

## Improved Key Frame Extraction Using Discrete Wavelet Transform with Modified Threshold Factor

Hussein Ali Aldelfy<sup>\*1</sup>, Mahmood Hamza Al-Mufraji<sup>2</sup>, Thamir R. Saeed<sup>3</sup>

<sup>1,2,3</sup> Department of Electrical Engineering, University of Technology, Iraq

<sup>1</sup> Faculty of Engineering, University of Mustansiriyah, Baghdad, Iraq

\*Corresponding author, e-mail: hussienaldulfy@gmail.com

### Abstract

Video summarization used for a different application like video object recognition and classification. In video processing, numerous frames containing similar information, this leads to time consumption and slow processing speed and complexity. By using key frames reducing the amount of memory needed for video data processing and complexity greatly. In this paper key frame extraction of Arabic isolated word using discrete wavelet transform (DWT) with modified threshold factor is proposed with different bases. The results for different wavelet basis db, sym and coif show the best result for numbers of key frames at the threshold factor value (0.75).

**Keywords:** Key frame, Video summarization, Wavelet

Copyright © 2018 Universitas Ahmad Dahlan. All rights reserved.

### 1. Introduction

The videos of the dynamic signs consist of large number of frames are not essential in order to determine the meaning of the performed sign, rather, only a few important frames from the video are sufficient. These most important and thus distinguishing frames are known as key frames [1]. Through the key frame sequence detection and identification, the sign language can be rapidly recognized. At the same time, sign language is a way for deaf people to communicate and exchange ideas through finger alphabet and gestures instead of language [2]. Extraction of key frames from the video and to analyze only these frames instead of all the frames present in the video can greatly improve the performance of the system. Analysis of these key frames can help in forming the annotations for the video.

Different techniques for key frame extraction were reported, a key frame extraction method for video copyright protection was presented in [3]. In this technique, a two-stage method is used to extract accurate key frames to cover the content for the whole video sequence. key frame extraction method based on unsupervised clustering and mutual comparison were developed in [4]. A method of key frame extraction using thresholding of absolute difference of histogram of consecutive frames of video data is proposed [5]. The brief representation and comparison of effective key frame extraction methods like cluster-base analysis, generalized Gaussian density method(GGD), General-Purpose Graphical Processing Unit (GPGPU), Histogram difference was presented in [6]. A square histogram based model using frame segmentation and automatic threshold calculation was developed in [7]. Key frame extraction method using wavelet statistics was proposed in [8] and the edge change ratio algorithm for detecting shots of the video and key frames are extracted from these shots [9].

Video summarization has been proposed to improve faster browsing of large video collections and more efficient content indexing and access. As the name implies, video summarization is a mechanism to produce a short summary of a video to give the user a synthetic and useful visual abstract of a video sequence, it can either be key frames or video skims [10]. By using the key- frame it is able to express the main content of video data clearly and reduce the amount of memory needed for video data processing and complexity greatly. so that could make the storage organization, retrieval and recognition of video information more convenient and efficient, thus key frame extraction is an efficient method for video summarization [5]. In this study, an improved method for key frame detection using discrete wavelet transform (DWT) with modified threshold factor is proposed to obtain near an optimum number of frames in order to decrease the required processing time. This paper is organized as

follows. Section 2 show the uses of key frame extraction. Section 3 presents the proposed method, and Section 4 briefly discusses the experimental results and conclusion.

## 2. Uses of Key-Frame Extraction

In this section, the applications that used for the key frame extraction are described as:

### 2.1. Video transmission

In order to reduce the transferred stress in the network and invalid information, the transmission, storage and management techniques of Video information become more and more important [11]. When a video is being transmitted, the use of key frames reduces the amount of data required in video indexing and provides the framework for dealing with the video content [12]. A key frame based online coding video transmission is proposed. Each frame can only choose the latest coded and reconstructed key frame as its reference frame. After coding and packetisation, compressed video packets are transmitted with differentiated service classes. Key frame along with difference values are sent from the source, using the key frame picture and difference values the picture is reconstructed at the destination [13].

### 2.2. Video annotation

Video annotation is the extraction of the information about Video, adding this information to the video which can help in browsing, searching, analysis .retrieval, comparison, and categorization. Annotation is to attach data to some other piece of data [14].

### 2.3. Video indexing

Key frames reduce the amount of data required in video indexing and provide framework for dealing with the video content. Before downloading any video over the internet, if key frames are shown beside it .users can predict it content of the video and decide whether it pertinent to his search. Other applications such as creating chapter titles in DVDs and prints from video [15].

### 2.4. Video Summarization

Video summarization is a compact representation of a video sequence. It is useful for various video applications such as video browsing and retrieval systems. A video summarization can be a preview sequence which can be a collect of key frames which is a set of chosen frames of a video. Key frame based video summarization may lose the spatio-temporal properties and audio content in the original video sequence, it is the simplest and the most common method [10].

## 3. Proposed Method

The proposed method comprises of four steps. The first, the video frames are read and, then two consecutive frames transformed with DWT to obtain four subbands, LL, HL, LH, and HH. Only three sub-bands, HL, LH, and HH are used to detect key frame. For each sub band different value is estimated by subtracting detail component values of current and next frame as shown in Equation 1.

$$D_l(k) = \sum_{i=1}^m \sum_{j=1}^n (cH_{l+1}(i,j) - cH_l(i,j)) \quad l = 1,2,3 \dots \dots \dots (1)$$

where  $i$  and  $j$  are numbers of row and column,  $cH_l$  are the HL, LH and HH band of gray images  $V_k$  and  $cH_{l+1}$  are the HL, LH and HH band of gray images  $V_{k+1}$ .

In the second step, mean  $\mu_l$  and standard deviation  $\sigma_l$  are calculated from step one using difference values of sub bands Equation 2 and Equation 3.

$$\mu_l = \frac{\sum_{i=1}^{N-1} (D_l(i))}{N-1} (2)$$

$$\sigma_l = \text{sqr}t \left[ \frac{\sum_{i=1}^{N-1} (D_l(i) - \mu_l)^2}{N-1} \right] (3)$$

The third step, threshold value  $\gamma_l$  for each sub band is computed by Eq.4. the last step a comparison between threshold and difference values and if the difference between the present values and the previous value greater than the related threshold ,the last frame considered as a key frame [7].

$$\gamma_l = \mu_l + \beta \cdot \sigma_l \quad (4)$$

where  $\gamma_l$  is the threshold value at level  $l$  and  $\beta$  is the modified threshold factor. The Figure 1 shows the flow chart for key frame extraction using modified threshold factor.

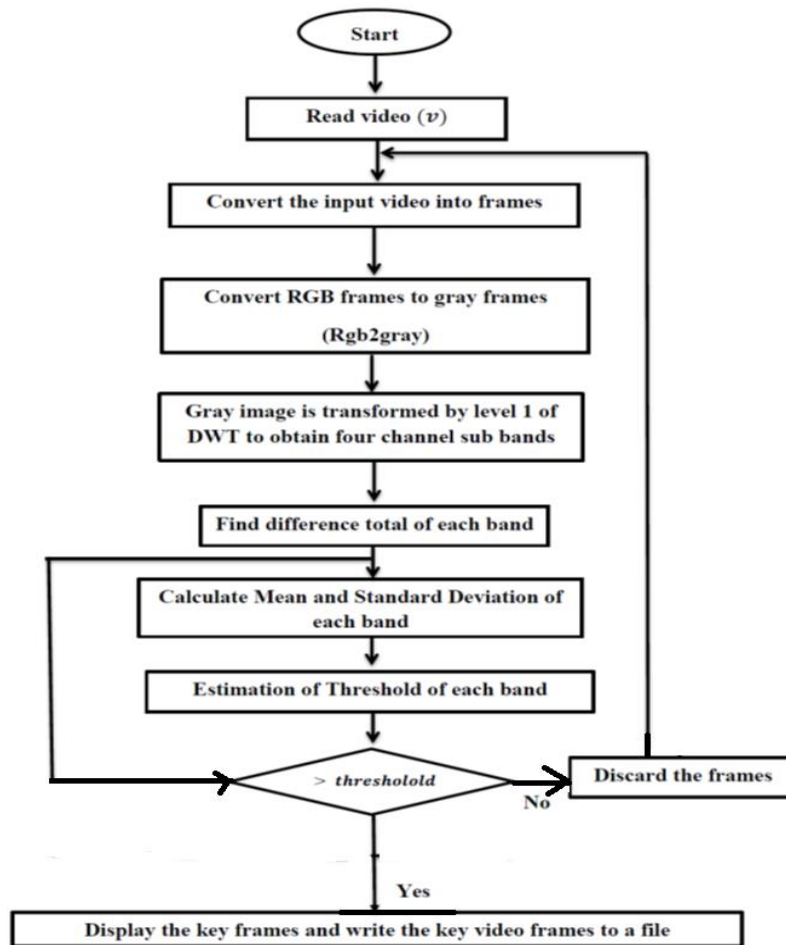


Figure 1. Flowchart for Key Frame Extraction using Modified Threshold Factor

#### 4. Experimental Results and Discussion

In this section of research is indicative words belong to Deaf and dumb people, these words were collected in Arabic language and Iraqi dialect and individual words in collaboration with Iraqi Ministry of labour and Social Affairs. More than forty words signalled task in dealing two words is used, water and drug are used as inputs to the proposed system also been using MATLAB program 2013b in the implementation of the proposed algorithm.

In Table 1 db wavelet basis is used with different values of modified factor and applied to the drug video that 2.96 Mhz and 30 frames/sec the contain 43 frames and the results show that as the  $\beta$  increased the numbers of key frames reduced otherwise if the threshold factor  $\beta$  smaller value chosen this lead to increase in key frames and for this reason at  $\beta$  equal to 0.75 can be considered suitable value. Furthermore, with different values of scale db1, db2 and db4 no effect appears in the numbers of key frames.

In the Table 2 Coif wavelet basis is used and observe the change in the number of key frames by changing the threshold factor value and obviously when the value is 0.75 the number of key frames an efficient. The key frames not influenced by changing the wavelet basis as there's not an apparent effect when changing the wavelet basis. In Table 3 sym wavelet basis is used ,the same results and conclusions as mentioned above for Table 1 and Table 2.

Table 1. Key Frame Numbers with a Different Value of Modified Threshold Value using db1, db2 and db3

$\beta$	No. of Frames N	Db1(keyframe t)	Db2(keyframe t)	Db4(keyframe t)
0.5	43	8	9	9
0.75	43	6	6	6
1	43	2	2	2
1.25	43	1	2	2
1.5	43	1	1	1

Table 2. Key Frame Numbers with a Different Value of Modified Threshold Value using Coif1, Coif2 and Coif3

$\beta$	No. of Frames N	Coif1(key frame t)	Coif2(key frame t)	Coif3(key frame t)
0.5	43	8	9	8
0.75	43	6	6	6
1	43	2	2	2
1.25	43	1	1	1
1.5	43	1	1	1

Table 3. Key Frame Numbers with a Different Value of Modified Threshold Value using Sym1, Sym2 and Sym3

$\beta$	No. of Frames N	Sym2(key frame t)	Sym3(key frame t)	Sym4 (key frame t)
0.5	43	9	8	9
0.75	43	6	6	6
1	43	2	2	2
1.25	43	2	1	2
1.5	43	1	1	1

In Figure 2 show the frames that contains in drug.mp4 video clip with the size of 2.96 MB and frame rate 30 frame/sec is 43 image. In Figure 3 show the key frame images after applying the algorithm based on wavelet bases db1 and  $\beta=0.5$  about 8 images.



Figure 2. Frames of the Input Video of Drug Word



Figure 3. Output key Frames of Drug Word Video db1 for  $\beta=0.5$

In Figure 4 show the frames that contain in a water.mp4 video clip with the size of 3.56 MB and frame rate 30 frames/sec is 52 image. In Figure 5 show the key frame images after applying the algorithm based on wavelet bases db1 and  $\beta=0.5$  about 15 images.

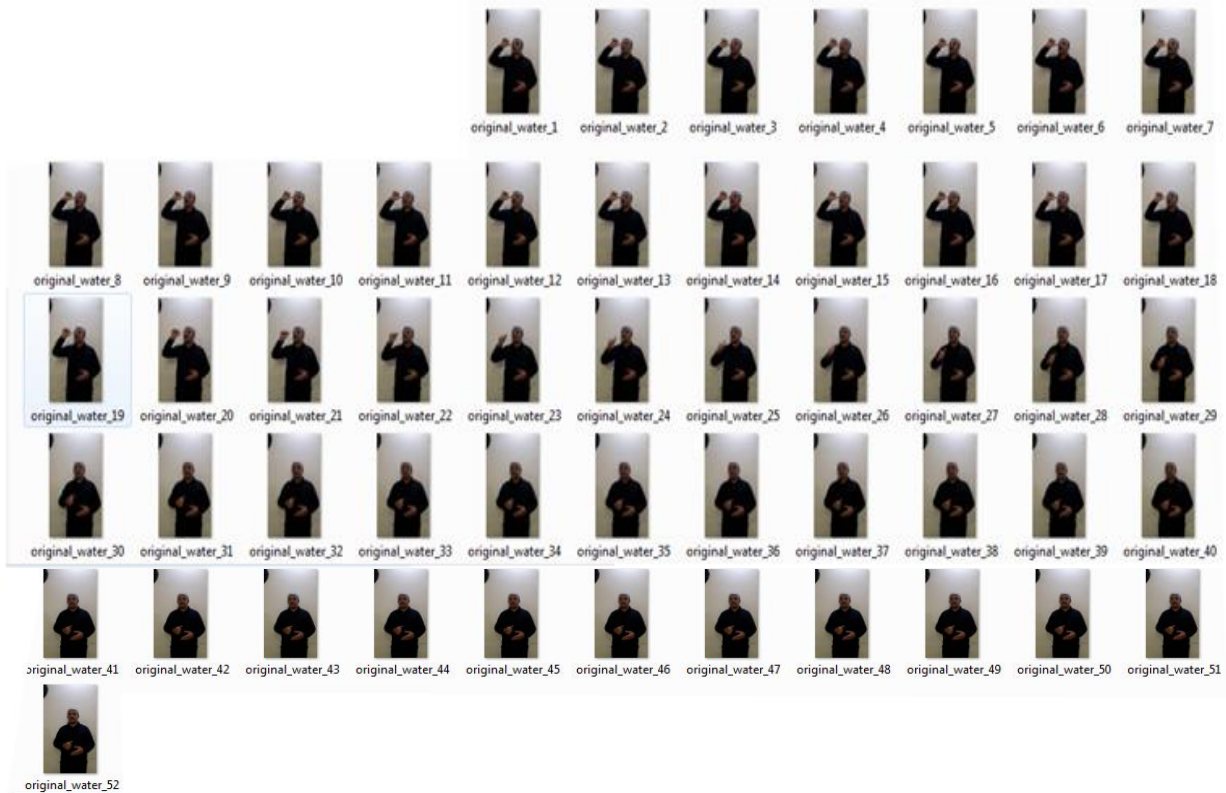


Figure 4. Frames of Input Video of Waterward



Figure 5. Output Key Frames Waterward Video db1 for  $\beta=0.5$

## 5. Conclusion

In this paper, an improved key frame extraction using discrete wavelet transform with modified threshold factor method with different wavelet basis db, sym and coif presented. The aim of this method is to reduce the redundant frames that can lead dimensionality reduction of feature vector classification of an isolated Arabic word for deaf and dumb peoples. According to the experimental results, key frames affected by the change values of threshold factor. As our future work, we will continue in our research for pattern recognition of sign language based on the resulted key frames from a used video clips.

## References

- [1] AS Jalal. Automatic Recognition of Dynamic Isolated Sign in Video For Indian Sign Language. 2015.
- [2] L Shurong, H Yuanyuan, H Zuojin, D Qun. Key Frame Detection Algorithm based on Dynamic Sign Language Video for the Non Specific Population. *International Journal of Signal Processing, Image Processing and Pattern Recognition*. 2015; 8: 135-148.
- [3] Y Shi, H Yang, M Gong, X Liu, Y Xia. A Fast and Robust Key Frame Extraction Method for Video Copyright Protection. *Journal of Electrical and Computer Engineering*. 2017; 2017.
- [4] NJ Janwe, KK Bhojar. Video Key-Frame Extraction using Unsupervised Clustering and Mutual Comparison. *International Journal of Image Processing (IJIP)*. 2016; 10: 73.
- [5] C Sheena, N Narayanan. Key-frame Extraction by Analysis of Histograms of Video Frames Using Statistical Methods. *Procedia Computer Science*. 2015; 70: 36-40.
- [6] A Makandar, D Mulimani, M Jevoor. Preprocessing Step-Review of Key Frame Extraction Techniques for Object Detection in video. 2015.
- [7] GI Rathod, DA Nikam. An algorithm for shot boundary detection and key frame extraction using histogram difference. *International Journal of Emerging Technology and Advanced Engineering*. 2013; 3: 155-163.
- [8] KT Tint, K Soe. Key frame extraction for video summarization using DWT wavelet statistics. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*. 2013; 2: 1829-1833.
- [9] A Nasreen, G Shobha. Key frame extraction using edge change ratio for shot segmentation. *International Journal of Advanced Research in Computer and Communication Engineering*. 2013; 2: 4421-4423.
- [10] P Mundur, Y Rao, Y Yesha. Keyframe-based video summarization using Delaunay clustering. *International Journal on Digital Libraries*. 2006; 6: 219-232.
- [11] G Liu, J Zhao. *Key frame extraction from MPEG video stream*. in Information Processing (ISIP), 2010 Third International Symposium on. 2010: 423-427.
- [12] T Liu, HJ Zhang, F Qi. A novel video key-frame-extraction algorithm based on perceived motion energy model. *IEEE transactions on circuits and systems for video technology*. 2003; 13: 1006-1013.
- [13] Q Zhang, G Liu. A key-frame-based error resilient coding scheme for video transmission over differentiated services networks. in *Packet Video*. 2007; 2007: 85-90.
- [14] K Khurana, M Chandak. Study of various video annotation techniques. *International Journal of Advanced Research in Computer and Communication Engineering*. 2013; 2: 909-914.
- [15] K Khurana, M Chandak. Key frame extraction methodology for video annotation. *International Journal of Computer Engineering and Technology*. 2013; 4: 221-228.