ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 3, March 2015

Movie Character Identification Framework Using Sift and Surf Matching Algorithm

Arumugam.K[1], Sangeetha.N[2]
P.G scholar, Dept. of M.E (Computer Science and Engineering), Alpha College of Engineering, Chennai, Tamilnadu, India.[1]
Assistant Professor, Dept. of M.E (Computer Science and Engineering), Alpha College of Engineering, Chennai, Tamilnadu, India.[2]

Abstract: - A large amount of short and single shot videos are created by video recorder from various environment every day such as short or large video clips. Solution to those videos for presenting and as well as managing is highly desired from perspective of artistry and professionalism long take shot video is also termed as one shot videos are able to present events and persons or scenario spots an informative manner. This paper presents a video composition system "movie character identification", which generates aesthetically to enhanced longshot video from short to large video clip. Automatically composite several related single shot into virtual long shot video with spatial, temporal consistency. Each videos searched frame by frame performed over entire pool to find start to end content correspondences through coarse-fine partial matching process. This content consistency of these correspondences enable to design several shot transition schemes to stitch one shot and multiple shot to another. The entire video comprises multiple single shots with consistent and fluent transitions. Those generated matching graph of videos in this system can also provide efficient video browsing mode. Multiple video albums and the results are effective and Useful in the proposed scheme.

Keywords: Surf, Sift, Movie Character, Object and Non Object Matching, face recognition.

I. INTRODUCTION

An image is an array, or a matrix of square pixels (picture elements) arranged in columns and rows.An **image processing** operation typically defines a new image *g* in terms of an existing image *f*.

$g(x,y) = t(f(x,y))$.We can transform either the range of *f*Or the domain of *f*:$g(x,y) = f(t_x(x,y), t_y(x,y))$.

1.1 PROPERTIES OF IMAGES:

Spatial Resolution: Width pixels/width cm and height pixels/ height cm

.

Intensity Resolution: Intensity bits/intensity range (per channel).

Number of Channels: RGB is 3 channels, grayscale is one channel.

Example Of Image: Digital image

An image can be expressed as rectangular array of integers and those integer represents the darkness or brightness of the image at that point.N: number of rows, M: number of columns, Q: gray levels N =, M =, Q = (q is the number of bits/pixel).

$$\begin{matrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ \dots & \dots & \dots & \dots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{matrix}$$

Solution based on desired features from a source video to the target video such as colorizing videos, reducing video blurs, and video rhythm adjustment studied automatic broadcast soccer video composition. There also exist studies on video texture which aims to provide a continuous and infinitely varying stream of images. To achieve these interactions, system first analyzes the video in a fully

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 3, March 2015

automatic preprocessing step that tracks the motion of image points across the video and segments those tracks into coherently moving groups.

II. Preliminaries And Related Work

Although reliably extracting objects in video and tracking them over many frames is a hard problem in computer vision, the manipulations support do not require perfect object segmentation and tracking, and can instead exploit low-level motion tracking and mid-level grouping information. Furthermore, aggregating point motion into coherent groups has a number of benefits that are critical for interaction: can select a moving region using a single click, estimate object motion more robustly, and to a limited extent.

A Risk of Missing Details. Possibly inaccurate summarization also may cause Inconvenience in browsing Solution targeting for only small display on mobile. Exact path finding approach Missed. The transition of video shots becomes meaningless. Either Human or Object Categorization. It can only deal with static background for foreground extraction.

The contribution of work can be summarized as follows: To propose a video puzzle scheme. It is able to extract video contents about a specific. An efficient method to find the content correspondences of multiple videos and then compose them into a clip with an optimized approach. To introduce two applications based on the video puzzle scheme, one about home video presentation and the other about landmark video generation. Specifically, the two specific applications introduced are:

Personal video presentation. Solution to those videos for presenting and as well as managing is highly desired from perspective of artistry and professionalism long take shot video is also termed as one shot videos are able to present events and persons or scenario spots an informative manner. This paper presents a video composition system "movie character identification", which generates aesthetically to enhanced longshot video from short to large video clip.

Automatically composite several related single shot into virtual long shot video with spatial, temporal consistency. Those generated matching graph of videos in this system can also provide efficient video browsing mode.

2.1PRE- PROCESSING:

Generating number of videos into single videos and as well as same sized resolution frames. All frame size should be same m x n. First Input short videos are converted into frames. . Then we eliminate some frames like information less frames .After we resize the each frames. Then all frames are merging into a single video for video categorization.

$H_1=H_2=H_3=\ldots\ldots\ldots\ldots\ldots=H_n$

$W_1=W_2=W_3=\ldots\ldots\ldots\ldots\ldots=W_n$

H – Resolution size of height

W – Resolution size of Width

N – Number of frames

If the frames are having less information must be eliminated by using mean formula. It will applicable for all sort of complex environment. 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.

2.2 CATEGORIZATION BASED ON TRANSITION CLUES:

Videos are categorized by using transition clues like human, object. Then 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. Videos are converted into same resolution size frames stored as resized frames database.

SURF algorithm are used to identify the human faces or not. If it has identified which is stored as human database otherwise stored as object. 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 video means a single shot that is with relatively long duration.

III. 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. Finally categorized frames are converted into Separate videos.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 3, March 2015

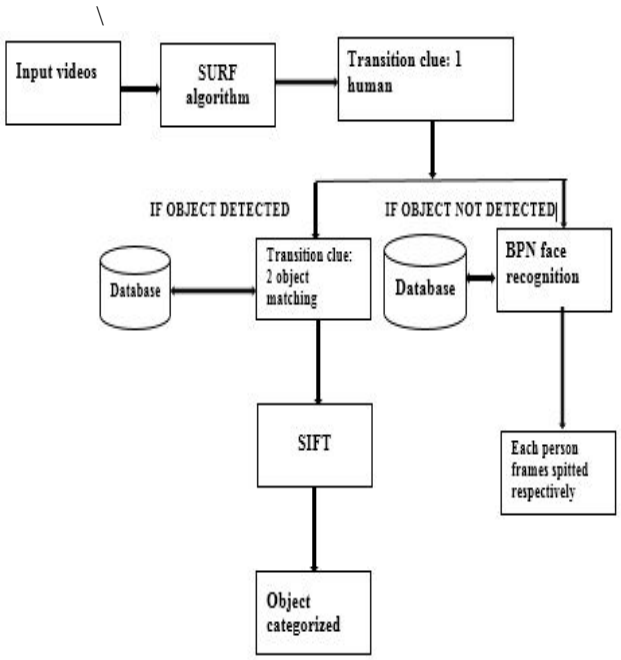

Fig: System Architecture

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, however, are not sufficiently applicable for video browsing and presentation.

The comparison with the video abstraction and presentation techniques, not only provide presentation approach but also further services such as editing. Objects and non-objects are identified by only predefined databases.

3.1 OBJECT& SEQUENCE MATCHING PROCESS:

Comparing given identified or selected frame of object or non-object with pre-defined databases. It is different from the traditional works. The transition clues are face any objects like chair, clock, and wall. It will look into their related databases for identification.

3.2 NON OBJECTS (HUMAN FACES)

Each person faces and objects are spitted respectively. Predefined databases can have n number of frames in different format and may be unique resolution. It does not affect current requesting process.

3.3OBJECTS

Depends on the database which has been stored n number objects and non-objects. But when databases are separated as human faces and objects are easily to retrieve which is about to be compared, it requires less time even the complex backgrounds.

IV.CONCLUSION:

Object & sequence matching process are replaced by using SURF algorithm (**Speeded up Robust Features**) for improving matching accuracy. It is able to extract video contents about a specific topic. The proposed system is flexible and as well as multiple components can be customized then applied to various applications. An efficient method to find the content correspondences of multiple videos and then compose them into a clip with an optimized

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 3, March 2015

approach. Two applications based on the video puzzle scheme, one about home video presentation and the other about landmark video generation. Specifically entire frames are computedthrough every pixel presence in the frames.

Videos are converted into same resolution size frames stored as resized frames database. SIFT Algorithm are used to identify the human faces or not. If it has identified which is stored as human database otherwise stored as object. Further investigate how to compose a content-consistent video from a video collection with an aesthetically attractive one-shot presentation. Converted frames are store in separate folder names as object and human faces.

REFERENCES

[1] G. Ahanger, "Automatic composition techniques for video production," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, pp. 967–987, Nov. 2006.

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

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

[4] 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 synthesis and editing," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 360–363, Aug. 2004.

[5] 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.

[6] Y. Caspi, A. Axelrod, Y. Matsushita, and A. Gamliel, "Dynamic stills and clip trailers," *Visual Comput.*, vol. 22, no. 9, pp. 642–652, Sep. 2006.

[7] P. Chiu, A. Girgensohn, and Q. Liu, "Stained-glass visualization for highly condensed video summaries," in *Proc. ICME*, 2004.

[8] T.Cootes, C. Taylor, and D. Cooper, "Active shape models-their training and application," *Comput. Vision Image Understand,* vol. 61, no. 1, pp. 38–59, Jan. 2007.

[9] C. Correa, "Dynamic video narratives," *ACM Trans. Graph.*, vol. 29, no. 4, Jul. 2010.

[10] J. E. Cutting, "Representing motion in a static image: Constraints and parallels in art, science, and popular culture," *Perception*, 2002.

[11] M. Everingham, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal visual object classes (VOC) challenge," *Int. J. Comput. Vision*, vol. 88, pp. 303–338, 2010.

[12] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal.Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[13] 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.

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

BIOGRAPHY

I did my UG in Sri BalajiChockalingam Engineering College at arni (tiruvannamalai) I am doing my P.G in Alpha College of Engineering at Chennai. Iwould like to know more and research about image processing domain. I have done some real time project as I have described about that above content.