스마트폰에서 촬영된 HEIF 파일의 디지털 포렌식 특징 분석

권영진*, 방수민**, 한재혁**, 이상진**
*고려대학교 컴퓨터학과
**고려대학교 정보보호대학원
{lucillek0603, yssumin, one01h, sangjin}@korea.ac.kr

Analysis of the HEIF files taken with a Smartphone for Digital Forensic Investigation

Youngjin Kwon*, Sumin Bang**, Jaehyeok Han**, Sangjin Lee**
*Dept. of Computer Science and Engineering, Korea University
**School of Cybersecurity, Korea University

요 약

HEIF (High Efficiency 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 있으며 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 HEIF 뷰어를 분석하는 포렌식 도구의 요구사항을 제시한다.

1. 서 론

JPEG (Joint Photographic Experts Group) 형식의 사진 파일은 거의 모든 디지털 기기에서 사용되고 있으며 사실상 오랫동안 표준 이미지 압축 방식으로 사용되어 왔다. 1992년에 개발된 JPEG 형식은 높은 해상도를 효율적으로 압축하여 표현하기에 기술적으로 부족하므로 BPG (Better Portable Graphics) 또는 WebP [1] 등 다른 형식의 이미지 파일로 대체하려는 시도가 있었으나 새로운 방식을 도입하려는 사용자를 충분히 확보하기가 어려웠다 [2].

하지만 2017년 애플은 이미지를 저장하는 포맷으로 HEIF를 채택하였고 아이폰의 카메라 어플리케이션은 이미지를 저장할 때 HEIF로 생성한다. 호환성을 위해 JPEG 형식은 사용자가 설정을 변경할 경우에만 생성되도록 하였다. 그리고 2019년 출시된 갤럭시 S10과 같은 안드로이드 기기에서도 HEIF를 지원하고 있다. 또한 카메라 촬영하기 전후 약 2초 동안의 상황을 녹화하여 움직이는 사진을 찍을 수 있도록 지원하는 라이브 포토(Live Photo)나 모션 포토(Motion Photo) 기능에서 생성된 파일도 HEIF가 조합된 형태로 생성된다.

이러한 이미지를 처리하는 기술의 변화는 사용자 간에 사진 파일의 공유가 제한적이거나 정상적으로 이미지를 열람할 수 있는 소프트웨어가 부족하다는 문제를 야기하였다. 심지어 EnCase와 같은 포렌식 도구에서조차 열람 기능을 지원하고 있지 않다. 특히, 디지털 포렌식 조사에서 HEIF의 열람은 물론이거니와 다중 이미지 저장, 버스트 샷 등 추가적으로 지원하는 기능들에 대한 이해가 미비하여 파일 내부에 포함된 메타데이터 정보(예: EXIF)를 파악하는 연구가 필요하다.

이 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.

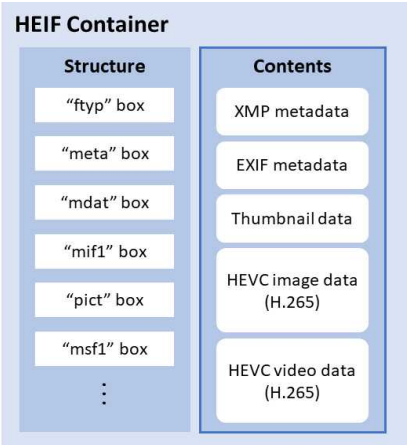
2. HEIF 파일 구조 분석

ISO/IEC23008-12 [3]에 정의된 HEIF 포맷은 멀티미디어 서비스를 위해 사용되는 표준 파일 포맷인 ISOBMFF (ISO Base Media File Format)를 기반으로 하며, 비디오와 오디오 콘텐츠를 위해 설계된 확장형 포맷인 HEVC (High Efficiency Video Coding)의 일부를 활용한 고효율 이미지 파일 형식이다. 동영상을 담을 수 있는 컨테이너 형식이므로 전통적인 이미지 파일 포맷보다는 MP4, MOV 파일과 같은 비디오 파일 포맷과 더 유사하며, 여러 이미지를 저장하거나 버스트 샷, 그리고 GIF와 같은 애니메이션 형태로도 이미지를 저장할 수 있다 (표 1).

HEIF은 박스(box)라는 기본 데이터 구조로 구성되며, 각 박스는 4바이트로 박스 크기와 ASCII으로 명명된 연상 기호(mnemonic)를 가지며 뒤에 페이로드(payload)가 저장된다. PNG (Portable Network Graphics) 파일에서 사용되는 청크 구조와 유사하지만, HEIF의 박스는 중첩되어 박스 간에 계층 구조를 이루거나 관계를 생성할 수 있다. [그림 1]과 같이 HEIF 구조는 첫 번째 박스로 파일에 대한 일반 인코딩 메타데이터를 포함하는 'ftyp'가 있고 모든 코덱이 사용될 수 있는 정지 이미지를 지정하는 'mif1'과 이미지 시퀀스에 해당하는 'msf1'을 지정한다. 그 외에도 여러 종류의 박스가 H.265 방식으로 인코딩된 이미지나 비디오 데이터를 저장하거나 메타데이터 정보를 표현하는데 사용된다. HEIF는 'ftyp' 박스의 길이를 지정하는 헤더로 시작하므로, 대상 파일이 HEIF 포맷을 가지는지 결정하기에 좋은 방법은 이 박스를 살펴보는 것이다.

3. HEIF 뷰어 기능 테스트 결과

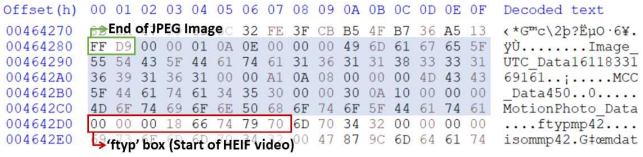
스마트폰 종류에 따라 생성되는 HEIF 구조가 상



(그림 1) HEIF 파일 내부 구조와 정보

이하다. 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 : 갤럭시 사진 파일
- MVIMG_YYYYMMDD_hhmmss.jpg : 갤럭시 모션 포토 (JPEG과 HEIF가 연결된 형태, [그림 2] 참고)



(그림 2) 모션 포토에서 JPEG와 HEIF이 연결되는 영역 분석

이 절에서는 HEIF 뷰어 소프트웨어에서 지원하는 기능을 테스트하여 디지털 포렌식 도구에서 필요로 하는 요구사항을 도출하고 논의한다 [4].

3.1. 일반 소프트웨어 지원 개요

HEIF은 Windows 10 (버전 1803 이상)에서 열람할 수 있으나 Microsoft 스토어에서 'HEIF 이미지 확장' 패키지를 추가로 설치해야 하며, 구글 크롬,

<표 1> 이미지 파일 형식의 기능 비교

구 분	HEIF	JPEG	PNG	GIF	BPG	WebP
기반 포맷	ISOBMFF	TIFF	-	-	-	RIFF
메타데이터	EXIF, XMP, MPEG-7	EXIF	-	-	EXIF, XMP	EXIF, XMP
다중 이미지 저장	○	×	×	○	○	○
썸네일 저장	○	○	×	×	○	×
무손실 수정	○	×	×	×	×	×

모질라 파이어폭스, MS 엣지와 같이 많이 사용되는 웹 브라우저에서도 안드로이드 9(Pie) 이상 버전에서 지원하며 기본으로 지원하지 않고 있다. 다수의 이미지 편집 소프트웨어에서도 아직 HEIF에 대한 열람이나 편집 기능에 대한 지원이 부족하다. 예를 들어, 어도비 포토샵(Photoshop)이나 라이트룸(Lightroom) 최신 버전은 macOS에서만 지원하고 Windows에서는 지원하지 않고 있으므로 HEIF를 편집하기 위해서는 macOS가 설치된 시스템을 사용하거나 Affinity Photo, GIMP, Paint.NET, Pixelmator, GraphicConverter, ImageMagick와 같은 이미지 편집 전문 소프트웨어를 설치해야 하는 상황이다. 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.

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 되지만, 윈도우 탐색기, 윈도우 사진 뷰어, 드롭박스의 웹 프리뷰, 포토샵(Mac), JPEG로의 변환 툴의 경우, 일반적으로 'pitm' 박스로 설정된 커버 이미지만 일반적으로 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: Windows 탐색기). 테스트에서 HEIF 파일에서 썸네일을 볼 수 있는 유일한 소프트웨어는 CopyTrans이며 내장된 썸네일을 표시할 수 있는 다른 소프트웨어는 없었다. 즉, 썸네일의 내용이 첨부한 마스터 이미지의 내용과 일치하지 않아도 되므로, 데이터를 숨길 수 있다.

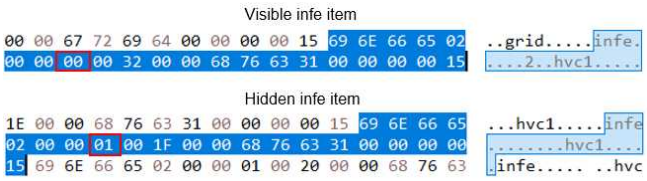
3.4. 무손실 수정 기능 및 보조 이미지 열람

HEIF의 주요 기능 중 하나는 기본 이미지의 무손실 수정으로, 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 편집 소프트웨어에서 그리드 보기, 자르기, 회전이나 반전 기능을 지원하는 경우는 거의 없었다. 또한 어떤 도구에서도 보조 이미지를 열람할 수 없었다. 일반 소프트웨어를 사용하는 조사관은 무손실 수정 기능으로 변경된 기본 이미지를 원래의 이미지라고 잘못 판단하기 쉬운 상황이므로, 이러한 특징을 이해하고 수정되기 이전의 이미지로 재구성할 수 있어야 할 것이다.

아이폰에서 촬영하고 자르기, 색감 조정 등의 편집이 반영된 사진 파일을 분석해보면 JPEG 형식으로 설정하고 촬영한 사진 파일은 수정한 내용을 포함하는 .AAE(Appleplist) 파일이 함께 생성된다. 만약 수정되지 않은 원래의 사진을 다시 추출하기 위해 변경 내용을 모두 되돌리면 HEIF로 추출되며 HEIF로 촬영한 사진 파일은 수정한 내용을 파일 내부에 자체적으로 포함시킬 수 있으므로 AAE 파일이 생성되지 않는다. 이러한 특징은 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 [그림 3]과 같이 'infe' 박스의 단일 비트 값에 의



(그림 3) 숨김 기능 설정 여부가 다른 HEIF의 'infe' 박스 비교

해 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.

4. HEIF 분석을 위한 포렌식 도구 요구사항

HEIF 사진 파일은 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. 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. 다음은 HEIF 파일을 분석하기 위한 포렌식 도구에 대한 요구사항이다.

- 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.

참고문헌

[1] Jyrki Alakuijala, "WebP Lossless Bitstream Specification.", 2012 (URL https://developers.google.com/ speed/ webp/docs/webp_lossless_bitstream_specification)

[2] G.K.Wallace, "The JPEG still picture compression standard," IEEE Transactions on Consumer Electronics, vol. 38, no. 1, pp. 18 - 34, 1992.

[3] ISO/IEC 23008-12:2017, "Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format".

[4] McKeown S, Russell G., "Forensic Considerations for the High Efficiency Image File Format (HEIF)." International Conference on Cyber Security and Protection of Digital Services (Cyber Security), IEEE. pp. 1-8, 2020.