(In English)



W coraz większym tempie przestawiamy się na fotografię cyfrową. To, co było kilkanaście lat temu nowinką staje się standardem, a aparaty na film staja się rzadkością. Możliwość natychmiastowego sprawdzenia wyniku, powszechność zapisu obrazu w telefonach, tabletach, coraz tańsza pamięć cyfrowa i sprzęt fotograficzny powoduje, że robimy teraz o wiele więcej zdjęć. Jednocześnie jednak fotografia stała się czymś bardzo przejściowym. Kiedyś wklejało się zdjęcia do albumów, kolekcjonowało w pudełkach, dziś siedzą one jako pliki na dysku komputera, a gdy dysk padnie (wszystkie dyski to czeka), nagle tracimy nasze zasoby. Pisałem już wcześniej o osobistych archiwach cyfrowych , tym razem bardziej szczegółowo o tym, jak zapakować i przechować obraz cyfrowy.

Zapis obrazu to nie tylko zdjęcia. Skany dokumentów w archiwum to także zapis cyfrowy, który powinien wiernie odzwierciedlać oryginalny dokument. Jak wybrać najlepszy format i sposób zapisu tak, aby przetrwał dla następnego pokolenia, aby nasze wnuki mogły oglądać albumy dziadków, a archiwa przechowały bezcenne już (bo papier się rozpadł) obrazy archiwaliów? Zapisany obraz przechowujemy w opakowaniu zwanym plikiem (file). W dalszym ciągu będzie o formatach tych kopert, do których wkładamy zdjęcia - plików komputerowych, kompresji i metadanych a także przekładaniu obrazu z jednej koperty do innej (konwersji).

Aparat cyfrowy to imitacja siatkówki oka. Imitacja ta nie jest zbyt dobra, gdyż oko pracuje zupełnie inaczej niż kamera, ale można to traktować jako przybliżenie. Obraz - zbierany przez soczewkę albo skanowany na płaskim skanerze, dzielony jest na małe elementy, zwykle kwadratowe (pixle), a kolor zapisywany jest osobno dla każdego kwadratu. Notowane są dane dla trzech kolorów, choć innych niż w oku człowieka. Po przetworzenu danych mamy więc prostokątną matrycę, a w każdej jej komórce zapisany jest kolor. Zapis charakteryzuje się wymiarami w pixlach (wysokość i szerokość) i trzecim wymiarem (zwanym głębokością), podającym jak szczegółowo rozróżnione są kolory. Najbardziej popularny zapis używa 8 bitów na każdy z 3 kolorów “podstawowych”(razem 24), co daje możliwość zapisania ponad 16 milionów odcieni. Tak zapisane dane pakowane są do plików w jednym z formatów zwanych rastrowymi.

Kryteria wyboru formatu

Do niedawna nie trzeba było żadnego narzędzia do obejrzenia malowidła, zdjęcia czy przeczytania książki. Dziś coraz częściej musimy użyć tego sprzętu (komputera lub urządzenia które spełnia tę sama funkcję, choć nazywa się inaczej - telefon, tablet itp.). Co gorzej, spotykamy się z dużą liczbą formatów, lepiej lub gorzej przystosowanych do naszych wymagań. Jakie one są?

1. Format powinien być publiczny, nie zamknięty. Niektóre, szczególnie starsze formaty, powstały w firmach zajmujących się obróbką obrazu, które zachowują prawa autorskie, ale zwykle format jest opublikowany i publicznie dostępny. Formaty zdefiniowane jako międzynarodowe standardy (np. ISO) mają dużo większą szansę pozostania użytecznym w przyszłości.

2. Format powinien być popularny (co może czasem być w sprzeczności z punktem 1). Standard który nie ma łatwo dostępnych narzędzi jest przydatny tylko w teorii.

3. Narzędzia do obróbki obrazu powinny być łatwo dostępne, a czytniki powinny być darmowe albo tanie, najlepiej otwartego oprogramowania. Dawanie komuś zdjęcia z adnotacją “możesz je sobie obejrzeć, ale musisz do tego kupić program za $500” jest w złym smaku. Podstawowa obróbka taka jak obracanie, przycinanie, zmiana rozmiaru itp. powinna być dostępna w popularnych, tanich i/lub otwartych narzędziach.

4. Formaty powinny posiadać możliwość zapisywania metadanych, patrz blog “ Co jest na odwrocie zdjęcia cyfrowego?
”.

Rozdzielczość i kompresja

Dla tych z nas, którzy parali się fotografią na kliszach srebrowych, znana jest ziarnistość filmu, połączona z jego czułością. Czym mniejsza czułość, tym mniejsze kryształki halogenku srebra, i tym drobniejsze szczegóły można zarejestrować. Kryształki zastąpiły w aparacie cyfrowym elementy światłoczułe - czym gęstsze tym drobniejsze szczegóły. Rozdzielczość czujnika podaje się zwykle w (mega) pixlach. Dla skanera ten wymiar podawany jest zwykle w pixlach na cal (lub centymetr), w skrócie ppi lub dpi.

Wielkość obrazu w pamięci komputera (szerokość x wysokość [w pixlach] x 3) może być znaczna. Dla oszczędności miejsca niektóre formaty stosują kompresję. Nie musimy się tu zastanawiać nad jej rodzajami, a jedynie nad tym, czy cykl kompresja - dekompresja pozostawia obraz bez zmian. Jeśli tak, mówimy, że kompresja jest bezstratna, jeśli nie, stratna. Stratna kompresja może być dużo bardziej skuteczna w przypadku zdjęć, ale w zależności od jej intensywności pozostawiać po sobie ślady.

Formaty

GIF

GIF (Graphic Interchange Program) został wprowadzony przez Compuserve w 1987 roku. Stosuje bezstratna kompresję, ale jest ograniczony do 8 bitów na trzy kolory (maks. 256 odcieni) i z tego powodu nie nadaje się do zapisu fotografii. Możliwości zapisu metadanych są bardzo ograniczone. Posiada jednak dwie bardzo pożądane funkcje. Można w nim zdefiniować kolor przezroczysty, co pozwala na tworzenie grafik (takich jak logo) które można nałożyć na istniejący już wzór graficzny. GIF posiada też możliwość zapisu wielu obrazów, które można wyświetlać jako krótkie filmy (animacje) - ta funkcja powoduje niemalejącą popularność tego formatu. Większość przeglądarek internetowych potrafi wyświetlać pliki w formacie GIF, włącznie z animacją, obsługuje go prawie każdy program graficzny. Pliki mają rozszerzenie .gif

PNG

PNG (Portable Network Graphics) powstał w celu przezwyciężenia problemów z GIF, do których należała ograniczona liczba kolorów i opatentowany sposób kompresji. Format ten został zatwierdzony do użycia w Internecie w 1996 i uzyskał status standardu ISO w 2004. PNG pozwala na zapis grafiki i zdjęć, z użyciem 24 lub 32 bitowego koloru, ma też możliwość stosowania koloru przezroczystego. Stosuje bezstratną kompresję, a więc nadaje się do archiwalnego przechowywania obrazu i jego obróbka nie wprowadza strat jakości. Możliwości zapisu metadanych są ograniczone: brak jest formatu EXIF stosowanego przez aparaty fotograficzne; jest możliwość zapisu danych w formacie XMP, ale popularne programy ani nie zapisują ani nie czytają tych danych. Popularność formatu PNG rośnie, jest odtwarzany przez przeglądarki internetowe i obsługuje go większość programów graficznych. Pliki maja rozszerzenie .png

TIFF

TIFF (Tagged Image File Format) stworzony przez firmę Aldus został wprowadzony do użycia w roku 1986. Mimo tego, że ma już ponad 25 lat, jest ciągle niezmiernie popularnym formatem wśród grafików, fotografów, w przemyśle wydawniczym. Może zapisać pliki o wielkości do 4 GiB, w pełnej gamie kolorów. TIFF posiada możliwość zapisu wielostronicowego (a więc może zapisać np. wszystkie strony faksu), posiada kompresję bezstratną i możliwość zapisu bez kompresji. Ma wiele dodatków i rozszerzeń, względnie standardowa jest wersja TIFF 6.0 - standard administrowany przez Adobe, oraz kilka wersji zarejestrowanych jako standardy ISO. Nie posiada animacji, przezroczystości, nie jest wyświetlany przez najczęściej stosowane przeglądarki internetowe. Jest popularny wśród archiwów jako format przechowywania archiwalnego obrazów i skanów. Można w nim zapisać metadane Exif i IPTC, użycie XMP, choć teoretycznie możliwe, nie jest powszechnie stosowaną opcją. Tiff jest bardzo popularny, obsługuje go prawie każdy program graficzny. Pliki mają rozszerzenie .tif lub .tiff

JPEG

JPEG (Joint Photographic Expert Group) jest bardzo popularnym formatem stworzonym dla zapisu zdjęć i innych obrazów półtonowych. Używa zawsze kompresji, która jest stratna, ale daje znaczące zmniejszenie rozmiaru obrazu. Przy tej samej wielkości pliku, obraz w formacie jpeg może mieć 25 lub więcej razy więcej pixli (5 razy więcej wymiarze liniowym) niż tiff, co w dużym stopniu kompensuje stratność kompresji. Przy zapisie archiwalnym pojawiają się jednak dwa problemy: po pierwsze błędy kompresji są najbardziej widoczne na ostrych - kontrastowych granicach elementów obrazu (np. na brzegach liter tekstu), po drugie każda przeróbka powoduje powstawanie dodatkowych błędów gdyż nie można całkowicie wyłączyć kompresji. Ten drugi problem można częściowo ominąć przy przechowywaniu zdjęć, jeśli używa się programu (takiego jak np. Picasa), który zapisuje tylko transformacje pozostawiając oryginał niezmienionym.

Format JPEG jest zarejestrowany jako standardy ISO, jest obsługiwany przez wszystkie programy do obróbki i wyświetlania obrazu, obsługują go przeglądarki internetowe - jest najpopularniejszym formatem do zapisu i wyświetlania zdjęć. W pliku JPEG można też zapisać metadane w Exif, IPTC i XMP, co zwiększa znacznie jego uniwersalność. Pliki maja rozszerzenie .jpg lub .jpeg, choć spotyka się czasem .jif, .jfif i inne.

JPEG 2000

Format JPEG 2000 (w plikach zaznaczany jako rozszerzenie .jp2) jest następną generacją formatu opracowanego przez Joint Photographic Expert Group. Posiada wszystkie zalety JPEG, lepszy algorytm kompresji, jest standardem ISO. Posiada możliwość zapisu bez kompresji, nadaje się więc do przechowywania materiałów archiwalnych. Zapis metadanych jest

tylko w formacie XMP. W sumie jest to bardzo dobry przyszłościowy format graficzny.

Mimo tego, że został wprowadzony ponad 10 lat temu, zbywa mu jednak na popularności. Wiele czytników i programów do obróbki albo nie obsługuje JPEG 2000 albo obsługuje go tylko w ograniczonym zakresie, używając wtyczek - ładowanie obrazu w tym formacie trwa znacząco dłużej. Popularny program do obróbki i organizacji zdjęć, Picasa, nie wspiera tego formatu, a zapis metadanych wymaga wyspecjalizowanych narzędzi. Jpeg 2000 nie jest też wyświetlany w przeglądarkach internetowych.

Inne formaty

Istnieje jeszcze wiele innych formatów poza wymienionymi powyżej. Z różnych powodów pozostawimy je jednak na uboczu.

RAW jest wspólną nazwą wielu formatów zapisujących ‘surowe’ dane z czujnika kamery - zawierają najbardziej szczegółowe dane obrazu, które następnie można dalej przerabiać. Choć wiele z nich używa elementów TIFF, są to formaty zamknięte, ograniczone do danego producenta kamery i jako takie nie nadają się do długoterminowego przechowywania ani wymiany obrazów.

BMP jest formatem rastrowym Microsoft, zbudowanym dla systemu Windows. Ponieważ jest bardzo popularny, spotyka się go często, szczególnie w starszych aplikacjach i jako format grafik w Windows.

PDF (Portable Document Format) nie jest formatem graficznym, ale może zawierać w sobie także grafikę. Jest formatem opisu dokumentu, zawierającym wszystkie elementy niezbędne do pokazania / wydrukowania jedno lub wielostronicowego dokumentu. Został stworzony przez formę Adobe w latach 1991-93 i spopularyzowany przez publikację przez tę firmę darmowych czytników PDF. Od 2008 jest standardem ISO, i nie jest już sterowany przez Adobe. W 2005 opublikowany został standard ISO o nazwie PDF/A, opisujący standard dokumentu z nastawieniem na przechowanie archiwalne, będącym pewnym podzbiorem formatu PDF.

PDF, a szczególnie PDF/A jest rekomendowany jako format dla długoterminowego przechowywania dokumentów. W istocie nadaje się on dobrze do tego celu, dając uniwersalny, względnie trwały zapis tekstu, sformatowanego na stronie, który może również zawierać grafikę,

zarówno rastrową jak i wektorową. PDF nie jest jednak formatem graficznym, i dlatego w przypadku zdjęć i skanów stanowi tylko drugą kopertę, w którą wkłada się już zapakowane zdjęcie. PDF nie jest bezpośrednio wyświetlany przez przeglądarki internetowe, ani przez programy do obróbki obrazu. Najnowsza wersja (PDF/A-2 z 2011 roku) oferuje kompresję JPEG 2000 i użycie metadanych, zarówno dla całego dokumentu jak i pojedynczych stron. Narzędzia do obróbki PDF (poza własnościowymi i raczej kosztownymi narzędziami Adobe) są jednak rzadkie, i nawet wykonanie prostych manipulacji takich jak dodanie, odjęcie czy obrócenie stron dokumentu wiąże się ze sporą gimnastyką. Jeśli chodzi o prezentację (a nie przechowywanie długoterminowe) dokumentów wielostronicowych, PDF jest wygodny do użycia, i konkuruje z innym formatem stworzonym do tego celu, DjVu.

Rekomendacje

Jakiego formatu używać do przechowywania zdjęć i skanów dokumentów, w domu i w archiwum? Z opisu powyżej wynika, że w przyszłości będziemy mieli doskonałe formaty dające archiwalny zapis obrazu, metadane, świetne narzędzia i możliwość obejrzenia naszych zasobów w dowolnym urządzeniu. Dziś jeszcze do tego etapu nie doszliśmy. Mamy formaty które zostały stworzone dawno i są popularne, oraz nowsze, które być może będą lepsze, ale brak narzędzi dyskwalifikuje je do użycia teraz. Jest więc prawdopodobne, że nasze dzieci a na pewno wnukowie będą musiały dokonać konwersji do ‘właściwego’ w 2050 roku formatu - przełożyć zdjęcia do nowych, lepszych kopert.

Co robić teraz? Fotografie można zapisywać w formacie JPEG, w największej możliwej rozdzielczości. Aparaty fotograficzne maja zwykle różne opcje, należy wybierać zawsze najlepszą jakość. Powoduje to zwiększenie rozmiaru pliku, ale pamięć jest tania i jej cena systematycznie spada. Należy przechowywać oryginały zdjęć, i nie modyfikować ich, a tylko robić kopie. Należy zapisywać metadane (popularne czytniki takie jak IrfanView lub XnView to potrafią, dużo danych zapisuje też program Picasa (Opcje/Tagi/Przechowuj tagi w zdjęciach). Skany, a szczególnie materiały archiwalne należy przechowywać jako pliki TIFF. W przyszłości będzie można je przenieść do formatu .jp2, kiedy stanie się on bardziej powszechny. Zapis metadanych jest też mocno polecany, choć do archiwaliów zwykle chcemy dodać więcej informacji, skąd pochodzą, jakie były ich losy, co zawierały itp. Do tego polecam zwykły arkusz rozliczeniowy lub dokument biurowy, albo wyspecjalizowany program archiwalny. Jeśli chcemy zachować dokumenty tworzone elektronicznie, format PDF nadaje się do tego celu bardzo dobrze.

Czytaj więcej

Artykuły w Wikipedii o formatach graficznych(w języku angielskim)

- GIF
- PNG
- TIFF
- JPEG
- JPEG 2000

Marek Zieliński

Artykuł ukazuje się (prawie) jednocześnie w *Blogu archiwistów i bibliotekarzy Instytutu Piłsudskiego*