Corrupted MP4 Carving Using MP4-Karver

Ahmed Nur Elmi Abdi, Kamaruddin Malik Mohamad, Yusoof Mohammed Hasheem, Rashid Naseem, Jamaluddin,
Muhammad Aamir

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia
Batu Pahat, Malaysia

Abstract—**In the digital forensic, recovery of deleted and damaged video files play an important role in searching for the evidences. In this paper, MP4-Karver tool is proposed to recover and repair the corrupted videos. Moreover, MP4-Karver extracts frames from video for automatically screen-video to detect illegal cases instead of targeting or watching complete video. Therefore, many existing approaches such as Scalpel's method, Garfienkel, Bi-Fragment Gap Carving, Smart Carving and Frame Based Recovery attempts to recover the videos in different ways, but most of the recovered videos are usually not complete playable. The proposed MP4-Karver focuses on recovery of video files and repair corrupted videos to complete and playable. Experimental results show that the proposed MP4-Karver effectively restores corrupted or damaged video for an improved percentage of the video restoration compared with existing tools.**

Keywords—Digital forensic; File carving; Repairing; Corrupted; Frame extraction

I. INTRODUCTION

Recovery of corrupted or damaged video files, is playing a crucial part in digital forensics for detecting illegal cases. In the last decade, for video recording the digital devices with built-in cameras like surveillance camera, CCTVs and mobile cameras are increased enormously [1]. Similarly, the large amount of recorded video data has been made due to the common usage of these digital devices. Although digital devices have their own internal memory or external devices like external hard disk, flash memory and multimedia card.

Besides the increase digital devices and data storage, the numbers of users of these devices are also increased, according to Pew Research Center Globalization review in which 26 out of the 36 countries surveyed [2]. The video data files stored in the digital storages can be easily fabricated and modified, thus the role of digital forensics has increased to collect, analyze and process the digital evidences.

To recover the video data from digital devices, File carving is a powerful method used in digital forensic by taking advantage of the video file's structure. Studies on the file carving have targeted the restoration of different file types like pdf, zip, and image files[3]. But there are small number of methods on carving of the file type like video and audio files [4]. These techniques uses header and footer to recover the video files[5]. However, when video file structure is not present or corrupted (fragmented), then by using the existing techniques the recovering possibility of that file is below than 50% [6].

After retrieving video data, each video file is to be screened manually in order to find the presence of any illegal cases. The screening of each individual video a very stressful process and also a time consuming job, because a video file is a series of several still images called frames [7]. A few minutes of video might consist of thousands of video frames.

In this paper, MP4-Karver tool is proposed to recover and repair the corrupted MP4 videos and make it playable. Even if the MP4 video is produced from CCTV or any other device. The proposed tool performs better than the existing tools and techniques by having three main features. 1) Retrieving video. 2) Repairing corrupted video. 3) Frames extraction. All other tools have the ability to retrieve video but they are unable to repair the corrupted one. This tool can extract frames from video into image files, for analyzing video. The rest of the paper is organized as follows; Section 2 discuss the related work consists of an overview of digital forensic, file carving and existing method video recovery while Section 3 defines MP4 restoration and repairing framework. Section 4 presents experimental and result discussion. Finally, Section 6 concludes this paper.

II. RELATED WORK

Digital forensics is an improving method and commonly refers to the process of investigating digital files. For digital forensic, file carving is powerful method to recover data from unallocated space and it depends upon the file structure instead of file's metadata [8]. Following are the different video recovery techniques.

The most common techniques are Scalpel's method, Garfienkel method, Bi-fragment gap carving technique, smart carving and Frame based recovery. Scalpel technique proposed a popular tool that does not depend on file system or metadata of the file for recovering a video file[9]. In this technique, it is necessary to have an indexing step to find the file header and footer from a whole disk image. Similarly, a restoration step is required to restore indexed header and footer. This technique can be used to restore non-fragmented files and cannot recover partially overwritten video files.

Another technique that is used to recover video is Garfienkel method [10]. This technique can restore video files even if the header or footer is fragmented. An extension of the signature based file restoration is used to recover damaged file. Like Scalpel technique, this technique does not also use the metadata of the file system.

In this technique, additional file size and length information is added to the header of the video. Therefore, based on the length information, a video file can be restored even when fragmentation happens and the file footer is removed. However, this technique fails when the allocated memory to the file that is to be recover, is overwritten.

Bi-fragment Gap Carving [11] is used to restore the video files. In this technique, a combination of region containing file header and footer are found and analyses. The difference between header and footer regions are measured and tested. If the differences do not qualify the default validation procedure, it is required to repeat the process until the gap passes the validation check. Nevertheless, this technique can only be used in a video file with two fragments and the main limitation is when the gap between the two file fragments is large.

Smart carving [12] is a technique proposed to restore files without considering the number of fragments in the file. It identifies the fragmented parts of the file, and then it performs permutation and combination process on that file. In this process the frames order of video identified and then restored it back to the original video file to check whether the recovered video is meaningfully and playable or not. However, there is a drawback of this technique. When a part of allocated memory to the video is overwritten, smart carving fails to retrieve that video.

A new technique known as Frame based recovery of damaged video file using video codec specification is recently proposed [13]. In this method a video file is separated into frames where every frame has at least a meaningful unit of a video. Each unit of video frame is encoded by using a codec description. Similarly, each frame has decoding header information at the start and end of a frame. Therefore, in this technique video carving is done by using video frames and decoding header information. However, the frame based recovery method can repair frames of the non-overwritten video files and partial overwritten video files [14]. Moreover, this technique cannot recover completely over-written videos.

In this paper, MP4-Karver tool is proposed to carving video files and repair corrupted videos. Then extract frames from videos for digital forensic analyzing.

III. MP4-KARVER VIDEO METHOD

This section describes about the benchmark data sets and compared tools which are used to evaluate the performance of the proposed MP4-Karver against other video carving tools. Also, this section explains the methodology of the proposed tool for recovering and repairing MP4 videos.

A. Data sets

The standard data sets are used to evaluate in a digital forensic investigation are commonly bitwise-copies of full hard drives. Table 1 presents the statistics of three data sets that are used in the proposed tool.

- Digital Forensic Research Workshop challenge data set (DFRWS-2007-challenge.img).

- Baseline Carving Data Set (BCDS.raw).

- Level_3-(Video.dd) from computer forensic tool testing (CFT)

TABLE I. STATISTICS OF THE DATA SETS

Data sets	Size	Number of files
DFRWS.img (**BFRWS**)	330 MB	8 MP4
BCDS.raw (**BCDS**)	237 MB	3 MP4
Level_3- Video.dd (**Level**)	48.5 MB	1 MP4

B. MP4 Video carving tool

There is a wide selection of video carving tools available ranging from expensive proprietary forensic tools (blade forensic) to open source tool (PhotoRec). A total of 3 file carving tools are selected for testing the proposed tool and they are listed in Table 2. The basis for tool selection criteria included: MP4 video file support, advanced carving features and tool availability. Each tool listed has the associated license, tool version number and tool platform details.

TABLE II. PROPERTIES OF THE CARVING TOOLS

Name	License	Version	Platform
MP4 -karver	Proprietary	v.1.1	Windows
PhotoRec	Open source	v 7.1	Multi
blade forensic	Proprietary	v1.11	Windows

C. Implementation Overview

The main aim of this research is to develop MP4-Karver tool for the carving of MP4 videos, repairing corrupted videos and extracting frames from videos. C-Sharp (C#) programming language is used for developing the proposed MP4-Karver tool. The hardware setup for this tool was Intel® Core i7 CPU and 4GB of physical memory with windows 7.

Figure 1 presents the steps of proposed MP4-Karver tool for recovering and repairing videos. Disk images are computer files that comprise two types, the contents and structure of whole data in the storage devices. Reading MP4 video in the disk image files, by using streaming files of MP4 video and this contains hexadecimal value. Hexadecimal format pattern is used by searching offsets to find the location of all the headers of the video file signature.

For the recovery MP4 videos, the signature of MP4 video header is first to be identified. The file header is about 4bytes size shown Figure 2. In this figure, the offset of 4 bytes (hex: 0x 66, 74, 79, 70) ftyp header signature of the MP4 video can be seen.

After verifying the header files of MP4 video using that pattern signature, it is possible to extract MP4 video data. Then the videos are recovered and the corrupted videos are passed to the second part of the method to be repaired. The repairing corrupted MP4 video files can be solved by taking advantage the structure of the video.
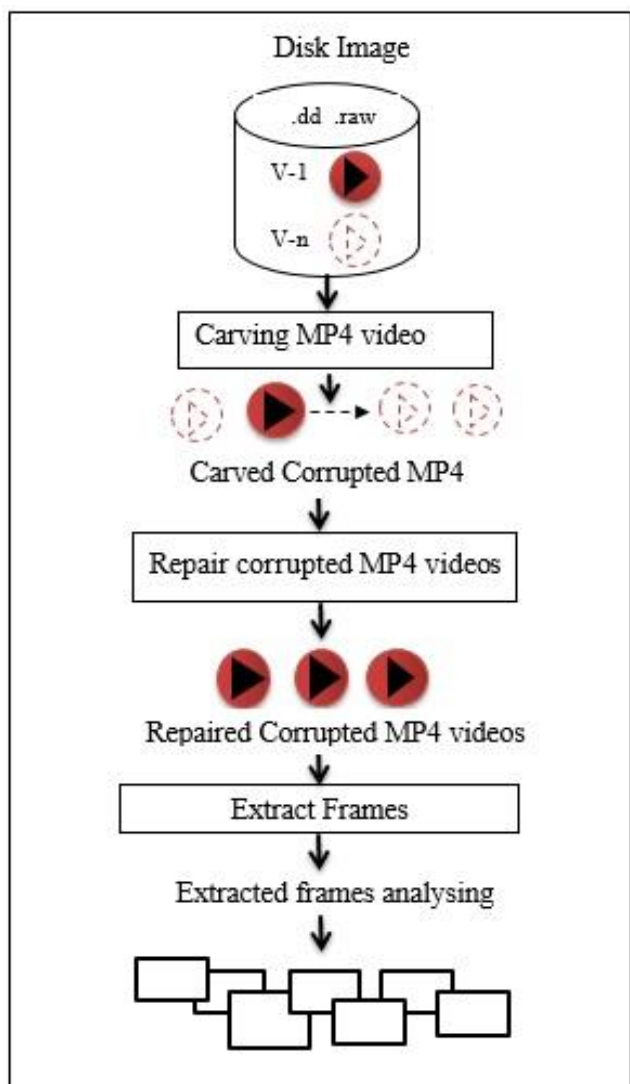
Fig. 1.   Methodology of Proposed MP4-Karver Tool
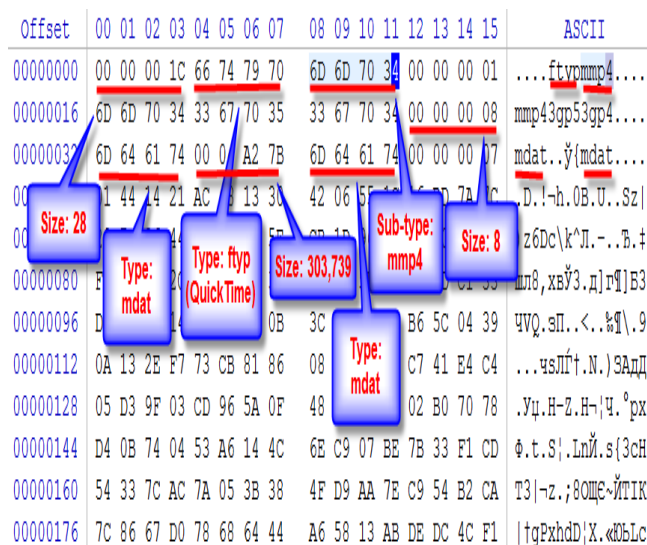

Fig. 2.   MP4   Video Signature

Figure 3 shows a proper MP4 -format file structure. It begins with file type box (ftyp) and Movie box (moov), followed by a Mdat (Media Data Box). Normally, this 'mdat' would contain raw video data (H.264 NAL units). Also, this figure shows the corrupt video file which looks a like the normal structure however, it contains the 'ftyp' and 'mdat' atoms only. Therefore, the file can be repaired simply by skipping over the initial 'ftyp' and 'moov' atoms, and the header for the 'mdat' atom, and copying the remaining data into a new file. Now, the videos are fully recovered and repaired.

MP4 video after recovery and repaired videos then are extracted the video frames for automatically screening. A video file is actually a combination of several still images which are called frames. Extracting frame starts with reading number frames of video contains which are still images and can be captured. Then, captured frames are converted into image formats jpg.
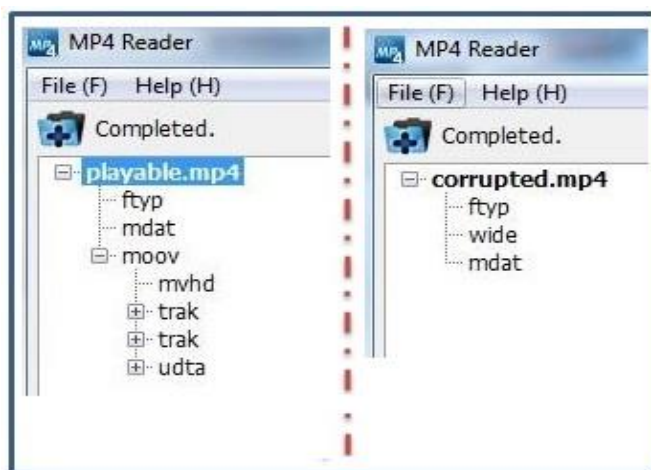

Fig. 3.   Comparing Normal and Corrupted MP4 Structure

IV.   EXPERIMENT AND RESULTS

To evaluate the performance of the proposed MP4-Karver tool for video carving and repairing corrupted videos, we use three data sets, i.e. 1) DFRWS 2) BCDS and 3) Level. These data sets contain corrupted and partial videos, which can be seen in Table 1. The experimental results of the total video carved by the proposed tool and the existing tools are given in Table 3.

TABLE III.   TOTAL MP4 VIDEO CARVED

DATA SETS	Total Videos	File Carving Tools	Total Carved
DFWRS2007.img	8	MP4-karver	**8**
		Photorec	**8**
		Blade forensic	1
bcds.raw	3	MP4-karver	3
		Photorec	2
		Blade forensic	2
L3_Video.dd	1	MP4-karver	1
		Photorec	1
		Blade forensic	1

We tested the tools and the results can be seen in Table 3. It can be seen that for the BFRWS data set, 100% videos are

recovered by MP4-Karver and Photorec tools. The Blade Forensic tool recovers only one video. For BCDS data set in Table 3, it shows that, total of three videos are recovered by MP4-Karver tool. The Blade Forensic and Photorec tools recover only two videos. However, Photorec recovered and Blade Forensic tools, two videos recovered are playable. Whereas Level data set, in Table 3 one video is recovered by the MP4-Karver tool. The Blade Forensic and Photorec tools also one video is recovered.

After the recovery of the videos, they may be playable or corrupted. According to table 4, In DFWRS dataset the recovered videos after repairing by MP4-Karver are all playable while Photorec recovered only one playable. On the other side Blade Forensic recovered only one video and that is playableOnce the videos are recovered, they could be playable or corrupted. The recovered videos by MP4-Karver is playable while Photorec and Blade Forensic recovered video is not playable. The rest videos of video recovered are partial videos mean not complete and playable video. BCDS data set, three videos are recovered through using MP4-Karver tool and 2 video is playable and one is partial is not playable. In the Level data set, one video is recovered by Photorec and Blade Forensic tools and that video is partial not playable.

TABLE IV. REPAIRED AND PLAYABLE

DATA SETS	File Carving Tools	Playable-Complete	Not playable	Partial Video
DFWRS2007.img	MP4-karver	**8**	0	0
	Photorec	1	7	0
	Blade forensic	1	7	0
bcds.raw	MP4-karver	2	0	1
	Photorec	2	0	0
	Blade forensic	2	0	0
L3_Video.dd	MP4-karver	**1**	0	0
	Photorec	0	0	1
	Blade forensic	0	0	1

The percentage improvement of the MP4-Karver over existing tool is given in Table 5. The percentage improvement is calculated as follows:

$$P_{im}MK(E) = \frac{VCE - VCM}{|Videos\ in\ a\ Data\ Set|} * 100 \quad (1)$$

Where $P_{im}MK$ is the **p**ercentage **im**provement of **MP4-Karver** over **e**xisting (**E**) tool. *VCE* is the **v**ideo **c**arved by **e**xisting tools and *VCM is* the **v**ideo **c**arved by **MP4-Karver** tool. The obtained value form VCE-VCM is divided by the total number of videos in a data set. To get the percentage 100 is multiplied. In next step the average of percentage improvement of MP4-Karver over existing tools is calculated by using the equation:

$$AP_{im}MK(E) = \frac{\sum_{i=0}^{|Data\ sets|} P_{im}MK(E)_i}{|Data\ Sets|} \quad (2)$$

The Equation 2 sums up all the values of Equation 1 obtained for all the data sets and divided it by the total number of the data sets.

TABLE V. AVARAGE IMPROVEMENT OF MP4-KARVER TOOL OVER EXISTING TOOLS

DATA SETS	File Carving Tools	Percentage Improvement
DFWRS2007.img	Photoric	88%
	Blade forensic	88%
bcds.raw	Photoric	0%
	Blade forensic	0%
L3_Video.dd	Photoric	100%
	Blade forensic	100%
Average	-	63%

Subsequently, the calculation of average improvement of MP4-Karver tool as compared to the existing tools. The Table 3 & 4 results obtained from it also used to view and presented a graphical way. Figure 4 shows the total of carved videos from data sets and by comparing the file carving tools. Also, Figure 5 demonstrates the result of repairing corrupted videos from a data set by using carving tools.

Next part highlights the results of the last step of the MP4-Karver video method of extraction frames. In Table 6 and Figure 6, the number of extracting frames from video recovered is presented, MP4-Karver tool is extracted 3522 frames from BFRWS data set where from BCDS data set extracted 19314 frames and the Level data set, and 262 frames are extracted. The Photorec and Blade Forensic tools they do not have the feature to extract frames from videos.

TABLE VI. RESULT OF VIDEO FRAMES EXTRACTED

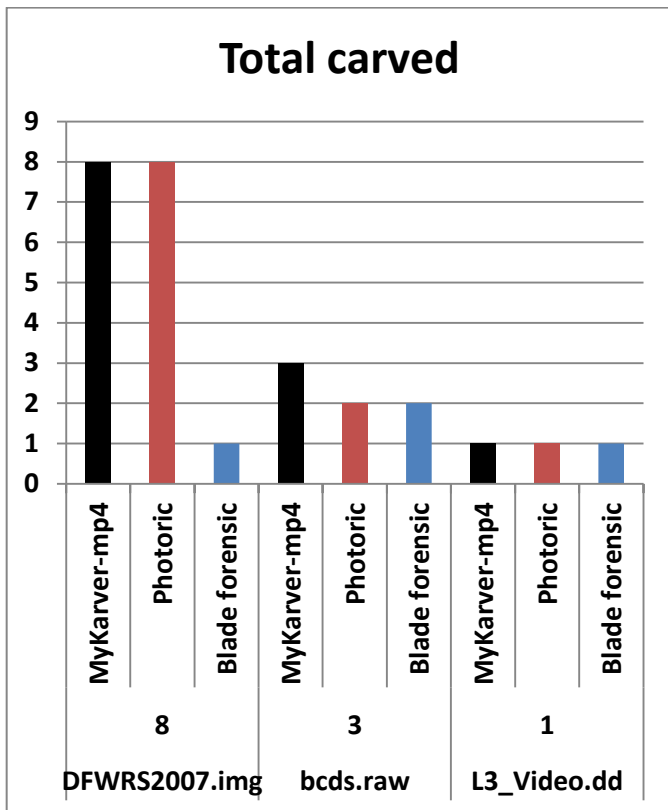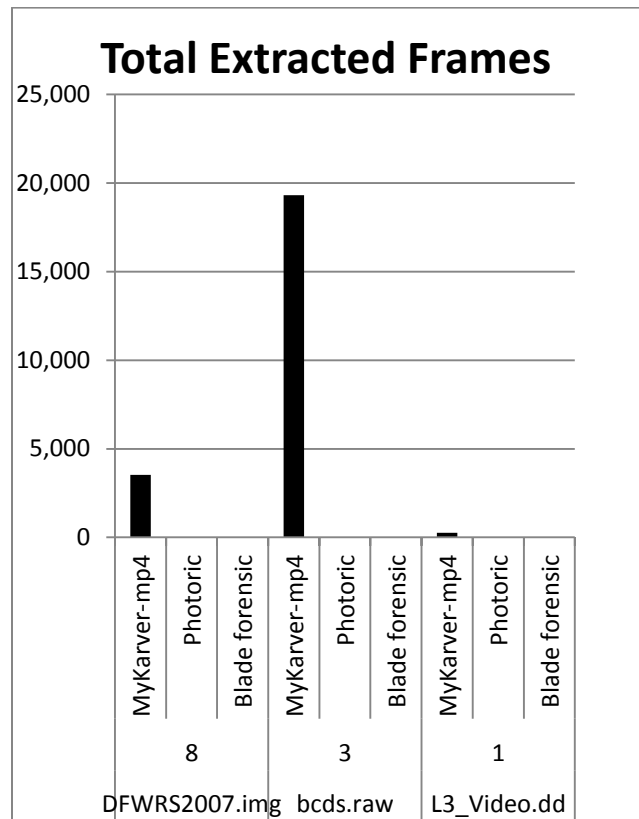| DATA SETS	Total Videos	File Carving Tools	Total Extracted Frames
DFWRS2007.img	8	MP4-Karver	**3,532**
		photorec	**0**
		Blade forensic	**0**
bcds.raw	3	MP4-Karver	**19,313**
		photorec	**0**
		Blade forensic	**0**
L3_Video.dd	1	MP4-Karver	**262**
		photorec	**0**
		Blade forensic	**0**

Fig. 4.   Total video Carved

Fig. 6.   Graphical representation of MP4 video frames
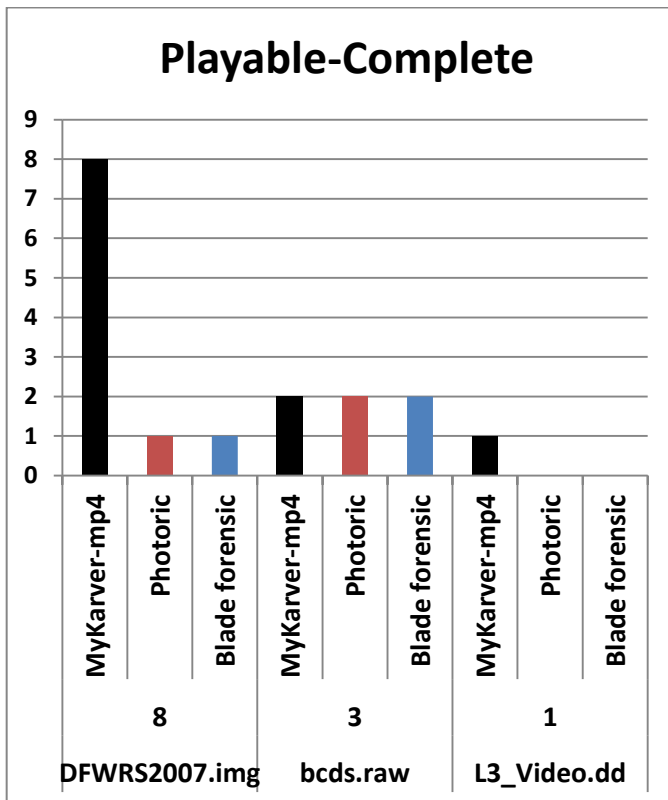
V.   CONCLUSION

This paper proposes a new MP4-Karver tool for video carving and repairing corrupted videos. The proposed MP4-Karver tool guarantees the efficient video restoration and repairing if it's corrupted or damaged and making complete playable videos and then frames can be extracted from videos. Moreover, MP4-Karver tool has the ability of restoring videos, if these are corrupted and overwritten or non-overwritten video files because many existing file-based techniques such as, Scalpel's method, Garfienkel, Bit fragment gap carving, smart carving and Frame based recovery cannot restore overwritten video files.

Most of the previous methods have been used to recover video files by searching for Metadata of the files, but the proposed tool recovers video file by taking advantage of the file structure and repairing corrupted video files without being affected by the corrupted parts. It is evaluated that the video restoration and repairing rate of MP4-Karver tool is higher than existing tools, i.e. Photorec and Blade forensic.

This research recommends doing further research on this tool and will add automatic video analysis for illegal activities detection such as child sexual graphic and also add enhancement, clarification and enlargement of the video in order to increase the filtering of evidence correctly to solve a criminal case.

Fig. 5.   Total video repaired

REFERENCES

[1] Deshmukh, A., & Desai, M. "A Survey on Corrupted Video Recovery Using CODEC Specifications", International Journal of Science and Research (IJSR), Volume 3 Issue 12, December 2014.

[2] G. H. Na, K. S. Shim, K. W. Moon, S. G. Kong, E. S. Kim, and J. Lee, "Frame-based recovery of corrupted video files using video codec specifications," IEEE Trans. Image Process., vol. 23, no. 2, pp. 517–526, 2014.

[3] S. L. Garfinkel, "Carving contiguous and fragmented files with fast object validation," Digit. Investig., vol. 4, no. SUPPL., pp. 2–12, 2007.

[4] B. Yoo, J. Park, and S. Lim, "A study on multimedia file carving method," 2011.

[5] S. Hand, Z. Lin, G. Gu, and B. Thuraisingham, "Bin-Carver: Automatic recovery of binary executable files," Digit. Investig., vol. 9, pp. S108–S117, 2012.

[6] V. Shoba and K. Gunasekaran, "An Integrated Approach for Recovering Corrupted Video Files using MPEG-4Visual Specifications," vol. 3, no. 1, pp. 249–254, 2015.

[7] Yannikos and Ashraf, "Automating Video File Carving and Content Identification," Adv. Digit. Forensics IX, pp. 195–212, 2013.

[8] K. M. Mohamad, T. Herawan, and M. M. Deris, "Detecting JFIF Header using FORHEADER 1," vol. 5, no. 4, pp. 17–26, 2011.

[9] S. Garfinkel, A. Nelson, D. White, and V. Roussev, "Using purpose-built functions and block hashes to enable small block and sub-file forensics," Digit. Investig., vol. 7, pp. S13–S23, 2010.

[10] W. Qiu, R. Zhu, J. Guo, X. Tang, B. Liu, and Z. Huang, "A New Approach to Multimedia Files Carving," 2014 IEEE Int. Conf. Bioinforma. Bioeng., pp. 105–110, 2014.

[11] Na, G. H., Shim, K. S., Moon, K. W., Kong, S. G., Kim, E. S., & Lee, J. (2014). "Frame-based recovery of corrupted video files using video codec specifications". Image Processing, IEEE Transactions on, 23(2), 517-526.

[12] Metz, J., & Mora, R. (2007). Analysis of 2006 DFRWS forensic carving challenge. 2007-8-28). http://sandbox. dfrws, org/2OO7/metz/dfrws20OT_carving_challenge, pdf.

[13] Thakre, C., Thakare, M., & Dhande, S. A. "Efficient Design of Frame Based Recovery Technique for Corrupted Video Files". INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING Vol. 3, Issue 6, June 2015

[14] Avinash Deshmukh, Manisha Desai. "A Survey on Corrupted Video Recovery Using CODEC Specifications". Volume 3 Issue 12, December 2014.