Sprite Aided Video Compression and Segmentation

With the rapid developments on multimedia and Internet, more and more images and videos are transmitted and viewed in digital form. Video representation and compression are at the foundation of such kind of multimedia applications. Object-based video coding, also known as the representatives of content-based video coding, has been proposed to represent and compress the foreground and the background separately, which aims at not only achieving high coding efficiency but offering many desired functionalities that never exist in frame-based coding schemes. Accordingly, one of the main problems of object-based video coding arises from how to efficiently compress the video objects represented with different motions. A further main problem of object-based video coding arises from how to segment the video into meaningful regions, i.e. the background and the foreground.

The sprite technique provides a concise and efficient method for the representation and compression of the video object. A sprite is an image composed of pixels belonging to a video object visible throughout the video sequence. Since the sprite image has contained all parts of an video object that were at least visible once in a video sequence, it can be used for direct reconstruction or predictive coding of the video object plane (VOP), e.g. the background VOP, at every time instant. In addition, the sprite image can provide the reference background for moving object extraction. Therefore, the sprite image is very useful for video coding and segmentation as well as the joint implementation. Towards this goal, this dissertation is focused on developing the novel video coding and segmentation algorithms based on the sprite technique.

Firstly, a fast and robust sprite generation algorithm is proposed in this dissertation. The off-line background sprite generation is taken as an example. Although MPEG-4 standard includes the static sprite coding tool, it still remains an open issue on sprite generation. The proposed algorithm aims at, on one hand, improving the sprite quality, and on the other hand, speeding up the sprite generation. Considering that video object segmentation is usually not accurate enough, we propose to use the reliability-based blending in sprite generation. This technique can be taken as the non-linear combination of the contribution pixels with different reliabilities achieved from the segmentation masks. Moreover, we present a rough image segmentation tool associated with background sprite generation in case that the auxiliary segmentation masks are unavailable. To speed up the sprite generation, a fast global motion estimation (GME) algorithm is developed, wherein the spatial and temporal feature point selection modules are integrated into the hierarchical implementation. The computing complexity of the proposed GME algorithm is controllable, so that we can achieve the best trade-off between the quality and speed of sprite generation.

Secondly, an efficient static sprite-based video coding algorithm is developed. Compared to the traditional algorithms, the major difference lies in how to efficiently compress the static sprite image. The proposed algorithm is based on arbitrary-shape

spatial prediction and DCT-like integer transform. Since the existing spatial prediction techniques cannot process the arbitrarily shaped video object, we propose a new padding technique based on rate-distortion optimization theory. The proposed technique is simple but very efficient, because it makes the texture of the padded contour block very similar to its neighbor blocks. To further improve the coding efficiency, we propose a new transform coefficient scanning strategy. Several scan tables are designed according to the distraction of transform coefficients. Compared to the static sprite coding in MPEG-4 video verification model, the proposed algorithm can improve the coding efficiency up to 3.0dB in PSNR at low bit rates.

Thirdly, an efficient dynamic sprite-based video coding algorithm is presented. The proposed algorithm employs some new coding techniques developed in the new H.264/AVC standard, based on which the advantages of sprite technique are fully utilized. The previous sprite generation algorithm is further developed to on-line generate the sprite image for the purpose of providing an optional reference (i.e. dynamic sprite) in predictive coding of the video. Accordingly, a new list of prediction modes is defined in terms of the dynamic sprite. Different from the traditional dynamic sprite coding algorithms, the proposed algorithm is equivalent to do two-stage motion compensated prediction. In other words, after the global motion is compensated with sprite warping, the local motion compensation is also utilized to compensate the slight distortion of global motion estimation. The rate-distortion optimization strategy is adopted so that the extra bits for coding motion vectors in terms of sprite prediction can be optimized.

Finally, to further extend the application of sprite technique, we present a video object segmentation algorithm by exploiting both the temporal information from the background and the spatial information from the foreground. The sprite image is generated with temporally recursive background blending, which includes all visible parts of the background throughout the sequence. Then the initial foreground regions are detected in each frame by subtracting the background derived from the sprite image. Meanwhile the raw frame is segmented into many homogenous texture regions containing the accurate edges using the watershed algorithm. Further merging and refining eventually determines the boundaries of foreground. Since background sprite image is formed with multi-parameter motion model across the sequence, the proposed scheme is more robust rather than the method simply utilizing differential information among adjacent frames.