International Journal of Cyber-Security and Digital Forensics (IJCSDF) 10(1): 10-16
The Society of Digital Information and Wireless Communications (SDIWC), 2021 ISSN: 2305-0011

A Proposed System Design of Deleted Image Metadata Recovery Software

Tabu S. Kondo
Department of Computer Science and Engineering, The University of Dodoma, Dodoma, Tanzania
chifulukwe@gmail.com

ABSTRACT

This paper describes the design and implementation features of Deleted Image Metadata Recovery Software which can automatically extract deleted image metadata so that important information of the senders of offensive image on the web can be identified. By so doing, the identified information will be used as evidence in a court of law. It is hoped that the software can be used by forensics professionals to recover deleted image metadata from the digital image. The proposed design of the Deleted Image Metadata Recovery Software is expected to use the algorithm for extracting deleted image metadata. Therefore, the proposed software will be able to analyze the authenticity of the digital image downloaded from the web and produce a report that can be utilized as evidence in a court of law.

KEYWORDS

Deleted image metadata, Image forensics, Digital image forensics, Metadata recovery software, Metadata recovery tool

1 INTRODUCTION

A basic advance in any crime scene investigation examination includes gathering data and deciding its reliability and how this data can include measurable incentive as far as choice that can be surmised. By and large, it includes data, for example, who did what, when and where. First, any collected data which is going to be used for forensic investigation needs to be reliable in terms of integrity and authenticity. It is clear that security services such as integrity and authentication are important. Then, there is the issue of metadata related with the information. Computer forensics has since a long time ago comprehended the significance of metadata. Metadata can be incredibly helpful in noting a portion of the fundamental inquiries of a computer forensic investigation, for example, who planned something for a record, when they did it and where it was finished. In a computer forensic investigation the accumulated data is utilized to break down the occasions that are the subject of an investigation [1].

As per [2], metadata is data that causes us use and comprehend other data. All the more especially, metadata is proof, normally put away electronically, that portrays the attributes, roots, use, structure, change and legitimacy of other electronic proof. Numerous occurrences of metadata in numerous structures happen in numerous areas inside and without computerized records. Some is supplied by the user, but most metadata is generated by systems and software. Some is essential proof and some is simply advanced mess. Valuing the distinction-realizing what metadata exists and understanding its evidentiary criticalness - are aptitudes fundamental to electronic revelation. Nonetheless, if proof is whatever will in general demonstrate or discredit a statement as truth, at that point plainly metadata is proof. Metadata reveals insight into the causes, setting, legitimacy, unwavering quality and dispersion of electronic proof, just as gives hints to human conduct. It's what could be compared to DNA, ballistics and unique mark proof, with an equivalent capacity to absolve and implicate.

Currently, media technologies have modified the responsibility of human beings from passive content consumers to active content creators/producers. Many metadata such as user-provided tags, comments, geo-tags, capture time and EXIF (exchangeable image file format) information, associated to multimedia resources, are available in the social media websites like Flickr and YouTube. The explosive growth of social multimedia content on the Internet is revolutionizing the way of content distribution and social interaction [3]. As such, it is normal to find offensive image in various online social networks these days. Offensive image are those that create a negative impact on a large group of people after viewing the content. To solve this problem, various techniques, algorithms and tools for extracting metadata of the image have

been devised so as to extract metadata that can be used as evidence in a court of law. Examples of these tools are Exiv2 and Exif viewer. However, when the image files are posted on websites or uploaded to social networking sites, the metadata is often removed so as to protect the privacy and reduce the size. This causes the use of the available techniques, algorithms and tools to fail as they works well with hidden metadata and not the removed metadata. To close this gap, there is a need to devise techniques, algorithms and tools to automatically extract deleted image metadata so that important information of the senders of offensive image on the web can be identified and used as evidence in a court of law.

2 PROBLEM STATEMENT

It is perceived that in an inconclusively huge asset space, powerful administration of organized data will progressively depend on successful administration of metadata. The requirement for metadata administrations is now clear in the present Web condition. Metadata isn't just key to disclosure, it will likewise be basic to compelling utilization of discovered assets (by building up the specialized or business structures in which they can be utilized) and to interoperability across convention areas [4].

Metadata accessible from records gives an objective zone to examiners to focus on to get data about the document or perhaps in any event, figuring out who possesses the document, or has been in contact with the document.

For intelligence and law enforcement communities, examining metadata can potentially save substantial time frequently tied to manual analysis [5]. In a digital forensics context, metadata available for examination frequently includes the file creator, creation times, modification times, titles, and number of revisions [5].

One of the basic features of metadata is that it is attached to the file, and is therefore independent of the file system. When the file (such as an image) is copied from system to system the metadata remains with the file. However, when the file may be uploaded to social networking sites such as Facebook, whatsapp, twitter, instagram, and Flicker things are different. That is, the metadata of the file is often lost when the files (such as photos, videos and audios) are posted on websites or uploaded to social media networks [6]. This causes the available algorithms and tools that work well with the hidden metadata to fail in working with the deleted metadata. As such, new techniques, algorithms and tools that will work well with the deleted metadata are highly needed so as to close this gap. This study therefore, aimed at proposing the design of deleted image metadata recovery software that can recover deleted or striped image metadata in a digital photography. The software is expected to enable its users to recover and view image metadata which has been deleted intentionally or unintentionally from the digital photography. As a result, image forensics professionals will be able to recover the deleted metadata and use them as evidence in a court of law.

3 LITERATURE REVIEW

3.1 Image Metadata Formats

Image metadata format is the standard protocols and techniques utilized to store image metadata within an image file. Several standardized formats of metadata exist, including: IPTC-IIM, Extensible Metadata Platform (XMP), Exchangeable Image File (EXIF), JPEG File Interchange Format (JFIF), Composite, ICC Profile, APP14, and Picture Licensing Universal System (PLUS) [7].

3.1.1 EXIF

EXIF represents Exchangeable Image File Format. The determination was initially delivered by the Japan Electronic Industry Development Association (JEIDA) which is presently the Japan Electronics and Information Technology Industries Association (JEITA), and spreads picture and sound documents utilized in advanced cameras. Since 2009, the standard has been mutually figured by JEITA and the Camera and Imaging Products Association (CIPA). The most recent rendition is EXIF 2.3 [8]. The standard depends on the TIFF document

structure and the JPEG pressure codec yet just as including the first TIFF metadata labels; it includes a wide scope of new ones identifying with different things important to picture takers and others managing the computerized media including ones for GPS. The two associations likewise mutually planned the Design rule for Camera File framework (i.e. Digital Camera Format (DCF)), the present form being 2.0. The standard characterizes the DCF essential fundamental picture as an EXIF/JPEG document and sets out point by point controls on how this is to be executed including a scope of required labels.

3.1.2 XMP

Extensible Metadata Platform (XMP) was first propelled by Adobe in 2001 as an endeavor to solidify the different metadata plans into a XML structure. In spite of the fact that XMP was initially a restrictive norm, it turned into an ISO standard [9] in 2012 and is successfully now open source. XMP has been intended to be applied to a wide scope of record types and offers a scope of outlines including the Dublin Center. In spite of the fact that EXIF/JPEG is the obligatory picture group for advanced cameras [10], XMP has an EXIF pattern and this regularly used to duplicate EXIF information to XMP when a document from a computerized camera is altered in a product picture preparing application.

3.1.3 IPTC-IIM

The International Press Telecommunications Council (IPTC) built up the IPTC-IIM (Data Trade Model) in the mid 1990s as a push to sort out, systematize and bring together the manner in which data was put away and shipped with pictures among news offices, photographic artists, photograph offices, libraries, exhibition halls, and other related businesses [11]. After the arrival of XMP by Adobe in 2001, IPTC worked with Adobe to build up a specialized execution of IPTC metadata in XMP position and that prompted the issue of IPTC Center Construction for XMP in 2005. This was extended by the IPTC Augmentation metadata construction in 2008. The present rendition of the two patterns is 1.2 [12].

3.1.4 PLUS

The Picture Licensing Universal System (PLUS) is a universal licensing language promoting the clear communication of image rights in all industries [13]. PLUS aims to simplify and facilitate the communication and management of image rights through its use of machine-readable coding and standardized language [14].

3.2 Digital Image Metadata Tools

Viewing digital image metadata requires extracting the information from the file. There are plenty of digital image metadata tools available. Some of these only support one file type (e.g., JPEG-only), while others support many file formats. A few examples of available digital image metadata tools are JPEGsnoop, Jeffrey's Exif Viewer, and Exiv2.

3.2.1 A Brief Description of JPEGsnoop

JPEGsnoop [15] is a detailed JPEG image decoder and analysis tool. It reports all image metadata and can even help identify if an image has been edited. JPEGsnoop is a free Windows application that examines and decodes the inner details of JPEG, MotionJPEG AVI and Photoshop files. It can also be used to analyze the source of an image to test its authenticity. It is a fake image detector via image signature analysis in an offline mode.

3.2.2 A Brief Description of Jeffrey's Exif Viewer

Jeffrey's Exif Viewer [16] is a free online Exif viewer that lets the user to view Exif data of images online. Any image from the computer can be uploaded in Exif Viewer, or through specification of URL of an image. Exif generates data which is metadata that is stored with each image. Exif Viewer basically tells various characteristics of the image such as; camera from which it was taken, file attributes, geolocation, date and time information, camera settings, and much more.

3.2.3 A Brief Description of Exiv2

Exiv2 [17] is a C++ library and a command-line utility to administer digital image metadata. This command-line utility is provided as an executable for Windows or source code for Linux and Mac. It provides fast and easy read and write access to the EXIF, IPTC and XMP metadata of digital images in various formats. It is available as free software and with a commercial license, and is used in many projects.

4 METHODS

In this research, we are interested in the technique, algorithm and tool for extracting deleted image metadata in social media networks and use the obtained information as evidence in a court of law. Therefore, systems development life cycle was used to propose the design of the required tool. As such, requirements analysis was carried out after the literature review in which the system was investigated. The requirements were then analyzed so as to propose the required design of the tool.

5 RESULTS

5.1 Functional Requirements

(i). The software shall be able to process the digital photography without altering it.
(ii). The software shall extract deleted EXIF, XMP and IPTC metadata stored in each digital image.
(iii). The software shall be used in various platforms in order to interoperate with technology evolution
(iv). The software shall provide offline and online modes to promote the use of metadata systems widely.
(v). The software shall provide the function of a report generator in order to generate related reports or statistic data based on specified criteria from metadata elements.

5.2 Nonfunctional Requirements

Usability
The software shall be provided with manual written in such a way that it is very easy to setup, install and run.
Performance
The software should be quick in retrieving the photograph file from the disk and display an output so as to minimize the time cost
Portability
The software shall work on most desktop operating systems and/or interfaced with other systems or products, as a library.
Reusability
The software's modular design shall allow other developers (and the community) extend the features of the software to improve the way it works or add new features.
Extensibility
Together with the software's modular design, the software's source code shall be provided with manuals to enable new (or other) developers extend the features of the project.
Repeatability
The software shall produce the same results when extracting deleted image metadata on the same digital image by the same operator
Reproducibility
The software shall produce the same results when extracting deleted image metadata on the same digital image in different laboratories with different operators.

5.3 The Proposed System Design

The proposed System design aims to address the gap in the current surveyed tools by implementing new functionalities capable of recovering deleted or striped image metadata. The tool will grow from the Exiv2's code. Figure 1 shows the architectural design of the proposed deleted image metadata recovery software and figure 2 shows the operations of Deleted Image Metadata Engine. Furthermore, figure 3 shows the basic flowchart of the proposed system. Generally, the proposed design of the tool will be able to work with tampered image and gives result in terms of true detection (presence of tampering) and false detection (no presence of

tampering).

The architectural design of the proposed deleted image metadata recovery software is shown in figure 1. Generally, the software through its Graphical User Interface (GUI) will accept digital image as an input, then it will process the image through its Deleted Image Metadata Engine. The Engine depends on Exiv2 for its operation and the Exiv2 is built on C++ code.

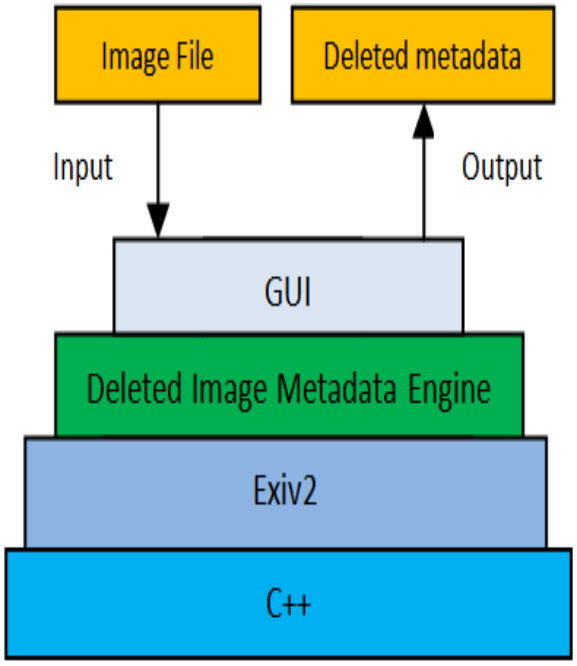

Figure 1. The Architectural Design of Deleted Image Metadata Recovery Software

The Deleted Image Metadata Engine of the software works as follows: The Image file is being accepted as an input through GUI of the software, and then the deleted image metadata are being extracted by using an appropriate Deleted Image Metadata Algorithm. As a result, the obtained deleted image metadata (if any) is saved in a storage device for further processing such as displaying or printing. In addition to that, the processed file is also saved in a storage device without alteration. Figure 2 depicts the general operations of the Deleted Image Metadata Engine and figure 3 shows the basic flowchart of the proposed system.

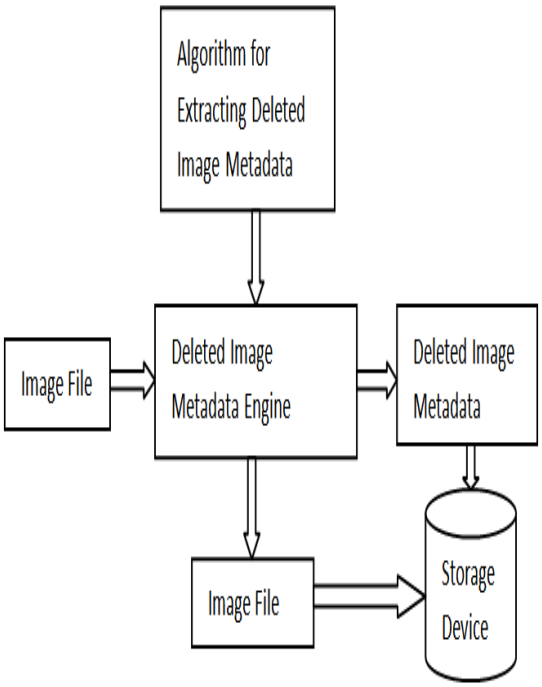

Figure 2. Operations of Deleted Image Metadata Engine

In order to make a detailed analysis of the deleted image metadata, the following steps are performed by the algorithm for extracting deleted digital image metadata.

Step 0: Start;
Step 1: Read the digital image file;
Step 2: Check for the digital image inputted in step 1 for deleted digital image metadata;
Step 3: Extract all deleted digital image metadata if exist;
Step 4: Identify the types of the existing deleted image metadata;
Step 5: Store the deleted image metadata in a file or display it on the screen;
Step 6: Maintain the original image file;
Step 7: End.

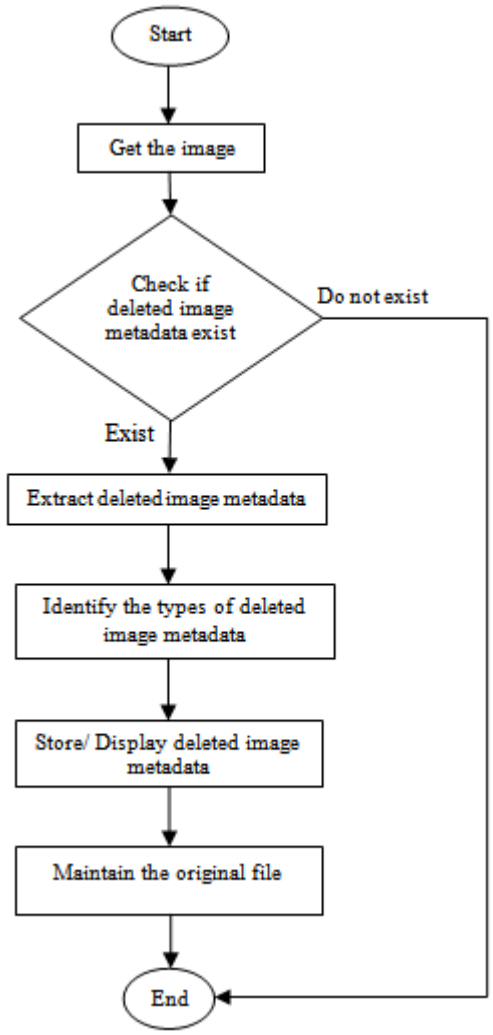

Figure 3. Basic flowchart of the proposed system

6 CONCLUSION

It is important to note that whether the metadata of the digital photography have been deleted intentionally or unintentionally, they can be recovered or restored back. Since those metadata on the digital photographs are not totally deleted or removed from it, instead they are likely to be hidden strongly. This process of hiding metadata of the digital photography is done mostly using special software. Therefore, special tools for recovering or restoring the deleted metadata from the digital photograph should be designed and developed. This study therefore proposes the design of deleted image metadata recovery software which can automatically extract deleted image metadata so that important information of the senders of offensive image on the web can be identified and used as evidence in a court of law.

It is important to note that an image may be offensive in one culture and nice in another culture. Therefore, the decision of whether an image is offensive or not will depend on the cyber laws of the nation.

Basically, the analysis of image metadata in digital investigation is considered important part to prove the offender and allows investigators to understand the timeline of a crime and the interpretation of the events. Metadata can also be used to prove the authenticity of digital photographs. Therefore, through careful inspection of the deleted image metadata by using the proposed design of the tool, it can be told whether an image has been tampered or not.

Future work will focus on coding the deleted image metadata recovery software.

REFERENCES

1. Salama, U., Varadharajan, V. and Hitchens, M (2012), Metadata Based Forensic Analysis of Digital Information in the Web, ANNUAL SYMPOSIUM ON INFORMATION ASSURANCE & SECURE KNOWLEDGE MANAGEMENT, JUNE 5-6, ALBANY, NY.
2. Ball, C. (2011), Beyond Data about Data: The Litigator's Guide to Metadata.
3. Kanellopoulos, D. (2015), Multimedia Social Networks, http://www.irma-international.org/viewtitle/113137/.
4. Sutton, S. A. (1999), Conceptual Design and Deployment of a Metadata Framework for Educational Resources on the Internet.
5. Migletz, J. (June 2008), AUTOMATED METADATA EXTRACTION, MASTER OF SCIENCE IN COMPUTER SCIENCE, NAVAL POSTGRADUATE SCHOOL.
6. Ashenfelder, M. (April 11, 2013), Social Media Networks Stripping Data from Your Digital Photos, [https://blogs.loc.gov/thesignal/2013/04/social-media-networks-stripping-data-from-your-digital-photos/], Site visited on 7/10/2017.
7. https://docs.oracle.com/cd/B28359_01/appdev.111/b28415/ch_metadata.htm.
8. CIPA, JEITA (2012) 'Exchangeable image file format for digital still cameras: Exif Version 2.3', http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf.
9. ISO (2012) Adobe Extensible Metadata Platform (XMP) Becomes an ISO Standard, http://www.iso.org/iso/home/news_index/news_archive/news.htm?refid=Ref1525.
10. CIPA, JEITA (2010) 'Design rule for Camera File system (DCF)', http://www.cipa.jp/std/documents/e/DC-009-2010_E.pdf.
11. IPTC (2014a) IPTC Photo Metadata Standard [online], available: https://iptc.org/standards/photo-metadata/iptc-standard/.

12. IPTC (2014b) 'IPTC Photo Metadata: Core 1.2/Extension 1.2' , [https://www.iptc.org/std/photometadata/specification /IPTC-PhotoMetadata.
13. https://www.iptc.org/std/photometadata/0.0/documentation/IPTC-PhotoMetadataWhitePaper2007_11.pdf, site visited on 11/2/2021.
14. http://cepic.org/app/uploads/2014/12/ITI_Metadata_Made_ Simple.pdf, site visited on 11/2/2021.
15. http://sourceforge.net/projects/jpegsnoop/, site visited on 10/12/2017.
16. http://regex.info/exif.cgi, site visited on 11/12/2017.
17. http://www.exiv2.org/, site visited on 5/12/2017.