*Human Factors in Virtual Environments and Game Design,* Vol. 50, 2022, 261–266
https://doi.org/10.54941/ahfe1002087

# Karaoke Game with Body Movement Tracking for Investigating Singing-motor Interactions: System Proposal

**Luiz Fernando Pedroso and Ryota Horie**

Graduate School of Engineering and Science, Shibaura Institute of Technology, Japan

## ABSTRACT

This paper presents and discusses the design, implementation and initial results of a game system using a singing-motor paradigm, using karaoke singing combined with a human-computer interface through body movement recognition. The purpose of this system is to collect experimental data on how karaoke singing and body movement interact and how they affect each other, in order to analyze the data from the perspective of human-machine interface, immersion, gameplay and entertainment. At the end, the results and outputs of the development are presented, and some follow-up proposals are discussed.

**Keywords:** Human-computer interface, Game design, Karaoke, Singing, Multitasking, Body tracking, Gestor sensor

## INTRODUCTION

Along with the constant evolution and growth of the games industry, the human interfaces for gaming also evolved significantly, from simple 2 button controllers to sophisticated multimodal gesture interfaces (da Luz, 2014). More specifically, the Music Games genre has taken advantage of the interface progress to innovate human interaction possibilities such as: real-like instruments, floor stepping sensor for dancing, hand movement in Virtual Reality and even full body tracking.

When it comes to music and rhythm game sessions, players usually are required to use a lot of focus, motor coordination and good timing (Salsabilla et al., 2020, Song et al., 2019). However, multitasking (also called dual-tasking in the cognitive science field) is a big challenge for the human brain, especially when it comes to tasks that require a high level of attention (Pashler, 1994). Therefore, considering the cognitive challenges of multitasking and a multimodal body movement interface, this study introduces a game system with a singing-motor paradigm, in which players have to sing karaoke and perform body movements simultaneously. While singing, the player will be challenged to break 3D obstacles that spawn in rhythmic patterns, using his own body to control an in-game avatar and perform poses, punches and kicks in the correct timing. The purpose of the system is to investigate how the singing-motor tasks interferes with each other, including how

 

musical aspects of the game design can be used to obtain different results in terms of difficulty, entertainment, and game experience.

## METHOD

The system configuration consists of 2 modules: The first module is a 3D game module developed in Unity 3D (Unity Technologies) environment, a game development engine, that handles the GUI and performs the body tracking tasks. For each song, the module receives as input a set of files that includes the music audio, lyrics, and the level timeline containing position and spawn timing of obstacles. As output, it saves a file containing the player's full body positions over time for each game session. The second module runs in a smartphone and it is responsible for the karaoke management, playing the background music and reading voice data from the player to evaluate the player singing skill. High scores mean the player has a good singing performance. The application receives as input the music audio and lyrics files, and saves the voice captured by the microphone during the game session. Additionally, the application sends music timing and score data in real time to the first module via TCP/IP communication.

During the game, a 3D humanoid avatar replicates player movements made in real time, while obstacles appear in the form of asteroids that are broken through the correct player interaction. The avatar resource asset is a 3D model named "Unity Chan" (© Unity Technologies Japan/UCL), and the obstacles asset is named "Breakable Asteroids", both from Unity 3D asset store, under the Standard Unity Asset Store EULA License agreement (Unity Technologies 2020).

The game user interface is performed by a Microsoft Azure Kinect (Microsoft Corporation). This device consists of a set of cameras and depth sensors that can retrieve image and 3D information from the environment and provide the position data of all joints from detected human bodies (Microsoft Corporation 2021). Thus, in the game, the Kinect device detects the player's full body position, controlling the in-game avatar in real time. In order to enable investigation of different motor coordination challenges, the game provides 2 different modes to be combined with singing: Posing mode and Fighting mode, to be detailed below.

### Game Mode: Posing

In this game mode, the player is challenged to break obstacles flying from the front. For break the obstacles, the player has to make instructed poses in particular timing while the keeping good performances of singing a karaoke song. Figure 1 shows the implemented game interface running in Pose mode (parts of showing lyrics are omitted in the figure).

For instructing poses, a horizontal bar guides the player to reproduce the body positions that are displayed, while obstacles (asteroids) fly in the avatar's direction. Each pose corresponds to 1 obstacle. The guide bar scrolls pictures from right to left of the screen and the player should be at the correct position when the picture reaches the red line in the middle. At this moment, the player's body joints orientations are registered and compared with the

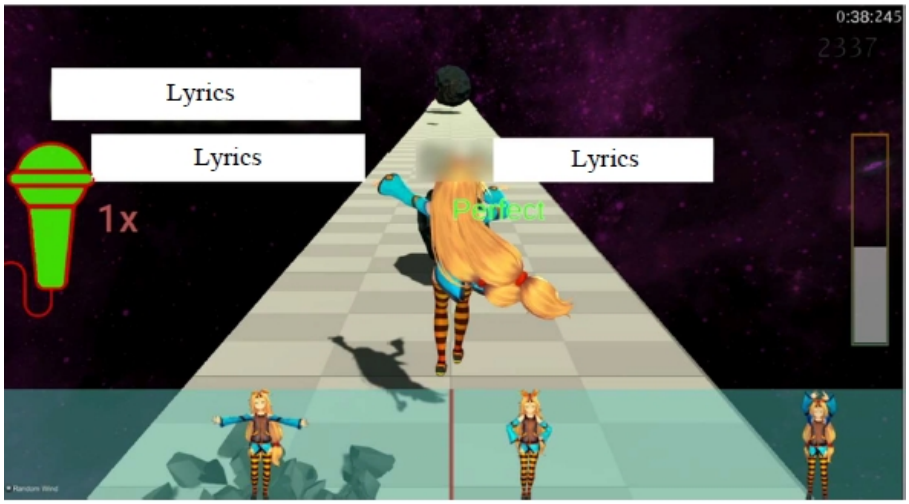


**Figure 1:** Game interface demonstration in pose mode (© Unity Technologies Japan/UCL).

target body position. Then, a precision score is generated from 0 to 100 as the evaluation of player's motion and classified as Perfect, Good, Fair or Bad. A text feedback corresponding to the precision score is presented.

For feedback of singing performance, the microphone icon shown at the left part of Figure 1 is introduced. The system receives the singing evaluation data from the smartphone module in real time, ranging from 0 to 100 points and displays it by filling the microphone icon. An empty icon means bad singing performance, while a full icon means perfect singing score.

In addition, there's a singing score threshold. Below the threshold, the microphone icon filling will be red, meaning disability of breaking obstacles even with making the pose successfully. Above the Threshold, the icon filling will be green and the avatar displays a particle aura, indicating that the singing score is good and breaking obstacles can be done by performing good posing.

Thus, the amount of points the player receives after each pose is calculated and in the following way, and shown in Table 1: If singing score is over the threshold and the pose score is classified as Good or Perfect, both scores will be multiplied and added to the game's total score. If the current singing score is bellow threshold, resulting score will be 0, no matter the posing. Similarly, if the pose is considered Fair or Bad, the resulting score will be 0.

Finally, in order to encourage the player to work hard both in singing and moving, a mechanism was introduced: The Power Bar, showed in the right part of the screen in Figure 1. In Pose Mode, the resulting score is added to the game's overall score and, also, fills the power bar with energy. At the end of each short song session, a big obstacle appears and can only be broken if the power bar has a minimum amount of energy. In other words, if the player focuses on only singing or only posing, the resulting score will be low, leading to a failure when breaking the big obstacle. The difficulty and amount

**Table 1.** Score calculation method for Posing mode.

| Singing Score input | Pose Score input | Result: Obstacle | Result: Obtained Score |
|---|---|---|---|
| Over threshold | Good or Perfect | Broken | Singing * Posing |
| Over threshold | Bad or Fair | Not broken | Zero |
| Below threshold | Good or Perfect | Not broken | Zero |
| Below threshold | Bad or Fair | Not broken | Zero |

of score/energy needed, as well as the threshold values are configurable to provide gameplay experience tweaks.

### Game Mode: Fighting

In this game mode, the player can move freely to control the avatar horizontally. The obstacles come flying in the avatar's direction, and the player should hit punches and kicks to break it and earn points. Similar to thje Posing mode, it is required to have a good motion score and a good singing score combined in order to be able to break obstacles. Minimum requirement of singing performance is indicated by the microphone icon and the particle aura is same as in the pose mode.

Additionally, in order to benefit the immersion and encourage wide movements (in other words, prevent players from "cheating the system"), the speed of blows is calculated in real time and taken in consideration, meaning that slow and weak punches/kicks aren't considered.

## RESULTS

After the requirements planning and designing phases, the game system was developed, implemented and tested successfully. It started with a functional and simple prototype and, as the system was being tested and used, needs for new mechanics and features emerged, resulting in the configuration currently presented in this study. For instance: The power bar and fighting aura mechanics were introduced to prevent players from focusing on only singing or only moving to earn points, while rewarding players who perform wide and strong gestures.

At last, for testing purposes, some subjects were invited to play the game, validating the functionality and other technical aspects of the system such as: game mechanics, communication between modules and computational performance.

The results are a playable and functional game system, correctly recording the data needed to carry out future experiments. A preliminary interview with the testers showed that, as expected, the multitask paradigm of singing-motor coordination is a hard challenge for the players, and the familiarity with the lyrics and the song is a key factor for the performance. Finally, a preliminary adaptation of the system for Virtual Reality (VR) environment was implemented. The game mechanics and architecture were not changed, but the deployment with a VR device using first person point of view was achieved successfully.

## DISCUSSION

Although pose and fight modes may seem a little similar, there's a crucial difference when evaluating it from the multitask paradigm perspective and its interaction with music: fighting means doing a quick movement at the right time, while posing means being still in an exact position at the right time. Because humans have a natural rhythmic connection with music through the pulse sensation named as beat (Ficth, 2011), the modes will result in different motor interactions between the body and the music, since fighting implies in a strong movement during the occurrence of the beat, while posing means be stopped in the target position during the occurrence of the beat. Thus, we believe that these modes can provide freedom and a wider range of possibilities regarding design of experiments.

Human being presents a natural interaction with rhythm, that can be exemplified in terms of the cognitive process by inferring beats from patterned acoustic stimulus and generating motor outputs such as dancing, clapping or playing drum patterns (Fitch, 2011). Therefore, with these concepts in mind, the next steps of this study will be towards the design of experiments. More specifically, we intend to study and analyze the singing-motor paradigm under the perspectives of game design and human interface, including musical aspects of the song, such as strong beats, weak beats, out of beat, complexity of rhythmic patterns and similarity with the melody sung in karaoke. With the study's results, we intend to seek answers for questions such as "How to make the singing-motor challenge more entertaining?", "How to design it easier or harder?", "What are the best methods to generate rhythmic motor challenges considering the musical aspects of a song?".

## CONCLUSION

This study showed the design and implementation of a game system that uses karaoke singing and performing body movements simultaneously in order to investigate the singing-motor relation between these tasks and music. The system architecture was presented and its modules were explained, followed by introducing the game mechanics and modes. The successful implementation of the system was presented and, to encourage players to strive in both singing and moving, two game modes were created. Moreover, the cognitive aspects of human and rhythm interaction are briefly discussed as a background for our motivation and the future design of experiments.

For next steps, we expect to design several singing-motor scenarios, conduct the experiments, further analyzing the collected data. For future improvements, it is planned to completely adapt the game system to Virtual Reality environment, increasing immersion and widening the system's application possibilities.

## REFERENCES

da Luz, A. (2014), "Evolution of the physical interfaces in videogames as a support to the narrative and the gaming experience."

Fitch, W. (2011), "The biology and evolution of rhythm: Unraveling a paradox." *Language and Music as Cognitive Systems*.

Microsoft Corporation 2021, "*Azure Kinect DK hardware specifications*", viewed 12 January 2022, <https://docs.microsoft.com/en-us/azure/kinect-dk/hardware-specification>.

Pashler, H. "*Dual-task Interference in Simple Tasks: Data and theory*." Psychol. Bull. 116(2), 220–244

Salsabilla, G. A., Fabroyir, H., Herumurti, D., Kuswardayan, I., and Hidayati, S. C. (2020). *"Tachyon: Multiplatform Rhythm Game with Automatic Beatmap Generation."* CENIM 2020 - Proceeding: International Conference on Computer Engineering, Network, and Intelligent Multimedia 2020, 162–167.

Song, D. H., Kim, K. B., and Lee, J. H. (2019). *"Analysis and evaluation of mobile rhythm games: Game structure and playability."* International Journal of Electrical and Computer Engineering, 9(6), 5263–5269.

Unity Technologies 2020, "*Asset Store Terms of Service and EULA*." viewed 21 December 2021, <https://unity3d.com/legal/as_terms>