Análisis de metadatos en vídeos digitales de dispositivos móviles

Carlos Quinto Huamán, Esteban Alejandro Armas Vega, Ana Lucila Sandoval Orozco, Luis Javier García Villalba

{cquinto,esarmas}@ucm.es, {asandoval, javiergv}@fdi.ucm.es

Grupo de Análisis, Seguridad y Sistemas (GASS)
Departamento de Ingeniería del Software e Inteligencia Artificial
Facultad de Informática, Despacho 431, Universidad Complutense de Madrid (UCM)
Calle Profesor José García Santesmases, 9, Ciudad Universitaria, 28040 Madrid, España

Resumen: *Actualmente la calidad, prestaciones y bajo coste de las cámaras fotográficas de dispositivos móviles hacen que éstas sean de uso común. Por tanto, el volumen de venta de dispositivos móviles con cámara fotográfica crece a un ritmo imparable desbancando poco a poco a las cámaras fotográficas digitales. Este escenario hace que el análisis forense de este tipo de vídeos cobre especial importancia y sea necesario y útil en multitud de situaciones (pruebas en casos judiciales, espionaje industrial, privación de la libertad de prensa, pederastia, etc). La mayoría de estos dispositivos insertan metadatos en el proceso de adquisición de la imagen y son de gran ayuda para una gran variedad de técnicas de análisis forense. Teniendo todo esto en cuenta, se estima necesaria la existencia de herramientas eficaces y robustas, que permitan la extracción de los metadatos de una forma veraz y consistente. En este trabajo se realiza un análisis detallado de vídeos con formato MP4 de dispositivos móviles. El análisis se realiza utilizando una herramienta personalizada para extraer la información almacenada en los átomos de los archivos de vídeo de 7 modelos de dispositivos móviles. Los resultados del análisis muestran considerables diferencias en los algoritmos de compresión de audio y vídeo, los parámetros de captura y en la estructura interna del archivo.*

Palabras clave: Análisis Forense, Cámara de Dispositivo Móvil, H.264, Metadatos.

Abstract: *Currently, the quality, performance and low cost of cameras in mobile devices make these are in common use. Therefore, the sales volume of mobile devices with camera grows at an unstoppable rate gradually supplanting digital cameras. This scenario causes the forensic analysis of such videos is particularly important and necessary and useful in many situations (evidence in court cases, industrial espionage, deprivation of freedom of the press, child abuse, etc.). Most of these devices insert metadata in the process of image acquisition. These are a great help for a variety of forensic analysis techniques. All this creates the need to have effective and robust tools that allow the extraction of the metadata of an accurate and consistent manner. This paper presents a detailed MP4 videos format analysis. The analysis is performed using a developed tool to extract the information stored in the atoms from the video files of 7 models of mobile devices. The analysis results show considerable differences in audio and video compression algorithm, parameters of capture and internal file structure.*

Keywords: Forensics, Mobile Device Camera, H.264, Metadata.

1 Introducción

En los países industrializados el 97% de teléfonos móviles incorpora una cámara digital integrada. Asimismo, la mayor parte del resto de tipos de dispositivos móviles también posee una cámara digital integrada. Estas cámaras, a diferencia de las cámaras digitales tradicionales (del inglés *Digital Still Camera* (DSC)), son llevadas por sus dueños gran parte del tiempo a la mayoría de los lugares a los que asisten [1]. En el año 2016 la venta de DSCs descenderá de un 47% de cuota de mercado sobre el total de cámaras digitales que obtuvo en el 2012 a un 27%. Asimismo, se prevé un incremento en las ventas de cámaras digitales integradas en teléfonos móviles, PC y tabletas, de un 31% de cuota de mercado sobre el total de cámaras digitales en 2012, a un 42% en el 2016 [2].

Prácticamente la totalidad de estas cámaras digitales tienen funciones de grabación de vídeo. Actualmente, existe una gran competencia entre los fabricantes por integrar una videocámara de alta definición al alcance del usuario en todo momento. Como consecuencia, y dada la gran cantidad de tiempo que una persona pasa junto a los dispositivos móviles, éstos se han convertido para muchas personas en el primer dispositivo de captura de fotografías y grabación de vídeos. Diariamente pueden verse imágenes y vídeos generados por dispositivos móviles en telenoticias, distintas aplicaciones, correo electrónico o en redes sociales. Todo esto hace que en ciertos casos existan restricciones legales o limitaciones a su utilización en distintos lugares, tales como: colegios, universidades, oficinas de gobierno, empresas, etc. Además, y como consecuencia de todo lo anterior, cada día las imágenes digitales generadas con dispositivos móviles son más utilizadas como testigos silenciosos en procesos judiciales (pornografía infantil, violencia callejera, redes sociales, etc.), siendo piezas cruciales de la evidencia de un crimen [3].

Por todas estas razones el análisis forense de imágenes y vídeos digitales de dispositivos móviles cobra especial fuerza en la actualidad. El estudio debe ser concreto para este tipo de dispositivos, ya que poseen características específicas que permiten obtener mejores resultados, no siendo válidas las técnicas forenses para imágenes y vídeos digitales generadas por otros tipos de dispositivos. En [4] se describe de forma clara y razonada la necesidad de técnicas de análisis forense específicas para dispositivos móviles.

Este trabajo está estructurado en 5 secciones, siendo el

primero la presente introducción. En la sección 2 se realiza un estado del arte del análisis forense para imágenes y vídeos generados por dispositivos móviles digitales haciendo énfasis en las técnicas basadas en metadatos. La sección 3se realiza una descripción de los principales sistemas de metadatos en imágenes y vídeos dando especial importancia a la especificación del contenedor multimedia MP4 con compresión H.264 que se detalla en la sección 4 por su alto grado de utilización en vídeos generados por dispositivos móviles. En la sección 5 se realiza un análisis de los metadatos en vídeos reales de varios teléfonos móviles. Por último, las principales conclusiones extraídas de este trabajo se presentan en la sección 6.

2 Técnicas de análisis forense

La mayor parte de las investigaciones realizadas en el campo de la identificación de la fuente se han realizado para imágenes fotográficas estáticas. Sin embargo, la investigación científica requiere soluciones a los temas forenses relacionados con las señales de video debido a sus peculiaridades y la amplia gama de posibles alteraciones que se pueden aplicar a ellos. La mayoría de las técnicas de análisis forense en este campo, que se pueden aplicar a una imagen pueden ser aplicadas a los diferentes fotogramas de un vídeo [5].

El análisis forense de imágenes digitales se puede dividir en dos grandes ramas [6]: autenticidad de imágenes digitales e identificación de la fuente de adquisición de una imagen.

La primera de las ramas trata de discernir si una imagen ha sufrido algún procesamiento posterior al de su creación, es decir, que no haya sido manipulada. La segunda de las ramas pretende identificar el tipo o clase de fuente que generó la imagen digital. Dentro de esta segunda rama puede realizarse una subdivisión en dos grupos: identificación del tipo de dispositivo fuente (cámara, escáner, generadas por computador, etc.) o identificación de la marca y modelo del dispositivo. Para el diseño de técnicas y algoritmos en cualquiera de estas ramas se aprovechan algunas características especiales de las imágenes o vídeos que sirven como herramienta para el análisis forense.

En [4] [7] se realiza un estudio de las características que pueden ser objeto de análisis forense en dispositivos móviles. El mayor problema con este enfoque es que los diferentes modelos de las cámaras digitales usan componentes de un número reducido de fabricantes y que los algoritmos que usan para la generación de las imágenes y vídeos también son muy similares entre modelos de la misma marca.

En [8] se realiza una comparación detallada de los principales grupos de técnicas de identificación de fuente de adquisición. Estas se dividen en cinco grupos y están basadas en: metadatos, características de la imagen, defectos de la matriz CFA e interpolación cromática, imperfecciones del sensor y las transformadas wavelet.

Aun teniendo en cuenta estas dos grandes ramas no se puede dejar pasar por alto la información de los metadatos que los dispositivos introducen en el proceso de adquisición de la fotografía. Suponiendo la veracidad de los datos contenidos en la imagen, es decir, que no se hayan dado manipulaciones mal intencionadas a posteriori, dependiendo de cada fabricante y dispositivo se arroja en una diversidad de formatos, una información útil para el analista forense (localización GPS, fuente de la foto, características técnicas de la imagen, etc.).

Los archivos de imágenes digitales pueden ser modificados además de una forma más o menos elaborada por cualquier usuario. Existen una gran cantidad de programas de edición accesibles a cualquier tipo de usuario que permiten modificar este tipo de contenido digital siendo muchas veces estos cambios imperceptibles para el ojo humano. Igualmente al caso anterior, estos cambios pueden ser intencionados o malintencionados, pero independientemente de la fe con la que se realizó el cambio, la imagen pierde su originalidad con respecto a la generación por parte de la fuente de adquisición. Estas situaciones pueden generar problemas o indefiniciones cuando las imágenes son utilizadas como evidencias en algún proceso, ya sea judicial o no, dado que no se puede garantizar la identificación de la fuente de adquisición del contenido o la no manipulación del mismo sin realizar un análisis forense previo.

A pesar de las debilidades de este tipo de técnicas, si existen en el archivo los metadatos y de alguna manera se logra comprobar que no han sufrido modificaciones externas, su uso es de gran utilidad para los analistas forenses. Existe información difícilmente inferible del propio contenido de la imagen como por ejemplo la información GPS o la fecha y hora de la toma de la imagen, entre muchas otras. Sin embargo, estas técnicas dependen en gran medida de los metadatos que los fabricantes deciden insertar cuando la imagen es generada y la corrección en seguimiento de la especificación o estándar de metadatos que utilice.

En [9] [10] realizan un estudio a fondo, donde se demuestra que los fabricantes no siguen fielmente la especificación Exif. Esto puede conllevar la extracción de información errónea o inválida para fines forenses.

Asimismo, este método es el más vulnerable a modificaciones malintencionadas, e incluso se puede dar el caso de la eliminación total de los metadatos, ya sea intencionadamente o de manera inconsciente. Ejemplos de ello son algunos programas de edición fotográfica, que al editar o comprimir una imagen, actualizan incorrectamente los metadatos o provocan la pérdida de los mismos.

En el caso del desarrollo de técnicas de análisis forense en vídeo, existen pocas referencias al respecto. Algunas se basan directamente en la secuencia de codificación y otras en la extracción de frames aplicando algún método de clasificación para imágenes fijas [11] [12].

3 Metadatos en vídeos digitales

Los metadatos o “datos sobre datos” registran información relacionada con las condiciones de captura de la imagen/vídeo, como fecha y hora de generación, presencia o ausencia de *flash*, distancia de los objetos, tiempo de exposición, apertura del obturador e

información GPS, entre otras. En otras palabras, información de interés que complementa el contenido principal de un documento digital. Los metadatos, entre otros usos, pueden llegar a ser una potente ayuda para la organización y búsqueda en librerías de imágenes.

Las imágenes digitales son almacenadas en una gran variedad de formatos como TIFF, JPEG [13] o PSD. Algunos de los distintos contenedores de metadatos para los distintos formatos son: IFD Exif, TIFF, Adobe XMP, e IPTC-IIM. La especificación Exif [14] es la más utilizada para identificación de la fuente por ser el contenedor de metadatos más común en las cámaras digitales 15. La especificación Exif incluye cientos de etiquetas, entre las que se encuentran *marca* y *modelo*, aunque cabe destacar que la propia especificación no hace obligatoria su existencia en los archivos.

Por su parte, los vídeos digitales son almacenados en una amplia variedad de formatos, denominados "contendores multimedia", que almacenan información de vídeo, audio, metadatos e información de sincronización y corrección de errores siguiendo un formato preestablecido en su especificación técnica. Como su propio nombre indica, contendor multimedia, es un archivo que contiene en su interior varios elementos, como mínimo las pistas de vídeo y audio [16][17]. Algunos contenedores también permiten incluir otros elementos como imágenes o subtítulos integrados, sin necesidad de archivos externos.

La Figura 1 muestra el formato del contenedor multimedia.

Figura 1: Formato del contenedor multimedia.

Las pistas de vídeo y audio normalmente están comprimidas a través de los diferentes códecs de cada uno de los contenedores multimedia. Estos códecs son los encargados de descomprimir la información para su posterior reproducción. Los códecs son la base para que todos los dispositivos actuales sean capaces de capturar o reproducir un archivo que contenga imágenes y sonido en su interior. Dependiendo del códec elegido se obtiene una mejor o peor calidad, así como, un mayor o menor tamaño. Asimismo, como ocurre con el vídeo, hay canales de audio incluidos en el archivo. También suelen estar comprimidos con un códec determinado para ahorrar espacio. La Tabla 1 presenta los contenedores multimedia más conocidos con los códecs de vídeo y audio que utilizan y especificado si contienen metadatos almacenados. En resumen, no sólo es necesario conocer el formato del contenedor para poder separar las pistas de vídeo y de audio, sino que también es necesario poder decodificarlas. Los contenedores más populares hoy en días son: MP4 (MPEG-4) [18], MOV (archivo Quicktime de Apple) [19], AVI (Audio Video Interleave de Microsoft) [20] y MKV (Matroska).

Tabla 1: Contenedores multimedia de vídeos.

Contenedor	Propietario	Formato de codificación de video	Formato de codificación de audio	metadatos
3GP	3GPP	H.263,MPEG-4 Part 2, H.264/MPEG-4 AVC	AMR-NB, AMR-WB,AMR-WB+, AAC, HE-AAC and HE-AAC v2	?
3G2	3GPP2	H.263,MPEG-4 Part 2, H.264/MPEG-4 AVC	AMR-NB, AMR-WB,AAC, HE-AAC, EVRC,EVRC-B, EVRC-WB, 13K (QCELP), SMV orVMR-WB	?
ASF	Microsoft	VFWor DMO	todos a través de ACM o DMO	Sí
AVI	Microsoft	VFW	todos a través de ACM	Sí
divx	DivX, Inc.	MPEG-4 Part 2	MP3, PCM, AC-3	?
EVO	MPEG	MPEG-2 Part 2,H.264/MPEG-4 AVC, VC-1	AC-3, E-AC-3, Dolby TrueHD, Linear PCM, DTS, DTS-HD, MPEG-2 Part 3	?
F4V	Adobe Systems	H.264/MPEG-4 AVC	MP3, AAC, HE-AAC[7]	Sí
FLV	Adobe Systems	Sorenson,VP6, H.264/MPEG-4 AVC	MP3, Nellymoser, ADPCM, Linear PCM,AAC, Speex	Sí
Matroska	CoreCodec	Virtually anything	Cualquiera virtualmente	Sí
MP4	MPEG	MPEG-2 Part 2, MPEG-4 ASP, H.264/MPEG-4 AVC, H.263, VC-1,Dirac	MPEG-2/4 (HE)-AAC, MPEG-1/2 Layers I, II, III (MP3), AC-3, Apple Lossless, ALS, SLS	Sí
MPG/MPEG	MPEG	MPEG-1,MPEG-2	MPEG-1 Layers I, II, III (mp3)	No
MXF	SMPTE	Virtually anything	Cualquiera virtualmente	Sí
Mov / QT	Apple	MPEG-2, MPEG-4 Part 2, H.264, H.263, H.261, Apple ProRes, Apple Pixlet, Cinepak, , DV, DVC Pro 50, Graphics, Motion JPEG, Photo JPEG, QuickTime Animation, Sorenson Video 2, Sorenson Video 3	AAC, HE-AAC, Apple Lossless, MP3, AMR Narrowband, MS ADPCM, QDesign Music 2, QCELP, IMA 4:1, MACE 3:1	Sí
RMVB	RealNetworks	RealVideo 8, 9, 10	(HE)-AAC, Cook Codec, Vorbis	?
VOB+IFO	DVD Forum	MPEG-2 Part 2, MPEG-1 Part 2	AC-3, Linear PCM,DTS, MPEG-2 Part 3,MPEG-1 Layer II	No

En este artículo se realiza un análisis detallado de los metadatos incluidos en el contenedor multimedia .mp4 por ser el más utilizado en la mayoría de los dispositivos móviles.

4 Especificación del contenedor multimedia MP4 con compresión H.264

MP4 [21] forma parte del estándar MPEG-4 parte 14 y se utiliza para distribuir vídeo y audio, por ejemplo H.264 AVC [18] para vídeo o AAC para audio, pero también puede almacenar otro tipo de datos (subtítulos, información de capítulos e imágenes fijas, entre otros). La extensión asociada a este contenedor es .MP4, pero no es poco frecuente encontrar archivos de audio que lleven la extensión .M4A, que es la extensión adoptada por Apple para distribuir música en iTunes y su reproductor iPod.

La gran mayoría de los archivos de audio de MP4 están comprimidos con el formato AAC (*Advanced Audio Coding*), aunque también admite compresión en MP3. En condiciones normales no supone un problema cambiar la extensión de un archivo .M4A a .MP4 manualmente si ello ayuda a trabajar más cómodamente con él. También podemos encontrarnos archivos de vídeo con las extensiones .M4V o .MP4V.

Asimismo, H.264 / MPEG-4 AVC [22] es el códec más utilizado en las videocámaras modernas y cámaras digitales de dispositivos móviles que almacenan los vídeos capturados en discos duros, tarjetas de memorias, etc. H.264 se ha convertido en el códec de vídeo más popular en los últimos años gracias a su relación calidad/tamaño, ideal para el uso de contenidos en calidad HD. Se usa igualmente en televisores, smartphones, lectores Blu-ray o vídeos de YouTube [22].

Debido al gran número de vídeos que se realizan hoy día y a la necesidad de saber su estructura para el análisis forense de los mismos, a continuación, se analiza la estructura del contenedor multimedia MP4 con códecs de compresión H.264 de vídeos capturados con dispositivos móviles Android. Este análisis es de suma importancia ya que conociendo la estructura del vídeo se puede obtener información que sirva de evidencia de manipulación del mismo.

4.1. Estructura de un atomo

Un vídeo está compuesto de un conjunto de átomos. Los tipos de átomos se especifican por un entero sin signo de 32 bits, típicamente interpretado como un código ASCII de cuatro caracteres normalmente en minúsculas. Los átomos son de naturaleza jerárquica. Es decir, un átomo puede contener otros átomos, que a su vez pueden contener otros, y así sucesivamente.

El formato de los datos almacenados en un átomo dado no siempre puede ser determinado sólo por el campo “type” del átomo; el tipo del átomo padre también puede ser importante. En otras palabras, un tipo de átomo dado puede contener diferentes tipos de información en función de su átomo raíz. Los átomos por lo general no siguen ningún orden particular. La cabecera del átomo contiene los siguientes campos:

- “size”: Un número entero sin signo de 32 bits que indica el tamaño del átomo. Sin embargo, este campo puede contener valores especiales que indican un método alternativo para determinar el tamaño del átomo. Estos valores especiales se utilizan normalmente sólo para átomos media data “mdat”:
  - 0: Se permite sólo por un átomo de nivel superior, designa el último átomo en el archivo e indica que el átomo se extiende hasta el final del archivo.
  - 1: Indica que el tamaño real del átomo está en el campo “extended size”.

  El tamaño real de un átomo no puede ser menor que 8 bytes.
- “type”: Un número entero de 32 bits que especifica el tipo de átomo. Por ejemplo ‘moov’ 0x6D6F6F76 para un átomo de película, ‘trak’ 0x7472616B para un átomo de pista. Conocer el tipo de un átomo permite interpretar sus datos.
- “extended size”: Un campo de 64 bits que es utilizado por átomos con datos que contienen más de 2^{32} bytes. En este caso el campo “size” se establece en 1.

Algunos átomos también contienen los campos “version” y “flags”. Estos atómos se denominan átomos completos y no son tratados como parte de la cabecera del átomo sino como campos de datos específicos para cada tipo de átomo que los contiene.

Los principales tipos de átomos son:

- ‘**ftyp**’: Tipo de compatibilidad de archivo, identifica el tipo de archivo y lo diferencia de los tipos de archivos similares, tales como archivos MPEG-4 y JPEG-2000. Contiene los siguientes campos: *Size*: Especifica el número de bytes en este átomo (32 bits); *Type*: “ftyp” (32 bits); *Major Brand*: Identifica el tipo de archivo de película. Si un archivo es compatible con múltiples marcas, éstas se especifican en los campos “Compatible_Brands”, y en este campo se identifica la marca preferida (32 bits); *Minor version*: Identifica el tipo de archivo de película, está representado en forma decimal codificado en binario (BCD) indicando año, mes y un código en binario cero. Por ejemplo “BCD 20 04 06 00” (32 bits); *Compatible Brands*: Lista de los formatos de archivo compatibles (32 bits).
- ‘**mdat**’: Muestras de datos media de la muestra de películas tales como marcos y grupos de muestras de audio de vídeo. Por lo general, estos datos se pueden interpretar sólo mediante el uso del recurso de película. Contiene los siguientes campos: *Size*: Especifica el número de bytes en este átomo (32 bits); *Type*: “mdat” (32 bits); *Data*: Contiene los datos de audio y video de la película.
- ‘**free**’: El espacio no utilizado disponible en el archivo. Contiene los siguientes campos: *Size*: Especifica el número de bytes en este átomo (32 bits); *Type*: “free” (32 bits); *Free Space*: Contiene los

bytes de espacio libre. Estos bytes valen todos 0 (32 bits).

- ‘**moov**’: Metadatos de recursos de la película (número y tipo de pistas, localización de datos de la muestra, y así sucesivamente). Describe donde se encuentran y cómo se interpretan los datos de la película. Contiene los siguientes campos: *Size*: Indica el número de bytes en este átomo (32 bits); *Type*: “moov” (32 bits). Adicionalmente, este átomo contiene los átomos presentados en la Figura 2. Como se observa en la figura, el átomo “moov” contiene 2 átomos tipo “trak”, uno para almacenar la información de la pista de video y otro para la pista de audio. Ambos átomos tienen una estructura similar.

5 Análisis de la especificación en vídeos de dispositivos móviles

Una vez presentada la especificación, se ha estimado oportuno realizar un análisis de vídeos reales capturados con dispositivos móviles. Este análisis tiene como objetivos profundizar en el conocimiento de la propia especificación y comprobar si ésta es seguida por los fabricantes. Obviamente dado el alto número de átomos que posee la especificación y que cada imagen sólo posee un subconjunto de ellos, se han elegido algunas estructuras para el análisis.

El análisis ha seguido un orden lógico de estructuras de mayor a menor nivel. Para el análisis se han seleccionado 10 vídeos grabados con 7 teléfonos móviles de diferentes fabricantes: Google (Nexus 5), Samsung (Galaxy Nexus, Galaxy S4 Mini, Galaxy S3 Neo), One Plus One (One Plus One) y Sony (Xperia M2, Xiomi Mi3).

Inicialmente se ha comprobado el orden y los átomos presentes en los vídeos, utilizando la herramienta HxD y así poder abrir y leer los vídeos en forma hexadecimal. Tras analizar los 70 vídeos se ha podido concluir que un mismo teléfono móvil siempre tiene los mismos átomos de vídeo y que estos mantienen el mismo orden. Se observa que los átomos encontrados concuerdan con lo descrito en la especificación.

El primer átomo encontrado es el “fytp” como indica la especificación. Posteriormente hay dos opciones: Si el vídeo tiene el átomo “free”, el siguiente átomo que se encuentra es el “moov” y todos sus átomos hijos, terminando los datos del vídeo con los átomos “free” y “mdat”. Pero si el vídeo no contiene el átomo “free”, el átomo que se encuentra es el “mdat”, seguido del átomo “moov” y todos sus átomos hijos.

Este orden no está establecido en la especificación y podría cambiar para un determinado modelo o fabricante, al igual que los átomos, ya que son opcionales y su orden también podría variar, aunque tras un análisis previo de los vídeos se observa que generalmente siguen el mismo patrón.

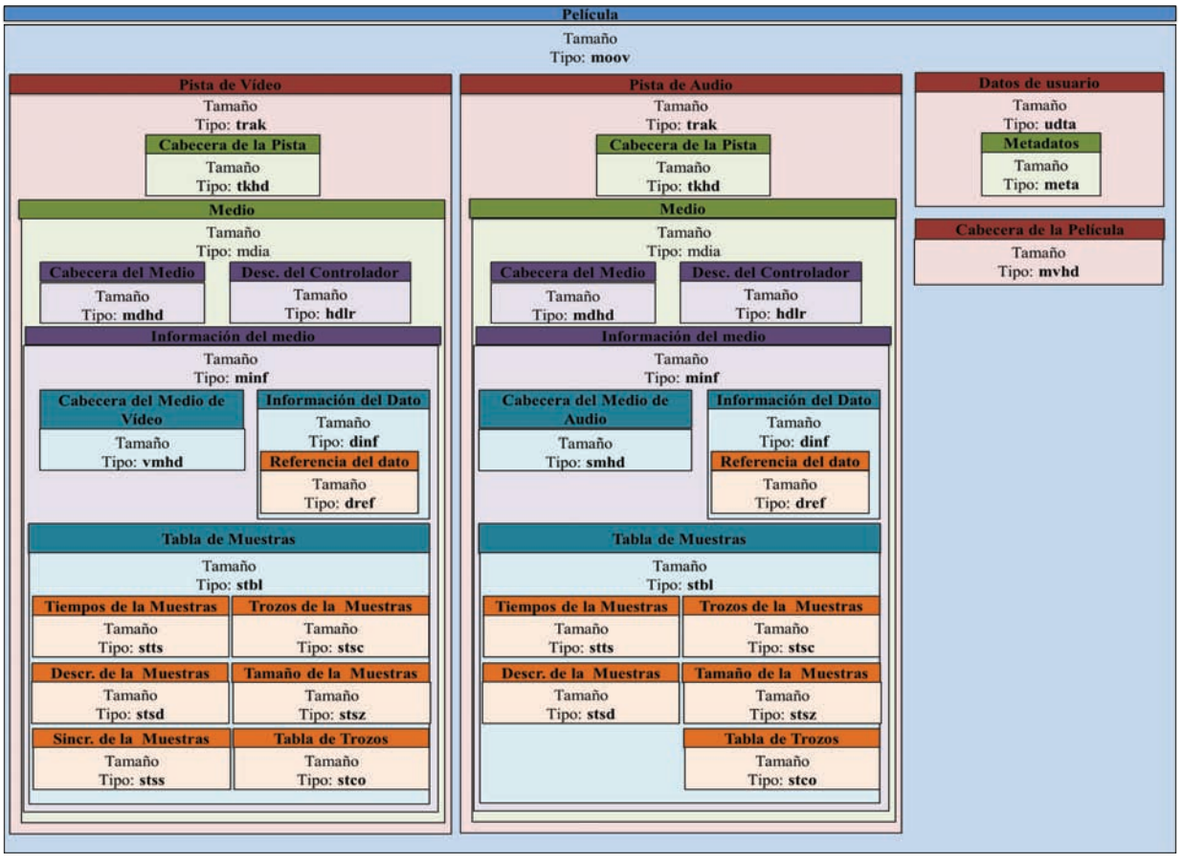

Figura 2: Estructura de los átomos de un vídeo MP4.

Se puede observar que el átomo “moov” siempre mantiene la misma estructura con los mismos átomos hijos. Se encontró que el primer átomo “trak” encontrado siempre se relaciona con la pista de vídeo y el segundo

con la pista de audio. Dentro del átomo “hdlr” del primer “trak” se encuentran los átomos “vide” y “VideoHandle”.

Asimismo, el átomo “hdlr” del segundo “trak” se encuentra los átomos “soun” y “SoundHandle”. A continuación se analiza el primer átomo “trak”, track de vídeo, en concreto su átomo hijo “minf”, ya que el resto es igual en todos los vídeos. El átomo “minf” tiene los átomos “vmhd”, “dinf”, que siempre tiene el átomo hijo “dref” y éste a su vez el átomo “url”, como se indica en la especificación.

Por último se encuentra el átomo “stbl”, que siempre posee los átomos “stsd”, “stts”, “stss”, “stsz”, “stsc” y el átomo “stco”. Algunas veces este átomo no está presente, y en su lugar se encuentra el átomo “co64”, que tras un análisis más profundo de los datos se concluye que el átomo “co64” y el “stco” son el mismo pero que reciben distinto nombre, dependiendo de la marca del dispositivo que grabó el vídeo.

Se puede concluir que el átomo “co64” lo utilizan principalmente los móviles de marca china, como “OnePlus” o “Xiomi” notando que, si el átomo es “co64” su tamaño es de 64 bits y si es “stco” su tamaño es de 32 bits. En el átomo “stsd” se encuentra el tipo de códec “avc1” y el átomo de extensión “avcC” y en algunos casos un átomo más, el átomo “pasp”, ya que éste es opcional como se indica en la especificación.

En el segundo átomo “trak”, pista de audio, en su átomo hijo “minf” encontramos que tiene a su vez los átomos hijos “smhd”, “dinf”, que es igual al del primer átomo “trak” descrito anteriormente, y el átomo “stbl”, que no contiene el átomo “stss” que solo se encuentra en el primer átomo “trak” ya que identifica los fotogramas clave del vídeo. Asimismo, el átomo “stsd”, que en este caso el tipo de códec es el “mp4a” y el átomo de extensión el “esds”. Finalmente, se observa que la mayoría de los modelos analizados tienen los mismos átomos y su estructura es similar. Sin embargo, los dispositivos móviles Nexus 5 del fabricante Google se encuentra el átomo “meta” que no está en los demás modelos.

Finalmente, se realizó un análisis de la información almacenada en los átomos, el resultado de este análisis de presenta en la Tabla 2.

Una vez analizados los átomos presentes en un vídeo, el orden en el que aparecen y la estructura de cada uno de los átomos para ver si corresponden con la especificación, se procede a realizar un análisis de las herramientas existentes para la extracción de metadatos en vídeos. Los programas utilizados son: Gspot 2.7, MediaInfo 0.7.81, ExifTools 10.13: En el análisis de estas herramientas se observa que las herramientas *Gspot* y *MediaInfo* extraen los mismos metadatos: El formato de vídeo, resolución, fechas de creación y modificación, códec de audio y de vídeo, velocidad, duración, número de frames y frames por segundo.

Adicionalmente, *MediaInfo* incluye el sistema operativo del dispositivo que generó el vídeo. Cabe destacar que Tabla 2: Estructura de los átomos de un vídeo MP4.

hay una diferencia notable e importante para el análisis forense: las fechas de creación y modificación son diferentes en ambas herramientas, lo que revela un error en el uso de las marcas de tiempo. Esta anomalía no es aceptable al realizar el análisis forense ya que disponer de esta información puede ser relevante en ciertos casos.

A diferencia de las dos herramientas analizadas *ExifTools* realiza un análisis más profundo sobre los metadatos de vídeos y se podría aproximar un poco más al alcance que se requiere de una herramienta de análisis forense de este tipo. Sin embargo, a pesar que *ExifTools* realiza un análisis más profundo de los átomos, la información que no muestra es aún numerosa.

Como se observa, en este análisis, las herramientas existentes para la extracción correcta de metadatos en vídeos digitales de dispositivos móviles tienen un alcance limitado pues no muestran todos los metadatos al usuario. Estas herramientas extraen la información de una serie de átomos concretos dentro de los metadatos de un vídeo

6 Conclusiones

El objetivo de este trabajo ha sido analizar los metadatos en vídeos digitales de dispositivos móviles, el análisis se centró especialmente en la especificación del contenedor multimedia MP4 con compresión H.264, ya que es la más utilizada por los fabricantes de dispositivos móviles. Se ha comenzado realizando una descripción de la estructura y elementos que lo conforman. Posteriormente, se realizó un análisis binario manual de los átomos almacenados en vídeos grabados de 7 modelos de dispositivos móviles de 5 fabricantes. Este análisis manual es lento y tedioso y hace ver la necesidad de herramientas para su tratamiento automático. Finalmente, se realizó un análisis de la información almacenada en los átomos del conjunto de vídeos. Como conclusión general se puede comentar que la inmensa mayoría de los fabricantes insertan información de gran utilidad en sus vídeos, aun teniendo en cuenta que muchas veces no siguen fielmente la especificación o que los datos almacenados carecen de uniformidad.

Agradecimientos

Los autores agradecen la financiación que les brinda el Programa Marco de Investigación e Innovación Horizonte 2020 de la Comisión Europea a través del Proyecto H2020-FCT-2015/700326-RAMSES (Internet Forensic Platform for Tracking the Money Flow of Financially-Motivated Malware).

.

Tabla 2: Estructura de los átomos de un vídeo MP4

Marca				Google	One Plus One	Samsung			Sony	Xiomi
Modelo				Nexus 5	One Plus One	Galaxy Nexus	Galaxy S4 Mini	Galaxy S3 Neo	Xperia M2	Mi3
Formato	Preferido			mp42	mp42	isom	isom	isom	mp42	isom
	Compatibles			['isom', 'mp42']	['isom', 'mp42']	['isom', '3gp4']	['isom', '3gp4']	['isom', '3gp4']	['isom', 'mp42']	['isom', '3gp4']
Película	Escala de tiempo			1000	1000	1000	1000	1000	1000	1000
	Volumen			1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Velocidad			1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Datos de Usuario			-	-	-	SDLN:SEQ_PLAY smrd:TRUEBLUE smta:saut	SDLN:SEQ_PLAY smrd:TRUEBLUE smta:saut	-	-
	Metadatos	Campo		com.android.version	-	-	-	-	-	-
		Valor		6.0	-	-	-	-	-	-
Pista de video	Escala de tiempo			90000	90000	90000	90000	90000	90000	90000
	Color			[0,0,0]	[0,0,0]	[0,0,0]	[0,0,0]	[0,0,0]	[0,0,0]	[0,0,0]
	Modo gráfico			0	0	0	0	0	0	0
	Manipulador			VideoHandle	VideoHandle	VideoHandle	VideoHandle	VideoHandle	VideoHandle	VideoHandle
	Subtipo del manipulador			vide	vide	vide	vide	vide	vide	vide
	Ancho			1920	1920	1920	1920	1920	1920	1280
	Alto			1080	1080	1080	1080	1080	1080	720
	ID			1	1	1	1	1	1	1
	Volumen			0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Códec	Tipo		avc1	avc1	avc1	avc1	avc1	avc1	avc1
		Resolución	Horizontal	72.0	72.0	72.0	72.0	72.0	72.0	72.0
			Vertical	72.0	72.0	72.0	72.0	72.0	72.0	72.0
		Profundidad		24	24	24	24	24	24	24
		Tabla de color		65535	65535	65535	65535	65535	65535	65535
		Ancho		1920	1920	1920	1920	1920	1920	1280
		Alto		1080	1080	1080	1080	1080	1080	720
		Extensiones		[avcC,pasp]	[avcC,pasp]	[avcC,pasp]	[avcC]	[avcC]	[avcC,pasp]	[avcC,pasp]
		Espaciado de píxeles	Horizontal	65536	65536	65536	-	-	65536	65536
			Vertical	65536	65536	65536	-	-	65536	65536
Pista de audio	Escala de tiempo			48000	48000	48000	48000	48000	48000	48000
	Manipulador			SoundHandle	SoundHandle	SoundHandle	SoundHandle	SoundHandle	SoundHandle	SoundHandle
	Subtipo del manipulador			soun	soun	soun	soun	soun	soun	soun
	Ancho			0	0	0	0	0	0	0
	Alto			0	0	0	0	0	0	0
	ID			2	2	2	2	2	2	2
	Volumen			1.0	1.0	1.0	1.0	1.0	1.0	1.0
	Códec	Tipo		mp4a	mp4a	mp4a	mp4a	mp4a	mp4a	mp4a
		Extensiones		[esds]	[esds]	[esds]	[esds]	[esds]	[esds]	[esds]

Referencias bibliográficas

[1] T. Ahonen and A. Moore, "Tomi Ahonen Almanac 2014: Mobile Telecoms Industry Annual Review," http://goo.gl/B1eX8, 2014.

[2] "Embedded Imaging Takes Off as Stand-alone Digital Cameras Stall," 2013. [Online]. Available: http://www.icinsights.com/data/articles/documents/484.pdf

[3] C. Y. Wen and K. T. Yang, "Image Authentication for Digital Image Evidence," Forensic Science Journal, vol. 5, no. 1, pp. 1–11, September 2006.

[4] V. L. L. Thing, K. Y. Ng, and E. C. Chang, "Live Memory Forensics of Mobile Phones," Digital Investigation, vol. 7, pp. 74–82, August 2010.

[5] P. Bestagini, M. Fontani, S. Milani, M. Barni, A. Piva, M. Tagliasacchi, and S. Tubaro, "An Overview on Video Forensics," in Proceedings of the 20th European Signal Processing Conference, Bucharest, Romania, August 2012, pp. 1229–1233.

[6] T. Gloe, M. Kirchner, A. Winkler, and R. Bohme, "Can We Trust Digital Image Forensics?" in Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, September 2007, pp. 78–86.

[7] T. Van Lanh, K. S. Chong, S. Emmanuel, and M. S. Kankanhalli, "A Survey on Digital Camera Image Forensic Methods," in Proceedings of the IEEE International Conference on Multimedia and Expo, Beijing, China, July 2007, pp. 16–19.

[8] A. L. Sandoval Orozco, J. Rosales Corripio, D. M. Arenas González, L. J. García Villalba, and J. C. Hernández Castro, "Techniques for Source Camera Identification," in Proceedings of the 6th International Conference on Information Technology, Amman, Jordan, May 2013, pp. 1–9.

[9] A. L. Sandoval Orozco, D. M. Arenas González, L. J. García Villalba, and J. C. Hernández Castro, "Anomalías en el Seguimiento de Exif en el Análisis Forense de Metadatos de Imágenes de Móviles," in Actas del XII Reunión Española sobre Criptología y Seguridad de la Información, Donostia-San Sebastián, España, Septiembre 2012.

[10] A. L. Sandoval Orozco, D. M. Arenas González, L. J. García Villalba, and J. C. Hernández Castro, "Analysis of Errors in Exif Metadata on Mobile Devices," Multimedia Tools and Applications, vol. 68, no. 1, pp. 1–29, January 2014.

[11] Y. Su, J. Xu, and B. Dong, "A Source Video Identification Algorithm Based on Motion Vectors," in Proceedings of the Second International Workshop on Computer Science and Engineering, vol. 2, Qingdao, China, October 2009, pp. 312–316.

[12] S. Yahaya, A. T. S. Ho, and A. A. Wahab, "Advanced Video Camera Identification Using Conditional Probability Features," in Proceedings of the IET Conference on Image Processing, London, UK, July 2012, pp. 1–5.

[13] E. Hamilton, "JPEG File Interchange Format. Version 1.02, September 1, 1992,"

http://www.w3.org/Graphics/JPEG/jfif3.pdf.

[14] S. Committee, "Exchangeable Image File for digital still cameras: Exif version 2.3, April 26, 2010," http://goo.gl/jgrCpC, 2013.

[15] R. Baer, "Resolution Limits in Digital Photography: The Looming End of the Pixel Wars," in Proceedings of the Imaging Systems Conference, Tucson, Arizona United States, June 2010.

[16] M. J. Kaur and N. Sharma, "Survey on the General Concepts of MPEG-Moving Picture Experts Group," PARIPEX-Indian Journal of Research, vol. 5, no. 2, 2016.

[17] B. G. Haskell, A. Puri, and A. N. Netravali, Digital video: an introduction to MPEG-2. Springer Science & Business Media, 1996.

[18] I. E. Richardson, H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia. John Wiley & Sons, 2004.

[19] Apple Computer, Inc., "Introduction to QuickTime File Format Specification," https://goo.gl/6rs8uB, 2016.

[20] Microsoft Developer Network, "AVI RIFF File Reference,"

http://msdn.microsoft.com/en-us/library/ms779636(VS.85).aspx, 2016.

[21] C. Lakshmanan, P. Mittal, S. Sehgal, and P. Sinha, "MP4 Container File Formats and Methods of Processing MP4 Container Files," November 2015, uS Patent 9,185,468.

[22] H. Schwarz, D. Marpe, and T.Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," Proceedings of the IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pp. 1103–1120, September 2007.