ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

Dynamic Composition Techniques for Video Production

M.R. Preethi [#1], S. Romy [#2], B. Ruth Angel [#3], M. Maheswari *[4]

UG Student, Dept. of CSE., Anand Institute of Higher Technology, Chennai, Tamilnadu, India[#1, #2, #3]

Assistant Professor, Dept. of CSE., Anand Institute of Higher Technology, Chennai, Tamilnadu, India[*4]

ABSTRACT: Image processing is a method to convert an image into digital form and perform some operations in order to get an enhanced image and also to extract some useful information from it. The amount of home video data is growing explosively because of the popularity of personal digital devices. Non-digital camcorder is used to record the digital videos which can be compared with former videos. Nowadays videos are usually captured 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, Users often need to maintain their own video clip collections captured at different time and locations. These unedited and unknown videos bring difficulties to their management and manipulation. The solution for presenting and managing the video clips is to create a large amount of short, single-shot videos by personal camcorder, such as the small video clips in family albums. From the short video clips, the proposed system presents a novel video composition system "Dynamic Video Production" which generates aesthetically enhanced long-shot videos. The main task of the proposed system is to automatically composite several related single shots into a virtual long-take video with spatial and temporal consistency. The entire long-take video thus comprises several single shots with consistent contents and fluent transitions. With the generated matching graph of videos, the proposed system can provide an efficient video for browsing. Various experiments are conducted on multiple video albums and the results demonstrate the effectiveness and the usefulness of the proposed scheme.

KEYWORDS: One-shot video, Viola-Jones, Back Propagation Network (BPN), Scale-invariant feature transform (SIFT), Speeded up Robust features (SURF).

I. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations in order to get an enhanced image and also to extract some useful information from it. Image processing system includes treating images as two dimensional signals while applying already set signal processing methods to them. The two types of **methods used for Image Processing** are **Analog and Digital** Image Processing.

The amount of home video data is growing explosively because of the popularity of personal digital devices. Non-digital camcorder is used to record the digital videos which can be compared with former videos. Nowadays videos are usually captured 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, Users often need to maintain their own video clip collections captured at different time and locations. These unedited and unknown videos bring difficulties to their management and manipulation.

The main aim of the proposed system is to automatically compose descriptive long-take video with content-consistent shots retrieved from a video pool that are converted into single shot video. When users want to share their story with others over video sharing websites, such as YouTube.com and Facebook.com, they may require 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. We propose a novel framework to compose descriptive long-take video with content-consistent shots retrieved from a video pool. Given a messy collection of video clips, Video Puzzle can select a clip subset with consistent major topic (similar with finding the clues and solving the Puzzle Games among the images [8]). The topics referred here are a person, scene or

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

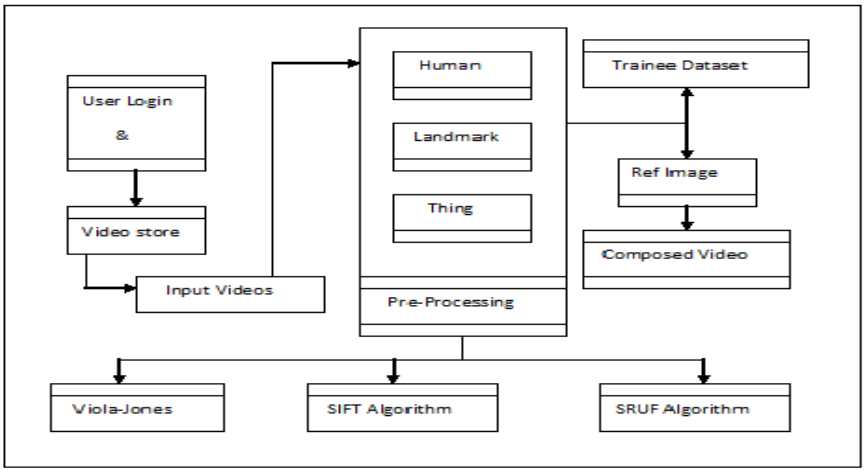
object. It is specified by users or found by an automatic discovery method. For each video, frame-by-frame search is performed over the entire pool to find start-end content correspondences through a coarse-to-fine partial matching. The content which is correspondence is general here and can refer to the matched objects or regions, such as human face and body. Content consistency of these correspondences enables us to design several shot transition schemes to seamlessly stitch one shot to another in a spatial and temporal manner. The entire long-take video comprises several single shots with consistent contents and transitions which is fluent. Meanwhile, with the generated matching of videos, the proposed system can also provide an efficient video browsing mode. Various experiments are conducted on multiple video albums and the results demonstrate the effectiveness and the usefulness of the proposed scheme.

II. RELATED WORK

The existing system provides a set of techniques for the production of music videos using computer graphics and animation. This application is illustrated using the examples of dance to the music/play to the motion. To produce the video Dance to The Music / Play to the Motion, they employed a plethora of techniques that are commonly used in digital special effects and computer animated films [11]. The existing system explores the use of tracked 2D object motion to enable novel approaches to interact with video. Features in the video are automatically tracked and grouped in an off-line pre-process that enables later interactive manipulation. It proposes a novel interfaces for three tasks that are used by present-day video interfaces: annotation, navigation, image composition [12]. It is a time based medium it has dual tracts of current video and current tool force users to work at the smallest level of detail. This technique conducts a preliminary user study to investigate the effectiveness of the system. The usual techniques to deal with complexity problems are abstraction, on-the-fly, compositional and equivalence techniques [13]. Service oriented computing and component based software engineering are benefited by behavioral protocol for the purpose of discovery, composition, composition correctness checking and adaptation [14]. Dynamic, real-time and cohesive video composition and customization are achieved by this technique. We also identify metrics for evaluation with respect to existing manually-produced video-based news. The evaluation shows the quality of automatic composition is better than, video composition. At last the result validate the assertions on which the automatic composition techniques [15].

III. PROPOSED ALGORITHM

A. *Design Considerations:*



- Viola-Jones
- Back Propagation Network (BPN)
- Scale-invariant feature transform (SIFT)
- Speeded up Robust features (SURF).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

B. *Description of the Proposed Algorithm:*

User Authentication & Video Storage:

Once the user needs to access the Video Puzzle Application, they have to register their information. Video Storage helps to secure videos kept by the user. So, proper administration control will be there to maintain a recognized users records and its personal information to keep is as privacy one.

Pre-processing

Viola- Jones

The humans and non-humans frames are separated by using Viola-Jones algorithm..Using this algorithm the frames will be categorized by human and non-human. The human frames are identified by ROI (Region of Interest) or else its comes under non-human frames.

$$h(\mathbf{x}) = \operatorname{sign}\left(\sum_{j=1}^{M} \alpha_j h_j(\mathbf{x})\right)$$

The following are the procedure for Viola Jones algorithm:

1) The first step of the Viola-Jones face detection algorithm is to turn the input image into an integral image. This is done by making each pixel equal to the entire sum of all pixels above and to the left of the concerned pixel.

2) The sum of all pixels inside any given rectangle using only four values. These values are the pixels in the integral image that coincide with the corners of the rectangle in the input image.
Sum of grey rectangle = D - (B + C) + A ----- (3)

3) The Viola-Jones face detector analyzes a given sub-window using features consisting of two or more rectangles.

Back Propagation Network

Videos are categorized by using transition clues like human, object. The human frame has to compose with and without reference image by using BPN algorithm and with the help of Trainee Database. After Preprocessing, Trainee database has to be created. It contains of list of human frames in a specific group with different angles. Each human frame is stored in a uniform manner. Initially each human frame has to be check one by one with Trainee Database Frames. If the first human Frame is match with first groups in Trainee Database, then a separate folder will create and store the frame on that. Repeating the process until all the human frames has to be completed in human categorized. Frames in the each folder will be composed when the reference image is empty and formed videos for all individual humans. If the reference image is not empty, frames will be compose depends upon the reference image for the specific human.

Scale-invariant feature transform (SIFT)

The non-human frames has to compose with and without reference image by using SIFT algorithm . Initially first frame has to be extracted from non-human frames and compare with the remaining frames consist in the non human frames. Comparing process has been done by object matching and their characteristics. Once the process has been completed collect all frames and store in a separate folder and composed in a one shot video. The second frame has been extracted and repeats the similar process for the previous frame. By repeating this process for all frames present in the non-human frames. If the reference image is empty, on that condition frames will be composed and formed videos for individual all non-humans. If the reference image is not empty, frames will be compose depends

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

upon the reference image for the specific non-human. Related object frames and related sequence frames are categorized into separate folder respectively. Finally categorized frames are converted into separate videos. The following are the procedure for the SIFT Algorithm:

3. Scale space extreme detection.
4. Key point localization
5. Orientation Assignment.
6. Key point descriptor.

Speeded up Robust features (SURF)

Object & sequence matching process are replaced by using SURF algorithm (Speeded up Robust Features) for improving matching accuracy and reducing time-consuming. Finally our proposed framework and three components are highly accurate under various conditions.

The SURF algorithm is same as the principles of SIFT algorithm, but it uses a different scheme and should provide better results. It works much faster. SURF uses a BLOB detector based on the Hessian to find points of interest. The Hessian matrix expresses the extent of the response and is an expression of a local change around the area. Given a point x = (x, y) in an image I, the Hessian matrix H (x, σ) in x at scale σ , is defined as follows:

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

----- (4)

where $Lxx(x,\sigma)$ is the convolution of second order derivative $\partial x/\partial x^2 g(\sigma)$ with the image in the point x, y similarly with $Lxy(x,\sigma)$ and $Lyy(x,\sigma)$.

VI. Conclusion and Future Work

In this system, we proposed “Dynamic Video”, which is used to generate a large amount of integrated system of video presentation, summarization and browsing, based on large amount of personal and web video clips. This system automatically collects content-consistent video clips and generates an one-shot presentation using them. It can facilitate family album management and website video categorization. Here we demonstrated the applications using “Video Puzzle” and the results show that it has great potential to be used in future video management systems. In the future, we aim to speed up the feature extraction step to further reduce the time cost and make it a practical system, and visualize the overall video graph with a hierarchical graph structure so that the browsing of the graph can be more efficient.

References

[1] Qiang Chen, Meng Wang, Zhongyang Huang, Yang Hua, Zheng Song, and Shuicheng Yan , “VideoPuzzle: Descriptive One-Shot Video Composition,” IEEE Transactions on multimedia, vol. 15, no. 3, april 2013.
[2] G. Ahanger, “Automatic composition techniques for video production,”*IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, pp. 967–987, Nov. 1998.
[3] G. Zhang, Z. Dong, J. Jia, and L.Wan, “Refilming with depth-inferredvideos,” *IEEE Trans. Visual. Comput.Graph.*, vol. 15, no. 5, pp.828–840, 2009.
[4] J. Lee and J. Oh, “Scenario based dynamic video abstractions using graph matching,” in *Proc. ACM MM*, 2005.
[5]J. Vendrig and M.Worring, “Systematic evaluation of logical story unit segmentation,” *IEEE Trans. Multimedia*, vol. 4, no. 4, pp. 492–499,Dec. 2002.
[6] J. Calic, D. Gibson, and N. Campbell, “Efficient layout of comic-like video summaries,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 7, pp. 931–936, Jul. 2007.
[7] C. Correa, “Dynamic video narratives,” *ACM Trans. Graph.*, vol. 29, no. 4, Jul. 2010.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

[8] K. Heath, N. Gelfand, M. Ovsjanikov,M. Aanjaneya, and L. J. Guibas, "Image webs: Computing and exploiting connectivity in image collections," in *Proc. CVPR*, 2010.
[9] 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.
[10] E. Bennett, "Computational time-lapse video," *ACM Trans. Graph.*, vol. 26, no. 102, Jul. 2007.
[11] Adriana Schulz, Marcelo Cicconet, Bruno Madeira, Aldo Zang, Luiz Velho, "Techniques for CG Music Video Production:," *IEEE Trans. VISGRAF Laboratory*, March 2010.
[12] Dan B Goldman, Chris Gonterman, Brian Curless, David Salesin, Steven M. Seitz, "Video Object Annotation, Navigation, and Composition," *IEEE Trans*. University of Washington, Jul 2006.
[13] Serge Haddad, Pascal Poizat, "Transactional Reduction of Component Compositions," *IEEE Trans.* Universit´e Paris Dauphine, France, March 2007.
[14] Juan Casares, Brad A. Myers, A. Chris Long, Rishi Bhatnagar, Scott M. Stevens, Laura Dabbish, and Albert Corbett, "Simplifying Video Editing with Intelligent Interaction," *IEEE Trans.* Human Computer Interaction Institute Carnegie Mellon University, May 2001.
[15] G. Ahanger and T.D.C. Little, "Automatic Composition Techniques for Video Production," *IEEE Trans*. Department of Electrical and Computer Engineering Boston University, 1998.