

중앙기록물관리기관의 종이기록물 영구보존용 마스터 파일로서 JPEG 포맷의 표준화에 대한 연구

A Study on the Standardization of JPEG Format as a Long-Term Preservation Master File for Paper Archives in the Central Archives of Korea

강 현 민(HyenMin Kang)*

〈 목 차 〉

I. 서론	2. 문제점
1. 개요	III. JPEG 포맷의 특징
2. 연구배경 및 방법	1. JPEG 포맷 개요
II. 국내 디지털화 표준 정의 현황 및 문제점	2. 마스터 파일로서의 JPEG 포맷 주요 특징
1. 국가표준 및 중앙기록물관리기관 원내표준 현황	IV. 결론

초 록

중앙기록물관리기관이 제정한 원내표준 NAK/G 8-1:2013(v.1.0)에는 종이기록물의 영구보존용 마스터 파일 포맷으로 TIFF, PNG, PDF/A, JPEG 2000 파일 등을 명시하고 있다. 본 연구는 기존 포맷들과 함께 마스터 파일의 요건을 갖춘 JPEG 포맷을 추가하는 표준화 가능성에 대해 검토하였다. 먼저 JPEG 포맷이 갖고 있는 고유의 특징을 소개하고, 대표적인 마스터 파일 표준 포맷인 TIFF 포맷과의 특징 등을 비교분석과정을 통하여, JPEG 포맷이 마스터 파일 포맷의 기능상의 요건이라는 대전제를 만족하므로 중앙기록물관리기관의 원내표준 내 JPEG 포맷을 추가하는 표준화 필요성을 주장하였다. 본 연구의 의의는 최근 가장 범용적으로 생산되어 지고 있는 JPEG 포맷을 영구보존용 마스터 파일 기본 요건을 충족하는 원내표준 포맷에 추가함으로써, 디지털화 된 종이기록물 원본의 진본성, 범용성, 편의성, 실용성 및 호환성 등 다양한 편익을 이용자들에게 제공할 수 있는 표준 환경을 마련했다는 것이다.

키워드: JPEG 포맷, 마스터 파일, 영구보존용 파일, TIFF 포맷

ABSTRACT

It is defined in the internal standard of the NAK namely a NAK/G8-1:2013(v.1.) such as TIFF, PNG, PDF/A, JPEG 2000 as a Long-Term Preservation Master File for internal Standardization of Paper Archives in the Central Archives of Korea. This study reviewed the possibility that JPEG format could be a standard as a Long-Term Preservation Master File for internal Standardization of Paper Archives in the Central Archives of Korea. I pointed to the unique characteristics of JPEG file format and analysed and compared with TIFF format that is a representative master file. Through this study, I suggested that JPEG format also should be additionally included in the standard, as a master file, because it meets a qualification and functional requirements of master file and can guarantee the authenticity, reliability, integrity and usability. The meaning of this study is that JPEG format, be added to master file standard, can set up a environment that has lots of benefit in the domain of usability, compatibility to user as well as the authenticity, reliability, integrity.

Keywords: JPEG format, Master file, Long-term preservation file, TIFF

* 행정자치부 국가기록원 부산기록관 사서주사(kanghm@korea.kr)

•논문접수: 2016년 11월 20일 •최초심사: 2016년 11월 28일 •게재확정: 2016년 12월 21일

•한국도서관·정보학회지 47(4), 489-510, 2016. [http://dx.doi.org/10.16981/kliss.47.201612.489]

I. 서론

1. 개요

종이기록물 원본을 매체수록 시 기록물의 결손, 찢김, 곰팡이, 세균, 충해, 황변, 구김 등 물리적, 생물학적, 생화학적 오염·훼손정도 및 내용의 판독가능 여부와 같은 물성정보, 매체수록 순간(瞬間) 기록의 내용(內容), 형상(形像), 열화(劣化) 정도에 대한 최대한의 상태정보 및 원본 재현 속성정보 등을 확보할 수 있어야 한다. 이렇게 저장된 파일은 종이기록물 원본의 멸실, 훼손, 도난, 망실 시 원본을 대신할 수 있는 유일한 최상위의 우선권을 갖는 파일이며, 매체수록 프로세스 수행 시 기록의 진본성, 무결성, 신뢰성, 이용가능성을 보장하는 최적(最適)의 파일로서 영구보존용 마스터 파일(이하 “마스터 파일”, master file)이라고 부른다. 따라서 마스터 파일은 아날로그 형태의 종이기록물에 대해 스캐닝을 통해 획득되어진 종이원본 기록물 내 최대한의 내용과 작업시점의 최선의 형상(形像)을 전자적 신호로 인코딩하여 매체에 담아내어, 원본 종이기록물의 진본성, 무결성을 보장하는 범위 내에서 디지털로 변환된 종이 원본기록에 대한 최초의 순열 개체(the first sequential entity)¹⁾이다.

마스터 파일의 기능상의 요건은 이미지의 장기보존성, 신뢰성 및 고재현성을 지원하는 파일 포맷으로, 특허권 문제에 자유롭고, 오픈소스이며, 넓은 색상공간과 ‘비압축’(uncompressed)이며 파일구조 일정부분 내에 기술메타데이터(technical metadata)가 획득될 수 있어야한다(BCR 2008, 23).

따라서 파일포맷의 압축 기능 지원여부나 압축 시 색상 속성정보의 ‘손실’ 또는 ‘무손실’ 여부는 마스터 파일의 기능상의 전제요건은 아니다. 오히려 최초 타임라인에서 스캔 시 이미지 파일에 대한 ‘압축의 여부’가 마스터 파일의 전제요건으로 판단 기준에 되어야 할 것이다. 즉, 마스터 파일이라함은 최초 생산된 ‘비압축’의 이미지 파일을 말한다.

마스터 파일을 확보하면 이후 기록물관리시스템에 의한 파일포맷 변환 자동화 프로세스를 거쳐 열람·활용용 접근 파일(access file)로, 미리보기용 썸네일 파일(thumb nail file) 등으로 자동포맷 변환을 할 수 있기 때문에 본고에서는 마스터 파일만을 대상으로 표준화 가능성을 논하고자 한다.

2. 연구배경 및 방법

각 급 기록관, 특수기록관, 영구기록물관리기관은 종이기록물의 매체수록 시 중앙기록물관리

1) 본고에서는 최초의 순열 개체인 영구보존용 마스터 파일을 J0이라고 명명하고 그 이후에 스캐닝 또는 복제 생산된 원본사본파일들을 순차적으로 J₁, J₂ ... J_n 이하하였다.

기관의 표준을 참조하여 기관의 기록관리정책과 시스템 상황에 맞는 표준 포맷을 결정하게 된다.

중앙기록물관리기관은 원내표준으로 NAK/G 8-1:2013(v.1.0) ‘기록물 디지털화 지침 - 제1부: 종이기록물 및 사진·필름(v1.0)을 제정하였고, 마스터 파일의 표준 포맷으로서 ‘TIFF, PDF/A, PNG, JPEG 2000’ 포맷만을 원내 표준으로 명기하고 있다(국가기록원 2013). 중앙기록물관리기관 마스터 파일로서 TIFF, PDF/A, PNG, JPEG 2000 등 4종류의 파일 포맷만을 표준으로 권장하고 있으며 JPEG 포맷은 그 권장 포맷에서 누락되어 있다.

특히 TIFF 파일은 고품질 고용량의 디지털이미지 속성정보를 획득할 수 있어, 영구기록물관리기관의 영구보존용 장기보존 포맷인 마스터 파일로 받아들여지고 있다(Kenney and Rieger 2000).

NARA도 종이기록물 마스터 파일로서 RGB 색상 공간, 24비트심도, 300DPI 이상의 해상도를 갖는 비압축 TIFF 개정 6.0 포맷을 권장하고 있다(Steven Puglia 1998). BCR도 종이기록물의 마스터파일로 TIFF 파일 포맷을 권장하고 있다(BCR 2008, 26). Steven Puglia(2004)는 마스터 파일의 파일포맷으로 ‘비압축’ TIFF(uncompressed TIFF), 무손실 압축(lossless compressed) LZW알고리즘이 적용된 TIFF와 함께 JPEG 파일포맷도 권장하였다.

JPEG 포맷은 최초 스캐닝이나 촬영 이후 어플리케이션에서 교보정을 하면 할수록 손실 압축을 진행한다. 이는 최초 생산된 JPEG 포맷의 파일만이 진본성과 무결성을 보장하는 마스터 파일로서의 요건을 충족시킨다는 것을 의미한다. 또한 최근 대부분의 스마트 폰과 디지털 카메라 뿐만 아니라 다양한 규격의 스캐너들도 원본 재현성이 뛰어나고 초고해상도 화질, 넓은 색상공간, 깊은 비트심도 및 1,680만 트루컬러 등을 지원하는 JPEG 포맷을 광범위하게 지원하고 있다. 이처럼 JPEG 포맷의 디지털 파일의 폭발적인 사용빈도의 증가와 생산량 급증 등 급격한 환경의 변화가 일어났다. 이것은 JPEG 포맷이 사실상(de-facto)의 정지화상의 산업표준으로 진본성, 무결성, 보편성, 호환성, 범용성, 활용성 측면에서 표준화 가능성 관점에서 많은 장점을 지니고 있음을 의미한다.

따라서 새로운 IT기술의 발전과 정보이용환경의 급격한 변화로 기록관리 업무프로세스 내에서도 다양한 이미지 포맷이 논의되고 검토되어야 한다. 이런 관점에서 현재 모든 정보기기 및 소프트웨어에서 광범위하게 지원되고 활용되고 있는 JPEG 포맷의 종이기록물 마스터파일 표준화 검토가 필요한 시점에 왔다. 왜냐하면 표준은 선제적으로 기술변화의 방향성을 제시하는 기능과 동시에 현재 널리 확산된 기술을 표준으로 선정하여 기술 및 포맷 채택이 확정함으로써 기록관리 실무 수행에 있어 혼돈을 줄여 주는 기능을 하기 때문이다.

본 연구 목적은 “영구보존용 마스터 파일의 기본 요건을 충족하는 종이기록물 디지털화 파일 포맷이라면 마스터 파일로서의 표준 포맷으로 삼을 수 있다”라는 대전제 하에 JPEG 포맷을 정식 원내 표준으로 추가 명시하는 것이다.

4 한국도서관정보학회지(제47권 제4호)

본고의 연구방법은 첫째, 국내 디지털화 표준 정의에 대한 국가표준 및 중앙기록물관리기관의 원내표준의 현황을 살펴보고 그 문제점을 지적하였다. 둘째, JPEG 포맷의 기본구조 등을 설명하고 이 포맷이 갖고 있는 고유의 특징과 장점을 제시하였고, 마지막으로 대표적인 표준 포맷인 TIFF 포맷과 비교·분석과정을 통하여, JPEG 포맷이 중앙기록물관리기관의 원내표준으로서 선제 조건인 마스터 파일 포맷의 기본 요건을 만족한다는 것을 증명함으로써, JPEG 포맷이 다른 포맷들과 함께 마스터파일의 표준으로 추가할 것을 주장하였다.

본 연구에서는 TIFF 이외에 PDF/A, PNG, JPEG200 등과의 비교분석을 실시하지 않았다. 그 이유는 이들이 현재 중앙기록물관리기관의 원내표준으로 적시되어 있고, 본고의 연구 목적은 이러한 포맷들과 함께 JPEG을 ‘추가’ 적시하자는 것이기 때문이다. 기존 여러 포맷 중 마스터 파일 포맷으로 가장 광범위하게 채택되어 마스터파일로서의 대표성이 강한 TIFF 포맷과의 비교만을 시도하였다.

본 연구의 의의는 최근 가장 범용적으로 생산되어 지고 있는 JPEG 포맷을 영구보존용 마스터 파일 기본 요건을 충족하는 원내표준 포맷에 추가함으로써, 디지털화 된 종이기록물 원본의 진본성, 범용성, 편의성, 실용성 및 호환성 등 다양한 편익을 이용자들에게 제공할 수 있는 표준 환경을 마련한 것이다.

II. 국내 디지털화 표준 정의 현황 및 문제점

1. 국가표준 및 중앙기록물관리기관 원내표준 현황

종이기록물 원본에 대한 영구보존용 마스터 파일 포맷 종류와 관련하여 직접적으로 언급된 국가표준에는 KS X ISO TR 13028:2010 “기록의 디지털화 이행 지침”이 있다. 이 국가표준은 5.2장에서 ‘디지털화 기록의 마스터 파일은 안전한 저장환경에서 독립적이고 불가침한 상태로 유지되는 기록이며, 보통 당대와 이후 파생산출물에 이용 가능한 최상의 기술규격에 따라 이뤄진다’고 정의하고 있다.

6.2.3장에서 디지털화 이미지에 대한 마스터 파일의 기술표준을 정할 때 필요한 기초적 수준의 기술규격요건을 아래와 같이 정의하고 있다.

1. 현실적으로 지원될 수 있는 최상의 기술규격을 디지털화 과정에 적용하는 것이 좋다
2. 포맷은 공개소스(즉, 소유권이 없는)이거나 개방형 표준을 채택하여야 하며, 공공영역에서 사용할 수 있고 공표된 기술규격을 가지고 있고, 관련 부문에서 광범위하게 사용되는 것이 좋다
3. 포맷은 임베디드(Embedded) 객체를 포함하거나 해당 포맷의 특정 버전 이외의 외부 객체와 링크되지

않는 것이 좋다

4. 다수의 응용 소프트웨어와 운영체제에 의하여 지원되는 포맷이 바람직하다
5. 모든 이용자가 특정 소프트웨어 제품을 이용하지 못 할 경우, 사용이 용이한 조화용 플러그 인을 활용하여 읽을 수 있는 포맷이 바람직하다
6. 결정에 도움이 될 수 있도록, 접근 가능하고 제품으로부터 독립적인 기술 지식을 이용하는 것이 좋다
7. 계속적 관리와 필요할 경우 마이그레이션 역량을 보장할 수 있는 적절한 기술 지원이 이루어지는 것이 좋다
8. 마스터 사본은 지원할 수 있는 최상의 기술표준에 따라서 생산하는 것이 좋다
9. 마스터 사본은 안전한 저장소에서 불가침의 상태로 보유하는 것이 좋다
10. 파생사본은 업무요건(예를 들면, 인터넷상의 배포를 위한 썸네일)에 가장 편리한 포맷으로 제작할 수 있어야 한다.

※ 출처: KS X ISO TR 13028:2010 ‘기록의 디지털화 이행 지침’ 6.2.3

중앙기록물관리기관 원내표준은 NAK/G 8-1:2013(v.1.0) “기록물 디지털화 지침 -제1부: 종이기록물 및 사진·필름(v1.0)”이다.

중앙기록물관리기관의 종이기록물 디지털화와 관련된 원내 표준을 살펴보면 아래와 같이 ① 5.2장에서 디지털화 파일형식, ② 6.4 컬러를 포함한 문서, ③ 부속서 C(참고) 서비스용 이미지 생성을 위한 디지털화 기준 등으로 나누어 정의하고 있다.

5.2장의 디지털화 파일형식에 따르면 기록물 디지털화를 위한 파일 형식은 국내외 표준화 기관 등에서 발행되는 표준에서 지정한 형식을 사용하여야 한다. 즉, 디지털화 파일 형식은 표준으로 제정된 ‘오픈소스 기반의 포맷으로 무손실 압축’할 수 있으며 압축 알고리즘은 ISO, IEC 등과 같은 국제 표준화 기구에서 표준으로 공개되어 있는 방식을 따르는 것이어야 한다고 정의하고 있다.

이러한 조건을 만족하는 디지털화 파일 형식의 예로 다음과 같이 들고 있다.

- TIFF : Tagged Image/Interchange File Format. 다양한 컴퓨터 시스템에서 활용되는 파일 저장 형식이며, 보통 장기보존 또는 마스터(master) 디지털화가 필요한 경우에 사용된다.
- PNG : Portable Network Graphics. 미국 CompuServe가 중심이 되어 개발한 이미지 포맷으로 인터넷 환경에서 빠른 전송 및 이미지 재생이 가능하다.
 - 비고 : ‘ISO/IEC 15948:2004 Computer graphics and image processing -Portable Network Graphics(PNG): Functional specification’ 기능사양을 충족한 포맷
- PDF/A : PDF/Archival. 전자문서 디지털 보존을 위해 특별히 고안된 포터블 문서 포맷으로 장기보존이 가능하기 때문에 디지털 프린팅 그리고 사진 등에 널리 이용된다.
 - 비고 : ‘ISO 19005-1:2005 Electronic document file format for long-term preservation - Part 1 : Use of PDF1.4(PDF/A-1)’의 PDF 1.4버전 기능사양을 만족하는 포맷
- JPEG2000 : 정지 디지털 이미지의 압축을 위한 국제표준으로 JPEG 보다 높은 압축률과 이미지 품질을 얻기 위해 개발된 포맷이다. 무손실 압축을 지원하며 한 이미지에 여러 해상도의 이미지를 포함할 수 있다.

※ 출처: KS X ISO TR 13028:2010 ‘기록의 디지털화 이행 지침’ 5.2

6 한국도서관정보학회지(제47권 제4호)

다음으로 ‘6.4장 컬러를 포함한 문서(종이기록물)의 디지털화 세부기준’으로 디지털 이미지의 해상도는 최소 400ppi, 비트 심도는 최소 24비트 컬러를 적용하며, 무손실 압축방식의 공개된 파일 형식으로 스캔하여 저장한다고 정의하고 있다.

<표 1> 종이기록물 디지털화 세부 기준

원본 규격	구분	세부 기준
문자만으로 구성된 흑백문서	해상도 및 비트 심도	1비트 이진이미지: 600ppi 8비트 그레이 스케일: 400ppi ※ 문서의 긴 면을 기준으로 4,000~6,000 픽셀에 맞춤
위터마크, 회색 음영 및 그래픽을 포함한 문서	해상도 및 비트 심도	8비트 그레이 스케일: 400ppi ※ 문서의 긴 면을 기준으로 4,000~6,000 픽셀에 맞춤
컬러를 포함한 문서	해상도 및 비트 심도	24비트 컬러: 400ppi ※ 문서의 긴 면을 기준으로 4,000~6,000 픽셀에 맞춤

※ 출처: NAK/G 8-1:2013(v.1.0) “기록물 디지털화 지침 -제1부: 종이기록물 및 사진필름(v1.0)”

마지막으로, 부속서 C (참고) 서비스용 이미지 생성을 위한 디지털화 기준으로 각급기관이 열람 및 웹서비스 등에 활용하기 위해 디지털 이미지를 생성할 경우는 아래와 같은 기준을 적용할 수 있다. 이 기준은 국제표준에서 제시한 권장사항으로서 각 기관은 각각의 상황과 조건에 따라 알맞게 적용할 수 있다고 정의하고 있다.

<표 2> 서비스용 이미지 생성을 위한 디지털화 기준

기록물 유형	해상도	비트 심도
문자만으로 구성된 흑백문서	최소 300ppi	1비트(2도(bi-tonal))
위터마크, 회색 음영, 회색 그래픽을 포함한 문서	최소 600ppi	8비트 그레이 스케일
컬러를 포함한 문서	최소 600ppi	최소 8비트 컬러

※ 출처: “기록물 디지털화 지침 -제1부: 종이기록물 및 사진필름(v1.0) 부속서 C(참고) 서비스용 이미지 생성을 위한 디지털화 기준”

2. 문제점

국가표준 KS X ISO TR 13028:2010 “기록의 디지털화 이행 지침”은 디지털화 이미지 포맷 종류에 대한 구체적인 언급은 없고 마스터 파일에 대한 기본요건을 명시하고 있어, 이는 마스터 파일의 요건을 갖춘 다양한 포맷을 수용하고자 하는 의도로 보여진다.

그러나 마스터 파일의 기능요건을 보다 더 구체적으로 명시하고, 이 마스터 파일의 요건에 맞는 대표적 포맷의 사례를 좀 더 세분화하여 국가표준으로서 명시하였다면, 각 급 기록관 및 영구기록물 관리기관에서는 이를 참고로 각 기관의 현실에 맞는 마스터 파일 포맷을 결정하고 일관된 매체수록 작업을 수행할 수 있게 됨으로써, 기록관리 영역에서 국가표준 본연의

기능과 역할을 할 수 있을 것이다.

또한 현재 스캐닝 산업계, 카메라업계 뿐만 아니라 스마트폰의 광범위한 보급으로 폭발적으로 생산되고 있는 초고해상도 파일로서 사실 상의 표준(de-facto standard)인 JPEG 포맷에 대한 검토는 이뤄지지 않았다.

이러한 이유로 다른 영구기록물관리기관, 특수기록관 및 각급 기록관 등에서 마스터 파일 포맷 결정을 위해 참조할 수 있는 유일한 표준은 중앙기록물관리기관인 국가기록원이 내부 기록관리 업무수행 과정상 적용하고 있는 원내표준인 『NAK/G 8-1:2013(v.1.0) “기록물 디지털화 지침 -제1부: 종이기록물 및 사진·필름(v1.0)』만을 참조 할 수밖에 없는 현실이다. 중앙기록물관리기관은 모든 지방기록물관리기관, 특수기록관뿐만 아니라 헌법기관기록물관리기관까지 기록관리 업무를 총괄·조정하는 최고의 기관으로서 중앙기록물관리기관의 원내표준이라 할지라도 그 파급효과와 영향력은 대단히 크다.

중앙기록물관리기관의 원내표준에서 명시하고 있는 마스터 파일 포맷 종류와 관련하여 살펴보면 다음과 같은 문제점을 발견할 수 있다.

첫째, 마스터 파일 포맷으로 ‘무손실 압축’만을 지원하는 포맷만 허용하고 있다. 그 포맷으로 TIFF, PNG, PDF/A, JPEG 2000만 한정하고 있다. ‘무손실 압축’이란 압축을 할 때 최초 획득된 원본의 색상 속성정보를 손실 또는 제거 없이 파일의 크기를 줄이는 기능을 말한다.

그러나 마스터 파일은 원본 종이기록물을 최초 스캐닝하여 생성될 때 획득되어 진본성과 무결성이 확보된 최초의 이미지 파일이다. 최초 스캐닝 또는 촬영 작업을 할 때에는 어느 파일 포맷이든 색상의 속성정보를 최대한 ‘무손실’로 획득되도록 작업한다. 즉, 마스터 파일은 초고해상도, 고비트 심도, 넓은 색상공간, 일관된 이미지 품질관리 등 최대한 원본의 색상정보를 획득하여 최고의 색재현력을 확보한 최초의 파일이다. 이 때 압축을 적용 하든 안하든 마스터 파일로서의 요건 충족여부에는 영향을 미치지 않는다.

무손실로 획득된 마스터 파일에 무손실 압축을 하던 손실 압축을 하던 포맷별 압축 기능의 적용여부는 마스터 파일을 획득한 다음에 압축여부를 결정할 문제이다. 일단 마스터 파일 확보되면 다양한 손실 또는 무손실 압축률을 적용하거나 포맷 변환 등을 통해 열람·활용용 접근 파일 또는 썸네일 파일로 자유롭게 변환할 수 있으면 된다. JPEG 포맷도 다양한 압축률(0%~100%)을 적용할 수 있는데 이때 압축률을 0%로 하여, 소위 ‘무손실 압축’ 기능을 적용한 무손실 마스터 파일이 획득되어 질 수 있는 것이다.

따라서 마스터 파일은 1차적으로 ‘무손실’로 획득된 최초파일로서 ‘압축’ 적용여부는 다음 이미지 처리단계에서 적용여부를 결정할 문제이므로 JPEG 포맷도 그 요건을 갖춘 포맷으로 볼 수 있다.

둘째, 보존용, 열람·활용용, 썸네일용 등 디지털화 된 이미지의 명확한 생산목적 및 이용

기준이 없다.

기록물의 훼손을 방지하기 위해서는 기록의 생애주기에서 오직 단 한번 만 스캐닝하는 것이 바람직하다. 이런 관점에서 본다면 스캐닝 시 ① 영구보존용 마스터 파일을 기본적으로 생산하고, 그 원본 종이기록물의 역사적·학술적 가치 등에 따라 마스터 파일로부터 ② 열람·활용용 접근 파일 ③ 썸네일 파일 등으로 자동 생성하여 파일의 신뢰성과 이용가능성을 높일 필요가 있다. NARA의 디지털화 로드맵을 보면 이미지 파일 기능을 분류하여 생성하는 사례를 참고할 필요가 있다(Jill Hurtst-Wahl 2010).

셋째, ‘서비스용 이미지 생성을 위한 디지털화 기준’의 비현실성과 기술오류이다.

중앙기록물관리기관의 원내표준 부속서 C에는 각급 기관이 열람 및 웹서비스 등에 활용하기 위해 디지털 이미지를 생성할 경우의 기준이 명시되어 있는데, ‘서비스용 이미지 압축방식과 파일형식은 TIFF, PNG, PDF/A, JPEG 등 무손실 압축방식을 지원하는 공개된 파일 형식을 사용하여야 한다.’ 정의되어 있다. 앞서 표준기술에선 무손실 압축 포맷으로 JPEG포맷은 제외되었는데 서비스용 이미지(즉 열람·활용용 접근파일) 압축방식과 파일형식에는 JPEG 포맷이 무손실 압축방식을 지원한다고 명시하고 있어 논리적 일관성을 잃은 표준의 기술이다. 최초 개발된 JPEG 포맷은 무손실 압축방식을 지원하지 않는다. 다만, 무손실 비압축(압축률 0%)만을 지원할 뿐이다.²⁾

마지막으로 부속서 C 서비스용 이미지 생성을 위한 디지털화 기준의 출처가 “KS X ISO TR 13028: 2010 ‘기록의 디지털화 이행지침’ 부속서 E”로 되어 있으나, 실제 “KS X ISO TR 13028: 2010 부속서 E (참고)”는 ‘직원들을 위한 권장기능 목록’이다. 이 사항은 표준 출처 기술의 단순 오류로 보여 진다.

Ⅲ. JPEG 포맷의 특징

1. JPEG 포맷 개요

JPEG(Joint Photographic Experts Group)은 1986년 ISO/TC97/SC2/WG8와 CCITT SGVIII와 공동협력으로 조직되어 표준화 작업을 진행하여, 1992년9월 ITU-T Recommendation T.81로 1994년 ISO/IEC 10918-1로 정식 승인된 국제 표준이다(BCR 2008). JPEG 이미지파일 포맷으로 가장 일반적인 파일 포맷은 JPEG/Exif이며, JPEG/

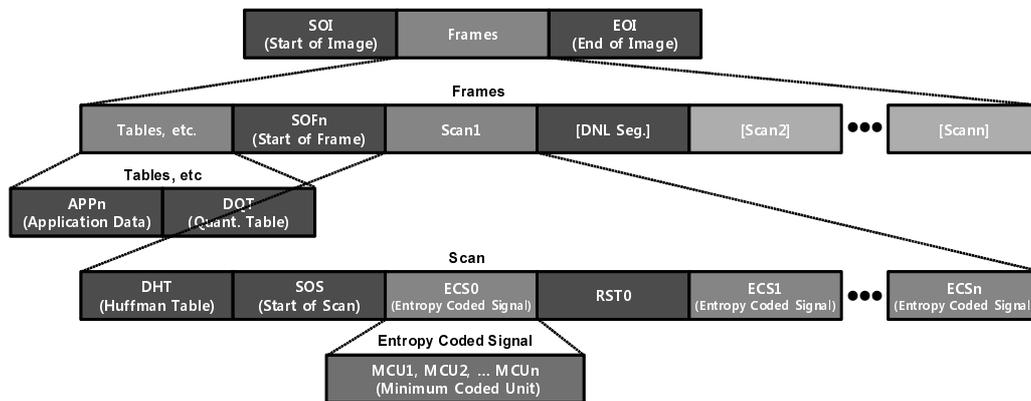
2) 1993년 무손실 압축표준은 정치화상전문가 그룹에 의해 JPEG 표준의 추가버전으로 개발되었다. 이 표준의 공식명칭은 ISO-14495-1/ITU-T.87이다.

JFIF 포맷도 인터넷 상에서 저장 및 전송되는데 사용된다. 이 두 가지 포맷유형은 잘 구분하지 않고 간단히 JPEG이라고 불러 지며, 일반적으로 사용되어지는 파일 확장자는 .jpg 또는 .jpeg이다. JPEG/JFIF는 65,535 x 65,535픽셀과 가로·세로비 1:1을 위해 4GB 픽셀까지 지원한다. JPEG 포맷은 다양한 색상영역을 갖는 자연 풍경 등 연속 톤으로 이뤄진 일상의 화상에서 인간의 시각이 정확하게 구분하기 힘든 부분에 대한 정보를 중심으로 DCT(Discrete Cosine Transformation) 압축알고리즘을 사용하여 화소를 제거하거나, 또는 그 영역을 하나로 합쳐 이미지의 크기를 줄이는 압축방식을 사용한다. <그림 1>은 JPEG 이미지 파일 구조를 나타낸 것인데 5단계 구조 중 4단계의 DHT(Huffman Table)의 호프만 알고리즘 즉, DCT 알고리즘을 이용하여 압축을 진행한다(강영재 2003).

정지영상(still image) 압축은 이미지 크기와 이미지 품질 사이에서 압축률(0%~100%)을 자유롭게 조정하여 압축품질의 범위를 사용자가 선택할 수 있다.

현재까지 JPEG 포맷은 손실 압축(lossy compression), 무손실 압축(lossless compression)을 모두 지원하는 단계까지 개발되었고, 이 두 가지 압축방식에 대해 살펴보자.

첫째, 손실 압축(lossy compression)은 디지털화 개체(예: J₀)을 압축하였을 때 원래 갖고 있던 정보량이 지정된 압축률에 따라 손실된 채로 압축(인코딩)되며 그 반대로 압축을 원래 크기(100%)로 디코딩하면 원래 갖고 있던 픽셀 정보는 잊어버린 채 이미지 품질이 원본에 비해 떨어지는 불가역적(irreversible)인 속성을 지닌다.



<그림 1> JPEG 이미지 구조

※ 출처: WIKIPEDIA 2016)

둘째, 무손실 압축(lossless compression)은 인접한 화소 값들의 차이를 손실 없이 부호화하는 것이다. 이를 ‘무손실 JPEG’이라하는데 1993년 정지화상전문가 그룹이 무손실 압축을 위해 만든 JPEG 표준의 신규버전이다. 연속적인 톤(continuous-tone) 이미지를 위한 무

손실 압축표준으로 공식적인 명칭은 ISO-14495-1/ITU-T.87이다. 화면 내에 존재하는 공간적 중복성을 제거하기 위해 화면 내의 화소 간 예측부호화 방식을 사용하며, 화질 저하로 인해 치명적인 결과를 방지해야 하는 의료 영상 등과 같은 분야에서 사용된다(박승철 외 2000). DICOM(Digital Imaging and Communication in Medicine) 표준에서 지정한 기본 JPEG 압축방법은 손실과 무손실에 따라 영상의 비트수가 <표 3>과 같이 나누어진다.

<표 3> DICOM의 JPEG 압축 전송구문 기본값

UID 이름	설명	DICOM 전송구문 UID
JPEG Baseline(Process 1)	Default Transfer Syntax for Lossy JPEG 8Bit Image Compression	1.2.840.10008.1.2.4.50
JPEG Extended (Process 4)	Default Transfer Syntax for Lossy JPEG 12Bit Image Compression	1.2.840.10008.1.2.4.51
JPEG Lossless, Non-Hierarchical, First-Order Prediction (Process 14) [Selection Value 1]	Default Transfer Syntax for Lossless JPEG Image Compression	1.2.840.10008.1.2.4.71

※ 출처: 박승철 외 2000

한편 미국의 산업표준국인 NTIS에서 정지화상 이미지 포맷에 대한 사용자 만족도를 조사하여 아래 <표 4>를 발표하였다(NTIS 2004).

<표 4> 주요 디지털화 이미지 포맷형식 장단점

포맷	지원성	사용성	장점	단점	비트심도	표준
JPEG	·모든 웹 브라우저 지원 ·저급에서 고급 이미지 응용프로그램에 모두 사용될 수 있음	·연속 톤의 피사체의 자연 색상을 갖는 모든 이미지 특화	·(무)손실 (비)압축 모두 지원 ·모든 컴퓨터 트루컬러 이미지 디스플레이 (1,680만 컬러) ·최고 100:1 압축과 점진적 디스플레이 (progressive display) 지원	·고압축률을 적용하면 이미지 품질 저하 ·직선 가장자리와 문자를 잘 다루기 어려움	·컬러데이터 픽셀 당 24비트	·IS10918-1 ·(ITU-TT.81) ·특허만료 (2006.10)
JPEG 2000	·W3C는 아직 지원 하고 있지 않음 ·지원을 늘리기 위해 많은 프로젝트가 진행 중	·웹 이미지, 시험인쇄, 의학이미지, 보안	·모든 JPEG의 장점과 손실 또는 무손실 압축 지원 ·위터마킹 같은 보안성 ·특징강화를 위한 웨이블릿 압축알고리즘 사용 ·상용 관심영역 코딩	·제한적인 지원	·픽셀 당 24단계까지 다양한 설정 가능	·JPEG그룹 ·특허 제한적 허용
TIFF	·많은 인쇄용 응용 프로그램이 지원 삽입 ·웹페이지에 할 수 없음	·고품질 인쇄용으로 대부분의 유형 지원	·무손실 압축 ·RGB 또는 CMYK, IBM 또는 Mac 비트 순으로 컬러 저장 가능	·지나치게 큰 파일 사이즈 ·웹 사용성과 보존성에 제한적	·1-64비트 부호불입정수형, 부호 없는 정수형 32 또는 64비트 빅 ICC 부동소수점 지원	·TIFF 6.0 ·2004년 특허 만료

※ 출처: NTIS 2004

위 <표 4>에서 보는 바와 같이 JPEG 포맷은 모든 웹 브라우저를 지원하고 “(무)손실 (비) 압축” 기능을 지원하고, 기록물과 같이 연속 톤을 지닌 자연의 색상에 특화되었으며, 특히 특허가 만료되어 상업적 종속성 없이 자유롭게 사용할 수 있는 포맷이다.

JPEG2000은 손실압축의 대표적인 JPEG 포맷의 단점을 극복하고자 나온 파일포맷으로서 웨이블릿 압축알고리즘을 사용하여 압축 시 손실이 발생하지 않는다. 그러나 JPEG 그룹의 특허 문제가 잠재적으로 존재하고 있어 상업적 종속성이 발생할 우려가 있으며, W3C에서 지원하지 않는 포맷으로 플러그인 기능이 지원되지 않으며, 웹페이지에 이미지 삽입이 불가능하며 인터넷 상에 자유롭게 전송될 수 없다. 또한 이미지를 모니터에 현시하려면 상업적으로 개발된 전용코덱이 임베디드된 어플리케이션을 구동해야 한다(NTIS 2004).

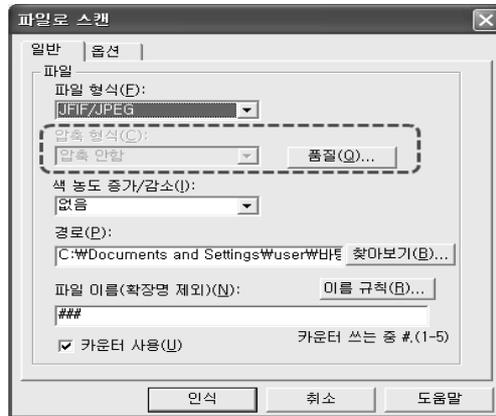
비록 JPEG2000이 손실압축이라는 ‘단점’이 아닌 고유한 ‘특성’이라 할 수 있는 JPEG 포맷에 대해 단점의 ‘극복’이 아닌 ‘무손실’ 압축이 가능한 새로운 버전의 형태로 보아야 할 것이며, 이는 마스터 파일의 기본 요건 중 하나인 ‘비압축’과는 무관한 새로운 독립된 기능을 하는 파일포맷이다.

BCR은 디지털화 개체를 생산하는데 있어서 마스터 파일(Master Files), 열람·활용용 접근 파일(Access files), 썸네일 파일(Thumb nail files) 세 가지 종류의 파일 생산을 권장하고 있다. 그 중 마스터 파일은 다양한 버전의 파일사이즈 추출과 다른 포맷 전환이 가능해야 하며, 또한 풍부한 정보, 연구적 품질 대체물, 미래의 요구를 수용할 수 있으며 다양한 응용프로그램에서도 작동되어야 하며, 집중적인 품질관리가 이뤄져야 한다는 것을 강조하고 있다(BCR 2008, 23).

2. 마스터 파일로서의 JPEG 포맷 주요특징

가. JPEG 무손실 비압축

국내 디지털화 마스터 파일 포맷의 기본요건으로 “무손실 압축”을 지원하는 포맷으로 정의하고 있다. 가장 대표적인 무손실 압축방식을 지원하는 TIFF 포맷의 ‘무손실 압축’방식은 LZW라는 무손실 압축알고리즘을 사용하며, 파일을 보정하거나 변형하였어도 다시 원래의 이미지 크기로 되돌렸을 때 이미지 품질의 손실이 없다. 즉, 이미지를 이루고 있는 화소의 색상의 속성정보가 사라지지 않는다. 따라서 이 포맷은 처음 보유했던 다양한 색상, 명도, 채도 등 색상 속성정보를 그대로 갖고 있으므로 사진작가나 그래픽디자이너, 인쇄출판업자처럼 이미지를 다양하게 처리하고 조작하고 변형하여 인쇄산업에서 가장 선호하는 파일 포맷 중 하나이다.



<그림 2> 스캐너 응용프로그램 내 JPEG 압축 캘리브레이션 설정

JPEG ‘무손실 비압축(압축률 0%)’가 갖는 의미는 무엇인가? <그림 2>는 JPEG 포맷으로 마스터 파일 획득을 위해 스캐너 응용프로그램에서 JPEG DCT 압축알고리즘에 대한 사용여부를 설정하는 화면³⁾이다. 일부 응용프로그램에서는 압축률 0%~100% 형식으로 또는 이미지 품질로 강-중-약 등 JPEG의 압축률 캘리브레이션을 자유롭게 설정하도록 이용자 인터페이스를 제공하고 있다. <그림 2>는 JPEG 포맷으로 스캐닝을 하되 압축률 0% 즉 비압축으로 스캐닝 캘리브레이션 값을 설정하는 화면이다. TIFF 포맷과 동일하게 JPEG 포맷 이미지는 화소의 손실이나 제거 없이 최초 모든 색상의 속성정보를 그대로 획득하게 된다. 즉 JPEG은 ‘비압축’ 하면 색상정보의 손실이 발생하지 않아 모든 색상정보를 갖고 있는 마스터 파일의 요건을 만족한다. 이는 TIFF 무손실 압축방식과 JPEG 무손실비압축 방식으로 생산된 파일은 마스터 파일로서의 요건에 동일하게 부합한다는 것을 의미한다.

따라서 이미지 품질로서 모든 디지털화 규격이 동일한 환경에서 TIFF의 무손실 압축파일이나 JPEG의 무손실 압축률 0%를 적용한 마스터 파일은 모두 색상 속성정보의 손실 없이 종이기록물 원본의 원형을 동일하게 재현해 낼 수 있다.

또한 JPEG 포맷의 손실 압축방식의 가장 큰 장점 중에서 열람·활용용 접근파일과 썸네일 파일의 관점에서 본다면 인간의 시각이 감지할 수 없는 이미지 속성정보에 대한 놀라운 손실 압축 알고리즘 기능을 가지고 있어 이미지의 크기와 이미지의 품질을 탄력적으로 조정할 수 있다. 따라서 JPEG 포맷은 열람용 디지털화 개체나 썸네일 디지털화 개체로 자유롭게 일괄 변형할 수 있고, 인터넷을 통한 기록정보 원본이미지 전송에 적합한 포맷을 손쉽게 생산해 낼 수 있는 장점이 있다.

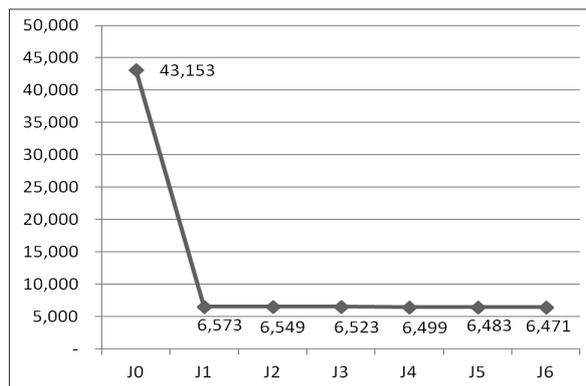
3) 이미지 품질을 좌우하는 중요한 요소 중 하나인 해상도 및 비트 심도 관련 설정 화면은 다른 메뉴에 있다.

나. 진본성 : JPEG포맷의 불가역성

JPEG의 무손실 비압축 방식으로 생산된 최초 마스터 파일 J₀ 파일만이 최고의 우선적 진본성을 확보한다. 즉, J₁에서 다시 J₀로 되돌릴 수가 없는 불가역적, 비영속적인 이미지 품질을 갖는 특징을 갖고 있다.

예를 들어 J₀ 파일을 어떠한 이미지 응용프로그램으로 열고 바로 저장하거나, 다른 이름으로 저장할 때, 응용프로그램은 DCT 손실 압축알고리즘이 내부적으로 자동으로 작동하여 J₁ 파일을 생산하게 된다. 이때 J₁파일은 이미 J₀파일 자체에 Exif와 같이 임베디드 된 메타데이터들이 이미 변경되었을 뿐만 아니라 본래 갖고 있던 색상의 속성정보까지 상실한 파일이다. 만약 이미지 처리 응용프로그램에서 J₀파일을 열기와 저장하기를 끊임없이 반복한다면 J₀파일은 점점 이미지의 해상도가 떨어져 J_n에서는 마스터 파일로서의 본래의 색 재현력을 잃게 된다. 기록물관리시스템 내 J₀에 임베디드 된 최초의 메타데이터가 J₁파일 자체에 임베디드 된 메타데이터 속성 값이 서로 상이한 점이 발견된다면 J₀파일과 J₁에 대한 진본성 판단여부에 중요한 단서를 제공하게 된다.

<그림 3>은 JPEG 포맷의 불가역적 이미지 특징을 잘 설명하고 있다. 특정 시점과 장소에서 생성된 JPEG 마스터 파일 J₀은 이미지 응용프로그램을 통해 최초로 파일을 열고 보정, 변형 등의 수정작업을 마친 후나 단순히 파일을 열고 이를 다른 이름으로 저장하려 할 때, 이미지 처리 응용프로그램은 임의의 설정 값에 따라 DCT 압축알고리즘이 자동으로 작동되어 원시이미지 정보가 손실된 파일 J₁이 생성된다. 여러 번의 수정과 저장과정을 반복 할수록 JPEG 이미지 크기(픽셀×픽셀)는 동일하나 이미지 품질과 파일 용량은 <그림 3>에서 보는 바와 같이 현격히 줄어든다.



<그림 3> 반복저장된 마스터 파일 J₀의 파일용량 변화

J₁→J₀로 역변환을 한다하더라도 J₁은 이미 J₀의 이미지 속성정보 값을 이미 상실한 상태에서 본래의 이미지 품질로 환원이 불가능한 불가역적(irreversible) 특성을 갖고 있다. J₁과

J_2, J_3, \dots, J_n 의 파일 용량과 품질이 서로 같지 않는 ($J_1 \neq J_2 \neq J_3, \dots$) 비영속성(non-consistent)이라는 두 가지 중요한 특징을 가지고 있다. 이처럼 JPEG 포맷의 불가역성, 비영속성이라는 이 두 가지 특징은 위변조를 원천적으로 불가능하게 하는 마스터 파일 J_0 이 갖는 중요한 특징이다.

결론적으로, 진본성과 무결성을 보장하는 시점의 개체라는 관점에서 본다면 최초 생산된 JPEG 무손실 압축률 0%를 적용한 J_0 이 종이기록물 원본이 갖고 있는 색상 속성정보를 최대로 보유한 유일한 마스터 파일이라고 할 수 있다.

이러한 JPEG 포맷의 '불가역적인 무손실 압축률 0%'라는 근원적 특징은 마스터 파일로서 진본성과 무결성을 보장하는 아이러니한 장점으로 작용하고 있다.

다. 무결성 : JPEG의 색상채널

JPEG 포맷의 색상 공간은 RGB와 CMYK 모드만 지원하고 레이어(layer)와 알파채널(Alpha channel)⁴⁾은 지원하지 않는다. 레이어와 알파채널을 이용하면 포토샵과 같은 이미지 처리 응용프로그램을 이용해 마스터 파일 이미지에 특정 문자나 조작된 이미지를 합성, 변경, 삽입이 가능하다(도유회 2010).

JPEG 포맷은 자유로운 변형·조작, 이미지 삽입 등 종이기록물 이미지 원본의 내용과 오염·훼손 상태 등을 교묘히 위변조할 수 있는 환경을 제공하지 않으므로 마스터 파일 포맷은 진본성과 무결성을 근원적으로 보장하는 특징으로 작용하고 있다.

라. 접근성 : 효율적 파일용량

JPEG 포맷은 마스터 파일로서 이미지 크기로 인한 저장 공간 점유의 효율성과 웹에서 기록정보 전달을 위한 범용성, 활용성, 호환성 등 이용가능성에 대한 장점이 있다.

<표 5>는 이미지 크기가 12,010×9,866 픽셀인 종이기록물 원본인 기록물에 대해 TIFF는 LZW 압축 알고리즘을 적용하지 않았고, JPEG은 DCT 압축알고리즘의 압축률을 0%로 설정하여 '원본 이미지 품질의 무손실'이라는 기준 하에 스캔하여 두 가지 포맷의 마스터 파일을 생성하였다.

사용한 스캐너 모델명은 WideTEK 25-600 Flachbett scanner이고 정밀도는 ±0.01mm, 광원은 White LED램프이다. 이미지 응용프로그램명은 Batch Scan Wizard를 사용했으며, 색상 공간은 sRGB, 밝기 127, 색대비 127, 비트심도 컬러 24bit, 감마값 22, 해상도 600dpi로 캘리브레이션을 설정하여 스캐닝 하였다. 작업 전 화이트 밸런스(white balance) 설정과

4) 알파채널(α channel)은 화소에 대한 색상표현의 데이터로부터 분리한 보조데이터로서 단색의 밝기, 농도 외에 직접 표시되지 않는 색 데이터를 넣는 채널을 말하며, 이미지를 서로 겹쳐서 보이게 하는 투명효과를 내기 위해 사용되는 화상마스크, 복구 및 화상 합성 등에 쓰인다.

〈표 5〉 마스터 파일 디지털화 개체 파일포맷(T₀, J₀)

구 분	TIFF	JPEG	비 고
이미지크기	12,010 × 9,866	12,010 × 9,866	
압축유무	비압축	비압축(0%)	비압축
색상공간	sRGB	sRGB	
밝기	127	127	
색대비	127	127	
획득값 R	0	0	
획득값 G	0	0	
획득값 B	0	0	
손실유무	무손실	무손실	
압축알고리즘	LZW	DCT	미적용
비트심도	컬러24bit	컬러24bit	
감마값	22	22	
해상도	600dpi	600dpi	
특허권	무	무	
파일크기	338MB	37.1MB	

스티칭(stitching) 작업을 통해 스캐너 작동환경도 동일하게 설정하였다.

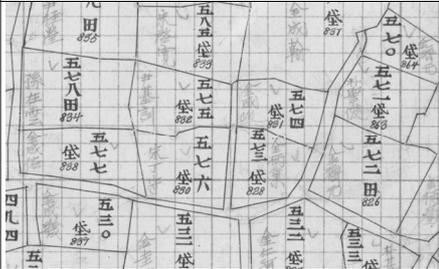
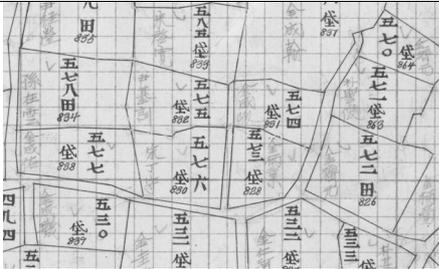
〈표 5〉를 보면 동일한 조건에서 생성된 마스터 파일로서 TIFF는 338MB, JPEG은 37.1MB의 파일크기를 나타냈다. 두 파일 포맷 간 파일크기가 약 10:1의 비율을 보였다. 같은 마스터 파일로서 기능을 수행한다면 보다 작은 저장공간을 차지하는 JPEG 이미지 포맷이 저장공간의 효율적 활용이라는 차원에서 보다 더 경제적인 포맷이라고 볼 수 있다.

마. 범용성 : 다양한 열람·활용용 어플리케이션 지원

마스터 파일로서 JPEG 포맷 파일을 이미지 처리 응용프로그램을 통해 모니터에 이미지를 현시하여 색 재현력과 기록 내용의 식별성과 시인성 정도를 육안으로 확인해 보았다. 〈표 6〉에서 보는 바와 같이, 동일한 조건에서 생산된 JPEG과 TIFF 파일 원본 이미지를 85%확대하여 1,691×1,067 픽셀 크기로 자르기를 한 후 색 재현력과 파일 용량을 살펴보았다. 두 포맷 모두 모니터 상에서 표현된 황변 등 열화 정도나 색 재현력, 그리고 확대 시 기록 내용의 식별성과 시인성이 모두 우수하였다. 마스터 파일을 생산하는 목적은 앞서 언급했듯이 스캐닝 시점 상의 종이기록물 열화 상태와 기록의 실제 내용 정보를 최대한 획득하는 것이었다. TIFF 포맷은 물론 JPEG 포맷도 모두 이러한 목적을 달성하기에 충분한 해상도와 이미지 품질을 나타내었다.

문제는 이미지 파일용량, 전용뷰어, PC 소요사양 부문에서 등이었다. 〈표 6〉에서 보듯이 파일 용량은 TIFF가 4.39GB, JPEG은 625KB로 약 7.2배의 용량차이를 보였다. JPEG은 인터넷 익스플로러, 크롬 등 일반 브라우저나 모든 기본 그래픽 응용프로그램에서 자유롭게

<표 6> TIFF와 JPEG 포맷의 마스터 파일의 열람·활용용 접근 파일 전환결과 비교

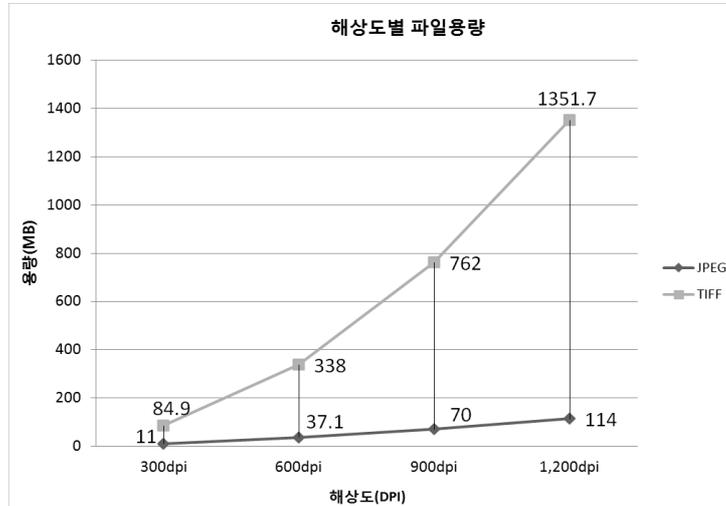
구 분	내 용	
파일포맷	TIFF	JPEG
파일용량	4.39GB	625KB
이미지크기	1,691 × 1,067 픽셀	1,691 × 1,067 픽셀
해상도	600 DPI	600 DPI
비트심도	컬러 24비트	컬러 24비트
확대율	85%	85%
색재현력	매우 만족	매우 만족
시인성	매우 만족	매우 만족
전용뷰어	필요	필요 없음
PC소요사양	고사양	중사양
스캔이미지		

열고, 닫고, 보정, 변형이 가능하였다. 반면 TIFF는 TIFF 코덱을 지원하는 전용뷰어 또는 플러그 인을 설치해야 모니터에서 현시할 수 있었고 인터넷 웹페이지에는 TIFF파일 자체로 삽입이 되질 않는다. 웹 페이지에 TIFF 파일을 삽입하려면 파일 포맷을 PNG, GIF 등의 이미지 포맷을 변경하여 삽입하여야 한다. 일부 응용프로그램에서는 TIFF 파일을 웹페이지에 삽입 시 자동으로 파일 포맷을 설정된 기본값(예: PNG)으로 변경하여 삽입되기도 한다. 이때 TIFF 포맷의 이미지 속성정보가 현저히 손상된다.

스캐닝 시점 상의 종이기록물 열화 상태와 기록의 실제 내용 정보를 최대한 획득한다는 것이라는 마스터 파일 생산 목적이라면 저장 공간 대비 비용, 웹 상에서 높은 이용성, 접근성, 범용성, 활용성 및 호환성 측면으로 보아 JPEG 포맷도 적합할 것으로 판단된다.

바. 경제성 : 저렴한 저장 및 유지관리 비용

이미지 크기가 12,010×9,866 픽셀인 종이기록물 원본 지적원도를 TIFF와 JPEG 포맷의 해상도별 이미지 파일용량 간 효율성을 좀 더 검토하기 위해, 동일한 스캐닝 작업환경에서 해상도를 300dpi, 600dpi, 900dpi, 1,200dpi로 다양하게 스캐닝을 실시하여 <그림 4>와 같은 결과 값을 얻었다.



〈그림 4〉 TIFF와 JPEG의 해상도별 파일용량 및 비용

구 분		300dpi	600dpi	900dpi	1,200dpi
JPEG	용량	11MB	37.1MB	70MB	114MB
	비용	5.5원	18.6원	35원	57원
TIFF	용량	84.9MB	338MB	762MB	1,351.7MB
	비용	42.5원	169원	381원	675.9원

〈그림 4〉에서 보듯이 TIFF와 JPEG 포맷의 4가지 해상도에 따른 파일의 용량은 평균 10:1로 JPEG 이미지 파일이 훨씬 적은 용량을 차지하였다. 그 이유는 TIFF는 원 이미지 픽셀 데이터를 그대로 보존하고 각각의 픽셀에 태그를 부여하여 값을 저장함으로써, 한 픽셀 당 태그정보가 상대적으로 많이 담길 수밖에 없는 파일의 구조적인 특징이 있기 때문이다. 반면 JPEG은 24비트심도 정보와 자연의 색상인 1,680만 트루컬러에 한 픽셀에 대한 색상 값만 담고 있다.

비용적인 관점에서 본다면 스토리지 단가가 1TB 당 53만원이라고 한다면 원가가 1MB 당 0.5원이다(강성호 2016). 기록물 1쪽 당 600dpi, 비트심도 24비트의 저장공간 비용은 JPEG은 18.6원, TIFF는 169원으로 약 9배 정도 차이가 난다. 이는 비용대비 저장공간 효율 성면에서 JPEG 포맷이 우수한 특성을 지니고 있다.

사. 지적재산권 : 특허문제

영구보존용으로 생산되는 마스터 파일은 모든 지적재산권 문제에 있어서 자유로워야 한다. 2002년 10월 27일 Forgent Networks사는 JPEG 기술의 특허권을 주장하였다. 그러나 2006년 2월 USPTO는 JPEG 특허가 이미 공표된 동일한 기술(prior art)에 기반 한 것임을 주장하였고, 이에 따라 2006년10월27일 미국 특허 20년 기간이 만료됨에 따라, 2006년11

월 Forgent사도 JPEG 표준 사용권에 대한 특허권리 주장을 포기하였다. 한편 Princeton Digital Image 사는 JPEG 이미지 압축표준이 '056특허'를 침해했다고 주장하여 수 많은 제조사와 도매상, 카메라 장비 제조사를 대상으로 소송 전을 벌였다. 그러나 그 특허는 General Electric사가 본래 소유하고 부여된 특허였다. 그 특허 또한 2007년11월에 만료되었다(Kawamoto 2009).

따라서 JPEG 포맷도 특허 종속성이 없어 영구기록물관리기관 뿐만 아니라 모든 개인, 기업, 단체에서 특허료 지불 걱정 없이 사용할 수 있다.

아. 무결성 : Exif 메타데이터

JPEG 포맷은 교환이미지파일형식(Exchangeable image file format)이라는 표준메타데이터 및 속성정보가 파일 자체에 임베디드 된다. Exif는 이미지, 소리 포맷을 정의한 부가된 태그 표준으로, 스캐너, 디지털 카메라 등에 사용된다.

Exif 표준 메타데이터 태그에는 아래와 같은 정보들을 저장한다.

- 시간정보: 이미지생성일자, 수정일자, 접근일자
- 스캐너 및 카메라 설정 값: 장비제조사명, 모델명, 기울기 값, 렌즈구경, 셔터스피드, 초점 길이, 미터모드, ISO 속도 정보 등
- 소프트웨어 정보: 미리보기 썸네일, 파일조작 소프트웨어 명
- 기술정보(Description Information)
- 저작권정보, 지리적 정보: 사진촬영 위치
- 스캐너·카메라 장비 고유 ID번호

Exif 데이터는 다양한 이미지 처리 응용프로그램이 이미지를 수정, 보정, 조작할 때 Exif 데이터의 정보를 인식하고 수정·저장한다. 현재 Exif 기술표준은 TIFF와 JPEG 이미지 파일에만 적용되고 있다. 특히 비트심도는 24비트를 유지토록 하고 있으며 Exif는 스캐너로 생산된 이미지에 대부분 사용되고 있다.

JPEG에 임베디드 된 Exif 메타데이터 속성값들은 파일 자체뿐만 아니라 기록물관리시스템 내에서도 획득되어 이중화 된 관리가 이뤄진다. 이는 마스터 파일과 기록물관리시스템 내의 기술메타데이터(technical metadata)와의 진본성, 무결성을 검증·확보하는 중요한 보안장치가 될 수 있다. 만약 한 곳의 메타데이터의 속성 값이 변경, 위조, 수정, 삭제 등이 발생하면 적절한 조치와 추적이 이뤄질 수 있도록 기본 데이터와 감사증적 값을 제공해 줄 수 있어, 진본성 및 무결성 강화에 중요한 특징으로 작용할 수 있다.

IV. 결론

종이기록물에 대한 마스터 파일 이미지 포맷 표준과 관련하여 제정한 국가표준 및 중앙기록물관리기관의 원내표준은 헌법기관기록물관리기관, 지방기록물관리기관, 특수기록관 및 각 급 기록관에게는 디지털 아카이빙 시대에서 명확한 마스터 파일 생산의 나침반 역할을 하는 매우 중요하고 영향력이 큰 표준이다.

JPEG 포맷은 앞서 살펴본 바와 같이 JPEG 포맷만이 갖는 다양한 특징들과 함께 국가표준에 명시한 마스터파일의 요건을 대부분 갖추고 있다. 또한 현재 스마트폰 카메라 파일포맷, 스캐너 파일포맷, 인터넷 웹페이지에 삽입된 이미지 포맷 등 마스터 파일요건 뿐만 아니라 열람·활용용 접근 파일, 썸네일 파일로도 가장 널리 사용되어 지고 있는 포맷이다.

따라서 JPEG 포맷은 BCR이 제시한 마스터 파일의 요건으로 지적재산권 문제가 없는 포맷 또는 오픈 소스, 넓은 색공간, 비압축(무손실), 파일 내 내장 기술메타데이터(Technical Metadata) 획득가능성 등 모두를 만족하고 있다.

결론적으로 말하자면, 마스터 파일로서의 JPEG 포맷은 종이기록물 원본의 내용과 형상을 최상의 해상도와 이미지 품질로 변환된 최적의 개체이다. JPEG 포맷은 무손실 비압축 방식과 불가역적 손실 압축 기능을 제공하고 있으며 다양한 기술메타데이터가 임베디드 되어 있다. 또한 특허권 만료로 지적재산권 분쟁의 가능성이 없으며, 열람용·썸네일용 파생포맷 파일 생산이 모든 이미지 응용프로그램에서 지원하는 호환성을 갖고 있으며, 융통성 있는 압축률 조절기능으로 이미지 크기와 파일 크기를 조절하여 비용대비 저장공간을 효과적으로 사용할 수 있는 경제성과, 모든 스마트폰 카메라, 스캐너, 웹 브라우저, 모니터, 프린터 등에서 광범위하게 사용되고 있는 범용성을 갖고 있어 사실상(de-facto)의 표준파일 포맷이다.

따라서 JPEG 포맷은 영구보존용 마스터 파일의 표준으로서 그 기능과 역할을 충분히 충족시킬 수 있는 파일이므로 추가 표준화 논의가 하루 빨리 이뤄져야 할 것이다.

참고문헌

- 강성호. 2016. 10. 8, 부산, [전화인터뷰].
- 강영재. 2003. 『인터넷 상에서 ICC 프로파일을 이용한 컬러 매칭』. 석사학위논문, 인하대학교 대학원 전기공학과.
- 국가기록원. 2013. 『기록물 디지털화 지침 -제1부: 종이 및 사진-필름(v1.0)』. 대전: 국가기록원.

- 노연숙. 2015. 『(사진가를 위한 포토샵) 디지털 컬러의 정석: 디지털 암실』. 서울: 디지털북스.
- 도유희. 2010. 사진 확장자에 따른 특성(JPEG, BMP, TIFF 등). 도유희 블로그.
<<http://yhdoh.tistory.com/entry/>>[인용 2016.12.12].
- 박승철, 박순만. 2000. 3가지 종류 CR 영상의 DICOM JPEG 압축률 비교와 시각적으로 손상이 되지 않는 비가역적 영상 압축에 대한 평가. 『대한PACS 학회지』, 2000(6): 91-95.
- 서은경. 2005. 디지털 자원의 보존 메타데이터 요소세트 개발에 관한 연구. 『한국정보관리학회지』, 22(3): 232-260.
- 이재영, 최주호. 2012. 전자기록 디지털컴포넌트의 포맷과 유효성 검증 기술 연구. 『한국기록관리학회지』, 12(3): 97-117.
- 아돌프 크놀. 2005. 도서관 희귀자료의 보존과 접근을 위한 디지털화의 표준화. 『한국기록관리학회지』, 5(2): 26-41.
- 한국기술표준원. 2010. 『기록물의 디지털화 이행지침』. 서울: 한국기술표준원.
- Adobe Developers Association. 1992. *TIFF Revision 6.0 final*. <<http://www.adobe.com/Support/TechNotes.html>>. [cited 2016.12.12].
- BCR' s CDP Digital Imaging Best Practices Working Group. 2008. *BCR's CDP Digital Imaging Best Practices Version 2.0*. <<http://www.BCR.org>>. [cited 2016.4.11].
- Besser, Howard. 1997. *The Changing Role of Photographic Collections with the Advent of digitization*. Washington: The Wired Museum. <<http://www.gseis.ucla.edu/~howard/Papers/garmil-eastman.html>>. [cited 2016.12.12].
- CISD. 2004. *Digital Media File Types : Survey of Common Formats*. Convergent Information Systems Division of NTIS. <<http://www.NTIS.org>>. [cited 2016.12.12].
- FLAIR. 2015. *Color Management System : chapter 1 CMS & ICC Profile*. Alberta Canada : Design Group Seminar. <<http://www.flairpackaging.com>>. [cited 2016.12.12].
- G. Parsons and J. Rafferty. 1998. *Tag Image File Format (TIFF) -F Profile for Facsimile*. Human communications Network Working Group. <<https://tools.ietf.org/html/rfc2306>>. [cited 2016.12.12].
- International Color Consortium. 1998. *Specification ICC.1:2001-04 : File Format for Color Profiles, [REVISION of ICC.1:1998-09]*. Washington: International Color Consortium.
- ISO/TC 130. 2002. *Graphic technology-Prepress digital data exchange-Tag image file format for image technology(TIFF/IT)*.
- Jill, Hurst-Wahl. 2010. *NARA's Technical Guidelines for Digitizing Archival Materials*

- for Electronic Access: Creation of Production Master Files—Raster Images*. Washington: NARA.
- Kawamoto, Dawa. 2005. “Graphics patent suit fires back at Microsoft.” *CNET News*. <<http://www.CNET.com>>. [cited 2016.12.12].
- Kenney, Anne R. and Rieger, Oya. 2000. *Moving Theory in to Practice: Digital Image for Libraries and Archives*. Research Libraries Group. <<http://www.rlg.org/preserv/mtip2000.html>>. [cited 2016.12.12].
- Light, sail. 2006. *USPTO: Broadest Claims Forgent Asserts Against JPEG Standard Invalid*. USPTO. <<http://www.groklaw.net/article.php?story=20060526105754880>>. [cited 2016.12.12].
- NARA. 2014. *Strategy for Digitizing Archival Materials for Public Access. 2015–2024*. NARA.
- Puglia, Steven. Reed, Jeffrey and Rhodes, Erin. 2004. *Technical Guideline for Digitizing Archival Materials for Electronic Access : Creation of Production Master Files—Raster Images*. <http://www.archives.gov/records_management/index.html> [cited 2016.12.12].
- Steven, Puglia. 1998. NARA Guidelines for Digitizing Archival Materials for Electronic Access. National Archives and Records Administration.
- Scanvec, Amiable. 2016. *SCANVEC AMIBLE: Color Profiler 5.0* <<http://www.scanvecamiable.com>> [cited 2016.12.12].
- USPTO. 2006. “USPTO: Broadest Claims Forgent Asserts Against JPEG Standard invalid.” *Groklaw.net* <<http://www.Groklaw.net>>. [cited 2016.12.12].
- WIKIPEDIA. 2016. *JPEG* <<https://en.wikipedia.org/wiki/JPEG>>. [cited 2016.12.12].
- WIKIPEDIA. 2016. *JPEG 2000* <https://en.wikipedia.org/wiki/JPEG_2000>. [cited 2016.12.12].
- WIKIPEDIA. 2016. *Tagged Image File Format* <https://en.wikipedia.org/wiki/Tagged_Image_File_Format>. [cited 2016.12.12].

국한문 참고문헌의 영문 표기

(English translation / Romanization of reference originally written in Korean)

- Adolph Knohl. 2005. “Standardizing Digitization for Preservation and Access to Rare Library Materials.” *The Journal of Korean Society of Archives and Records*
- Doh, You-Whae. 2010. *The characteristics of photo extensions(JPEG, BMP, TIFF*

- etc.). <<http:yhdoh.tistory.com/entry/>>. [cited 2016.12.12].
- Kang, Sung-Ho. 2016. 10. 8, Busan, [Interview].
- Kang, Young-Jae. 2003. *Color Matching over the Internet Based on ICC Profile*. M.A. thesis., Inha Univ, South Korea.
- Korean Agency for Technology and Standards. 2010. *The Implementation Guide for Records of Digitization*. Seoul: Korean Agency for Technology and Standards.
- Lee, Jae-Young. Choi, Joo-Ho. 2012. "Validation and the Format of the Electronic Record Digital Component Technology Research." *The Journal of Korean Society of Archives and Records Management*, 12(3): 97-117.
- Management*, 5(2): 26-41.
- National Archives of Korea. 2013. The Guide line of Archives Digiziation—the first section : paper, photograph and film(v1.0).DaeJeon: National Archives of Korea.
- Noh, Yeon-Sook. 2015. *(Photoshop for Photographer) the Code of Digital Color: Digital Darkroom*, Seoul: Digital Book Ltd..
- Park, Seung-Chul. Park, Soon-Man. 2000. Comparison of DICOM JPEG Compression ratio with 3 kinds of CR Images, and Assessment of Visually Lossless Irreversible Image Compression; Seoul: *The Journal of Korean PACS*. 2000(6):91-95.
- Seo, Yeun-Kyung. 2005. "A study on Preservation metadata for Digital Resources." *The Journal of Korea Society for Information Management*. 22(3): 232-260.