International Journal of Scientific & Engineering Research, Volume 7, Issue 3, March-2016 10
ISSN 2229-5518

An Analysis of the Current Challenges in Efficient Video Processing

Deepti.C[1],Renuka Devi M.N[2] and Madhavi G[3]

[1]Asst.Professor, Department of ISE, PESIT Bangalore South Campus,deeptic@pes.edu
[2]Asst.Professor, Department of ISE, PESIT Bangalore South Campus,renukadm@pes.edu
[3]Shri Shiridi Sai Institute of Science and Engineering, Ananthapuramu

ABSTRACT

Video in digital format is common in professional and home appliances. Video can be thought of as a sequence of image frames which may be complemented by audio information. An individual video has distinctive motion features, colour histograms, motion histograms, text features, audio features and objects present in its frames. Video processing systems involve a stream processing architecture in which video frames are processed one or more at a time. This type of processing is crucial in systems that have live video content where loading the entire set into the workspace is wasteful. . Thispaper takes a look at four distinctive video processing tasks namely Video Quality Improvement, Video Segmentation, Video Search and Retrieval, Video Summarization all of which are significant issues in nature. However these tasks do not have 100 % efficient solutions .In our paper we present an overview of the challenges that need to be tackled in this regard.

Key Words: Video Processing, Segmentation, Retrieval, Summarisation

1. INTRODUCTION

In todays' world, many state-of-the-art video applications and new technologies in different fields such as biology, medicine, engineering, and entertainment have led to emergence of new digital areas of video applications. The videos are compressed, transmitted, captured, and stored in various digital forms with different types and amounts of impaired artifacts. Speckle noise and coding artifacts are most common causes of noise in the field of bio-medical applications, Entertainment and engineering applications are considered to be dominated by digital coding and transmission artifacts. Furthermore, each emerging application and technology has introduced different kinds of specific, correlated/structured distortions. The removal of such distortions, i.e., denoising, has not adequately been recognised and needs to be further discovered. This is especially true when stable growth of new emerging digital multimedia applications and services are observed. Most current aspect in video processing is video quality assessment and improvement techniques. Visual quality is critical in systems which are intended for currently challenging applications. Video segmentation is an important technique for improvement of video quality.

2. VIDEO QUALITY IMPROVEMENT

The depth of a camera's motion introduces motion blur. Precise assessment of motion blur parameters is non-trivial. Two kinds of motion blurring effects are prominently found in videos, one is the blurring effect caused by movement of the hand where the camera is positioned, and the other is the blurring effect caused by a fast moving object in the video.
The PSF (point spread function) of a video frame is estimated by comparing stable area in blurry frame and non-blurry frame after which the Richardson-Lucy algorithm[1] can restore the blurry frame by non-blinddeconvolution. An additional method to deblur the objects which move fast in the video is by reconstructing the background in each frame, so an accurate matte of blurry object can be extracted using alpha matting [2] as shown in Figure 1.

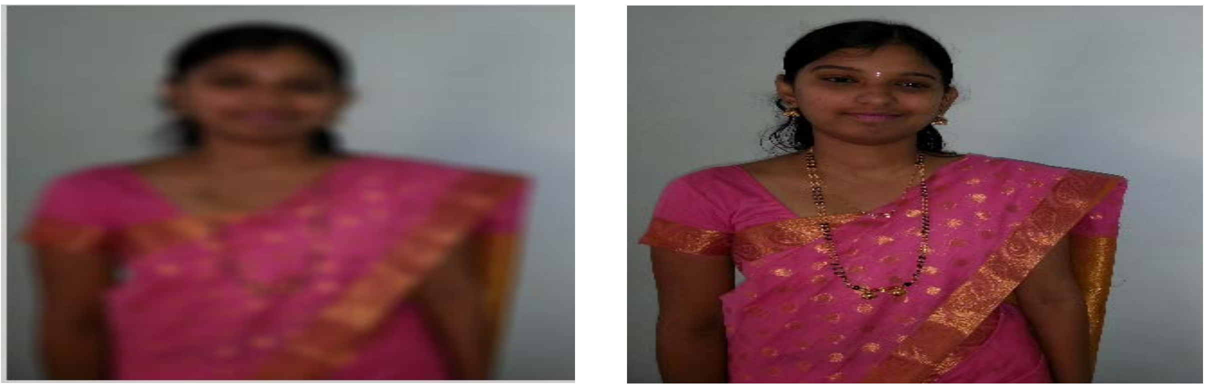

Figure 1 : Comparison of Blurred and Deblurred Video Frame

2.1 Video Inpainting

Video inpainting aims to remove objects or restore missing or tainted regions present in a video sequence by utilizing spatial and temporal information from neighbouring scenes.

Inpainting algorithms focus on much smaller regions that are categorized by linear structures such as lines and object contours.

The foremost objective is to create an inpainted area that is merged effortlessly into the video as seen in Figure 2.

This requires preservation of visual consistency and also no distortion in the selected area should be noticeable to the human eye when the video is played as a sequence.

Original Video Frame

Video Frame with user supplied restoration mask

Inpainted Video Frame

Restored Video frame

Figure 2 :Illustration of inpainting in videos[3]

2.2 Error Concealment

Error concealment is another technique used for video quality improvement that minimizes the deterioration of signals caused by packet loss when the video signal is conveyed via wireless networks.

The error concealment techniques for video transmission over noisy channels include

1) Spatial error concealment [4] is a method to compensate a lost macro block in intra-coded frames in which no valuable time-based information exists. This method can efficiently recover both smooth and edge areas based on selective directional interpolation.

2) Dynamic mode-weighted error concealment [5] is used for trading missing pixels in a missing macro block of inter-coded frames. This method adopts a decoder-based error tracking model and combines several concealment modes adaptively to reduce the mean square error of each pixel. The method is capable of concealing lost packets and reducing error propagation effect.

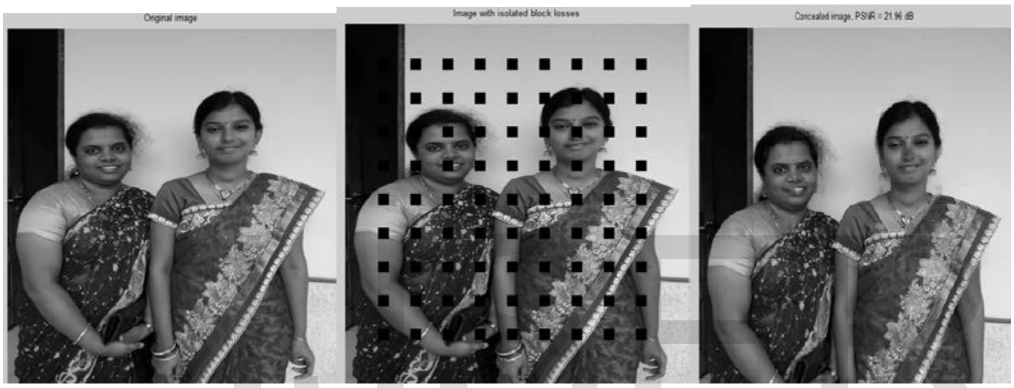

Figure 2: Example of Distorted and Concealed Video Frame Sequence

3. VIDEO SEGMENTATION

Video segmentation divides a given video in to its constituent objects. Segmentation of non-trivial objects is one of the most difficult tasks in video processing.
The challenge is that segmentation should stop when required objects have been isolated. Two prominent techniques with regard to video segmentation are discussed below.

3.1 Video Matting

Figure 3 denotes the problem of precise extraction of foreground objects in images and video. There is growing interest in using matting techniques to create novel transformations and enable other editing tasks such as photo enhancement, red-eye correction, foreground object retargeting and non-photorealistic translation.

Figure 3: Single Frame matting

3.2 Background Subtraction

There are many challenges in developing a good background subtraction algorithm. For an outdoor environment, a background subtraction algorithm should adjust to various levels of illumination at different times of the day and adverse weather conditions that can alter the background. It should avoid detecting non-stationary background objects such as swinging leaves and shadow cast by moving objects.

Presently background subtraction algorithms include simple techniques such as frame differencing [6], adaptive median filtering [7] and more sophisticated probabilistic modelling techniques. The performance of these algorithms can be analysed based on how they differ in pre-processing, background modelling, foreground detection, and data validation as illustrated in Figure 4.

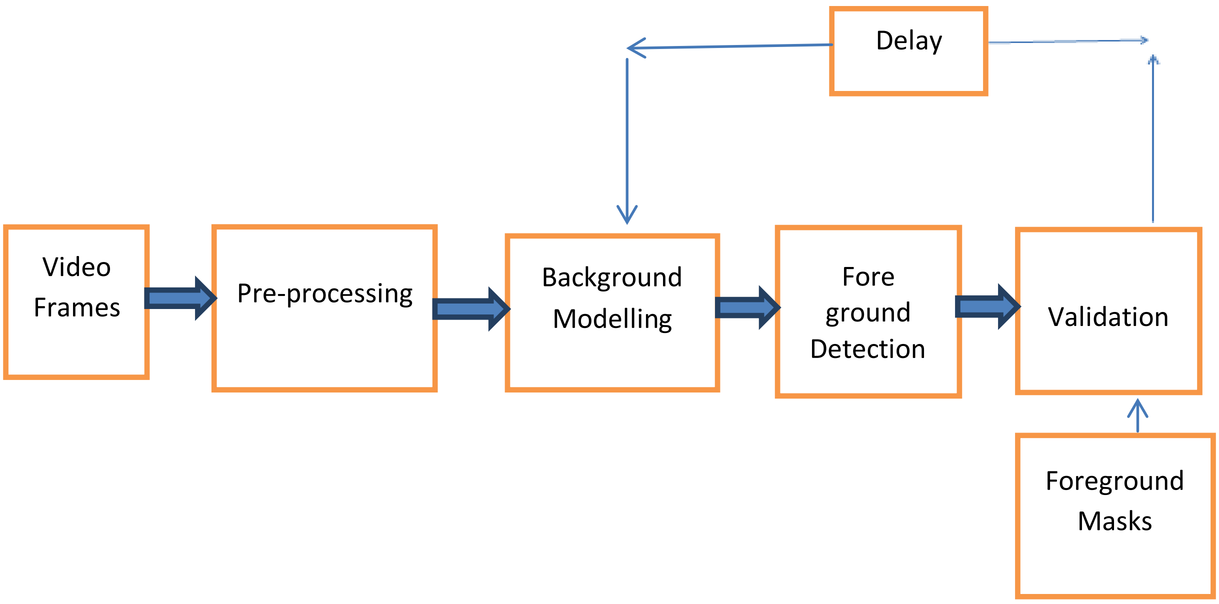

Figure 4: Background subtraction algorithm

Pre-processing consists of a collection of simple image processing tasks that change the raw input video into a format that can be processed subsequently. Background modelling uses a new video frame to calculate and update a background model which provides a statistical description of the entire background scene. Foreground detection then identifies pixels in the video frame that cannot be sufficiently clarified by the background model and outputs them as a binary candidate foreground mask. Finally, Data Validationinspects the candidate mask, rejects those pixels that do not match with actual moving objects and yields the final foreground mask. While complicated techniques often produce superior performance, simple techniques such as adaptive median filtering can also produce good results with much lower complexity.

4. VIDEO SEARCH AND RETRIEVAL

An application where the user wants to search his/her private database using video queries while keeping the query and database content secret from the server can be considered as video search. Video search and similarity comparison also depends on comparing visual features which are extracted from key frames.
Visual feature extraction [8] is performed on the video as part of pre-processing on the user-side. These visual features are properly encrypted and stored on the server as back up. During search, visual features from the query image or video are mined and encrypted before being sent to the server. The server compares the encrypted query features to encrypted features in the database and returns videos with the highest similarity as shown in Figure 5. Similarity comparison can be performed over different sub sequences of two videos and accumulated to get an overall similarity score.

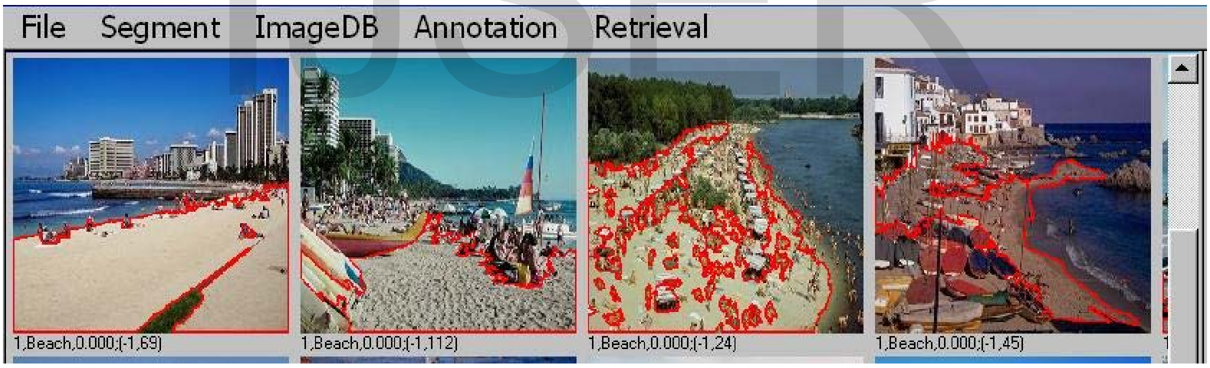

Figure 5: Representation of Video Search and Retrieval [9]

5. VIDEO SUMMARIZATION

Video summarization is the task of extracting a set of key frames to characterise the video content. *Key frames* are single images or video frames spontaneously extracted from video sequences. There are diverse ways to extract key frames from videos. Sampling based method obtains key frames by random or uniform sampling from the video. Sampling based extraction [10] is simple but may not capture dynamic content of the video. Since private videos are stored in encrypted form on the server, browsing video content requires real-time decryption.
Real-time browsing of encrypted video can be made possible by restricting the decryption to only key frames. A limited number of motion vectors may not be sufficient to reconstruct video content, but can be used by the server as seen in Figure 6 to extract key frames that contain large motion or abrupt transition.

ISSN 2229-5518

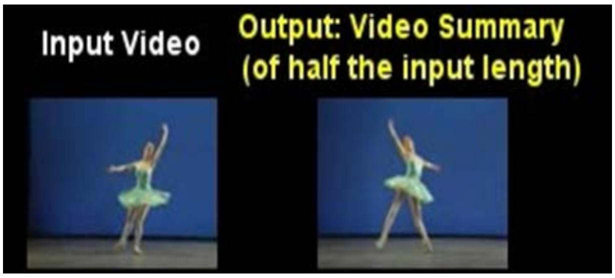

Figure 6: Video Summary [11]

Content Based Video Retrieval(CBVR) [12] is hence a progressive process of retrieving desired videos from a large collection. To perform content-based key frame extraction, two general approaches are pictured. One is to apply partial encryption of a video and make certain information available after encryption.

Another approach is to pre-process the video and identify key frames on the user side. The user then applies format compliant encryption on the video and sends the video along with the key frame numbers to the server. In this case, the key frame selection algorithm has to be very efficient to reduce computational cost on the user side.

6. VIDEO CLASSIFICATION

Another exciting video processing task is to automatically allocate tags such as beaches, portrait, indoor to videos and categorise the collection. Privacy-preserving video classification [13] and annotation is desirable because it can organize and present the private video collection as shown in Figure 7 for users.

One approach would be to let the server compare private videos with already labelled public videos and assign ranks based on most similar public videos.

In order to reduce difficulty in communication, only a limited set of encrypted distances should be sent back to the user. Therefore, the server needs to be able to define which public videos are more like the private video. Storing the visual features without any protection exposes content .One likely approach that can trade-off some security for improved efficiency would be to store a rough representation of visual features and allow the server to quickly overlook candidates that are unlikely to be similar.

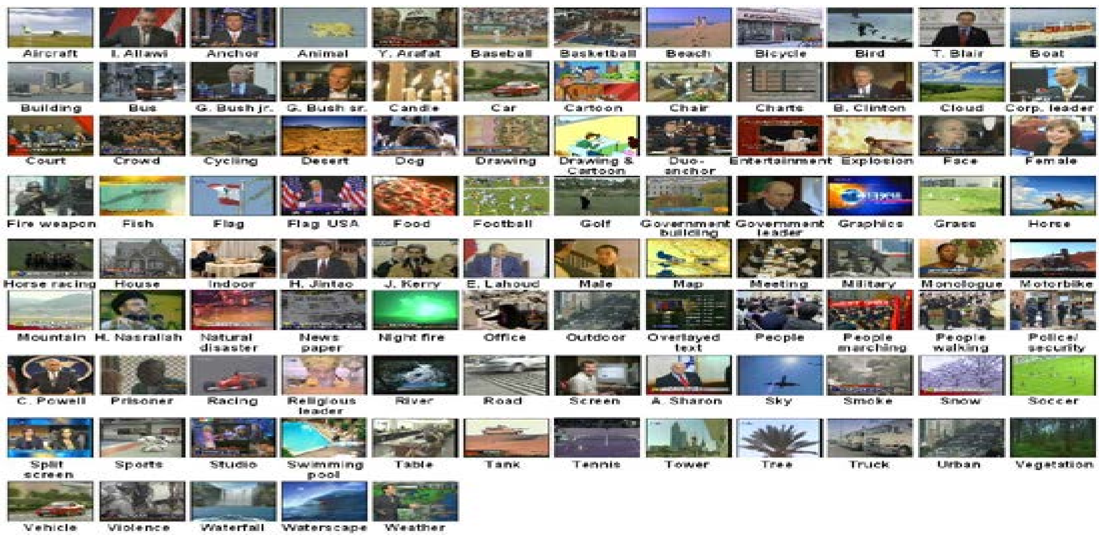

Figure 7: Snapshot of a Video Collection[14]

7. VIDEO ENCRYPTION

Privacy protection and secure processing techniques for video becomes a demanding need in the case of secure online video management, where users store their private videos in encrypted form on remote servers as seen in Figure 8 and the server performs processing tasks over the encrypted videos.

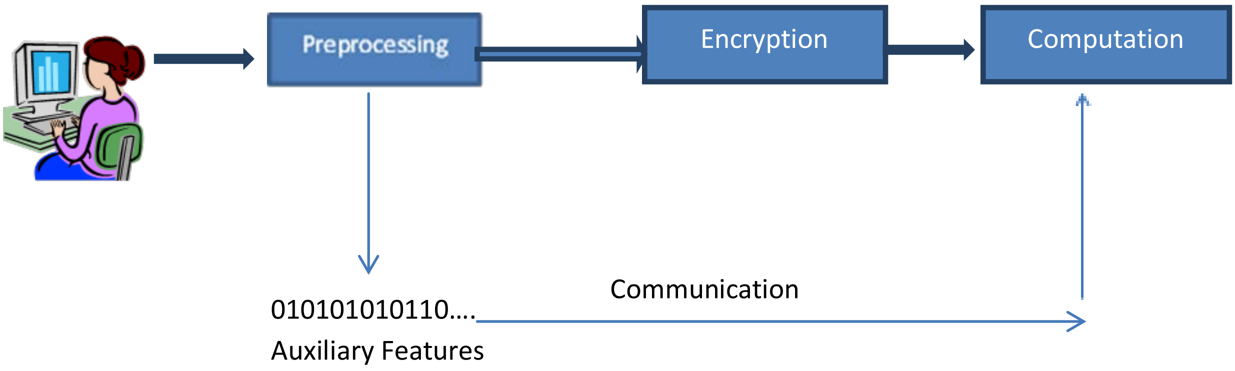

Figure 8 : Video Encryption System

Videos can be encrypted before or after compression, with various levels of protection.Efficient and secure video processing techniques will certainly have a compromise between efficiency and security. It is therefore a vital task to analyse the proper protection level required for each specific application.

One of the cryptography tools that can be used to compare encrypted information with public ones is homomorphic encryption [15]. Homomorphic encryption allows the computation of addition or multiplication between encrypted values. Homomorphic encryption is the conversion of data into cipher text that allows complex mathematical operations to be performed on encrypted data without compromising the encryption. It is often the semantic content and visual appearance of the data rather than the exact signal values and statistics that require to be protected. Using the Pallier additive homomorphic encryption [16] as an example, the summation of two values can be computed by the product of their encrypted values and the result is also homomorphicly encrypted.

8. VIDEO COMPRESSION

Video compression is used in networks to reduce transport time and storage costs. Transport bandwidth and latency requirements are critical for mobile or wireless access devices. Mobile access devices rely heavily on video compression in order to view streaming video. Consumer demand for higher quality video is focused on High Definition (HD) TV and the high graphics quality of video game platforms. The elementary concept of video compression is to decrease the amount of bits for video representation by exploiting spatial and temporal correlations in video sequences. Compression performance can be enhanced by employing motion-compensated prediction, which predicts each frame block wise from the previous frame. The prediction error can be more effectively compressed than the original frame data. Transcoding [17] is a fundamental network-based video-processing operation in which one video format is decompressed to a raw stream, which is then compressed into another format.. These transcoding and transrating operations are optimally done in the network, minimizing the bandwidth required to transmit the content

CONCLUSION:

The above discussed techniques are an attempt to give an overview of video processing challenges and solutions that are relevant to today's real time issues in video processing. In this article, we have considered the scenario of identifying reasonable issues andcurrent day challenges in video processing. Given the enormous and large size and affluent information of video data, it is important to design highly efficient and effective video processing techniques.

The effort of the authors in this paper will be fruitful if it could inspire the readers towards working on video processing challenges .

REFERENCES

[1]Yu-Wing Tai; Ping Tan; Brown, M.S., "Richardson-Lucy Deblurring for Scenes under a Projective Motion Path," in Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.33, no.8, pp.1603-1618, Aug. 2011

[2] Ehsan Shahrian Varnousfaderani, Medical University of Vienna.,MargritGelautz, Vienna University of Technology, Brian Price, Adobe Systems Inc.,Ji-Ho Cho, Vienna University of Technology, . "Image and Video Matting (ICCV 2013 Tutorial)", International Conference on Computer Vision (ICCV 2013), Sydney, Australia

[3]Matsushita, Yasuyuki, et al. "Full-frame video stabilization with motion inpainting." Pattern Analysis and Machine Intelligence, IEEE Transactions on 28.7 (2006): 1150-1163.

[4]Wang, Yao, and Qin-Fan Zhu. "Error control and concealment for video communication: A review." Proceedings of the IEEE 86.5 (1998): 974-997.

[5]Kung, Wei-Ying, Chang-Su Kim, and CC Jay Kuo. "Spatial and temporal error concealment techniques for video transmission over noisy channels." Circuits and Systems for Video Technology, IEEE Transactions on 16.7 (2006): 789-803.

[6]Kastrinaki, V., Michalis Zervakis, and Kostas Kalaitzakis. "A survey of video processing techniques for traffic applications." Image and vision computing 21.4 (2003): 359-381.

[7]Sen-Ching, S. Cheung, and Chandrika Kamath. "Robust techniques for background subtraction in urban traffic video." Electronic Imaging 2004. International Society for Optics and Photonics, 2004.

[8]Shen, Bo, and Ishwar K. Sethi. "Direct feature extraction from compressed images." Electronic Imaging: Science & Technology. International Society for Optics and Photonics, 1996.

[9]Smoliar, Stephen W., and HongJiang Zhang. "Content-based video indexing and retrieval." IEEE multimedia 2 (1994): 62-72.

[10] Wang, Zhou, Ligang Lu, and Alan C. Bovik. "Video quality assessment based on structural distortion measurement." Signal processing: Image communication 19.2 (2004): 121-132.

[11] Ferman, A. Müfit, and A. Murat Tekalp. "Two-stage hierarchical video summary extraction to match low-level user browsing preferences." IEEE Transactions on Multimedia 5.2 (2003): 244-256.

[12] Petkovic, Milan. "Content-based video retrieval." (2000): 74-77.

[13] Fan, Jianping, et al. "A novel approach for privacy-preserving video sharing." Proceedings of the 14th ACM international conference on Information and knowledge management. ACM, 2005.

[14] Brezeale, Darin, and Diane J. Cook. "Automatic video classification: A survey of the literature." Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 38.3 (2008): 416-430.

[15] Gentry, Craig. A fully homomorphic encryption scheme. Diss. Stanford University, 2009.

[16] Lagendijk, Reginald L., Zekeriya Erkin, and Mauro Barni. "Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation." Signal Processing Magazine, IEEE 30.1 (2013): 82-105.

[17] Zhang, Xianguo, et al. "A background model based method for transcoding surveillance videos captured by stationary camera." Picture Coding Symposium (PCS), 2010. IEEE, 2010.