Live Video Partitioning And Video Indexing Using Spatio-Temporal Coordination

Iswarya V[1], Sudhir shenai[2]
[1, 2] Dept of Computer science and Engineering
[1, 2] E.G.S Pillay Engineering College, Nagapattinam, India

Abstract- *To sustain an ongoing rapid growth of live video information, there is an emerging demand for a sophisticated content-based video indexing system. The video segmentation is an imperative technique used for the improvement of video quality on the basis of segmentation. The function of video segmentation is to segment the moving objects in video sequences. Video segmentation is the first step towards automatic annotation of digital video sequences. Its goal is to divide the video stream into a set of meaningful and manageable segments (shots) that are used as basic elements for indexing. Each shot is then represented by selecting key frames and indexed by extracting spatial and temporal features.However, the model is also semi-schema based as it allows additional declared elements in the instantiated objects as compared to its schema definition. Moreover, not all elements in an object need to be instantiated at one time as video content extraction often requires several passes due to the complexity and lengthy processing. The main idea behind domain-specific approach is the use of domain knowledge to guide the integration of features from multi-modal tracks in single video segments. The main purposes are to reduce manual annotations and to summarize the lengthy contents into a compact, meaningful and more enjoyable presentation. Indexing scheme is essential to determine the methods by which users can access a video database. The query language that can generate dynamic video summaries for smart browsing and support user-oriented retrievals in single video. The goal of this work is to automatically partition a video into topical segments which are then presented to the user in a customized video player. The approach taken in this work is to identify topics based on text similarities across the video. Results are presented from surveys showing a high level of satisfaction among student users of automatically segmented videos.*

Keywords- Domain specific approach, Video segmentation, Multimodal tracks, manual annotations

I. INTRODUCTION

The video coding standard provides many new features for future multimedia applications and enables interactivity with objects in video sequences via content based representation .

With the introduction of video object, automatic object segmentation becomes a prerequisite for video encoding.

However, this is a very difficult task because physical objects normally are not homogeneous with respect to some features such as color, luminance, or optical flow. .In a coding algorithm that combines a spatial analysis of the sequence with motion compensation has been proposed. On one hand, the spatial analysis is used to obtain a general scheme to deal with any kind of sequences and scene changes. On the other hand, motion information is used to compensate the spatial information. Motion information is generally the most important cue in motion segmentation since physical object is often characterized by coherent motion that is different from the motion of the background. Most video object segmentation techniques separate the video object from the background based on optical flow fields or change detection masks. In different methods are proposed to extract the objects from the background by using motion information. However, motion estimation is often not very accurate due to occlusion and aperture problems, and the sensitivity to noise. In a morphological motion filter algorithm is proposed to resolve edge pixels localization along the motion boundary. However, this algorithm partitions the image into different regions only according to the gray-level. Thus, it is sensitive to noise and the computational load is very high. In , a spatial-temporal segmentation algorithm is proposed, but there are many thresholds that need to be set By using this criterion, some object regions are merged and several video object regions are obtained. In the third step, these object regions are projected into the next frame and then tracked in subsequent frames. Some experimental results on the segmentation of the test sequences .

A. Segmentation and Tracking Algorithm

In the proposed algorithm, four efforts have been made to get the final segmentation results: (1) Intra-frame segmentation is performed using hierarchical morphological segmentation algorithm. In this stage, the image is partitioned into many flat zones with accurate contour information. (2) Based on the motion criterion, a binary image is derived from the partitioned regions and several object regions are obtained and their affine model parameters are also re-estimated. (3) In the tracking stage, these video objects are projected into the next frame according to their affine model parameters that have been obtained in stage 2. Uncovered and overlapped regions are regarded as uncertain regions. (4) To deal with the uncertain region, a two-scan Zig-Zag algorithm is used to absorb the pixels in the uncertain region according to the displaced frame difference.

B. Intra-frame segmentation stage

A hierarchical morphological segmentation that preserves the contour of the objects in the image has been described . This method is composed of three steps: image simplification using morphological operators, markers'extraction, and region growing using watershed algorithm with a hierarchical queue. The simplification goal is to make the signal easier to segment. By using morphological connectivity operator, the image can be segmented into constant gray-level regions with sharp contours corresponding precisely to those of the original signal. In the next step, the interior of each homogeneous region is "marked" by a label, that is, a constant gray level value, which is unique for this region. Finally, these labeled markers are input to watershed algorithm with a hierarchical queue and a region growing procedure is performed. After this intra-frame segmentation, the image is partitioned into many regions with different labels. Thus, the image is composed of a number of regions each with a unique label value.

II. LITERATURE REVIEW

Development and application explosion of multimedia technologies bring forward active research on the subject of Content Based Video Retrieval (CBVR) with digital video libraries and archives of immense size. A video can be analyzed at different levels of granularity . For object tracking, the low-level is an individual frame that is usually used for extracting static visual features like color, texture, shape, or interest points. So, we may talk about double-ply segmentation i.e. video temporal segmentation on the base of spatial frame segmentation matching. Morphological filters are nonlinear filters suitable for filtering task. Video segmentation on the basis of spatial segmentation analysis of separate shots or so-called "double" segmentation (spatio-temporal) is considered. Initial video data is analyzed with its granulation of separate segments, characterizing fragments with single meaning. Consequently, this approach permits conducting video data segmentation in near to real-time mode that accordingly allows implementation for video data analysis. . The aim of experimental researches is studying of video segmentation, specific for determination of segments with group correlation between objects.

Robust, accurate visual object tracking is fundamental to a variety of computer vision applications including robotics, human tracking and biometric identification, intelligent transportation systems, smart rooms, and military targeting systems. In addition, the noise and background clutter that are almost always present in real-world video sequences make the visual tracking problem particularly challenging. After partitioning, the samples are propagated through a parallel filter bank and the estimated states are generated as a weighted sum of the filtered samples. In order to take maximum advantage of the state estimates derived from the particle filter, the final estimates from the combined approach are recursively fed back. These results, which are typical of those we have obtained, demonstrate the performance advantage that can be gained approach by explicitly considering the most recent observation when constructing the proposal distribution.

III. PROPOSED SYSTEM

A novel technique for the segmentation and tracking of moving objects is proposed. The proposed scheme extracts moving objects based on both motion and spatial information. The goal of the research presented in this paper is to automatically partition a lecture video into topical segments, which can then be presented to the users as visual index points. However, automatically segmenting a video lecture by topic or subtopic is a very challenging problem as the precise meaning of a topic is subjective. The main reasons for errors in automatic segmentation are also presented based on a manual analysis of some of the videos employed for evaluation. Indexing is the task of dividing a video into segments that contain different topics. A video is composed of a sequence of thousands of images (or frames). In order to process video data efficiently, a video segmentation technique should detect scene changes and find the unique images. Such spatio-temporal segmentation is discussed and experimentally investigated. Partitioning a video source into meaningful segments is an important step for video indexing. We present a comprehensive study of a partitioning system that detects segment boundaries. The system is based on a set of difference metrics and it measures the content changes between video frames.

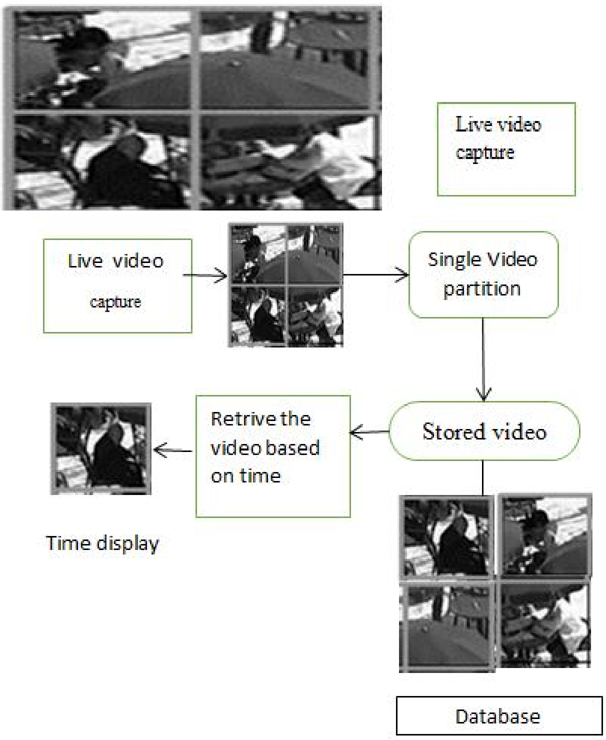

Fig 1:Architecture

IV. METHODOLOGY

Since video contains spatial, temporal, semantically rich and huge numbers of correlated frames, managing and processing them for further stages of CBVR that is feature extraction and indexing is the complex task and needs very large storage and computing time.

Algorithm: SBD Algorithm to detect shots or shot-segments of the video

Input: An live single video

Output: v1, v2,v3 and v4 //Number of video frame

Procedure:

1. For all frames f1, f2, ,f3,f4
2. Select f1, f2 ,f3,f4 // f1=first frame, f2= second frame, f3= third frame, f4= fourth frame
3. For all blocks of frames f1 ,f2,f3 and f4
4. Select b1, b2,b3,b4 // b1, b2 are non overlaping blocks of size 16x16
5. Apply wavelet transform to b

$$F_{ij}(x,y) = X(x,y) f_{ij}(x,y) X^{-1}(x,y) \quad 0 \le y \le N-1 \quad 0 \le x \le M-1$$

7. Store the distance in a vector

$$L2_{ij} = \sqrt{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left(| F_{ij}(x,y) - F_{ij-1}(x,y) | \right)^2}$$

8. Repeat the same process for all the blocks of two frames
9. Compute mean value of distance vector
10. Repeat this process for every consecutive frames.
12. Stop

V. CONCLUSION

Video abstraction is an integral part of single video applications, including video frame. Shot boundary detection is the process of identifying the significant content change in the video. A Square histogram based model is developed using frame segmentation and automatic threshold calculation. The keyframe is extracted by using a reference frame approach per shot. A total of around videos of different types are tested on this model and the model is able to detect all frame boundaries and is storing the suitable frames as keyframes to represent the video summary. An efficiency of almost 95% to 98% is observed using this algorithm.

REFERENCE

[1] Ali Amiri, Mahmood Fathy, Atusa Naseri , Computer Engineering Department, Iran University of Science and Technology, Tehran,Iran ― Keyframe extraction and video summarization using QRDecomposition‖,IEEE 2007.

[2] Anastasios D. Doulamis, Nikolaos D. Doulamis and Stefanos D.Kollias , National Technical University of Athens, Department of Electrical and Computer Engineering , Heroon Polytechniou, 157 73 Zografou, Greece ―Relevance feedback for content-based retrieval in video databases: a neural network approach‖.

[3] D.Besiris, F. Fotopoulou, N. Laskaris, G. Economou, Department of Physics, Electronics Laboratory, University of Patras ―Key frame extraction in video sequences: a vantage points approach‖, IEEE 2007.

[4] EMRAH AŞAN, M.S, Department of Electrical and Electronics Engineering, The Graduate School Of Natural And Applied Sciences of Middle East Technical University ―Video Shot Boundary Detection by Graph Theoretic Approaches‖, September 2008.

[5] Eung Kwan Kang, Sung Joo Kim, and Joon So0 Choi, Dept.of Physics Chung-Ang University, 221 Huksuk-Dong, video retrieval based on scene change detection in compressed streams", iit assessment.