

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Seminarski rad u okviru predmeta „Računalna forenzika“ 2016/2017

TIFF format (.tif)

Mislav Granić

Zagreb, siječanj 2017.

Sadržaj

Uvod	3
Značajke i opcije	3
Osnovni TIFF	4
Višestраниčne sudatoteke (multiple subfiles).....	4
Trake.....	4
Sažimanje	5
Tipovi slika.....	5
Redosljed Byte-ova	5
Ostala TIFF polja	5
TIFF proširenja.....	5
Sažimanje	6
Tipovi slika.....	6
Stabla slika	6
Kvadratici (Tiles)	6
Ostala proširenja.....	7
Privatne oznake	7
Tip Internetskog medija (Internet Media Type).....	7
TIFF oznake sažimanja	8
Metapodaci u TIFF datotekama [2].....	8
Zaključak.....	10
Literatura	11

Uvod

Tagged Image File Format ili skraćeno TIFF ili TIF je računalni format datoteke koji se koristi za skladištenje rasterskih grafičkih slika. Popularan je među grafičkim umjetnicima, fotografima i u objavljačkoj industriji. TIFF format je široko podržan u aplikacijama za manipulaciju slike, aplikacijama za objavu i izgled stranice, u faksiranju, obradi teksta, optičkom prepoznavaju znakova itd.

Format je izrađen od strane Aldus Korporacije u svrhu desktop publishinga. Zadnja verzija je izdana 1992. nakon čega je ažurirana s Adobe sustavom. Izdano je nekoliko manjih proširenja formatu uključujući TIFF/EP (ISO 12234-2), TIFF/IT (ISO 12639), TIFF-F (RFC 2306) i TIFF-FX (RFC 3949).

TIFF je nastao kao pokušaj da se prodavači desktop skenera sredinom 1980-ih dogovore za zajednički format skenirane slike umjesto mnoštva različitih formata. U početku TIFF je bio samo binarni format slike (samo dvije moguće vrijednosti za svaki piksel), jer su desktop skeneri podržavali samo to. Kako su skeneri postajali napredniji, a tvrdi diskovi obilniji, TIFF se razvio tako da podržava slike u nijansama sive boje, a zatim slike u boji. Danas, zajedno s JPEG i PNG, TIFF je popularan format za slike visokih boja i dubina.

Značajke i opcije

TIFF je fleksibilan, prilagodljiv format datoteke za obradu slika i podataka u jednoj datoteci, zbog uključivanja oznaka za zaglavlja (veličina, definicija, raspored slika-podaci, primjenu sažimanja), koji definira geometriju slike. TIFF datoteka može biti npr. spremnik koji sadrži JPEG (lossy) i PackBits (lossless) sažete slike. Osim toga, TIFF datoteka može uključivati vektorski isječen put (okolne linije, izreze, okvire slika). Sposobnost za pohranu podataka o slikama u negubljevom (lossless) formatu čini TIFF datoteku korisnom arhivom za slike jer za razliku od standardnih JPEG datoteka, TIFF datoteka se koristeći negubljevu kompresiju (ili nikakvu) može uređivati i ponovno spremiti bez gubitaka kvalitete. To nije slučaj kada se koristi TIFF kao spremnik koji sadrži sažeti JPEG. Ostale TIFF opcije su slojevi i stranice. TIFF nudi mogućnost koristiti LZW sažimanje kao vrstu sažimanja podataka bez gubitaka. Korištenje ove mogućnosti bilo je ograničeno zbog patenata na LZW tehniku do njihovog isteka 2004. godine.

TIFF 6.0 specifikacija sastoji se od sljedećih dijelova:

- Uvod (sadrži informacije o TIFF administraciji i korištenju privatnih polja i vrijednosti, itd)
- Osnovni TIFF
- TIFF proširenja
- Dodaci

Osnovni TIFF

Kada je TIFF uveden, njegova proširivost izazvala je probleme s kompatibilnošću. Fleksibilnost u kodiranju izazvala je vic da TIFF znači „Thousands of Incompatible File Formats“. Da bi izbjegao te probleme, svaki TIFF čitač je bio dužan pročitati osnovni TIFF. Između ostalog, osnovni TIFF ne uključuje slojeve, ili sažete JPEG ili LZW slike. Osnovni TIFF je formalno poznat kao „TIFF 6.0, Part 1: Baseline TIFF“.

Slijedi nepotpun popis potrebnih značajki osnovnog TIFFa:

Višestranične sudatoteke (multiple subfiles)

TIFF čitači moraju biti spremni za više/višestranične slike (subfiles) po TIFF datoteci, iako se ne zahtijeva da se zapravo išta radi sa slikama nakon prve.

Postoji više od jednog direktorija sa slikama (Image File Directory, IFD) u TIFF datoteci. Svaki IFD definira sudatoteku. Jedna od primjena sudatoteke je opisivanje povezanih slika, kao što su stranice duplikata dokumenta. Osnovni TIFF čitač nema potrebu čitati ostale IFD osim prvog.

Trake

Osnovna TIFF slika se sastoji od jedne ili više traka. Traka (ili vrpca) je pododjeljak slike sastavljene od jednog ili više redova (horizontalni red piksela). Svaka traka može biti sažeta neovisno o cijeloj slici, i svaka počinje na granici byte-a.

Ako visina slike nije ravnomjerno djeljiva s brojem redaka u traci, zadnja traka može sadržavati manje redaka. Ako su izostavljene oznake definicija traka, pretpostavlja se da cijela slika sadrži jednu traku.

Sažimanje

Osnovni TIFF čitači moraju omogućiti sljedeće tri sheme sažimanja:

- Bez sažimanja
- CCITT Group 3 1-dimenzionalni modificirani Huffman RLE
- PackBits sažimanje – oblik run-length kodiranja

Tipovi slika

Tipovi slika osnovnog TIFF-a su: dvorazinska, tonovi sive, paleta boja, i RGB slike u boji.

Redoslijed Byte-ova

Svaka TIFF datoteka započinje s dvobitnim indikatorom o redoslijedu byteova: „II“ za little-endian ili „MM“ za big-endian. Sljedeća dva bytea označavaju broj verzije formata, koji je oduvijek bio 42 za svaku verziju TIFFa. Za sve riječi, dvostruke riječi, itd., u TIFF datoteci se pretpostavlja da su u označenom poretku byteova. TIFF 6.0 specifikacija navodi da usklađeni (compliant) TIFF čitači moraju podržavati oba tipa poretka byteova.

Ostala TIFF polja

TIFF čitači moraju biti spremni na susret i ignoriranje privatnih polja koja nisu opisana u TIFF specifikaciji. TIFF čitači ne smiju odbiti pročitati TIFF datoteku ako neka opcionalna polja ne postoje.

TIFF proširenja

Mnogi TIFF čitači podržavaju dodatne oznake uz one iz osnovnog TIFF-a, ali svaki čitač ne podržava svako proširenje. Kao posljedica toga, značajke osnovnog TIFF-a su postale najmanje zajedničke karakteristike TIFF formata. Značajke osnovnog TIFF-a su se proširene u TIFF proširenja, ali se proširenja mogu također odrediti u privatnim oznakama.

TIFF proširenja su formalno poznata kao „TIFF 6.0, Part2: TIFF Extensions“.

Slijedi nekoliko primjera TIFF proširenja definirana u TIFF 6.0 specifikaciji:

Sažimanje

- CCITT T.4 dvorazinsko kodiranje
- CCITT T.6 dvorazinsko kodiranje
- LZW shema sažimanja
- Sažimanje bazirano na JPEG koristi DCT (Discrete Cosine Transform (Diskretna Kosinusna Transformacija)) uvedenu 1974. godine od strane N. Ahmed, T. Natarajan i K. R. Rao.

Tipovi slika

- CMYK slike
- YcbCr slike
- HalftoneHints
- Tiled slike
- CIE L*a*b* slike

Stabla slika

Osnovna TIFF datoteka može sadržavati niz slika (IFD). Tipično, sve slike su povezane, ali predstavljaju različite podatke, kao što su stranice dokumenta. Kako bi se izričito podržavalo više pogleda na iste podatke, uvedena je SubIDF oznaka. To omogućuje slikama da budu definirane po strukturi stabla. Svaka slika može imati niz djece od kojih je svako dijete zasebna slika. Tipična upotreba je pružanje sličica (thumbnails) ili nekoliko verzija slike u različitom prostoru boja.

Kvadratići (Tiles)

TIFF slika se također može sastojati od niza kvadratića. Svi kvadratići u slici imaju iste dimenzije i mogu biti sažeti neovisno o cijeloj slici, slično kao trake. Kvadratne sličice su dio TIFF 6.0 Part 2: TIFF proširenja, tako da podrška za kvadratne sličice nije potrebna u osnovnim TIFF čitačima.

Ostala proširenja

Prema TIFF 6.0 specifikacijama, sve TIFF datoteke koje koriste predložena TIFF proširenja koja nisu odobrana od strane Adobe tvrtke kao dio osnovnog TIFFa, ne bi trebale biti zvane TIFF datoteke ili bi trebale biti označene na neki način kako ne bi bile pomiješane s uobičajenim TIFF datotekama.

Privatne oznake

Programeri mogu podnijeti zahtjev za blokom „privatnih oznaka“ kako bi omogućili uključivanje vlastitih informacija u TIFF datoteku bez uzrokovanja problema miješanja među datotekama. TIFF čitači trebaju ignorirati oznake koje ne prepoznaju, i registrirane privatne oznake programera se garantirano ne smiju podudarati s tuđim oznakama ili sa standardnim skupom oznaka definiranim u specifikaciji. Privatne oznake su numerirane u rasponu 32, 768 nadalje.

Privatne oznake su rezervirane za informacije značajne samo nekim organizacijama ili za eksperimente s novom shemom sažimanja unutar TIFFa. Na zahtjev, TIFF administrator (trenutno Adobe) će izdvojiti i registrirati jednu ili više privatnih oznaka za organizaciju, kako bi izbjegli mogući sukob s drugim organizacijama.

Organizacije i programeri su obeshrabreni odabirati svoje vlastite oznake sami jer to može dovesti do ozbiljnih problema s kompatibilnošću. Međutim, ako postoji mala ili nikakva šansa da će TIFF datoteke izaći iz područja privatnog okruženja, organizacije i programeri se potiču da koriste TIFF oznake u ponovno iskoristivom rasponu „65,000 – 65, 535“. Nema potrebe za kontaktiranjem Adobe-a ako se koriste brojevi u tom rasponu.

Tip Internetskog medija (Internet Media Type)

Vrsta MIME slike/tiff (definirana u RFC 3302) bez aplikacijskog parametra se koristi za osnovne TIFF 6.0 datoteke ili za naznaku da nije potrebno identificirati specifični podskup TIFF proširenja. Opcionalni „aplikacijski“ parametar (Primjer: Content-type: image/tiff; application =foo) je definiran za sliku/tiff kako bi se identificirao određeni podskup TIFF proširenja za kodirane podatke slike, ako je poznat. Prema RFC 3302, specifični TIFF

skupovi ili TIFF proširenja korištena u aplikacijskim parametrima moraju biti objavljeni kao RFC.

Vrsta MIME slike/tiff-fx (definirana u RFC 3949 i RFC 3950) temelji se na TIFF 6.0 s TIFF Tehničkim bilješkama TTN1 (Stabla) i TTN2 (Zamjena TIFF/JPEG specifikacija). Koristi se za Internet faks kompatibilan s ITU-T Recommendations za Group 3 crno-bijeli, nijansi sive (grayscale) i faksove u boji.

TIFF oznake sažimanja

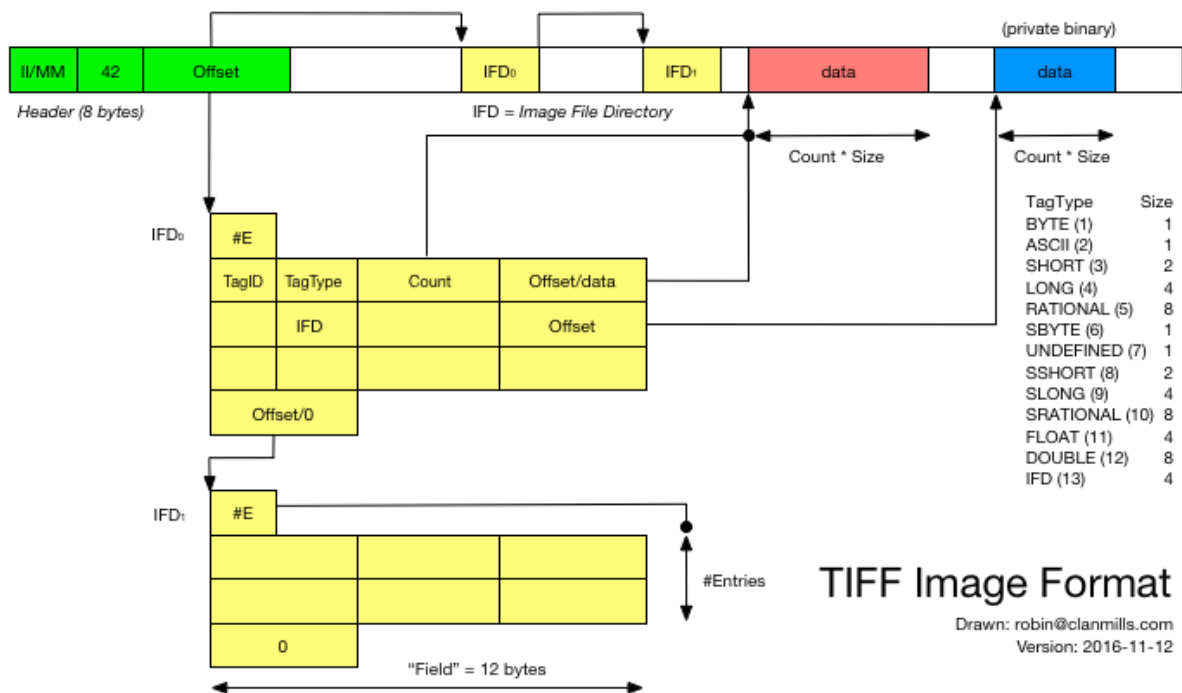
TIFF oznaka 259 (010316) pohranjuje podatke o metodi sažimanja. Uobičajena vrijednost je 1 = bez sažimanja.

Većina TIFF pisara i čitača podržava samo neke TIFF sheme sažimanja.

Metapodaci u TIFF datotekama [2]

TIFF je kao spremnik. Vrlo je fleksibilan i može se nositi s više stranica, različitim prostorom boja i konfiguracijama okvira, kao i metapodacima. TIFF spremnik je toliko fleksibilan da se koristi kao struktura za većinu sirovih formata uključujući Adobeov DNG. Dodatno, TIFF spremnik se koristi za spremanje metapodataka koji se nalaze unutar Exif blokova ugrađenih u JPEG i PNG datoteke.

Postoji 64-bitna verzija TIFF-a koji se zove BigTiff, ali trenutno nije podržan od strane exiv2.



Slika 1. format TIFF datoteke [2]

Format se sastoji od zaglavlja od 8 byteova od kojih su 2 bita za oznaku endian zastavice (II ili MM), 2 bytea „magic number“ (42) i 4 bytea pokazivač na direktorij slike (IFD).

Direktorij slike (Image File Directory) ima zaglavlje od kojih 2 bytea sadrže duljinu direktorija iza kojih slijedi 12 bytea „polja“ ili „oznaka“. Zatim slijedi 4 bytea koja pokazuju na sljedeći direktorij ili nula u slučaju da nema sljedećeg.

Oznaka se sastoji od 12 byteova: „TagID, TagType, Count and Offset“, koji imaju 2, 2, 4 i 4 bytea. TagID definira svrhu zapisa (širinu, visinu, prostor boja, itd.) i TagType definira format datoteke. Count i Offset se koriste za sadržavanje sirovih podataka ili za osiguravanje pokazivača u datoteci gdje čitati sirovi podatak za tu oznaku.

Zaključak

Sposobnost za pohranu podataka o slikama u negubljevom formatu čini TIFF datoteku korisnom arhivom za slike jer za razliku od standardnih JPEG datoteka, TIFF datoteka se koristeći negubljivu kompresiju može uređivati i ponovno spremati bez gubitaka kvalitete. Osim toga, korisna je u forenzici jer se puno podataka iz nje može izvući, no nije najčešće korišteni tip datoteke za slike.

Literatura

[1] URL, wikipedia [https://en.wikipedia.org/wiki/TIFF#Part_2: TIFF Extensions](https://en.wikipedia.org/wiki/TIFF#Part_2:_TIFF_Extensions)

[2] URL, exiv2 http://dev.exiv2.org/projects/exiv2/wiki/The_Metadata_in_TIFF_files