1 Introduction

In this paper, we outline two new 3DCG technologies, MoCaToon and AniFace, which have been developed to improve the efficiency of the movie production process in the labor-intensive world of Japanese Anime, where much of the work is still done by hand. MoCaToon is a technology that enables motion capture systems to be used in the animation production process, and AniFace enables animators to customize 3DCG characters according to their specific requirements, especially in terms of character style and motion.

2 MoCap data distillation for cartoon

Using our system, the motion capture (MoCap) data distillation process automatically extracts key-poses while preserving the characteristics of the original motion. A minimum number of postures are selected so that they can imitate the style of limited animation and then animators can modify them easily and intuitively. Our approach regards each clip of MoCap data as a motion curve in a pose space of high dimension, and then estimates the preservation of motion expressed by those of the curve's shape, whose variation is approximately estimated by the total lengths, computed with the pose distances metric as \( \sum_{i=0}^{M-1} D(i,i+1) \rightarrow \max \) where \( f_i \) is the frame index of \( i \)-th key-pose, \( D(i,i+1) \) denotes the pose distance between the \( f_i \)-th and \( f_{i+1} \)-th frames, and \( M \) denotes the total number of key-poses decided by an animator. This maximization is efficiently computed using dynamic programming. This data-distillation technique is more suited to key-frame extraction for cartoon production than other relevant existing methodologies.

3 Key-frame re-arrangement for lip-sync

Lip-sync animation systems based on blend-shapes (linear shape interpolation), normally generate the key-frames for each mouth shape by analysis of target speech. In the style of low frame-rate animation, the simple phoneme blend-shape mapping approach has been implemented a multi-phoneme assignment problem where several few phonemes are, at times, assigned to the same frame. AniFace implements a key-frame re-arrangement mechanism to prevent this problem. The overall process of AniFace is as follows: (1) target speech is divided into the phoneme segments using our original speech recognizer, (2) the phonemes are converted to the key-frames of each mouth shape using a heuristically pre-defined blend-shape mapping table, and (3) the key-frames are rearranged to fit the user-defined frame-rate.

4 Experimental result

The usability of our two systems was evaluated based on an experiment in which an original anime sequence from The Galaxy Railways was reproduced in 3DCG. Animation of faces, clothing and hair was produced by hand this time. Some of the background images were 2D and toon-rendering was introduced to the characters. Five CG creators and one MoCaToon operator were engaged in this production. Figure 1 shows a shot from the anime movie. In the original 2D animation production of 330 cuts, 10 days were required for scene layout, 17 days for original key-frame drawing and 5 days for motion and effect. Using our system, however, only 2 days were required for motion capturing, 5 days for data editing, 5 days for character layout, 5 days for 3D animation, 7 days for face, clothing and hair animation and 4 days for final rendering. Although it is somewhat difficult to compare these two processes, the mass production of anime becomes possible when motion data and 3D character data are accumulated in database. Complex scenes, such as action or crowd behavior in particular, can be realized through the introduction of motion capture systems. The overall quality of animation improves because the anime production workflow drastically changes in which feedback production is possible like animation.

Figure 1: A scene from The Galaxy Railways is reproduced, appearing like 2D hand-drawn anime.

5 Conclusion

MoCaToon and AniFace combine to create a revolution in a cartoon anime production. Future work will include the more efficient anime production by introducing physics based facial expressions, clothing motion and hair dynamics also by adding creator's directable and editable functions in 3D simulation.

Acknowledgements

This research is supported by Japan Science and Technology Agency, CREST project.