IADIS International Journal on Computer Science and Information Systems
Vol. 15, No. 2, pp. 137-147
ISSN: 1646-3692

HEAD MOVEMENTS IN THE IDLE LOOP ANIMATION

Maja Kocoń
West Pomeranian University of Technology in Szczecin. Poland

ABSTRACT

This paper describes the synthesis of virtual human head movements while waiting for a response in human-machine interaction systems. The described approach uses a three-dimensional human head model and kinematic chain of rigid elements. We proposed a method to perform a motion sequence of a virtual head in the idle mode. This mode is a state when a virtual person or robot performs subtle movements during the waiting phases between different interaction intervals. This form of action removes the problem of character freeze in motion and has a higher degree of user acceptance. This research covers issues related to creating trajectories of motion to rotate the head model's rigid objects. The proposed method has been used to trajectory generation for loop animation and executes smooth movements in the idle mode. Since the used algorithm has low computational requirements and introduces low interaction delays, we can use it in real-time operation in the HMI framework.

KEYWORDS

Idle Loop Animation, Human-Machine Interaction (HMI), Head Gestures, Virtual Agent, Motion Synthesis

1. INTRODUCTION

In recent years, human-machine interaction research (HMI) has shown increasing popularity in the use of virtual reality (i.e. serious games, systems with human-like interfaces). Therefore, in HMI systems, communication between the user and the machine can be performed using a virtual representation of a human, commonly known as an avatar. The primary goal of HMI researchers is to design an interaction close to interpersonal communication. Thanks to movable avatars, information can be transmitted using non-verbal signals such as facial expressions, head movements and body language, which may provide intuitive interaction and improve communication.

In nonverbal interaction, the main communication channel is face and head gestures, because it provides information about the person's emotional condition or intentions (Boker 2011). The face transmits a person's emotional state, but it does not properly consider the intensity of emotions (Cohn 2004) and the situational context. For this purpose, it is essential to consider the movement aspects of the head to full interaction. There are many publications on social behaviour emphasized the meaning of head motion in interaction. Harrigan (2005) and Graf (2002) examined the effects of head movement in non-verbal communication. For example, the basic head movements such as tilting and nodding are essential in conversation, in active listening. As a consequence, head-nodding can be used instead of verbal information like "yes"/ "no" (Boker 2011, Heylen 2005, Munhall 2004) or can be used for the point at something: "this one"/ "that one" (Tojo 2000) in addition, the meaning of the spoken word is determined by the head's movement (Sun 2011, Greenwood 2017).

In the works on the human motion, most research concerns the expression of emotions, and the concept of idle movements is often neglected. In real life, people perform a lot of movements, even when they don't do anything, but avatars and robots usually are “frozen” in standby mode. Because the main objective of this study's is to model human behaviour, we should consider the situation when the virtual person is waiting for interactions with the user. Idle movements are interpreted as subtle actions that occur while the avatar or robot is waiting for interactions with the environment and does not perform any tasks. The set of these subtle movements include head-swaying, small facial grimaces and eye blinking. This form of activities eliminates the moments when the avatar does not move (Kocoń 2014) what appears many times in HCI applications or games. As a result, the realism of the animation improves, and the chance of the uncanny valley (MacDorman 2006) effect decreases. Several applications still ignore this issue, which results in unnatural behaviour for the human like constant position or repetitive sequence. The significance of using such gestures has been presented in (Salem 2011), where two interactions schemes were compared. The gesture-based robot was perceived more positively by users because it displayed social competencies. Head gestures complete verbal communication and can increase the realism of the actions of virtual characters. Therefore three-dimensional human is popular in HMI applications such as social robotics (Ligthart 2018), driver assistance systems (Schwarz 2017) where we have user-friendly interfaces as custom avatars, or in serious games (Cai 2018).

Idle aspects have been characterized in (Jung 2013), where Jung et al. demonstrated the significance of body motion in no interaction conditions by testing five selected roles of movement used when one of the subjects of observation does not speak, and another one is idling. The goal of their research was to find rules to generate improved motion system but they did not include the anthropomorphic aspects of motion. In (Egges 2004) an approach based on Principal Component Analysis is presented, where they generate two layers of subtle motions like variations in body posture and change of balance when a person is standing. This approach is used to obtain the continuity between used animations without additional motion sequences. Based on a dedicated database of full body postures and different sequences of motion, idle animation is generated to represent a specific person.

An interesting application of idle movements is presented in (Song 2009). Idle motions are extracted from video data, where participants were observed at the information centre. They recorded ten people, and recording time was about 10 minutes per each subject of observation. They analyzed idle motion and selected six behavioural patterns which were applied to the robot with the 3-DOFs neck. The head and neck motion was used to the interaction like eye contact and looking around for the user. In (Asselborn 2017) Asselborn explored the effect of a robot's idle type movements during interaction with children. In their approach, they used a set of gestures with the three level intensities such as low, medium and high. Results show that the user perception of humanity and friendliness increases with the level of intensity in the idle mode. Cafaro et al in (Cafairo 2009) describe gaze behaviour for animated avatars that are simply idling. To design gaze behaviour they have observed the subject of observation in two situations, first: the main bus terminal for waiting behaviour and the main shopping street downtown for walking behaviour. Additonaly, they have constructed a virtual model of the space to explore their testing behaviours. Base on it, they observed the relation between gaze targets and proxemics for waiting person.

In this work, we propose an approach to provide idle motion in virtual space. Because the idle mode is not a primary movement in the interaction scheme and has an effect on continuously performed actions, it is desirable to use a technique with a low computational cost. Since for the humans the most readable form of information is the message communicated by another person, we decided to focus on motion synthesis using the human-like virtual character.

2. HEAD MOTION MODELING

For the head animation, a three-dimensional model was used. The model was built using the properties of mid-age woman's head, which was created based on basic information on facial topology, including aspects of the anatomical structure of the flexible face areas and rigid head. In our visual representation of the head, we have taken the essential anatomical features of the human face and head like head proportions, and facial shape into consideration. The main proportions of the head like the distance between the jaw and the top of the head, division of the head and the position of the ear is shown in Figure 1.

Initial motion analysis was made on a test group of 30 people in the situation of waiting for an interview. Sequences were recorded in the front, side, and top view to obtaining ranges of motion for three rotations. Waiting for the interaction took from 15 to 30 seconds, and a total of 180 sequences were obtained with different types of waiting for situations at each development phase of our avatar.

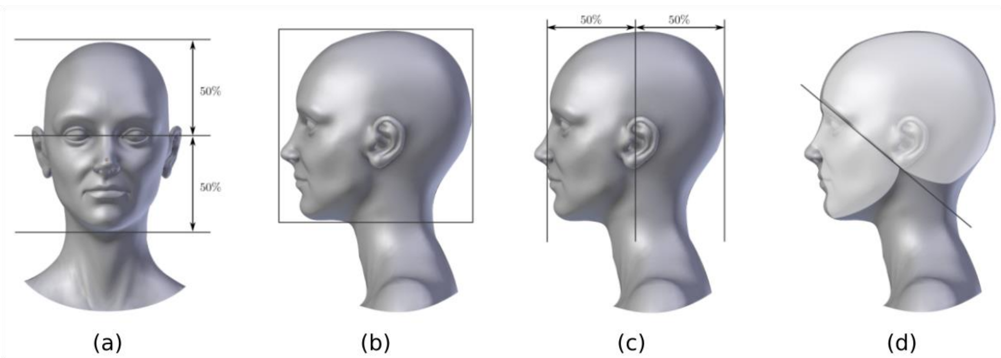

Figure 1. The main proportions of the head: the distance between the jaw and the top of the head (a); the head viewed in profile (b); the position of the ear (c); division into upper skull and facial part (d)

Based on the recorded sequences, types of motion for three rotations in the idle situation were observed, what is shown in Figure 2. The ranges of motion obtained for selected types of rotation were computed using the analysis of recorded sequences for three rotations in the real situation. Obtained ranges for selected gestures of the human head are presented in Table 1.

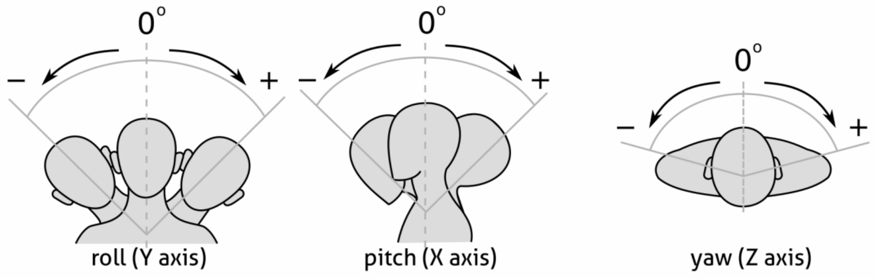

Figure 2. Selected gestures of the human head

Having anatomical aspects of a human motion, we decided that head should be analyzed as a rigid body; therefore, for motion description, we have used the set of non-deformable elements connected by joints. The localization of the elements that correspond to the head and neck is depicted in Figure 3. The presented kinematic chain contains three elements, where element C_1 indicate pitch motion, element C_2 roll rotation and element C_3 yaw motion. The obtained ranges of motion expressed in degrees for discussed types of rotation are shown in Table 1.

Table 1. The ranges of motion expressed in degrees for selected gestures of the human head

Roll		**Pitch**		**Yaw**	
Min	**Max**	**Min**	**Max**	**Min**	**Max**
-14	19	-12	20	-23	32

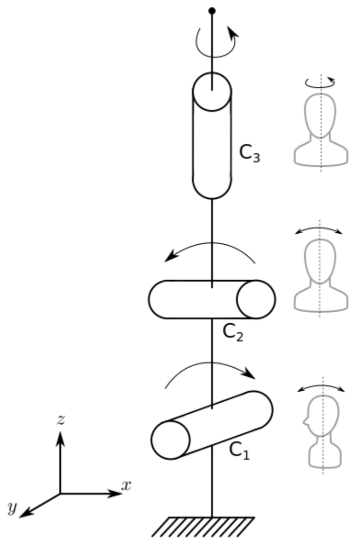

Figure 3. The chain of rigid elements defined for head movements

3. IDLE MOTION SYNTHESIS

In idle mode, the subtle head gestures can be modelled by slight rotations of rigid elements, where the movements should be random for every axis in three-dimensional space. To generate such movements, we have decided to create rotation trajectories where every angle change is performed by correcting the actual angle by the trajectory value in each animation frame. For this purpose in this work, we propose a scheme to generate trajectories for modelling a human head model by rotating rigid elements using pseudo-random source.

The procedure of trajectory generation can be summarised in the following steps:

1. Generate a sequence of numbers where any value $\mathrm{r}(n)$ has to satisfy the condition $|\mathrm{r}(n) - r(n-1)| = |\mathrm{r}(n) - r(n+1)| = 1$, where $n = 1, \dots, N-2$ and N is the length of the trajectory.
2. Smoothing obtained trajectory by performing the convolution of the data with the smoothing window kernel.
3. Normalization and adjustment the smoothed trajectory to the expected range of changes.
4. Modify the trajectory by introducing the split the trajectory into two parts, exchange them and join together to make a seamless trajectory.

The scheme of the proposed procedure is presented in Figure 4. First, as a pseudo-random source of numbers, we have decided to use a linear congruential generator (LCG) algorithm *(Press 2007)* due to its simplicity and ease of obtaining the same series of numbers for a given configuration. Such a generator can be described as:

$$r(n+1) = [a \cdot r(n) + c] \, mod \, m,$$

where a, c and m are constants which define the generated sequence of pseudo-random numbers, $n = 0, \dots, N-1$ and $r(0)$ is the seed value. In our experiments, we have use $a = 3101517285$, $c = 672311$ and $m = 2^{32}$ which have been selected during experiments by tweaking the values, applying them in trajectory generation and assessment of the final head movements. They can be selected arbitrarily, and the proposed values were suitable for our application. The values have no important meaning in this context but can be used to reproduce experiments. The trajectory is generated using the formula $u(0) = 0$, where $u(0) = 0$, $n = 0, \dots, N-1$ and δ - changes between consecutive values.

In the next step, smoothing is performed to remove rapid changes which are undesirable for natural-looking animation. The smoothing process has been carried out by using low-pass filtering with a smoothing window kernel (*Smith 1997*). The filtering operation is realized by convolution of the generated trajectory with kernel function (*Press 2007*), what is depicted in Figure 4, where W is the window size and $u(n-m) = 0$ for $n - m < 0$.

Many window functions can be used for smoothing, and the level of smoothness is mainly dependent on kernel size. The selection of the window is not significant for the considered application, so we decided in our study to use the Hanning window *(Poularikas 1998)* defined as you can see in the proposed scheme, where: $m = 0, \dots, W-1$.

In the last phase, the smoothed trajectory is adjusted to the desired range change where all values should be in the range $[\alpha_1, \alpha_2]$ and $\alpha_1 < \alpha_2$. Therefore, a new stage has been applied where the smoothed data is scaled and adjusted to the final range based on $\hat{u}(n)$, where s_{min} and s_{max} denote the minimal and maximal value of $s(n)$ respectively.

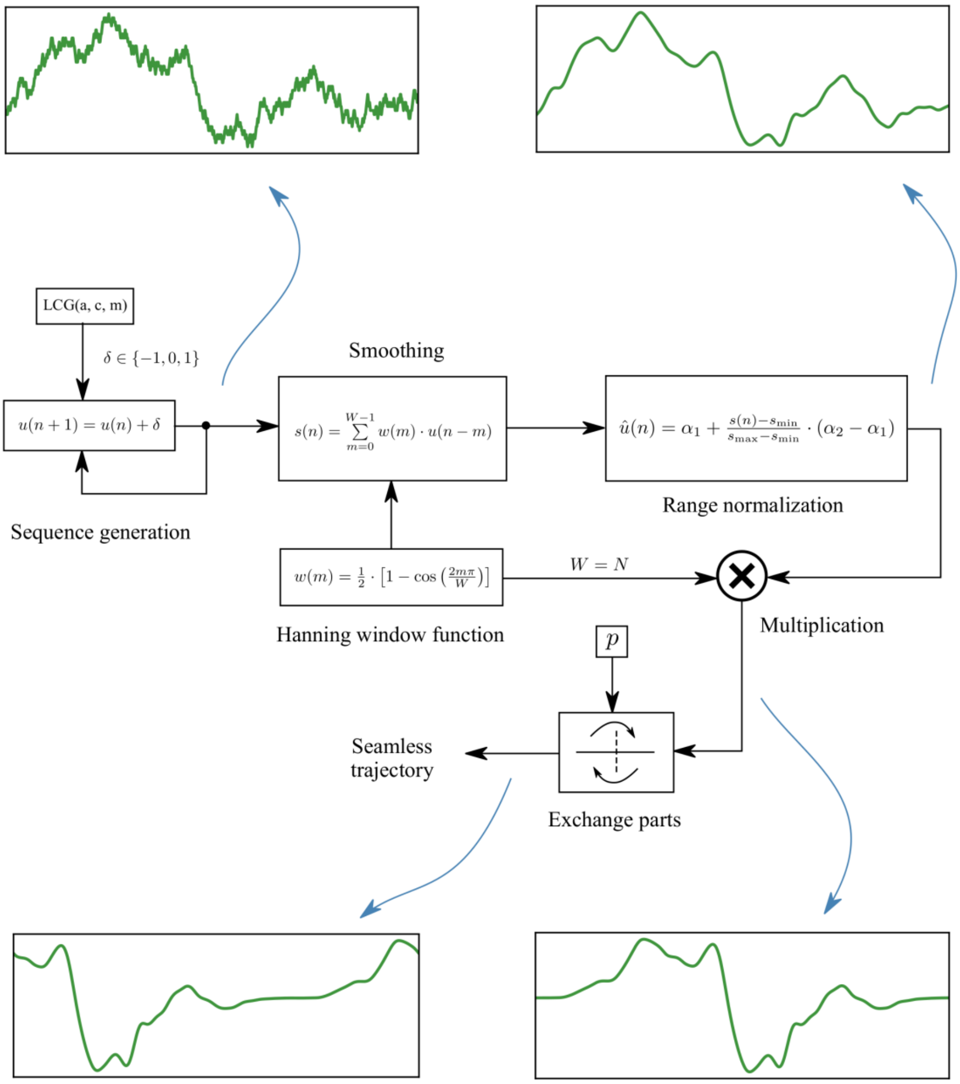

Figure 4. Diagram of the proposed scheme

As an example of a final $\hat{u}(n)$ trajectories for head animation in idle mode, trajectories generated for each rigid object are presented in Figure 5, where the parameters are shown in Table 2.

Table 2. Parameters of example trajectories

N	W	$r(0)$	α_1	α_2
900	59	127	-4	12
900	59	259	-13	13
900	59	8821	-9	9

Figure 5. Example seamless trajectories for 3D rotation of three elements of the model

The advantages of the proposed approach include the simplicity, low computational cost and reproducibility of trajectories for a given length $\boldsymbol{N}$ and random seed $\boldsymbol{r(0)}$. The configuration parameters allow us to get a whole range of functions for animation process modelling.

The problem with the obtained trajectories is the fact that despite they are seamless, the boundary values are different in every case as depicted in Figure 3, and their parameters are shown in Table 2. Since the rotation starts from any head position, it is important to use the trajectory as an incremental value. Therefore, the trajectory should start from zero value. To modify the trajectory, we have introduced the split point position p, where we split the trajectory into two parts, exchange them and join together. We can select the splitting position arbitrarily, but it has to be selected at a value equal to zero in this case. In Figure 6 you can observe split point p and all process of the trajectory looping.

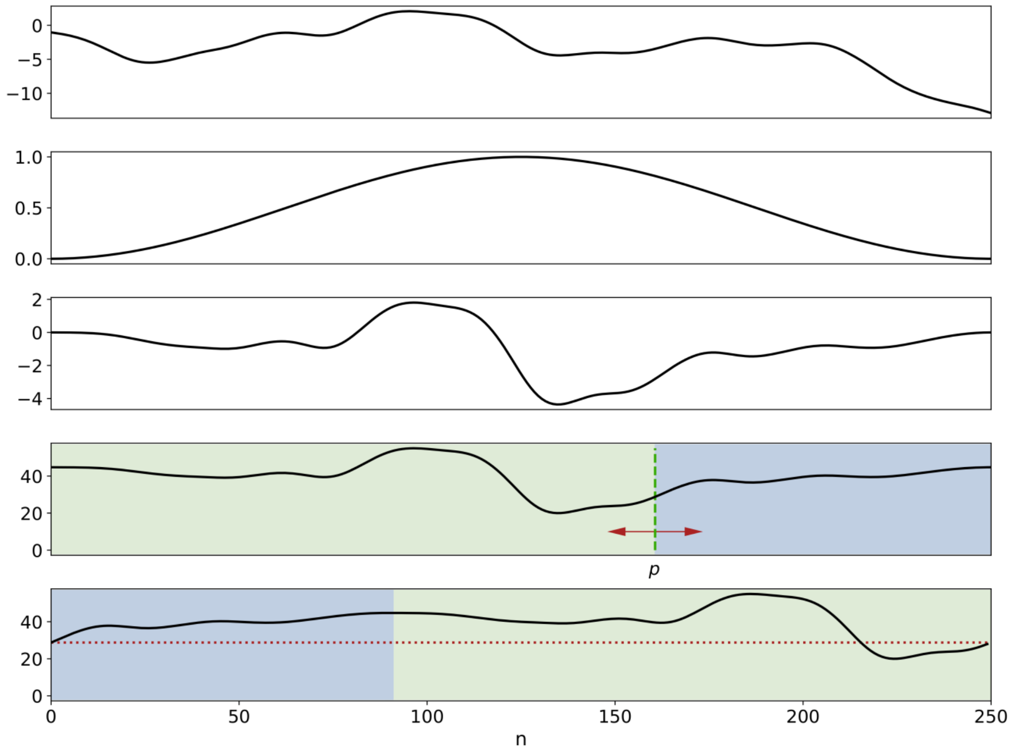

Figure 6. Process of the trajectory looping (from top to bottom): input trajectory, Hanning window function, input trajectory multiplied by the window function, exchange two parts of the trajectory around the split point p, the final seamless trajectory

Assuming that the angle may change in two directions, the function has zero-crossing points. To find out the zero crossing positions, we have used the function:

$v(n) = sign[\hat{u}(n+1)] - sign[\hat{u}(n)]$ and if the $v(n) \neq 0$ then the zero crossing occurs at position n. In our approach, when there are more zero-crossings than one, we select it randomly. The modification of trajectories from Figure 5 are depicted in Figure 7 where zero-crossings selected as split points are marked, and their values are presented in Table 3.

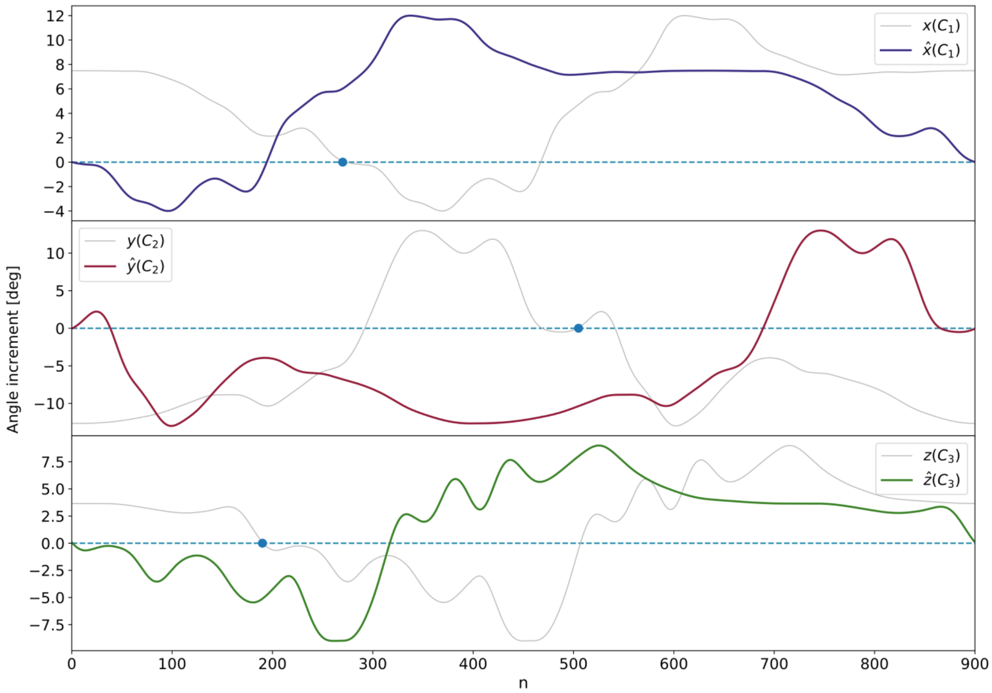

Figure 7. Modified trajectories using marked zero-crossings

Table 3. Parameters of final seamless trajectories

N	W	$r(0)$	α_1	α_2	p
900	59	127	-4	12	273
900	59	259	-13	13	503
900	59	8821	-9	9	190

Finally, based on the proposed approach, we have obtained motion sequences for idle mode, an example of a few frames of animation is shown in Figure 8.

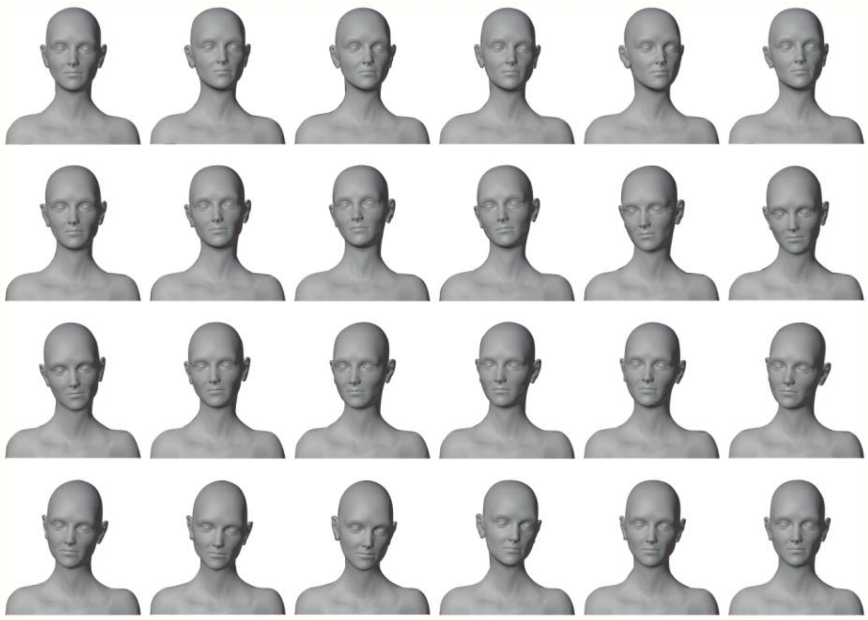

Figure 8. Example animation frames for the person in idle mode

4. CONCLUSION

The primary part of our study is to improve the situation when the virtual model of a person is inactive in terms of external events. This paper introduces the method of modelling head movements of artificial avatars in the state between interactions.

The method suggested in this paper makes it possible to accurately simulate subtle head movements by incorporating motion trajectories for the modification of the angle of rigid bodies used to move the human head model. To obtain final looped trajectory, we have proposed algorithm which exploits the basic sequence generator, smoothing and normalizing stages and trajectory correction phase. The most time-consuming part of our technique is the smoothing stage which incorporates the filtering operation in the time domain.

REFERENCES

Asselborn, T. et al, 2017. Keep on moving! Exploring anthropomorphic effects of motion during idle moments. *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).*

Boker, S. et al, 2011. Something in the Way We Move: Motion Dynamics, Not Perceived Sex, Influence Head Movements in Conversation. *Journal of experimental psychology. Human perception and performance, Vol. 37*, pp. 874-91.

Cafaro, A. et al, 2009. Animating idle gaze in public places. IVA '09: Proceedings of the 9th International Conference on Intelligent Virtual Agents, Springer-Verlag, 250-256.

Cai, Y. et al, 2018. *VR, Simulations and Serious Games for Education* (1st. ed.). Springer Publishing Company, Incorporated.

Cohn, J. F. et al, 2004. Multimodal coordination of facial action, head rotation, and eye motion during spontaneous smiles. *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, pp. 129- 138.*

Egges, A. et al, 2004. Personalised Real-Time Idle Motion Synthesis. *12th Pacific Conference on Computer Graphics and Applications* (PG 2004), Seoul, Korea, pp. 121-130.

Graf, H. P. et al, 2002, Visual prosody: facial movements accompanying speech. *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, Washington, DC, USA, pp. 396-401.

Greenwood, D. et al, 2017. Predicting Head Pose from Speech with a Conditional Variational Autoencoder. *Proceedings of Interspeech 2017, ISCA,* pp. 3991-3995.

Harrigan, J. et al, 2005. *The New Handbook of Methods in Nonverbal Behavior Research.* Oxford University Press, Series in Affective Science.

Heylen, D, 2005. Challenges Ahead: Head movements and other social acts during conversations. *Proceedings of the Joint Symposium on Virtual Social Agents.* The Society for the Study of Artificial Intelligence and the Simulation of Behaviour, pp. 45-52.

Kocoń, M, 2014. Facial expressions modeling for interactive virtual environments. *19th International Conference on Methods and Models in Automation and Robotics,* Miedzyzdroje, Poland, pp. 744-747.

Ligthart, M. et al, 2018. Reducing Stress by Bonding with a Social Robot: Towards Autonomous Long-Term Child- Robot Interaction. ACM/IEEE International Conference on Human-Robot Interaction, Chicago, United States.

Jung, J., et al, 2013. Guidelines for Contextual Motion Design of a Humanoid Robot. *Int J of Soc Robotics, Vol.*5, pp. 153-169.

Macdorman K., F., 2006. Subjective Ratings of Robot Video Clips for Human Likeness, Familiarity, and Eeriness: An Exploration of the Uncanny Valley. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science.*

Munhall, K.G. et al, 2004. Visual Prosody and Speech Intelligibility Head Movement Improves Auditory Speech Perception. *Psychological science, Vol.* 15, pp. 133-137.

Poularikas, A., 1998. *Handbook of Formulas and Tables for Signal Processing. Electrical Engineering Handbook.* Springer Berlin Heidelberg.

Press,W. H. et al, 2007. N*umerical Recipes 3rd Edition: The Art of Scientific Computing.* Cambridge University Press, USA, 3 edition.

Salem, M. et al, 2011. Effects of Gesture on the Perception of Psychological Anthropomorphism: A Case Study with a Humanoid Robot. Social Robotics. ICSR 2011, Lecture Notes in Computer Science, Vol. 7072, Springer, Berlin, Heidelberg.

Schwarz, A., 2017. DriveAHead - A Large-Scale Driver Head Pose Dataset. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, pp. 1165-1174.

Smith, S. W., 1997. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, USA.

Song, H. et al, 2009. Design of idle motions for service robot via video ethnography. RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication 2009, pp. 195-199.

Sun, X et al, 2011.Towards visual and vocal mimicry recognition in human-human interactions. *2011 IEEE International Conference on Systems, Man, and Cybernetics*, Anchorage, AK, pp. 367-373.

Tojo, T. et al, 2000. A conversational robot utilizing facial and body expressions. *2000 ieee international conference on systems, man and cybernetics,* Nashville, TN, Vol. 2, pp. 858-863.