A Novel Video Object Segmentation Approach for Noisy Video Sequences towards Effective Video Retrieval

S.Padmakala and Dr.G.S.AnandhaMala

***Abstract*—In modern times, video object segmentation has emerged as one of the most imperative and challenging area of research. The principal objective of video object segmentation is to facilitate content-based representation by extracting objects of interest from a series of consecutive video frames. Recently, a number of video object segmentation algorithms have been discussed and unfortunately most existing segmentation algorithms are not adequate and robust enough to process noisy video sequences. Competence of most segmentation techniques is affected by the presence of noise in frames which is a critical issue of edge preservation. This paper presents a novel video object segmentation approach for noisy color video sequences towards effective video retrieval. Initially, the noisy video frames are denoised using a strategy based on an enhanced sparse representation in transform domain. Afterwards, the background is estimated from the denoised frames using the Expectation Maximization (EM) algorithm. Then, the foreground objects i.e.) moving video objects are segmented with the aid of the novel approach presented. The biorthogonal wavelet transform and the L2 norm distance measure are employed in the foreground object segmentation. The experimental results demonstrate the effectiveness of the presented approach in segmenting the video objects from noisy color video sequences.**

***Index Terms*—Background Estimation, Biorthogonal wavelet transform, Denoising, Expectation Maximization (EM) algorithm, Foreground object, L2 norm, Moving Object, Video, Video Objects (VO), Video Object Segmentation.**

I. Introduction

Video has largely turned into a necessary component of today's multimedia applications; consisting of VCR (Video Cassette Recording), Video-on-Demand (VoD), virtual walkthrough and more with the speedy development in video technology [1]. A huge amount of video data is now extensively available owing to the technology's progress in multimedia, digital TV and information systems [2]. Due to the sudden growth in digital video content, an efficient way to access and manipulate the information in a huge video database has become a difficult and timely issue [3]. Therefore, the need for developing tools and systems that can effectively search and retrieve the desired video content has gained enormous popularity among researchers.

The development of novel sophisticated algorithms for proficient description, segmentation and representation of the image visual content will greatly determine the success of emerging multimedia applications, like video editing, video summarization, Content-Based Image Retrieval (CBIR), Content-Based Video Retrieval (CBVR) and video surveillance. Moving (or Motion) Picture Experts Group – 4 (MPEG-4) standard has launched the concept of Video Objects (VOs) for content-oriented coding of video sequences specifically to address all the aforesaid needs concerned with huge amount of video data [5].

Video content can be informally defined as a group of objects or activities, which appear in the particular video at few frame ranges. An object may be an item, a person or generally something concrete. An activity can be considered as an action or a relation among certain objects. There is no differentiation in the way of handling objects and activities and therefore generally it refers only to objects [6]. Video object segmentation has been recognized one of the vital issues in object-based video analysis and retrieval [10]. The emerging MPEG-4 standard has introduced the concept of content-based representation and content-based Visual query. The novel features of the MPEG-4 standard include enhanced coding efficiency, content-based interactivity and great error robustness for a large range of bit rates. In order to apply the MPEG-4 visual coding standard, each image of video sequences needs to be represented in terms of video objects [11], that is, the objects to be encoded in video sequences should be identified before the encoding process starts. So, video object segmentation has become more and more important [12].

The MPEG group has espoused the concept of Video Objects (VOs) and Video Object Planes (VOPs) for enhancing the coding efficiency and offering multimedia functionalities to the upcoming encoders. In a typical digital video stream, the meaningful entities of arbitrary shape (semantic objects), such as a human, a chair, a building and so on correspond to Video Objects, whereas the projection of video objects into a plane is represented by VOPs [8]. The objective of video object segmentation is to enable an object-based description of the scene by extracting objects of interest from a series of successive video frames. Most vision and multimedia applications have video object segmentation as an initial step [7]. Different applications, such as scene understanding, object-based video encoding, surveillance

S.Padmakala, Research Scholar, St.Joseph's College of Engineering, Anna University, Chennai, India. (e-mail: spadmakala1975@yahoo.com, spadmakalaphd@gmail.com)

Dr.G.S.AnandhaMala, Professor and Head, St.Joseph's College of Engineering, Chennai, India. (e-mail: malamanosuki@gmail.com)

applications and 2D-to-pseudo-3D video conversion, depend on the ability to extract objects from video sequences [13]. Moreover, it navigates through many application domains from security to medical imaging, with its major areas, being, though in no way limited to, surveillance and object tracking, content based video retrieval and analysis, video footage analysis for a variety of investigation purposes, traffic systems, video coding and medical diagnosis [14]. It is probable to access and perform object manipulation in the video with a good segmentation. It facilitates high-level image analysis like object recognition and scene interpretation [7].

The two classes of objects present in a video sequence are: Background objects and Foreground Objects. The background objects correspond to all those objects that are found in the scene during the entire sequence or longer than a predefined period of time. All other objects materializing in the scene represent the foreground objects [13]. The task of object segmentation is principally difficult when dealt with natural scenes where the background contains shadows, moving objects, and undergoes illumination changes. The video object segmentation techniques available in the literature can be grouped into three classes namely, 1) Region-based methods utilizing a homogeneous color or texture criterion, 2) Motion-based approaches using a homogeneous motion criterion, and 3) Object tracking [9]. Moreover, the proposed approaches for video object segmentation will be either supervised or unsupervised. In supervised approaches, the number of objects present in the sequence or more often the grouping of homogeneous regions to semantic objects [15]. [24], [32] is determined based on human interaction [31], while the unsupervised approaches necessitate no such interaction.

Several approaches have been proposed by researchers for video object segmentation [16 – 19, 21 - 27]. A brief review of image and video segmentation algorithms that have been presented for multimedia applications is provided in [16]. Every segmentation method emphasizes on different issues. A few methods are computationally expensive but provide; in general, accurate results while others have low computation but fail to offer reliable segmentation. A large number of video shots are used to effectively test a few of these methods and are examined throughout large shots. Moreover, many methods work in correspondence with fine tuned parameters developed by experts. A principal disadvantage of most methods is that they are not tested on noisy videos and videos with artifacts. In general, an object segmentation algorithm works by classifying the pixels of a video image into a certain number of classes that are homogeneous in correspondence to some characteristic (e.g., texture or motion). It creates objects by aggregating the image pixels. Some methods work based on color features and others on motion features. Some methods incorporate various features in expectation of better results. However, the use of additional features does not guarantee better result because some features can turn into erroneous or noisy data and complicate the achievement of a good solution [20].

In this paper, we propose a novel approach for segmentation of moving video objects in noisy color video sequences towards effective video retrieval. Initially, the frames are extracted from the noisy color video sequences and converted into grayscale. Afterwards, the noisy grayscale frames are denoised using the approach presented by Kostadin Dabov et al. [30]. Then, the background is estimated from the denoised frames with the help of the Expectation Maximization (EM) algorithm. Subsequently, the moving video objects in the foreground are segmented using the following steps: First, the estimated background image and the current frame are divided into micro blocks of size 2x2. Then, biorthogonal wavelet transform is applied on the micro blocks and block matching is carried out using L2 norm distance measure. The calculated distance measures are compared against a pre-defined threshold to segment the moving objects in the foreground. A black mask image of current frame size is created and the segmented foreground object is superimposed at the corresponding location in the mask image with white pixels. The above steps are carried out for all the frames extracted from the video sequence. Finally, the morphological operator “close” is applied on the resultant mask image to obtain the exact foreground object. The experimentation on the presented approach shows its efficiency in segmenting the moving objects from noisy color video sequences.

The rest of the paper is organized as follows: Section II presents a brief review of some of the recent researches in video object segmentation. A concise description of the concepts utilized in the presented approach is given in Section III. The proposed novel approach for noisy video object segmentation is presented in Section IV. The experimental results are given in Section V. Finally, the conclusions are summed up in Section VI.

II. Review of Related Researches

A handful of researches are available in the literature for video object segmentation in color video sequences. These researches have motivated us to do this research. A brief review of some of the recent researches is presented below:

Drelie Gelasca and Ebrahimi [21] have proposed a perceptually driven objective metric for segmentation quality evaluation, based on psychophysical experiments on synthetic artifacts. The perceptual metric proposed was tested by conducting a study on the real artifacts produced by typical video object segmentation algorithms. Seven recent segmentation algorithms were chosen and analyzed both objectively and subjectively. Initially, the real artifacts introduced were classified based on a subjective perception. Secondly, they proposed a perceptual objective metric capable of predicting the subjective quality as perceived by human viewers. The results showed an improved performance when compared to the commonly adopted MPEG and the wqm, mqm metrics and better adaptability when considered with various segmentation applications. Determination of optimal perceptual parameters for specific segmentation applications such as video compression and mixed reality was achieved.

The design procedure employed for generating a ground truth for the evaluation of motion-based algorithms for video-object segmentation has been described by Tiburzi *et al.* [22]. In general, motion-based segmentation algorithms are either based on a few pixel intensity models and, particularly on some optical flow estimation of the scene when dealt with moving cameras. Since specific settings for this factor can considerably increase (high complexity settings) or decrease (low complexity settings) segmentation accuracy. The algorithm's tuning and results evaluation were made possible by the extensive range of critical factors considered and the development of an accurate ground-truth for their presented corpus. Unger *et al.* [23] have proposed a novel method for the generation of artificial background images that allows background subtraction from video sequences with unconstrained camera movement. They have made use of optical flow to estimate the global motion models. For every video frame, an artificial background image was produced by the accumulation of compensated global motion and weighted neighboring frames. Their major contribution was the statistical modeling of the deviation of optical flow and global motion, which were used to determine the weights for accumulation. The proposed Rayleigh-based method delivered superior results when estimated quantitatively. The only downside of their work is that the background pixels in succeeding video frames must be extremely correlated.

Fatih Porikli and Yao Wang [9] have proposed an Automatic Video Object Segmentation approach using Volume Growing and Hierarchical Clustering technique that blends the advantages of color, texture, shape, and motion of various segmentation methods in a computationally feasible way. Initially, a spatiotemporal data structure has been built for each group of video frames, in which each pixel is assigned a feature vector on the basis of low-level visual information. Subsequently, using an adaptive, three dimensional, centroid-linkage method, the least homogeneous components called as volumes, are expanded from selected marker points. The self descriptors are then determined by assessing the boundary, trajectory, and motion of the volumes. A self descriptor characterizes every volume and the relational descriptors that capture the mutual properties between pairs of volumes. These descriptors are utilized to compute the similarity between different volumes and the volumes are further grouped into objects based on that similarity value. The segmentation output of a fine-to-coarse clustering algorithm was a multi-resolution object tree representation.

A Segmentation method for deformable objects in monocular videos has been presented by Peng Tang and Lin Gao [25]. Initially, they have introduced a dynamic shape to symbolize the prior knowledge about object shape deformation in a manner of auto-regressive model, which treats the shape as a function of subspace shapes at previous time steps. A framework of Markov random field energy was developed by fusing both spatial-temporal image information and model predictions. The devised framework of Markov random field energy can be effectively minimized by the graph cut algorithm so as to achieve a global optimum segmentation. Both the orthogonal basis and the auto regressive model parameters are updated on-line, using final segmentation results in order to capture model variations, thereby forming an effective closed loop system. Lastly, the impending of the proposed segmentation method with respect to noise, clutter, and partial occlusions were demonstrated by means of promising experimental results.

Dubravko Culibrk *et al.* [26] have presented a background modeling and subtraction approach for video object segmentation. Also, in correspondence to their application domain, they have proposed a Neural Network (NN) architecture with the intention of building an unsupervised Bayesian classifier. The segmentation in natural-scene sequences with complex background motion and changes in illumination have been efficiently handled by constructed classifier. The weights of the proposed NN act as a model of the background and are updated temporally to reflect the observed statistics of background. Based on the formerly published test pool containing diverse surveillance-related sequences, the segmentation performance of the proposed NN was examined both qualitatively and quantitatively, and was compared with two existing probabilistic object segmentation algorithms. The proposed algorithm was parallelized on a subpixel level and designed to facilitate efficient hardware implementation. Aree A. Mohammed *et al.* [27] have developed an object extraction method and proposed efficient algorithms for characterization of object motion. The set of their proposed tools serves as the base for development of object based functionalities for manipulating video content. The estimators produced by different algorithms were compared with respect to quality and performance and was also tested on real video sequences. Their proposed method will aid for the most recent encoding standards and description of multimedia content – MPEG4 and MPEG7.

III. Prerequisites

In this section, a concise description of biorthogonal wavelet transform and L2 norm distance measure is presented.

A. Biorthogonal Wavelet Transform

The biorthogonal wavelet transform falls under the class of invertible transforms. Since the biorthogonal wavelets possess two sets of low-pass filters (for reconstruction), and high-pass filters (for decomposition), the properties of perfect reconstruction and symmetric wavelet functions prevail in them. One set is considered as the other one's dual. In contrast, the orthogonal wavelets comprise only one set. In case of biorthogonal wavelets, the decomposition and reconstruction filters are acquired from two dissimilar scaling functions related to two multi resolution analyses in duality. An additional beneficial characteristic of biorthogonal wavelets over the orthogonal wavelets is that the former kind has higher embedding capacity when utilized in the decomposition of the image into diverse channels [28]. The aforesaid characteristics of the biorthogonal wavelets make them potential candidates in the wavelet domain.

Biorthogonal wavelets initiate two wavelets. One of them, φ , aids in the analysis while the other ψ , aids in synthesis as is found in (a) and (b) respectively

$$C_{j,k} = \int s(x)\varphi_{j,k}(x)dx \qquad (1)$$

$$s = \sum_{j,k} c_{j,k} \psi_{j,k} \qquad (2)$$

Here, s represents the signal while j,k are integers.

B. L2 Matrix Norm

The L2-norm of a matrix A is given as [29]:

$$\|A\|_2 = \max_{(1<=i<=N)} (sqrt(\lambda_i))$$

Where λ_i is an (always real) eigenvalue of $A^T A$; or

$$\|A\|_2 = \max_{(1<=i<=N)} (mu_i)$$

Where mu_i is a singular value of A.

IV. Novel Approach for Video Object Segmentation In Noisy Color Video Sequences

The proposed novel approach for video object segmentation in noisy color video sequences is presented in this section. Noisy color video sequences in Audio Video Interleaved (AVI) and Moving (or Motion) Picture Experts Group - 1 (MPEG -1) formats can be segmented using the proposed approach. To start with, the frames in the color video sequence are extracted and converted into grayscale. Gaussian noise is then added to the grayscale frames. Afterwards, the noisy grayscale frames are denoised using the image denoising strategy presented by Kostadin Dabov *et al.* [30]. Subsequently, the background is estimated from the denoised frames using the Expectation Maximization (EM) algorithm. Consequently, the foreground objects i.e.) moving video objects are segmented using the novel approach presented. A description of the de-noising algorithm employed, background estimation and foreground object segmentation are given in the following sub-sections. Fig 1 shows the block diagram of the proposed video object segmentation approach.

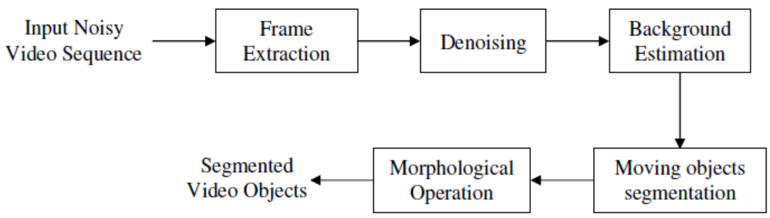

Fig. 1. Block diagram of the proposed video object segmentation approach

A. Denoising Algorithm

In the proposed video object segmentation approach for denoising the noisy video frames, we have employed the image denoising strategy presented by Kostadin Dabov *et al.* [30]. Their strategy is based on an enhanced sparse representation in transform domain. The similar 2D fragments (e.g. blocks) of the images are grouped into 3D data arrays for enhancing the sparsity. Collaborative filtering, a special procedure, is also presented to manage these 3D groups. The denoising strategy consists of three successive steps namely: 3D transformation of a group, shrinkage of transform spectrum, and inverse 3D transformation. Precisely, block matching and shrinkage in a 3D transform domain are utilized for the accomplishment of grouping and collaborative filtering respectively. The general procedure involved in the algorithm is as follows.

Processing of the input noisy image is done by successively extracting reference blocks from it and for every block:

- Locate blocks that are identical to the reference block (block matching) and stack them together to form a 3D array (group);
- Collaborative filtering of the group is carried out and the resulted 2D estimates of all grouped blocks are returned to their respective original locations.

The obtained block estimates can overlap, hence there are multiple estimates for each pixel and these estimates are aggregated to obtain the estimate of the whole image. The detailed steps in the algorithm can be found in [30]. The strategy achieves denoising at reasonable computational cost and allows for effective complexity/performance trade-off. In addition, the employed strategy can be used for denoising various noises such as additive colored noise, non-Gaussian noise and more by modifying the calculation of coefficient's variances in the algorithm.

B. Background Estimation

The estimation of background from the denoised frames is explained in this sub-section. The Expectation Maximization (EM) [4] algorithm is employed in the estimation of the background. Let there be, F individual frames in the video sequence and a background B_I. The background can be estimated by maximizing the log likelihood

$$P_{B_I} = D_1 * e^{\left(\prod_{i=1}^{|F|} (-(F_1 - F_i)), \left[D_2 * \sum_{i=1}^{|F_i|} \prod \left(F_{i-1}, P_{B_I}^{i-1} \right) \right]^{-1} \right)}$$

Where D_1 and D_2 are constants, P_{B_I} be the probability of the video sequence. The uniform probability obtained of the background is used to update the estimated background. The update function of the background B_I is

$$B_I \leftarrow \sum_{i=1}^{n} \left(\prod P_{B_I}^i, F_i \right) . \left(P_{B_I}^i \right)^{-1}$$

Obviously, a video sequence is likely to contain moving foreground objects along with the background. Therefore, our primary concern is to eliminate those foreground objects to the maximum. When considering the successive frames in a video sequence, it is believed that the foreground object moves about the fixed background. So, we correlate the consecutive frames in the video sequence to determine the background by maximizing the above log likelihood function. The maximum log likelihood function when applied iteratively enables a better portrayal of the background. For estimating the background effectively, we have used an update function which improves upon the value of 'B_I'. We repeat the maximization of log likelihood function to eliminate all possible disturbances by the foreground object.

C. Foreground Object Segmentation

This sub-section describes the novel approach proposed for the segmentation of foreground objects. The background estimated from the denoised frames is utilized in the segmentation of the foreground objects. In addition, black mask images are used to mark the segmented foreground object. Initially, the background image is divided into micro blocks of size 2 x 2. Similarly, the candidate frames (denoised video frames) are divided into micro blocks of size 2 x 2. Then, the foreground objects in the candidate frames are segmented by computing the similarity between micro blocks of each candidate frame and the background using the L2- norm distance method. The steps involved in the computation of similarity between two micro blocks are as follows: First, the biorthogonal wavelet transform is applied on the two micro blocks. Subsequently, the distance between the two micro blocks is calculated by employing the L2-norm distance method. The equation describing the above steps is as follows:

$$d(Y_{BI}, Y_{CB}) = \| Bior(Y_{BI}) - Bior(Y_{CF}) \|_2$$

Where $\|\cdot\|_2$ denotes the L2-norm, Y_{BI} denotes a micro block in the estimated background image, Y_{CB} denotes a micro block in a candidate frame and $Bior()$ represents the biorthogonal wavelet transform. On the basis of the computed L2-norm distance, the foreground object is segmented and superimposed on the mask image. The micro blocks having distance less than a predefined threshold are marked with white pixels on the mask image. On completion of these steps, a rough sketch of the foreground object is obtained in the mask image, which is bound to contain noises. Finally, to eliminate the noise, the morphological operator "Close" is applied on the formed mask image. Based on the formed mask image, the actual foreground object from the original color video frame is extracted and superimposed at the corresponding locations in the mask image.

V. Experimental Results

This section presents the results obtained from the experimentation on the proposed video object segmentation approach. The presented approach has been validated by experimentation with several kinds of color video sequences in AVI and MPEG-1 formats. The proposed approach is programmed in Matlab (Matlab 7.4). The intermediate results obtained in segmenting the video objects in a color video sequence are presented here. The Gaussian noise (σ = 20) added video frames, denoised video frames, estimated background, mask image with segmented foreground objects, mask image with segmented foreground objects after applying morphological operator, and segmented foreground objects are portrayed along with the original color video frames and their corresponding grayscale frames in the following figures.

Fig. 2 Sample Frames of input color video sequence

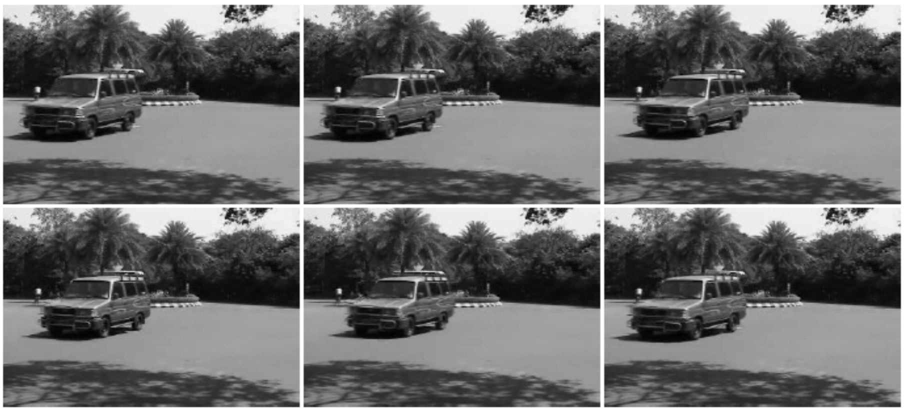

Fig. 3. Grayscale converted frames

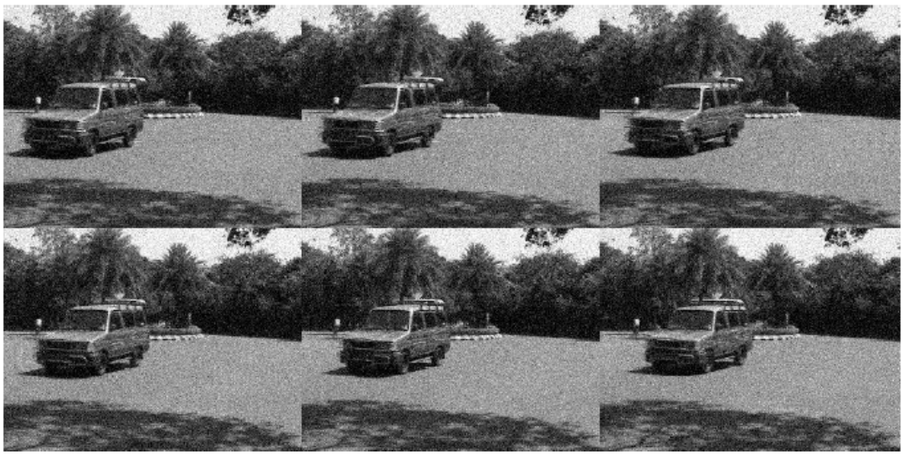

Fig. 4. Grayscale frames with Gaussian noise (σ =20)

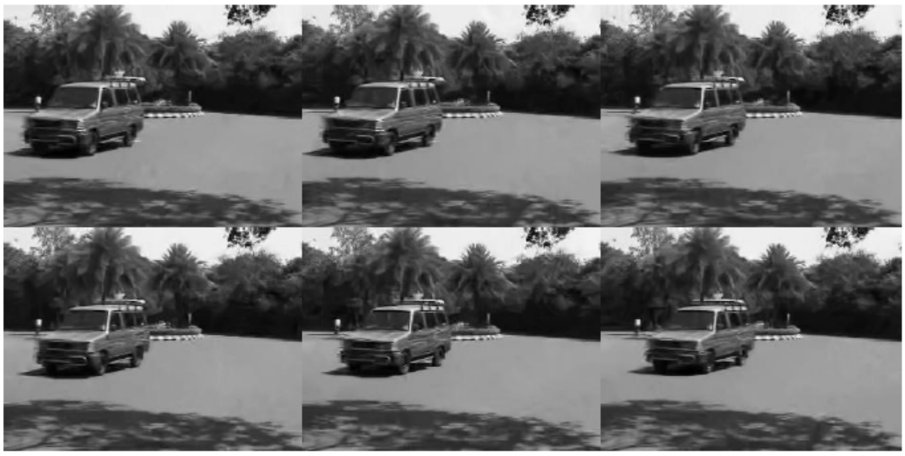

Fig. 5. Denoised video frames

Fig. 6. Estimated Background

IACSIT
International Association of
Computer Science and Information Technology
WWW.IACSIT.ORG

Fig. 7. Mask image with segmented foreground objects

Fig. 8. Mask image with segmented foreground objects after applying morphological operator

Fig. 9. Segmented foreground objects

VI. CONCLUSION

In digital video processing and computer vision, one of the most challenging and active research areas is video object segmentation. A significant issue for the successful use of many video sequences is video object segmentation that accentuates partitioning the video frames to semantically meaningful video objects and backgrounds. Video object segmentation is a vital operation for content-based video coding, multimedia content description, intelligent signal processing and more. In this paper, we have presented a novel and efficient video object segmentation approach for noisy color video sequences towards achieving effective video retrieval. An image denoising strategy based on an enhanced sparse representation in transform domain has been utilized in denoising the noisy frames. Expectation Maximization (EM) algorithm is utilized in the estimation of background and foreground objects have been extracted with the aid of the novel approach presented. The novel approach presented for foreground object segmentation made use of biorthogonal wavelet transform and L2-norm distance measure. The experimental results have demonstrated the effectiveness of the proposed approach in segmenting the video objects from noisy color video sequences.

REFERENCES

[1] Rynson W.H. Lau, Qing Li, Antonio Si, "VideoMAP: A Generic Framework for Video Management and Application processing", 33rd Hawaii International Conference on System Sciences (HICSS), pp.8044, Volume 8, 2000.

[2] Walid G. Aref, Ann C. Catlin, Ahmed K. Elmagarmid J. Fan, Moustafa A. Hammad, Ihab Ilyas, Mirette Marzouk, and Thanaa Ghanem, “Video Query Processing in the VDBMS Testbed for Video Database Research”, 1st ACM International Workshop on Multimedia Databases, (ACM MMDB), PP 25-32, 2003.

[3] Chih-Wen Su, Hong-Yuan Mark Liao, Kuo-Chin Fan, “A Motion-Flow-Based Fast Video Retrieval System”, 7th ACM SIGMM international workshop on Multimedia Information retrieval, November 10-11, 2005.

[4] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). “Maximum likelihood from incomplete data via the EM algorithm,” Journal of the Royal Statistical Society, B, 39(1):1-38.

[5] K. N. Ngan, S. Panchanathan, T. Sikora and M.-T. Sun, “Guest Editorial: Special Issue on Representation and Coding of Images and Video,” IEEE Trans. on CSVT, Vol. 8, pp. 797-801, Nov. 1998.

[6] C. Makris, K. Perdikouri, S.Sioutas, A.Tsakalidis, K. Tsichlas, "Time and Space Efficient Content Queries for Video Databases", MDDE 2001 Workshop (Multimedia Data and Document Engineering), Vol. 43, pp 1-8, July 2001.

[7] Porikli, F.M., "Video Object Segmentation by Volume Growing Using Feature-Based Motion Estimator", In proceedings of 16th International Symposium on Computer and Information Sciences (ISCIS), November 2001.

[8] T. Sikora, “The MPEG-4 Video Standard Verification Model,” IEEE Trans. on Circ. and Syst. for Video Techn. vol. 7, no. 1, pp. 19-31, Feb. 1997.

[9] Porikli, F.M. and Wang, Y., "Automatic Video Object Segmentation Using Volume Growing and Hierarchical Clustering", EURASIP Journal on Applied Signal Processing, ISSN: 1536-1276, Vol. 3, Issue 2, pp. 442-453, March 2004.

[10] Lifang Wu and Xianglong Meng, “A Robust Object Segmentation Method”, International Journal of Innovative Computing, Information and Control, Vol. 4, No. 11, pp. 3059-3065, November 2008.

[11] Çigdem Eroglu Erdem, A. Murat Tekalp, Bülent Sankur, “Video Object Tracking with Feedback of Performance Measures”, IEEE Transactions Circuits and Systems for Video, vol. 13, no. 4, April 2003.

[12] Ming Zhao, Jiajun Bu, and Chun Chen, "Semi-Automatic Video Object Segmentation Basing On Hierarchy Optical Flow", in SPIE: Electronic Imaging and Multimedia Technology III, vol. 4925, pp. 307-316, October 2002.

[13] Dubravko Culibrk, Vladimir Radenkovic, Daniel Socek, "Enhancing Video Object Segmentation Results Through Biologically Inspired Postprocessing",8th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services (TELSIKS 2007), Serbia, pp. 329-332, 26-28 September 2007.

[14] Ahmed, R.; Karmakar, G.C.; Dooley, L.S., “Probabilistic Spatio-Temporal Video Object Segmentation Incorporating Shape Information”, IEEE International Conference on Acoustics, Speech and Signal Processing, 14-19 May 2006.

[15] S. Herrmann, H. Mooshofer, H. Dietrich, and W. Stechele, “A video segmentation algorithm for hierarchical object representations and its implementation,” IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp.1204–1215, Dec. 1999.

[16] Salembier, P. and F. Marques, Region-based Representations of Image and Video: Segmentation Tools for Multimedia Services, IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 8, pp. 1147-1169, 1999.

[17] "Special issue on segmentation, description, and retrieval of video content." IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 5, Sept. 1998.

[18] A. Amer and E. Dubois, "Image segmentation by robust binarization and fast morphological edge detection," in Proc. IAPR/CIPPRS Int. Conf. on Vision Interface, pp. 357-364, (Montreal, Canada), May 2000.

[19] "Workshop on image analysis for multimedia interactive services," Proc. COST211ter, Louvain-la-Neuve, Belgium, June 1997.

[20] Aishy Amer, "Memory-based Spatio-Temporal Real-Time Object Segmentation for Video Surveillance", Real-time imaging. Conference No7, Santa Clara CA, ETATS-UNIS, vol. 5012, pp. 10-21, 22 January 2003.

[21] E. Drelie Gelasca and T. Ebrahimi, “On Evaluating Video Object Segmentation Quality: A Perceptually Driven Objective Metric”, IEEE Journal of Selected Topics in Signal Processing, vol. 3, no. 2, pp. 319-335, 2009.

[22] Tiburzi, F. Escudero, M. Bescos, J, Martinez, J.M.,” A Ground Truth for Motion-Based Video-Object Segmentation”, 15th IEEE International Conference on Images Processing (ICIP), pp 17-20, Oct. 2008.

[23] Unger, M. Asbach, M. Hosten, P., “Enhanced Background Subtraction Using Global Motion Compensation and Mosaicing”, 15th IEEE International Conference on Image Processing (ICIP), pp 2708-2711, Oct. 2008.

[24] D. Gatica-Perez, C. Gu, and M.-T. Sun, “Semantic video object extraction using four-band watershed and partition lattice operators,” IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 603–618, May 2001.

[25] Peng Tang and Lin Gao, “Video Object Segmentation Based On Graph Cut With Dynamic Shape Prior Constraint”, 19th International Conference on Pattern Recognition (ICPR), pp. 1 – 4, Dec. 2008.

[26] Dubravko Culibrk, Oge Marques, Daniel Socek, Hari Kalva and Borko Furht, "Neural Network Approach to Background Modeling for Video Object Segmentation", IEEE Transactions on Neural Networks, pp. 1614-1627, November 2007.

[27] Aree A. Mohammed, “Region-Based Segmentation of Generic Video Scenes Indexing”, Proceedings of World Academy of Science, Engineering and Technology, Vol. 32, August 2008.

[28] Suhad Hajjara, Moussa Abdallah and Amjad Hudaib, "Digital Image Watermarking Using Localized Biorthogonal Wavelets", European Journal of Scientific Research, Vol. 26, No. 4, pp. 594-608, February, 2009.

[29] Horn, R. A. and Johnson, C. R. "Norms for Vectors and Matrices" Ch. 5 in Matrix Analysis. Cambridge, England: Cambridge University Press, 1990.

[30] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," IEEE Transactions on Image Processing, vol. 16, no. 8, August 2007.

[31] L. Patras, E. A. Hendriks, and R. L. Lagendijk, “Video segmentation by MAP labeling of watershed segments,” IEEE Trans. Pattern Anal. Machine Intell. vol. 23, pp. 326–332, Mar. 2001.

[32] R. Castagno, T. Ebrahimi, and M. Kunt, “Video segmentation based on multiple features for interactive multimedia applications,” IEEE Trans. Circuits Syst. Video Technol., vol. 8, pp. 562–571, Sept. 1998.

S.Padmakala received the B.E. degree, from the Department of Computer Science, Bharth Institute of Technology, University of Chennai, Chennai, India and M.E. degree from the Department of Computer Science, Anna University, Chennai, India, in 1997 and 2006, respectively. She is currently pursuing the Ph.D. degree in Anna University, Chennai, India, working closely with Prof. G.S. Anandhamala. Presently, she is working as Associate Professor, at St.Joseph's College of Engineering, Chennai, India from 2000 to till date.

Dr.G.S.Anandha Mala received B.E degree from Bharathidhasan University, Trichy, India in Computer Science & Engineering in 1992, M.E degree in University of Madras in 2001 and Ph.D degree from Anna University, Chennai, India in 2007. Currently she is working as Professor in St.Joseph's college of Engineering, Chennai, India, and heading the department of Computer Science and Engineering. She has published 20 technical papers in various international journal / conferences. She has 15 years of teaching experience on graduate level. Her area of interest includes Software Engineering and Grid Computing.