japanese Google Image Search.

3.2.2 Get Face Landmark Index

In order to morph the many photographs, baby images, and corresponding points in the feature points, it is necessary to obtain the index. In the proposed method, we use the Dlib of open source library, with results that have been learned by the data set iBUG-300-W, to get the points of each part of the face. As a result of fig.2, the eyes, noses, mouths, eyebrows, the feature points, and the 68 points with index composed of contour, are realized by automation morphing. Here it can be seen that morphing are concentrated in the mouth and even parts of the face.

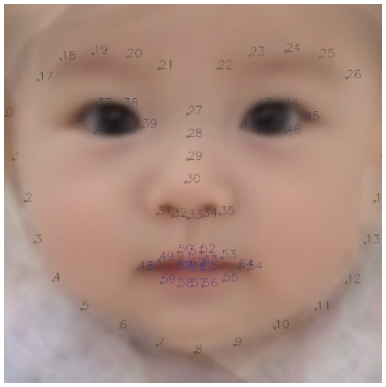

Figure 2: Average baby face and landmark index

3.2.3 Inverse Age Progression

Inverse age progression is a process of generating an averaged image from the features obtained by the face image and (2), without collapse baby face images obtained in (1). The implementation uses a Dlib and OpneCV. Obtaining feature points is performed at Dlib, performs image generation by passing the coordinate data of the feature points OpenCV side. Dlib automatically sorts the feature points obtained, treated as a landmark. Its index is unchanged in all of the face image. By setting the contribution ratio with respect to the index number, it is possible to perform processing for each selectively parts. The color components are produced by (1), the present process is a concept that only features of the specified part is inherited. If carried out well as deformation of the 3D modeling, it should also be considered, such as contour around the chin.The generated face image is used by Live2D.

Inverse age progression is the process of generating a final image from the averaged age of the facial images at Get Face Landmark Index. This is all done without collapsing baby face images obtained in Average Baby Face Generation. The implementation uses a Dlib and OpenCV. It obtains feature points obtained by Dlib, performs image generation by passing the coordinate data of the feature points OpenCV side. Dlib automatically sorts the feature points

obtained and treated them as a landmark. Its index is unchanged in all of the facial images. By setting the contribution ratio with respect to the index number, it is possible to process each selective parts. The color components are produced by Average Baby Face Generation, the present process is a concept that only features the specified parts inherited. If carried out well as deformation of 3D modeling, contour around the chin and other aspects should also be considered. The generated facial image is used by Live2D.

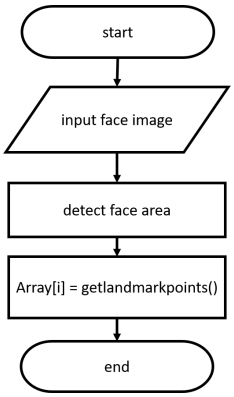

Figure 3: Get landmark index

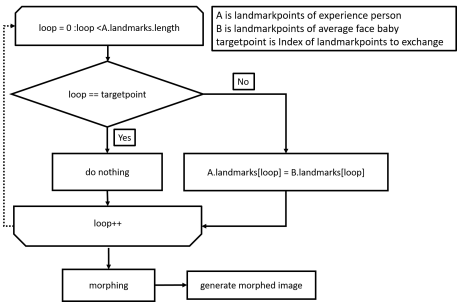

Figure 4: Inverse age progression

3.2.4 Dynamic texture generation

In the generated images only pasted to the 3D model has a large discomfort, it is necessary to generate a facial expression that varies vividly dynamic. Therefore use the “2D image deformation software” “Live2D” for character animation, to prepare the animation set in advance, and real-time generated in the VR video space in Unity. In creating the model data of Live2D, the image cut out the part of the eyes and the mouth of the“average face of the baby face image” in fig.2 as a template.

3.3 Display of Haptics Synchronized with Voice

When holding the baby doll, the doll is strengthening the interaction by vibro of Vibro transducer Vp2 (Vp210) which was mounted on the baby doll. This vibro is controlled by voice. Only by passing the band-pass filter using a Fourier transform effective frequency band (5Hz 200Hz) and extracted as the vibration data onto an output audio can it perform real-time tactile presentation through HMD attendant controller (Vive Controller). However, since it is the only vibrator that is built in the standard Vive controller, we felt its electric current output which vibrates the baby doll was insufficient after many exhibitions. To fix this problem we mounted the Vp210 onto Vive controller to improve the oscillating capability.

Figure 5: Inverse age progression

3.4 Content of Experience

This work allows for single player mode and two player collaboration. By making the photo of the players look younger and then combine the photographs together, the baby's face was born. Through taking care of the generated baby, players can experience the importance of parental love and family ties with this VR project (fig.7). During the virtual world, one of the the Vive controllers take on the shape of a milk bottle for the baby. Sensory clues such as the increase in milk consumption by the baby and a sense of touch have all been implemented. After the experience, a copy of the mother's pocket book contains the image of the baby's face was issued to the player as a souvenir.

Figure 6: exhibition booth

Figure 7: Picture of experience

Figure 8: Maternal and Child Health Handbook

3.5 Real-time Facial Expression Change of Baby

Fig.10 is a block diagram of the system for generating a baby’s facial textures in real time. The essence of this work is to generate a texture of baby face that changes according to the players’ input image. Because of this, there is a major difference between our achieved research result and the below commonly used animation methods: Using the frame by frame animation technique that animates objects using subtle differences between each image and changing facial expression by changing the location of vertices of 3D models. Therefore, We implemented real-time facial expression changes by the use of RenderTexture which is one of Live2D and Unity features. In preparing for vertex deformation data for use in Unity In Live2D, it was prepared as follows: Material texture to be used as the face of the template players are using (right eye, left eye, mouth, and foundation). Read a file that Live2DManager has generated by the inverse age progression first on the Unity side, to draw using a method called the Graphics class DrawMeshNow. The texture was then rendered using a camera component drawing that have been Live2DModelUnity. Finally, the rendered texture was used to set the material of the face of the baby 3D model to the texture. The different states of the baby model: cry, drink milk, and sleep are monitored by BabyEmotion to manipulate the Animator of animation state. Furthermore, in the Animator, the animation state was operated by parameters of

Live2D, then the facial expression was changed. Live2DModelUnity's drawing, the rendering of RenderTexture, a Unity Reference, can change the texture in real time because it is performed every frame. It is generating the visual texture of baby face image in real time by the above operation.

Figure 9: Live2D Template

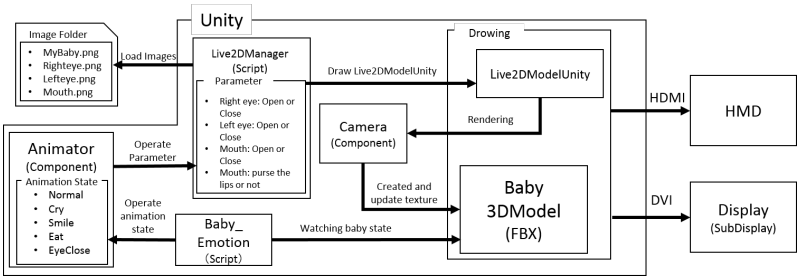

Figure 10: System of baby emotion

3.6 Interaction using image, voice and haptic

Fig11 is a diagram that presents baby voices and vibro tactile associated with it. BabyEmotion reads wave files such as a baby's voice, which was prepared in advance from the audio folder, including: moody, laugh, cry, start drinking, stop drinking, waiting, and sleeping. This BabyEmotion sends the audio files corresponding to the parameter information to AudioSource where the audio file is outputted to a headphone via the HMD as the audio information is presented as haptic. Sampling period in this case is the same as drawing updated frequency (or 100 FPS). Vibration data is what tells TriggerHapticPulse how much should the ViveController vibrate. TriggerHapticPulse is a function within SteamVR_Controller which in term is a function within SteamVR_Plugin. This makes the vibrator controllable by Vive controller via the speech wave-

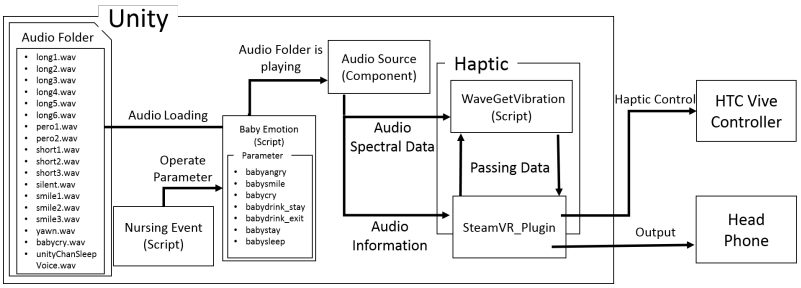

Figure 11: System of haptic by sounds

form. In parallel with the process of the spectrum of the voice data are WaveGetVibration and SteamVR_Plugin located inside Unity. In order to vibrate the controller mounted onto baby, WaveGetVibration uses SteamVR_Controller. After that, it gives the spectrum extracted from the above-mentioned audio data to the argument durationMicroSec of TriggerHapticPulse which is a method of SteamVR_Controller. The above processing is performed at the same frequency as the drawing update. In addition, while performing a process of the spectra of audio data, the process of giving the spectrum to the controller was at the same sampling period. Accordingly, the vibration waveform is to be reproduced from speech waveform and Vive controller to be reproduced from the headphones, approximately becomes the same waveform. On the other hand, the parameter information of BabyEmotion (are drinking milk, milk strikes in addition to mouth) is the event the baby is intended in response to.

This work had you experience towards the more than 300 people through the finals from qualifying IVRC 2016. In addition, maternal and child health handbook was distributed after the experience the end, experience person each other each other to show the baby face image, share the experience that occurred in the VR in the VR outside, became the new communication tool that can be compared.

3.7 Feedbacks for Baby Face Generation

Although this system is capable of creating a baby's facial image that is similar to the players', in the questionnaire, which recorded the responses of 12 people comprised of both game developers and students, with "similar" on one end, "not similare" on the other end, and were asked to evaluate in four stages with the maximum score being 4 points symbolizing most similar. The average score was 2.27 points which can still be improved. Furthermore, it is necessary to generate a baby's facial image of different races especially for international exhibitions in the future.

3.8 Haptic Feedback

Compared to the standard installed vibrator Vive controller, although the Vive controller fitted with a Vp210 gives out a stronger vibration, if players shift their positions or change their postures while holding the baby, our system will have a hard time adjusting to the change. For future challenges we plan on enhancing the baby mockup.

3.9 Possibility of Nursing Simulator using VR

In recent years, the nursing educations in Japan, in order to foster nurses with practical skills that can support a diverse of clinical practices, are looking into simulation educational systems such as the objective structured clinical examination (OSCE). The practical simulator nursing education using virtual patients has attracted much attention. Our Virtual Reality work, "Real Baby - Real Family", as hinted by its name, was implemented using a realistic baby and is accessible to the general public where all can experience the love of a new founded family. This type of VR project remains unrealized until now where we constructed a virtual baby that inherits facial features of the players, have a tactile sensation, and allows new and inexperienced parents to quickly become accustomed with raising a child. For future challenges, we hope to simulate other essential actions relating to raising a child such as breast feeding as well as giving the virtual baby a heartbeat and body temperature.

3.10 Application to Entertainment System

When exhibited at IVRC 2016, this VR system has an option that allows players to create a two-dimensional image of the baby from actual photographs of the players or substitution photos. This Virtual Reality system gives the illusion that the players really have created a baby whether the players are friends, heterosexual couples, or homosexual couples. The believable 3D baby model inside the virtual world and the 2D baby photo created from the players‘ image further enhances the experience. In the future we would like to tailor this entertainment system to the subculture market that makes two-dimensional images and babies.

3.11 Toward an International Exhibition

This project "Real Baby-Real Family" was awarded the Laval Virtual Prize in IVRC 2016. We are planning on improving this project for future international exhibitions. Some of the improvements we would like to make include enhancing the single and two player experience for this project. For this purpose, it is necessary to share the experience with a partner while wearing HMD. We are also considering the sharing of high quality no delay position information between both the UDP communication.

4 Conclusion

Throughout the IVRC 2016 Tokyo qualifying round and later the final competition, the baby team continued to refine the system until it delivers a satisfying baby themed experience. Using the knowledge gained from exhibitions, the weight and intensity of vibration of the baby doll were adjusted to deliver a believable sense of touch when interacting with the baby. Furthermore, we succeeded in creating a baby face that borrows features from the players' photographs and created a virtual baby capable of fostering deeper connections with the players than previous works such as YOTARO. We also experimented with a nursing simulator child care technique in this project and the result showed the potential of a nursing simulator technique for VR system. In the future we are hoping to create more VR entertainment systems featuring a complete VR world and real life physical object such as 2D picture and baby doll. The system will be updated to develop educational opportunities and it also develop the new media service like current photo media for family relationships.

Acknowledgement: Special thanks to Fujitsu Social Science Laboratory for helping us realizing this project.

References

[1] Laerdal Medical Corporation. "SimNewB". http://www.laerdal.com [Accessed 25 December 2016].

[2] Junichi IDO, Etsuko UEDA, Yoshio MATSUMOTO, and Tsukasa OGASAWARA. 2P2-C20 Robotic Telecommunication System based on Facial Information Measurement. *JSME annual Conference on Robotics and Mechatronics (Robomec)*, 2006.

[3] H. KOZIMA. A Robot that Learns to Communicate with Human Caregivers. *Proceedings of the First International Workshop on Epigenetic Robotics, 2001*, 2001.

[4] Giorgio Metta, Giulio Sandini, Lorenzo Natale, and Francesco Panerai. Sensorimotor interaction in a developing robot. 2001.

[5] Yuya MOCHIZUKI and Tatshya NISHIKIZAWA. "Real baby / Real family". Available at: http://ivrc.net/2016/ [Accessed 25 December 2016].

[6] Chiyoko ONO, Hiroki KUNIMURA, Madoka HIRAI, Tetsuya Wagner MATSUZAKI, Masatada MURAMOTO, Toshiaki UCHIYAMA, Kazuhito SHIRATORI, and Junichi HOSHINO. Design Consideration in Developing Baby Type Robot. *Bulletin of Japanese Society for Science of Design*, (56), 2009.

[7] B. SCASSELLATI. Theory of mind for a humanoid robot. *Autonomous Robots*, 2001.

[8] Naoko TOSA. Neuro Baby. *Journal of the Society of Instrument and Control Engineers*, 1995.