ABSTRACT OF DISSETATION

TIME-VARYING MESH PROCESSING FOR SEGMENTATION, SUMMARIZATION, INTERPOLATION, AND COMPOSITION

(３次元ビデオ処理：分節化、要約、補間、編集)

Jianfeng Xu, Aizawa-Yamasaki Lab., Dept. of Electronic Engineering, The University of Tokyo

Time-Varying Mesh (TVM), which is composed of a sequence of mesh models, captures the realistic and dynamic scene of the real world including a human's shape and appearance from multiple synchronized videos. This thesis is to propose key techniques of TVM processing including motion segmentation, sequence summarization, frame interpolation, and motion composition. There are many challenges in the TVM data. For instance, the raw data in mesh models are in low level without any semantic information, which causes the semantic gap. Because each frame is generated independently in the generating system, the topology and the number of vertices may vary frame by frame, which causes correspondence information is difficult to obtain. There is some noise that comes from the generating system, which requires the proposed algorithms are robust to noise. Because the topology in different frames may change, many existing algorithms cannot be employed directly. Another issue is the efficiency to deal with the huge data that come from TVM.

Management of TVM is essential to use the contents efficiently especially when the contents become very huge. Fundamental techniques are proposed to manage and re-use the TVM data. For example, a similarity measure is necessary in many applications like retrieval, which is useful in TVM management. If a TVM sequence is segmented into single motions, it will become much easier to process further. And summarization of a long sequence is helpful to reduce the viewing time. We also propose an editing system to re-organize the frames for a new purpose. Frame interpolation is an important method to generate new frames from the original frames. By these techniques, it is possible to manage and re-use the TVM data efficiently.

Similarity measure is a powerful tool to reflect the motion degree in a TVM sequence, which is the base of motion segmentation, summarization, and composition. Three types of similarity measures are proposed based on statistical methods such as histogram and

mutual information. Those statistical methods are very suitable for huge and noisy data. By analyzing the similarity measures, a TVM sequence is segmented into single motions, which is the pre-requisite of our other algorithms including sequence summarization and motion composition. Two approaches are proposed for two different types of motions in TVM sequences, namely, periodic motions and non-periodic motions. Our experimental results demonstrate high precision and recall are achieved.

Summarization of a TVM sequence is necessary in those applications with limited storage, bandwidth, and other resources. For this purpose, we propose an efficient scheme to extract key frames in each single motion or shot. However, it is conflict to summarize a sequence compactly and completely. Therefore, we consider the trade-off between the rate and distortion in our cost function. Unlike many algorithms in 2D video, it is not required to decide the key frame number in advance in our scheme, which is more friendly to the user.

Currently, one serious problem of TVM is that it is time-consuming and expensive to generate TVM sequences. Therefore, it is necessary to re-use the original data to create new sequences. An editing system is presented for this purpose. A motion graph, which includes all the motions in the database, is defined and constructed considering the smoothness of transitions between motions. The user selects some desired motions from the motion database. Then, an optimized path is searched between every two desired motions by a modified Dijkstra algorithm. The edited sequence is rather smooth.

In this thesis, an approach is presented for frame interpolation in TVM in high level, which is very challenging due to the absence of semantic information in mesh models. A semantic human model is employed to estimate the motion vectors of the object. The mid-frames are interpolated linearly by the estimated motion vectors. Our experimental results demonstrate the effectiveness of the approach. This technique can be applied in many areas such as frame rate up-conversion and motion blending, and can support the motion analysis in high level.

The proposed algorithms in this thesis are key techniques in managing and re-usage of TVM data, which involve not only low level processing but also high level processing in mesh models.

Jianfeng Xu, June 18, 2007