KIPS Trans. Softw. and Data Eng.
Vol.10, No.10 pp.421~428
ISSN: 2287-5905 (Print), ISSN: 2734-0503 (Online)

https://doi.org/10.3745/KTSDE.2021.10.10.421

Forensic Analysis of HEIF Files on Android and Apple Devices

Youngjin Kwon[†] · Sumin Bang[††] · Jaehyeok Han[†††] · Sangjin Lee[††††]

ABSTRACT

The High Efficiency Image File Format (HEIF) is an MPEG-developed image format that utilizes the video codec H.265 to store still screens in a single image format. The iPhone has been using HEIF since 2017, and Android devices such as the Galaxy S10 have also supported the format since 2019. The format can provide images with good compression rates, but it has a complex internal structure and lacks significant compatibility between devices and software, making it not popular to replace commonly used JPEG (or JPG) files. However, despite the fact that many devices are already using HEIF, digital forensics research regarding it is lacking. This means that we can be exposed to the risk of missing potential evidence due to insufficient understanding of the information contained inside the file during digital forensics investigations. Therefore, in this paper, we analyze the HEIF formatted photo file taken on the iPhone and the motion photo file taken on the Galaxy to find out the information and features contained inside the file. We also investigate whether or not the software we tested support HEIF and present the requirement of forensic tools to analyze HEIF.

Keywords : H.265, HEIF, HEIC, Motion Photo, Digital Forensics

스마트폰에서 촬영된 HEIF 파일 특징 분석에 관한 연구

권 영 진[†] · 방 수 민[††] · 한 재 혁[†††] · 이 상 진[††††]

요 약

HEIF(High Efficiency Image File Format)는 MPEG에서 개발된 이미지 포맷으로써, 비디오 코덱인 H.265를 활용하여 정지된 화면을 하나의 이미지 형태로 저장할 수 있도록 개발된 컨테이너이다. 아이폰은 2017년부터 HEIF를 사용하고 있으며, 2019년부터는 갤럭시 S10과 같은 안드로이드 기기도 해당 포맷을 지원하고 있다. 이 포맷은 우수한 압축률을 가지도록 이미지를 제공할 수 있으나, 복잡한 내부 구조를 가지고 있어 기기나 소프트웨어 간 호환성이 현저하게 부족하여 일반적으로 사용되는 JPEG(또는 JPG) 파일을 대체하기에는 아직 대중적이지 못한 상황이다. 하지만 이미 많은 기기에서 HEIF를 사용하고 있음에도 불구하고 디지털 포렌식 연구는 부족한 상황이다. 이는 디지털 포렌식 조사 과정에서 파일 내부에 포함된 정보의 파악이 미흡하여 잠재적인 증거를 놓칠 수 있는 위험에 노출될 수 있다. 따라서 본 논문에서는 아이폰에서 촬영된 HEIF 형식의 사진 파일과 갤럭시에서 촬영된 모션 포토 파일을 분석하여 파일 내부에 포함된 정보와 특징들을 알아본다. 또한 이미지 뷰어 기능을 지원하는 소프트웨어를 대상으로 HEIF에 대한 지원 여부를 조사하고 HEIF를 분석하는 포렌식 도구의 요구사항을 제시한다.

키워드 : H.265, HEIF, HEIC, 모션포토, 디지털 포렌식

1. 서 론

JPEG(Joint Photographic Experts Group) 형식의 사진 파일은 거의 모든 디지털 기기에서 사용되고 있으며 사실상 오랫동안 표준 이미지 압축 방식으로 사용해오고 있다. 하지만 1992년에 개발된 JPEG 형식은 현대의 높은 해상도를 효율적으로 압축하여 표현하기에 기술적으로 부족하다는 단점이 있다. 이를 보안하기 위해 BPG (Better Portable Graphics) [1], WebP[2] 등 다른 형식의 이미지 파일로 대체하려는 시도가 있었으나 오랜 기간 사용된 JPEG 형식의 범용 소프트웨어 지원 기반이 매우 견고하여 새로운 방식을 도입하려는 사용자를 충분히 확보하지 못하였다[3].

그러나 2017년 애플이 이미지를 저장하는 포맷을 HEIF으로 변경하면서 아이폰을 비롯한 애플 기기의 카메라 어플리케이션은 이미지를 저장할 때 HEIF를 기본으로 한다. JPEG 형식은 사용자가 설정을 변경할 경우에만 저장되고 있다. 이어서 삼성이 2019년 출시된 갤럭시 S10에 HEIF 기능을 추가하면서 지금은 iOS뿐 아니라 안드로이드 기기에서도 HEIF를 지원하고 있다. 또한 카메라 촬영하기 전후 약 2초 동안의

※ 이 논문은 2021년 한국정보처리학회 춘계학술발표대회의 우수논문으로 "스마트폰에서 촬영된 HEIF 파일의 디지털 포렌식 특징 분석"의 제목으로 발표된 논문을 확장한 것임.
† 비 회 원 : 고려대학교 컴퓨터학과 정보보호융합전공 학사과정
†† 비 회 원 : 고려대학교 정보보호대학원 석사과정
††† 준 회 원 : 고려대학교 정보보호대학원 박사과정
†††† 종신회원 : 고려대학교 정보보호대학원 교수
Manuscript Received : June 25, 2021
Accepted : July 5, 2021
* Corresponding Author : Sangjin Lee(sangjin@korea.ac.kr)

상황을 녹화하여 움직이는 사진을 찍을 수 있도록 지원하는 라이브 포토(Live Photo)나 모션 포토(Motion Photo) 기능에서 생성된 파일 역시 HEIF가 조합된 형태이다.

이렇게 이미지를 처리하는 기술의 변화가 짧은 기간에 도래하여 2020년 기준으로 갤럭시 기기와 아이폰의 점유율이 세계 스마트폰 출하량의 37%를 차지할 정도로 HEIF가 대량으로 생산되고 있음에도 불구하고 이에 대한 지속적인 연구가 진행되지 않아, 사용자 간 사진 파일 공유가 제한적이고 정상적으로 이미지를 열람할 수 있는 소프트웨어가 부족하다는 문제를 야기하였다[4]. 심지어 EnCase와 같은 주요 포렌식 도구에서조차 아직 열람 기능을 지원하고 있지 않다. 특히, 디지털 포렌식 조사에서 HEIF의 열람을 비롯해 다중 이미지 저장, 버스트 샷 등 추가적으로 지원하는 기능들에 대한 이해가 미비하여 파일 내부에 포함된 메타데이터 정보(예: EXIF)를 파악하는 연구가 필요하다.

본 논문에서는 HEIF 형식의 사진 파일과 모션 포토 파일을 분석하여 파일 내부에 포함된 정보와 특징들을 알아보고, HEIF 뷰어 기능을 지원하는 소프트웨어를 조사하고 디지털 포렌식 조사에서 고려해야 할 사항을 살펴본다.

2. HEIF 파일 분석

2.1 개요

ISO/IEC23008-12에 정의된 HEIF 포맷은 ISO Base Media File Format (ISOBMFF)을 기반으로 하여 MPEG-4 파트 12 사양의 일부로 비디오와 오디오 콘텐츠를 위해 설계된 확장형 포맷이다[5]. 여러 이미지를 한번에 저장할 수 있을 뿐만 아니라 버스트 샷이나 애니메이션과 같은 영상 시퀀스도 저장할 수 있다[6]. 또한 Table 1과 같이 HEIF는 각 이미지에 대한 역할을 부여할 수 있고 이미지 속성 저장 공간(ItemPropertyContainerBox)을 제공한다. 속성은 크게 두 종류인데, 수정 사항을 적용하지 않고 이미지 항목에 대한 정보를 제공하는 '설명 속성(Descriptive property)'과 이미지 항목에서 수행해야하는 변환 수정에 대한 정보를 제공하는 '변환 속성(Transformative property)'이 있다[7].

Table 1. Roles of Image Formats

Role	Description
Cover image	A representative image of the image items and image sequence tracks of the file.
Thumbnail image	A smaller-resolution representation of a master image.
Auxiliary image	An image that complements a master image
Master image	An image that is not a thumbnail image or an auxiliary image.
Hidden image	An image that should never be displayed
Pre-derived coded image	A coded image that has been derived from other images
Coded image	A coded representation of an image
Derived image	An image that is represented in a file by an indicated operation to indicated input images and can be obtained by performing the indicated operation to the indicated input images.

2.2 기본 이미지 데이터 구조

HEIF는 '박스(box)'라는 기본 데이터 구조로 구성된다. 박스는 각 4바이트로 이루어진 크기(size)와 ASCII 문자로 명명된 기호(mnemonic)를 가지며 뒤이어 페이로드(payload)를 저장한다.

파일에 나타나는 첫 번째 박스는 'ftyp'로, 파일에 대한 일반 인코딩 메타데이터를 포함하며 브랜드라 표시된다. 이 브랜드는 코딩형식에 제한을 두는 역할을 한다. 모든 코덱이 사용될 수 있는 정지 이미지에 해당하는 'mif1'과 이미지 시퀀스에 해당하는 'msf1'을 지정하며, 둘 다 .heif 확장자에 해당한다. .heic 확장자에 해당하는 브랜드는 특정 인코딩 프로필의 사양과 함께 HEVC 인코딩을 사용한다. 'ftyp'는 HEIF 파일의 첫 번째 박스이기 때문에 이는 파일의 첫 데이터가 'ftyp'의 크기를 나타내는 값임을 뜻한다. 따라서 파일의 HEIF 여부를 결정하는 가장 좋은 방법은 'ftyp'의 헤더를 살펴보는 것이다.

'meta'는 핸들러 지정자 'hdlr'로 시작하여 나머지 메타데이터를 포함한다. HEIF에서 핸들러 지정자는 'pict'이지만 다른 파일 형식에서는 다를 수 있다. 각 이미지 또는 이미지 시퀀스의 위치는 파일과 함께 'iloc'를 사용하여 매핑 되고 이는 항목의 오프셋 및 길이를 지정하는 인덱스 역할을 한다. 이후 파일의 모든 이미지가 각 이미지의 item entry를 나타내는 'infe'를 포함하는 'iinf'에 저장된다.

'iref'는 이미지의 썸네일('thmb')을 연결하는 것과 같이 이미지들 사이의 관계를 형성할 때 사용된다. 또한 파일에 여러 개의 이미지가 있을 때는 단일 이미지를 대표 항목으로 태그해 뷰어에 표시할 기본 이미지로 지정할 수 있고, 이는 'pitm'에 저장된다. 파일의 맨 끝에는 이미지/항목 데이터 박스인 'idat' 또는 미디어 데이터 박스인 'mdat'에 포함된 압축 데이터가 있다.

2.3 변환된 이미지 데이터 구조

HEIF 이미지에는 색에 대한 정보를 저장하는 'colr'와 가로 세로 비율을 저장하는 'pixi'과 같이 서술적인 속성과 더불어 기본 이미지를 조작하는 과정에서 생성되는 변환적인 속성도 저장한다. 이는 실제 원본이 수정되지 않고 변환된 정보를 저장하는 방식으로, 무손실 편집이 가능하며 수정된 버전을 사용자에게 제공한다는 것을 의미한다. 그러나 이러한 변환은 임의적이지 않으며 파일 판독기에서 지원해야 한다.

변환 데이터를 저장하는 박스의 종류에는 회전 'irot', 자

르기 ‘clap’, 반전 ‘imir’ 등이 있다. 또한 여러 이미지를 함께 렌더링 하여 그리드 ‘grid’ 또는 오버레이 ‘iovl’와 같은 단일 이미지로 표시할 수 있다. 이 밖에도 파생된 이미지에는 그리드 같은 파생된 이미지의 유형을 지정하는 고유한 ‘infe’ 항목이 있다. 이러한 파생된 이미지 ‘iref’ 박스의 적절한 참조를 사용하여 마스터 영상 및 보조 영상과 연결된다.

2.4 영상 시퀀스 데이터 구조

기본 구조는 일반적으로 PNG (Portable Network Graphics) 파일에서 사용되는 청크 구조와 유사하지만, HEIF의 구조는 Fig. 1의 왼쪽과 같이 다수의 박스가 중첩되어 상자 간에 계층 구조를 이루거나 관계를 생성하고 있다. 이는 동영상을 담을 수 있는 컨테이너 형식이므로 전통적인 이미지 파일 포맷보다는 MP4, MOV 파일과 같은 비디오 파일 포맷과 더 유사하다. 여러 이미지를 저장하거나 버스트 샷, GIF (Graphics Interchange Format)와 같은 애니메이션 형태로도 이미지를 저장할 수 있으며, 이미지 파일 포맷의 특징을 Table 2와 같이 비교할 수 있다[8].

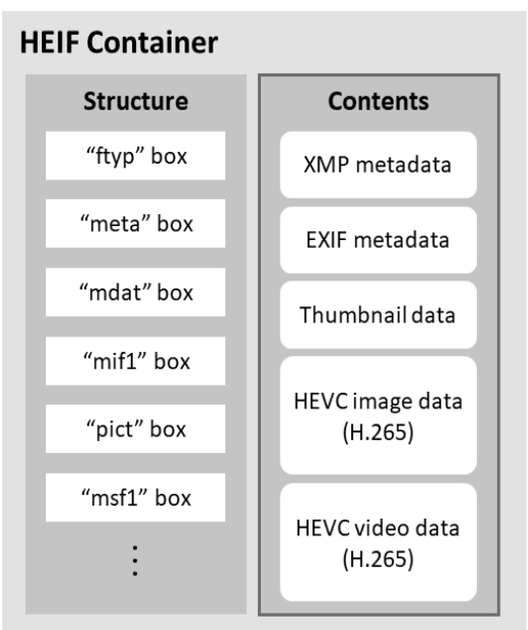

Fig. 1. HEIF File Internal Structure and Information

영상 시퀀스 데이터는 주로 ‘moov’에 포함되어 있으며 다른 파일 형식에서는 비디오 콘텐츠를 포함하는 데 사용된다. ‘moov’는 메타 박스 내에 중첩되지 않으며 하나 이상의 ‘trak’을 포함해 각 박스에는 이미지 시퀀스가 포함된다. ‘trak’에는 ‘hdlr’가 있으며 HEIF에서는 ‘pict’로만 설정될 수 있기 때문에 다른 형식의 비디오 파일과 내용을 구별할 수 있다. 이 밖에도 ‘trak’에는 미디어 메타데이터 및 데이터 처리와 관련된 많은 다른 요소가 있다.

또한 HEIF 파일에서는 이미지 타일링이 지원된다는 특징이 있다. 타일을 사용하면 각 영상 세그먼트를 병렬로 디코딩하거나 전체 영상을 디코딩하지 않고도 고해상도 영상의 작은 부분을 분리하여 디코딩할 수 있다. 즉, 타일을 사용하면 더 빠른 처리와 메모리 부하 감소가 가능하다[8].

3. HEIF 뷰어 기능 테스트 결과

스마트폰 종류에 따라 생성되는 HEIF 구조가 상이하다. 먼저 아이폰 12 (iOS 14.4.2)와 삼성 갤럭시(안드로이드 11)에서 기본적으로 설치되어 있는 카메라 어플리케이션을 이용하여 HEIF 샘플 이미지를 수집하였고, 종류별로 파일 이름을 구분한 결과는 다음과 같다.

- IMG_0000.HEIC : 아이폰 사진 파일
- IMG_0000.MOV : 아이폰 라이브 포토
- YYYYMMDD_hhmmss.heic : 안드로이드 사진 파일
- MVIMG_YYYYMMDD_hhmmss.jpg (Fig. 2)
  : 안드로이드 모션 포토(JPEG과 HEIF가 연결된 형태)

이 절에서는 HEIF 뷰어 소프트웨어에서 지원하는 기능을 테스트하여 디지털 포렌식 도구에서 필요로 하는 요구사항을 도출하고 논의한다.

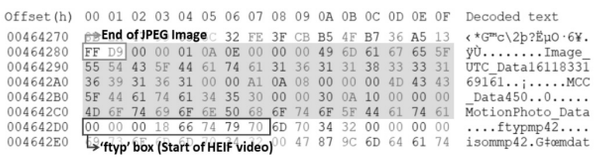

Fig. 2. Analysis of the Area Where JPEG and HEIF are Connected in Motion Photos

Table 2. Feature Comparison of Image File Formats

Category	HEIF	JPEG	PNG	GIF	BPG	WebP
File format	ISOBMFF	TIFF	-	-	-	RIFF
Metadata	EXIF, XMP, MPEG-7	EXIF	-	-	EXIF, XMP	EXIF, XMP
Multiple images	○	×	×	○	○	○
Thumbnail	○	○	×	×	○	×
Lossless compression	○	×	×	×	×	×

3.1 일반 소프트웨어 지원 개요

HEIF은 Windows 10 (버전 1803 이상)에서 열람할 수 있으나 Microsoft 스토어에서 'HEIF 이미지 확장' 패키지를 추가로 설치해야 한다. 구글 크롬, 모질라 파이어폭스, 마이크로소프트 엣지와 같이 대중적으로 사용되는 웹 브라우저에서도 안드로이드 9(Pie) 이상 버전에서 지원하며 기본으로 지원하지 않고 있다. 다수의 이미지 편집 소프트웨어에서도 아직 HEIF의 열람이나 편집 기능에 대한 지원이 부족하다. 예를 들어, 어도비 포토샵(Photoshop)이나 라이트룸(Lightroom) 최신 버전은 macOS에서만 지원하지만 Windows에서는 지원하지 않고 있어 HEIF를 편집하기 위해서는 macOS가 설치된 시스템을 사용하거나 Affinity Photo, GIMP, Paint.NET, Pixelmator, GraphicConverter, ImageMagick와 같은 이미지 편집 전문 소프트웨어를 설치해야 한다. 포렌식 도구인 Encase 20.4, X-Ways Forensics 19.8, FTK Imager 4.2.1.4, Autopsy 4.13.5에서도 지원 여부를 확인해보았으나 열람 기능을 지원하지 않았으며, Magnet AXIOM 4.10만 지원하였다[9,10].

3.2 다중 이미지 저장 기능

다른 이미지 포맷과 구별되는 HEIF의 특징은 여러 개의 이미지 또는 시퀀스를 각각의 관련 메타데이터와 함께 저장할 수 있는 컨테이너라는 점이다. 이것은 GIF나 APNG (Animated PNG)에서 가능한 다중 프레임 애니메이션과 달라 HEIF의 항목이 동일한 이미지 스트림의 일부가 될 필요가 없고 완전히 독립적인 이미지들의 갤러리를 나타낼 수 있다.

따라서 HEIF는 단일 정적 이미지로 접근해서는 안 된다. 하지만 윈도우 탐색기, 윈도우 사진 뷰어, 드롭박스의 웹 프리뷰, 포토샵(Mac), JPEG로의 변환 툴의 경우, 일반적으로 'pitm' 박스로 설정된 커버 이미지만 표시하는 것으로 나타났다. 이는 HEIF 파일에 여러 개의 이미지가 포함되어 있고 이러한 도구 중 하나를 사용하여 보거나 JPEG로 변환된 경우 하나의 이미지만 표시되어 불법 미디어를 쉽게 숨길 수 있다는 것을 의미한다. 이러한 문제는 메타데이터에도 적용되며 Exiftool, CopyTrans 모두 기본 이미지(커버 이미지)에 대한 메타데이터만 표시하는 것을 확인하였다.

3.3 파일 내 썸네일 저장 기능

HEIF 컨테이너 내에 하나 이상의 이미지 또는 이미지 시퀀스가 썸네일을 포함할 수 있다. 기본 이미지에 썸네일이 포함된 경우 전체 크기의 이미지에서 썸네일을 새로 생성하는 대신 썸네일을 미리 보기로 표시하는 옵션이 있다(예: 윈도우 탐색기). Fig. 3과 같이 HEIF 파일의 썸네일을 볼 수 있는 유일한 소프트웨어는 CopyTrans이며 내장된 썸네일을 표시할 수 있는 다른 소프트웨어는 없었다. 즉, 썸네일의 내용이 첨부한 마스터 이미지의 내용과 일치하지 않아도 되어 데이터를 숨길 수 있다.

Fig. 3. Windows Explorer Displaying .HEIC Files with (bottom) and Without (top) CopyTrans

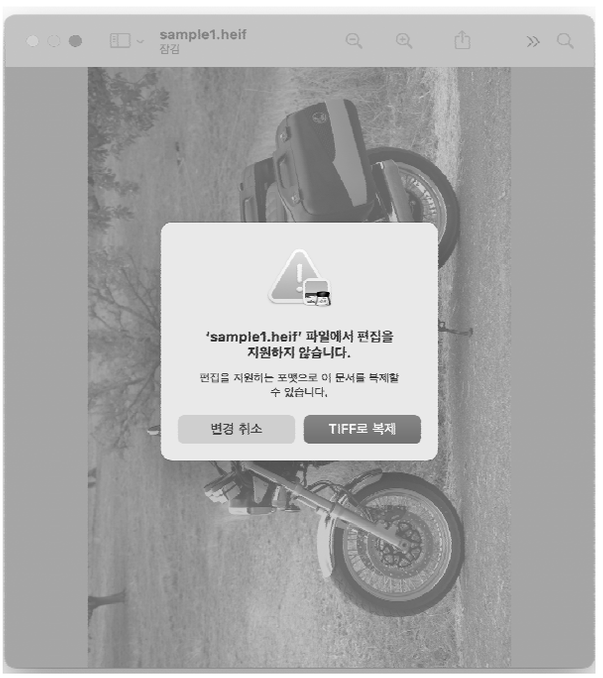

Fig. 4. macOS Preview doesn't Support .HEIF Editting

3.4 무손실 수정 기능 및 보조 이미지 열람

HEIF의 주요 기능 중 하나는 기본 이미지의 무손실 수정이다. 기본 이미지의 자르기, 회전, 반전 또는 보조 이미지를 통한 깊이 적용, 투명도 마스킹 등이 가능하다. 모든 경우 파생 이미지가 생성되며 선택적으로 기본 이미지로 설정될 수 있다. 그러나 Fig. 4와 같이 편집 소프트웨어에서 그리드 보기, 자르기, 회전, 반전 등의 기능을 지원하는 경우는 거의 없었다. 또한 어떤 도구에서도 보조 이미지를 열람할 수 없었다. 일반 소프트웨어를 사용하는 조사관이 무손실 수정 기능으로 변경된 기본 이미지를 원래의 이미지라고 잘못 판단할 수 있어, 이러한 특징을 이해하고 수정되기 이전의 이미지로 재구성할 수 있어야 할 것이다.

현재 아이폰에서는 HEIF 파일을 자르거나 회전, 반전을 하면 원본 HEIF 파일을 그대로 저장해 둔 채 변경 사항이 적용된 JPEG 파일과 변경 내용을 포함하는 .AAE(Appleplist) 파일을 생성하여 저장한다. 만약 수정되지 않은 원래의 사진을 다시 얻기 위해 변경 내용을 모두 되돌리면 새로 생성됐던 JPEG 파일과 .AAE 파일만 삭제되는 것이다.

하지만 HEIF 파일은 자체적으로 수정한 내용을 파일 내부에 자체적으로 포함시킬 수 있다. 이러한 특징으로 인해 HEIF 분석 과정에서 무손실 편집 기능이 적용되었는지가 확인되어야 하고 조사관이 기본 이미지 외에도 보조 이미지나 편집 기능으로 파생된 이미지의 저장 여부에 대해 관심을 가져야 한다는 점을 시사한다.

3.5 기타 부가 기능

그 외에 부가적으로 HEIF는 이미지 시퀀스(버스트 샷이나 애니메이션)를 지원하고, 커버 이미지(첫 번째 프레임)와 더불어 보조 이미지(두 번째 이후 프레임)를 저장할 수 있으나 이러한 내용을 읽고 쓸 수 있는 소프트웨어는 확인할 수 없었다. HEIF 뷰어나 플레이어에서 특정 이미지가 출력되지 않고 숨겨지도록 하는 기능도 지원하는데, 이 기능을 활성화시키는 플래그(hidden)가 설정되었을 경우에는 macOS의 미리보기나 GIMP에서만 설정 상태와 이미지를 확인할 수 있었다. '숨김' 설정이 된 이미지는 Fig. 5와 같이 'infe' 박스의 단일 비트 값에 의해 설정된다. 또한 HEIF는 외부 참조 기능을 지원하고 'dinf' 박스에 URL(Uniform Resource Locator) 형식의 주소를 저장해놓을 수 있다.

HEIF는 유연한 컨테이너 형식이기 때문에, 동일한 콘텐츠가 다양한 방법으로 표현될 수 있다. 특히, 여러 개의 이미지(프레임)가 저장된 경우에 이미지 시퀀스에 따라 인터 코딩(inter-coding)이나 인트라 코딩(intra-coding) 방식으로 압축시킬 수가 있는데, 이는 정확하게 동일한 픽셀이 서로 다른 이진 표현을 가질 수 있음을 의미한다. 인터 코딩을 사용하여 새 프레임을 삽입하면 인코딩 프로세스에 상당한 영향을 미치며 시퀀스의 모든 이미지에 대해 대량의 바이너리 변화를 초래할 수 있다. 즉, HEIF는 다중 이미지를 지원하므로 전체 컨테이너의 해시 값만으로 불법 미디어를 탐색하는데 한계가 있다. 조사관은 이러한 특징에 유의하여 무결성 검증을 위해 HEIF에서 산출된 해시 값을 ZIP과 같은 압축 파일처럼 취급해야할지를 고민해야 한다.

3.6 테스트 결과

주요 포렌식 도구들(Encase 20.4, X-Ways Forensics 19.8, FTK Imager 4.2.1.4, Autopsy 4.13.5)과 이미지 뷰어 소프트웨어들(Dropbox Preview, Photoshop Mac, Exiftool, Copytrans, Win10 HEVC Free, macOS Preview, FFmpeg, Gimp)를 대상으로 HEIF의 주요 기능(보조 이미지 저장 기능, 다중 이미지 저장 기능, 파일 내 썸네일 저장 기능, 무손실 수정 기능)을 지원하는지를 macOS Big Sur와 Windows 10에서 테스트 해 본 결과는 Table 3과 같다. 모두 Fig. 4에 보이는 파일(sample.heif)을 대상으로 진행하였다. 표에 나타난 바와 같이 HEIF의 주요 기능을 지원하는 소프트웨어는 거의 없음을 확인할 수 있다.

Table 3. Test Result of Softwares Regarding Supporting HEIF Files

Function	Software (Forensic Tool)	
	Supported	Not supported
Cover image view	AXIOM 4.10	Encase 20.4 X-Ways Forensics 19.8 FTK Imager 4.2.1.4 Autopsy 4.13.5
Additional image view	N/A	Dropbox Preview Photoshop Mac Exiftool CopyTrans
Thumbnail view	CopyTrans	Win10 HEVC Free Dropbox Preview Photoshop Mac macOS Preview
Lossless compression / Auxilary image view	Gimp	Win10 HEVC Free Dropbox Preview Photoshop Mac macOS Preview FFmpeg CopyTrans

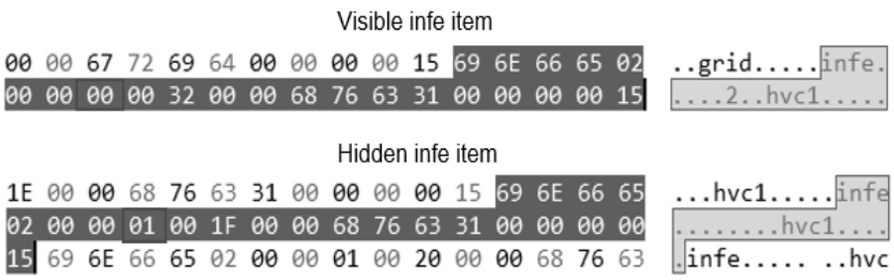

Fig. 5. Comparison of 'infe' Boxes of HEIF Files with Different Hidden Function Settings

4. HEIF 파일의 디지털 포렌식 특징

4.1 HEIF 파일 분석의 제한사항

현재 HEIF에 대한 포렌식 도구 지원이 미흡하다는 점을 감안할 때 핵심 증거를 누락할 가능성을 줄이기 위해 수사관에게 이 형식을 분석할 수 있는 출발점을 제공하는 것이 중요하다. 앞서 언급했듯이, HEIF 파일은 전통적인 파일과 다르게 파일의 시작 부분에 시그니처를 가지고 있지 않으므로 알려져 있는 확장자로 설정되어 있지 않고 관련된 배경지식이 부족하다면 최초에 파일을 식별하는 것이 간단하지 않을 수 있다. 이에 ftyp 박스에서의 브랜드(heic임을 명시한)와 hdlr의 'picture' 타입을 통해 파일을 식별해야 한다. 만약 디지털 포렌식 조사 과정에서 HEIF 파일이 식별된다면 헥스 뷰어나 파서를 이용해 파일의 박스 구조를 1차적으로 분석해야 한다. 기존 ISOBMFF 파서들은 일부 HEIF 이미지에서 문제가 있기 때문에 알려진 ISOBMFF/HEIF 박스에 대한 문자열과 필터를 추출하는 간단한 파이썬 스크립트가 더욱 알맞을 것으로 보인다. 이 단계에서 중요한 것은 파일의 head에 있는 'meta' 박스의 infe 항목 수와 종류를 확인하고, 이미지 시퀀스를 위한 'moov' 박스가 있는지 확인하는 것이다.

'infe' 박스를 분석할 때, 'hvc1HEVC Image'와 같은 이미지 유형은 마스터 이미지를 나타내며, 파생 이미지에는 그리드의 경우 'grid'와 알파 및 오버레이의 경우 'iovl'과 같이 '파생 이미지' 텍스트가 포함되어 있다는 사실을 알고 있어야 한다. 또한 동일한 수의 이미지가 있는지 확인해야 한다. 그렇지 않은 경우, 박스의 플래그 구성요소에 해당하는 각 infe 항목의 8번째 바이트를 검사하여 숨겨진 비트가 설정되었는지 확인해야 한다. 이 비트가 설정된 경우 파일 복사를 해야 하며, 숨겨진 모든 항목에 대해 '01' 대신 '00'을 입력하여 항목이 보이도록 변경해야 한다. 아이폰 영상의 경우 모든 타일이 숨겨지고 기본 파생 영상만 표시된다. 그러나 이는 불법 이미지를 숨기기 위한 연막 역할을 할 수 있으므로 모든 숨겨진 이미지를 볼 수 있도록 전환한 후 검사해야 한다.

'moov' 박스에 저장된 영상 시퀀스는 하위 박스와 데이터 구조가 많아 훨씬 복잡하며 트랙이 비활성 상태로 설정될 가능성도 있다. 이 영역에서 더 많은 작업을 수행해야 하지만 MPC와 같은 미디어 플레이어에서 비디오로 시퀀스를 미리 보거나 FFmpeg를 통해 프레임을 내보내는 것이 현재로서는 가장 효과적인 방법인 것 같다. 테스트에서는 트랙을 비활성 상태로 설정하는 것에 영향을 받지 않는 것처럼 보였지만, 특히 여기서 테스트되지 않은 다중 시퀀스 파일에 대해 이를 보다 완전하게 탐색하기 위해 더 많은 작업이 필요하다.

정지 영상과 이미지 시퀀스의 경우 메타데이터 및 미리 보기를 위한 옵션이 모두 부족한 것으로 나타난다. Windows 10에서 HEVC 확장의 유료 버전을 사용하거나 Windows에서 CopyTrans 유틸리티를 사용하여 커버의 썸네일을 미리 볼 수 있다. 둘 다 탐색기 썸네일에 내장된 썸네일을 사용한다. 그러나 파일에 여러 개의 미리 보기가 있거나 시퀀스에 대한

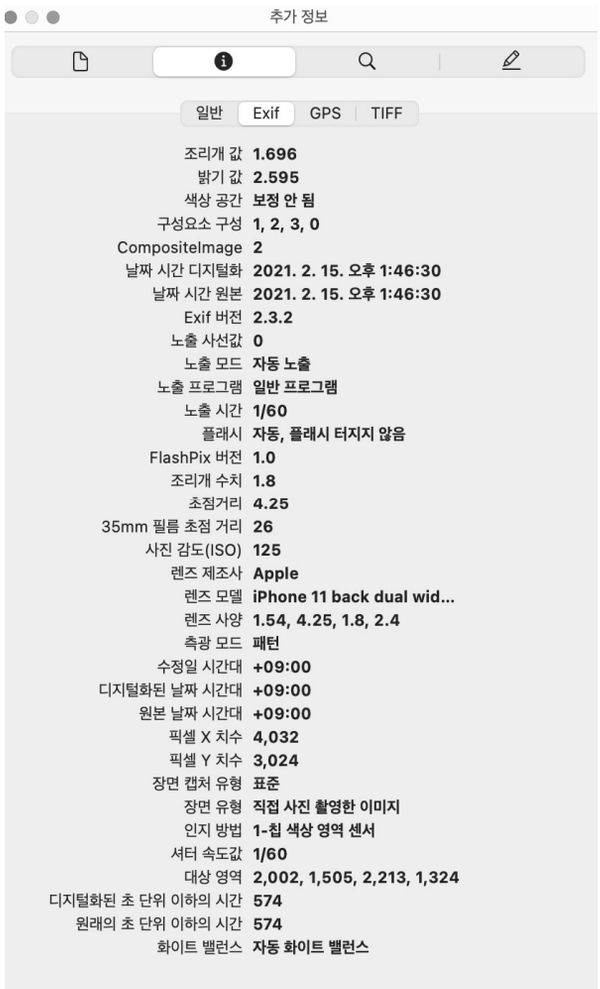

Fig. 6. EXIF Metadata View on macOS Preview (Only Contains Metadata of Cover Image)

미리 보기 트랙이 있는 경우 C++ API 중 하나를 사용하여 파일을 재구성하지 않고 현재 이러한 미리 보기를 볼 수 있는 좋은 방법은 없는 것으로 보인다. macOS Preview 또한 Fig. 6에서 볼 수 있듯이 파일의 기본 항목에 대한 메타데이터를 미리 볼 수 있지만 다른 메타데이터 요소는 탐색하기 어렵다. 그러나 이 경우 파일에서 텍스트 항목으로 추출할 수 있다.

더욱 복잡한 것은 보조 파일을 열람하는 것이다. 테스트된 도구 중에서는 이러한 기능을 제공하는 것은 없었다. 미리 보기 위해 파일을 재구성하기 위해서는 프로그래밍 지식이 필요할 것이다. 마지막으로, URL/URN 값을 수동으로 확인하여 'dref' 박스에 외부 파일 소스가 지정되었는지 확인하는 데 주의해야 한다. 이렇게 하면 HEIF 파일에 포함되지 않은 소스에서 증거를 빼낼 경우 누락되는 것을 방지할 수 있다.

4.2 HEIF 파일 분석을 위한 포렌식 도구 요구사항

HEIF 사진 파일이 JPEG 형식을 대체할 수 있을지는 더 지켜봐야겠지만 앞으로 널리 사용될 것으로 예상된다. 특히, 스마트폰에서 촬영되는 사진 파일의 상당수는 HEIF로 생성될

것이다. HEIF의 분석은 이미지 시퀀스, 메타데이터 구조, 미리 보기 및 보조 파일 등 분석 대상에 따라 분명한 차이가 있다. 그럼에도 불구하고 HEIF의 기능을 지원하는 소프트웨어는 현저히 적은 현황이다. 실제로 포렌식 조사관들은 HEIF를 개별적으로 분석하는데 지나치게 많은 시간을 소비하고 있다. 따라서 HEIF 파일의 파싱, 미리 보기 및 분석을 위한 전용 도구가 필요한 것은 분명하다. 특히, 디지털 포렌식 조사에서는 지금까지 확인한 제한사항들이 보완이 되어야 한다. 따라서 다음과 같이 HEIF 파일을 분석하기 위해 도구가 갖추어야 할 요구사항을 제시한다.

- 비활성/숨김 여부나 역할에 관계없이 HEIF 파일의 모든 이미지 및 이미지 시퀀스를 표시한다. 특히 미리 보기와 보조 항목이 포함되어야 한다.
- HEIF 파일의 모든 이미지, 트랙 및 구성 요소에 대한 메타데이터를 표시해야 한다. 여기에는 항목에 대한 HEIF 박스 구조 내에 관련 플래그와 메타데이터도 포함되어야 한다.
- 'dref' 박스의 URL의 참조를 통해 HEIF 파일의 외부 소스에서 얻어낸 콘텐츠를 검색하고 표시할 수 있어야 한다.
- 파일 구조에 대한 시각적 개요를 제공해야 한다. 예를 들면, 어떤 이미지들이 있는지와 그것들의 관계, 사용된 인코딩, 데이터 오프셋, 첨부된 메타데이터, 썸네일 항목이 있다.
- 인코딩 차이를 감안한 이진 및 픽셀 수준 해시를 포함하여 파일에 있는 모든 이미지의 해시를 제공해야 한다. 인터 코딩 시퀀스에 삽입된 중복 프레임을 사용하여 공격을 방지하기 위해 이미지 시퀀스의 개별 파일을 재구성하고 해시해야 한다.

5. 결 론

HEIF 형식은 디지털 포렌식 분야에서 새롭게 분석해야 하는 대상이며, 이 형식의 증거를 처리하기 위한 새로운 도구와 접근 방식이 필요하다. HEIF는 기존의 스틸 이미지 파일 형식과 유사하지 않은 발전된 컨테이너 형식이며, 여러 개의 이미지와 시퀀스가 다양한 방식으로 포함되고 배열될 수 있다. HEIF는 많은 임베디드 항목과 숨겨진 콘텐츠도 허용하는데, 이는 현재 포맷에 대한 지원이 미흡하여 미리 보기가 어렵다.

본 논문은 HEIF의 데이터 은닉 잠재력과 포맷에 대한 포렌식 분석을 수행했다. 조사자가 중요한 증거를 놓칠 위험이 있으므로 향후 연구에서 HEIF 파일 내용에 대한 상세한 분석을 용이하게 진행하기 위한 새로운 도구를 개발하는 것이 요구되는 바이다.

References

[1] U. Albalawi, S. P. Mohanty, and E. Kougianos, "A hardware architecture for better portable graphics (bpg) compression encoder," In *2015 IEEE International Symposium on Nanoelectronic and Information Systems*, pp.291-296, 2015.

[2] J. Alakuijala, "WebP Lossless Bitstream Specification," 2012, [Internet], https://developers.google.com/ speed/webp/docs/webp_lossless_bitstream_specification.

[3] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Transactions on Consumer Electronics,* Vol.38, No.1, pp.18-34, 1992.

[4] "Counterpoint," Counterpoint Research, Last Modified APRIL 30. 2021, accessed JUNE 24. 2021, [Internet], https://www.counterpointresearch.com/global-smartphone-share/

[5] Information technology — high efficiency coding and media delivery in heterogeneous environments. Standard, International Organization for Standardization, Geneva, CH, December 2017.

[6] "HEIF Image Format," nokiatech.github.io, accessed JUL 7, https://nokiatech.github.io/heif/technical.html.

[7] S. McKeown and G. Russell, "Forensic considerations for the high efficiency image file format (heif)," In *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, IEEE, pp.1-8. 2020.

[8] M. J. Bennett, "Assessing the potential useof high efficiency video coding (hevc) and high efficiency image file format (heif) in archival still images," *Code4Lib Journal*, Vol.41, 2018.

[9] "MAGNET FORENSICS," Magnet Forensics Blog, Last Modified JAN 25. 2021, accessed MAR 20. 2021, [Internet], https://www.magnetforensics.com/blog/android-motion-photos-in-magnet-axiom/

[10] "High Efficiency Image File (HEIF) Format, MPEG-H Part 12," loc.gv, last modified JUN 17. 2021, accessed JUL 5, [Internet], https://www.loc.gov/preservation/digital/formats/fdd/fdd000525.shtml

권 영 진

https://orcid.org/0000-0001-7907-9484
e-mail : lucillek0603@korea.ac.kr
2018년 ~ 현 재 고려대학교 컴퓨터학과 정보보호융합전공 학사과정
관심분야 : Digital Forensics, Cloud Computing, Multimedia Processing

방 수 민

https://orcid.org/0000-0003-2537-5506
e-mail : yssumin@korea.ac.kr
2020년 서울시립대학교 수학과(학사)
2020년 ~ 현 재 고려대학교 정보보호대학원 석사과정
관심분야 : Digital Forensics, Memory Forensics

한 재 혁

https://orcid.org/0000-0001-5724-0775
e-mail : one01h@korea.ac.kr
2011년 서울시립대학교 수학과(학사)
2016년 고려대학교 정보보호학과(석사)
2016년 ~ 현 재 고려대학교
정보보호대학원 박사과정
관심분야 : Digital Forensics, Data Hiding, File System

이 상 진

https://orcid.org/0000-0002-6809-5179
e-mail : sangjin@korea.ac.kr
1987년 고려대학교 수학과(학사)
1989년 고려대학교 수학과(석사)
1994년 고려대학교 수학과(박사)
1989년 ~ 1999년 ETRI 선임연구원
1999년 ~ 2001년 고려대학교 자연과학대학 조교수
2001년 ~ 현 재 고려대학교 정보보호대학원 교수
2008년 ~ 현 재 고려대학교 디지털포렌식연구센터 센터장
관심분야 : Digital Forensics, Steganography, Hash Function