IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014
ISSN: 2320 – 8791 (Impact Factor: 1.479)
www.ijreat.org

Advanced Descriptive One-Shot Video Puzzle Composition

M.Kavitha[1], B.Balamurugan[2]

[1]Assistant Professor, Department of Computer Science and Engineering, Sriram Engineering College, Chennai-35

[2]M.E. II Year, Department of Computer Science and Engineering, Sriram Engineering College, Chennai-35

Abstract

This paper is mainly based on video composition for reducing the size of a video by extracting the unwanted information with suitable algorithm. It will widely focus on easy uploading the data in a common network. This is very helpful to reduce the time of a data transferring to the server. Hence the user can easily access the data in a public network with less time period. Video can be processed by using SURF algorithm with assuming threshold value. This value can be calculated by mathematical mean calculation and thus the video can be converted into number of frames and those frames are separated for making a single shot video. Then the video can be further compared with already loaded database with suitable face recognition method. By using this method, the hidden face can be easily predicated. This paper can be widely applicable in military, research and banking applications.

Keywords *- SURF algorithm, threshold value, video authoring, image retrieval.*

1. INTRODUCTION

The popularity of personal digital devices, the amount of home video data is growing explosively. These digital videos have several Characteristics, compared with former videos recorded by non-digital camcorder, nowadays videos are usually captured more casually due to the less constraint of storage, and thus the number of clips is often quite large many videos may only contain a single shot and are very short and their contents are diverse yet related with few major subjects or events. Users often need to maintain their own video clip collections captured at different locations and time. These unedited and unorganized videos bring difficulties to their management and manipulation. For example, when users want to share their story with others over video sharing websites and social networks, such as YouTube.com and Facebook.com, they will need to put more efforts in finding, organizing and uploading the small video clips. This could be an extremely difficult “Puzzle” for users. Previous efforts towards efficient browsing such large amount of videos mainly focus on video summarization. These methods aim to capture the main idea of the video collection in a broad way which is not sufficiently applicable for video browsing and presentation. This paper, further investigate how to compose a content-consistent video from a video collection with an aesthetically attractive one-shot presentation. One-shot videos or long-shot video1 also known as long-

take video (this will exchangeable use them hereafter), means a single shot that is with relatively long duration.

Long shot has been widely used in the professional film industry, MTV video2 and many other specific video domains owing to its uniqueness in presenting comprehensive content in a continuous and consistent way. This paper introduces a scheme, "Video Puzzle", which can automatically generate a virtual one-shot presentation. This proposed seamless video composition technique is inspired from these works and this also integrate the object-level matching into the video composition procedure. Although this work and, can provide aesthetically pleasing form videos among the user's video collection, the targets and methodologies used are totally different. This system aims to automatically discover content-consistent video shots and compose into a virtual long-take video with spatial and temporal consistency while aims to provide a tree structure collection with temporal smoothing for ease of video browsing.

2. SYSTEM ORGANISATION

2.1. System Architecture

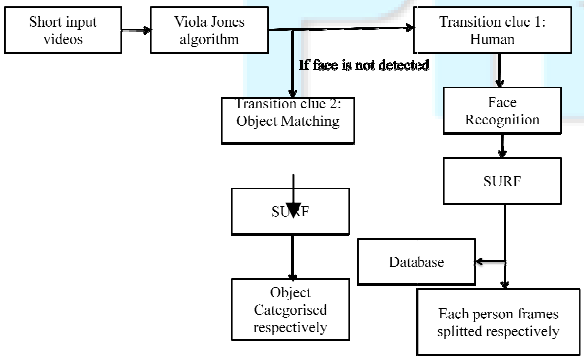

Fig: System Architecture Diagram.

This paper consists of three modules.

- Pre-processing
- categorization – Transition clues
- Video composition

The major technique which is used in the system is given by,

Pre – Processing

First This Input short videos are converted into frames. Since the videos are consists number of frames. Then this eliminate some frames like information less frames i.e., unnecessary to retrieve it. That should be calculated by using the mean value of an input (Mean of Input frame<15). After this resize the each frame. Then all frames are merging into a single video for video categorization. That should be very helpful to implementing different category of objects and videos. The overall scheme "Video Puzzle" aims to discover content- consistent video shots and composes them into a virtual long-take video. To this end, this proposes a novel graph-based visualization and path finding approach. The graph is constructed based on geometry matching (homograph mapping) and object matching (human, face). Based on the multi-cue content matching, the transition of video shots becomes meaningful and seamless. The video correspondence pair one by one.

This proposes a space-temporal morphing-based transition through matched local patterns, i.e., matched local common pattern, matched human or face. The produced transition is more natural than the traditional transitions such as fade-in, fadeout, wipes, and dissolve. Since both image-level and sequence level matching for video pairs are available, this can accomplish a content-based continuous transition. The proposed content- based transition produces

virtually consistent link for the final composition. The Video Puzzle system, which can automatically generate a virtual one-shot presentation from multiple video clips, provides a novel presentation of video contents and enables users to have a deeper impression of the story from the video collection.

Categorization – Transition clues

Videos are categorized by using transition clues like human, object. From that the detected value can be processed for further computation. Thus the commutated value further classified to holding the different characteristic value can be treated in respective process. Then we are taking human clue for first categorization by using Viola-Jones algorithm, if faces are not detected in frames that frames are separated into another process for object matching clue. Implement a coarse-to-fine partial matching scheme to generate a matching graph of the video collection. The matching scheme serves as a three-level matching, i.e., video pair selection, sequence-sequence correspondence finding, and frame-level exact matching. The video pair selection acts as an evidence for ensuring the non-redundant and complete quality of the generated one-shot video. It uses a hashing-based method to quickly obtain the video similarity measurement. We then find sequence correspondence of the selected video pairs through local key points matching. The final frame-level matching aims to find different matched objects to provide variant and rich clues for video transition generation. We implement three object matching methods in this part, i.e., salient object matching using local visual pattern discovery, and human and face appearance matching based on automatic human and face localization.

Video Composition

Object & sequence matching process are done by using SIFT algorithm (Scale-invariant feature transform). Related Object frames and related sequence frames are categorized into separate folder respectively. The ultimate video clip is expected to contain only one shot; we compose two videos only in their starting or ending part in order to keep the storyline within the clips. Thus the video sequence correspondence is also found within the starting and ending part.

For detail, the video is first partitioned into two parts with equivalent duration. The sequence correspondence represents a sequence in the second part of the video is matched with a sequence in the first part of the video. Then, the sequence similarity is scaled by a preference factor which is set to 1 when the two sequences are close to the start or end of the videos and gradually turns to 0 when one of the sequences is far from the video border. For each video pair and, we obtain the sequence correspondence by finding the sequence pair with the largest sequence similarity. The video similarity in the graph is also replaced by this sequence similarity. This process of searching for sequence correspondence is critical in our system. It determines whether video clips can be composed with the other videos. We will also discuss how to use the video matching score. Finally categorized frames are converted into Separate videos.

SURF (Speed Up Robust Features) algorithm, is base on multi-scale space theory and the feature detector is based on Hessian matrix. Since Hessian matrix has good performance and accuracy. In image I, x = (x,

y) is the given point, the Hessian matrix H(x, σ) in x at scale σ, it can be define as

$$\mathbf{H(x, \sigma)} = \begin{bmatrix} Lxx(x,\sigma) & Lxy(x,\sigma \\ Lyx(x,\sigma) & Lyy(x,\sigma) \end{bmatrix}$$

Where Lxx (**x**, σ) is the convolution result of the second order derivative of Gaussian filter

$\partial^2/\partial x^2 = (\sigma)$ with the image I in point **x**, and similarly for Lxy (**x**, σ) and Lyy (**x**, σ).

The transition generated by the proposed method looks much more smooth and seamless than the fade-in and fade-out method. The main objects are still clear enough to identify when using our method in the transitions 1 to 4. The fade-in and fade-out method brings a lot of blurring and ghost. The main advantage of proposed method is that we first automatically focus to the main object within the frames in order to avoid the displacement. The morphing between the matched objects is visually pleasing. The feathering method can reduce the ghosting effect. The result of transition 1 is without feathering effect. We can still get some blurring and ghost effect, especially on the border area where there exists most disagreement among the two frames.

SURF creates a "stack" without 2:1 down sampling for higher levels in the pyramid resulting in images of the same resolution. Due to the use of integral images, SURF filters the stack using a box filter approximation of second-order Gaussian partial derivatives, Since integral images allow the computation of rectangular box filters in near constant time. In descriptors, SIFT is good performance compare to other descriptors. The proposed SURF descriptor is based on similar properties. The first step consists of fixing a reproducible orientation based on information from a circular region around the interest point. And second construct a square region aligned to the selected orientation, and extract the SURF descriptor from it. In order to be invariant to rotation, it calculate the Haar-wavelet responses in x and y direction.

Face Recognition

The matching graph brings a new tool for video album browsing. Unlike the traditional video summarization methods which try to extract key content of video album and present in an concise manner, the matching graph has several characteristics:

(1) It presents the video album in a wide-range manner and gives the user the direct impression over the large numbers of clips.

(2) More importantly, the matching graph gives the Correspondence information of two videos. The linking edge introduces the browser to explore the whole set with continuous content rather than one by one.

(3) It also highlights the "key" clip in the video. The "key" clip has the most number of edges connecting from and to it.

(4) It provides an interactive interface for user. When user clicks certain video clips, the system can display the linked compact set of video clips that contains this video clip. If the user prefers certain person on the graph edge, the system can generate a consistent one-shot video presentation focusing on the person. The interaction is flexible and has many potential variation. It can be shown that the video album contains several compact sets, which can

generate one-shot videos, and some isolated nodes. Three examples of the generated one-shot videos are demonstrated in which respectively shows the human appearance, face appearance and object key point induced matching clues. It is worth noting that these One-shot Videos are obtained by automatic path finding the video, timestamps are available, and thus we first prune the graph edges with the time stamp information. Multiple one-shot videos are then generated by the method introduced. It is worth noting that the One-shot Video is obtained by fixing the matched object. And One-shot Video is obtained by automatically.

3. CONCLUSION

The long video can be converted into single shot video by neglecting the non informative frames. Before the conversion of single shot video it can be converted into number of frames because the video consist of number of frames and then the frames are resized to reduce the memory of the device. Then the separated frames are merged to form a new video. The merged video can be categorized by using viola – Jones algorithm in the way of objects and humans and then it can be compared to the system database to detect the face of the unauthorized person i.e., face recognition.

REFERENCES

[1] Video Puzzle: "Descriptive One-Shot video Composition", Qiang Chen, Meng Wang.

[2] G. A hanger, "Automatic composition techniques for video production ,"*IEEE Trans. Knowl. Data Eng., vol. 10, no. 6, pp. 967–987,*Nov. 1998.

[3] C. Barnes, D. Goldman, E. Shechtman, and A. Finkelstein, "Video tapestries with continuous temporal zoom," in *Proc. SIGGRAPH, 2010.*

[4] E. Bennett, "Computational time-lapse video," *ACM Trans. Graph.,*vol. 26, no. 102, Jul. 2007.

[5] K. S. Bhat, S.M. Seitz, J. K. Hodgins, P. K. Khosla, K. S. Bhat, S.M.Seitz, J. K. Hodgins, and P. K. Khosla, "Flow-based video synthesisand editing," *ACM Trans. Graph., vol. 23, no. 3, pp. 360–363, Aug.*2004.

[6] The One-Shot Videos[Online]. Available: http://www.youtube.com/the one shot video.

[7] K. Heath, N. Gelfand, M. Ovsjanikov,M. Aanjaneya, and L. J. Guibas, "Image thisbs: Computing and exploiting connectivity in image collections," in *Proc. CVPR*, 2010.

[8] X. S. Hua, L. Lu, and H. J. Zhang, "Optimization-based automated home video editing system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 5, pp. 572–583, May 2004.

[9] C. Huang, H. Ai, Y. Li, and S. Lao, "High-performance rotation invariant

[10] multiview face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 671–686, Apr. 2007. I. Kemelmacher-Shlizerman, E. Shechtman, R.Garg, and S.

[11] Seitz, "Exploring photobios," in *Proc. SIGGRAPH*, 2011.