

ROLA KOMPRESJI OBRAZÓW W INTERNETOWYCH PUBLIKACJACH KARTOGRAFICZNYCH

THE IMPORTANCE OF IMAGE COMPRESSION FOR INTERNET GIS

Krystian Pyka

AGH w Krakowie

**Słowa kluczowe: kompresja, internetowa publikacja kartograficzna, ortofotomapa, wielo-
rozdzielczość, piksel, wizualizacja**

Key words: compression, internet cartography, orthophotomap, multiresolution, pixel, visualization

Streszczenie

Dzisiejsi użytkownicy internetu oczekują aby działające on-line, interaktywne publikacje kartograficzne prezentowały mapy o wysokiej jakości. Jedną z barier stosowania rastrowych map podkładowych w przeglądarkach GIS są ich duże rozmiary przy relatywnie małej prędkości transmisji. Artykuł analizuje przydatność kompresji obrazów dla hybrydowych publikacji kartograficznych. Porównuje się dotychczasowe metody kompresji z metodą falkową. Wskazuje się zalety i wady tej kompresji oraz proponuje rozwiązanie łączące system piramid obrazowych z kompresją falkową.

Wprowadzenie

Udostępnianie informacji geograficznej jest dzisiaj traktowane przez internautów jako usługa standardowa, powszechnie dostępna. Potwierdza to mnogość adresów internetowych które serwują publikacje kartograficzne. Ale duża jest nie tylko liczba stron internetowych „z mapami” lecz zaskakuje także różnorodność rozwiązań pod względem merytorycznym i technicznym.

Konkurencja pomiędzy grafiką wektorową a rastrową zakończyła się zwycięstwem użytkownika, który wymaga aby informacja przestrzenna była pokazana w sposób komunikatywny a przy tym estetyczny, nie interesując się jaką drogą uzyskano te przymioty.

Jednym z powodów atrakcyjności obrazów rastrowych dla kartografii internetowej jest rozwój technik kompresji obrazów. Właśnie jesteśmy świadkami udanych implementacji kilku nowych rozwiązań w dziedzinie kompresji obrazów, z których za najciekawsze należy uznać kompresję falkową.

Artykuł jest próbą analizy przydatności kompresji metodą falkową dla potrzeb udostępniania publikacji kartograficznych w trybie on-line.

Internetowe publikacje kartograficzne

Termin internetowa publikacja kartograficzna tylko pozornie ogranicza przedmiot zainteresowania do klasycznych map wyświetlanych na ekranie monitora komputerowego. Można wyróżnić następujące rodzaje internetowych publikacji kartograficznych:

- a) ilustracyjne,
- b) ilustracyjne z elementami interaktywności,
- c) serwisy informacji geograficznej.

Celem publikacji a) jest tylko prezentacja uprzednio przygotowanego obrazu kartograficznego, który przemawia klasycznymi środkami wyrazu. Jedyną innowacją w stosunku do map papierowych jest (powinno być) dodatkowe okienko nawigacyjne, pomagające przemieszczać się po mapie, której rozmiary są z reguły dużo większe niż wymiary ekranu.

Publikacje b) dodają do a) elementy reagujące na akcję użytkownika, ujawniając w odpowiedzi nowe, uprzednio niewidoczne informacje w formie tekstowej lub graficznej.

Publikacje c) w podstawowej postaci są zupełnym przeciwieństwem a). Zamiast wyświetlenia zredagowanego obrazu kartograficznego oferują listę tematów do wizualizacji lub nawet pozwalają zadawać pytania o konkretne obiekty geograficzne, wybierane według określonych kryteriów. Użytkownik ma możliwość komponowania treści „mapy”, zawsze będzie to jednak obraz o relatywnie niskich walorach estetycznych a jego komunikatywność jest zróżnicowana.

Punkt ciężkości w powyższych trzech typach publikacji przemieszcza się od prezentacji gotowej mapy w a) do wyszukiwania danych i wizualizacji uzyskanej odpowiedzi w przypadku c).

Zdaniem autora, najwięcej zwolenników zyskują rozwiązania hybrydowe, które wykorzystują i zalety klasycznych publikacji kartograficznych i zalety rozwiązań bazodanowych. Jak zwykle problemem jest znalezienie kompromisu zarówno w aspekcie merytorycznym – przekazywanie informacji, estetycznym – wysoka jakość grafiki oraz technicznym – krótki czas dostępu.

Mapa podkładowa

Mapa podkładowa, jeden z klasycznych elementów map tematycznych, wcale nie traci na znaczeniu. W ujęciu tradycyjnym stanowi tło dla głównego tematu mapy, spełniając rolę orientującą obiekty tematyczne w przestrzeni geograficznej. W publikacjach internetowych mapa podkładowa służy jako tło do wyświetlenia wyników zapytań do bazy danych przestrzennych (c) albo jako tło i jednocześnie podstawa geograficznej lokalizacji elementów interaktywnych (b).

Mapa podkładowa może być przygotowana w postaci rastrowej lub wektorowej. Ograniczenie dalszych rozważań do mapy podkładowej zapisanej rastrowo powoduje, że wybieramy rozwiązanie o funkcjonalności podobnej jaką ma podkład w opracowaniach analogowych. Konsekwencją tego wyboru jest uwzględnienie a nawet promowanie fotomapy w roli mapy podkładowej.

Użytkownicy internetu łatwo znajdują informacje o nowych technikach rejestracji obrazów, przekonując się o powszechności wykorzystywania fotomap, głównie w USA. Stopniowo zwiększa się pokrycie Polski ortofotomapami, na razie ze zdjęć lotniczych w skali ok. 1:30 000 i satelity IKONOS, a wkrótce ze zdjęć lotniczych w znacznie większej skali (obszar południowo-wschodni). Istnieją przesłanki, aby prognozować dla całego kraju pełne, odnawialne pokrycie ortofotomapami o nominalnej skali 1:10000 (lub większej). Czyli mapa fotograficzna ma szansę stać się podstawowym produktem kartograficznym a przez to jest predysponowana do spełniania funkcji mapy podkładowej.

Stratna kompresja częstotliwościowa

Kompresja obrazów o strukturze rastrowej polega na usunięciu redundancji danych, poprzez wykrycie i usunięcie korelacji w dziedzinie przestrzennej (pomiędzy sąsiednimi pikselami) lub spektralnej (pomiędzy składowymi triadowego modelu barw lub kanałami obrazów wielospektralnych).

Kompresja bezstratna ma istotne znaczenie dla archiwizacji danych ale jej przydatność dla potrzeb wizualizacji na ekranach monitorów ogranicza się tylko do obrazów o małej liczbie barw. Natomiast dla obrazów wielotonalnych, zawierających (nominalnie) np. 2^8 stopni szarości czy 2^{24} barw, kompresja bezstratna jest praktycznie nieprzydatna. W tym wypadku skuteczniejsza jest kompresja stratna, która zapewnia niewiarygodnie duży stopień redukcji objętości obrazu przy niezauważalnym spadku jakości.

Człowiek w sposób specyficzny, choć nie do końca znany, odbiera bodźce świetlne. Stwierdzono m.in. że oko ludzkie charakteryzuje się logarytmicznym typem nieliniowości w przetwarzaniu natężenia odbieranego światła oraz, że przekształca odbierane częstotliwości przestrzenne jak filtr górnoprzepustowy (Oppenheim, 1982). Kompresja bazująca na analizach częstotliwościowych uwzględnia sposób percepcji obrazu przez człowieka. Stąd tak duży sukces kompresji metodą JPEG, której ważną część algorytmu – tablica wag decydująca o kwantyzacji – powstała na drodze eksperymentów psychowizualnych (Wallace, 1991).

Każdy obraz rozumiany jako uporządkowany zbiór jasności poszczególnych pikseli można opisać w przestrzeni częstotliwości stosując transformację Fouriera. Transformacja z przestrzeni miejsca do przestrzeni częstotliwości może być odwrócona co skutkuje pełną rekonstrukcją obrazu pierwotnego. Natomiast kompresja w dziedzinie Fouriera polega na pominięciu pewnych współczynników, zwłaszcza z zakresu dużych częstotliwości (związek z właściwością oka jako filtru).

Kompresja JPEG realizuje transformację z przestrzeni miejsca do przestrzeni częstotliwości poprzez dyskretną transformację kosinusową (DCT – Discrete Cosine Transform) przy czym proces odbywa się w blokach obrazu o wymiarach 8×8 pikseli (Wallace, 1991). Skutkiem „blokowości” algorytmu jest widoczna w obrazie skompresowanym sztuczna tekstura, która ujawnia się tym bardziej im większy stopień kompresji. Stosowane do oceny jakości obrazu skompresowanego miary statystyczne np. błąd średni kwadratowy czy miara PSNR (Peak Signal to Noise Ratio), nie potrafią wykazać spadku jakości pomimo istotnej degradacji obrazu (Mikrut, 2003). Miary te są liczone dla całego obrazu, nie oddają rzeczywistego, obserwowanego wizualnie spadku jakości.

Kompresja falkowa wywodzi się również z korzenia fourierowskiego, czyli ma charakter analizy częstotliwościowej. Istotą przekształcenia falkowego jest dekompozycja sygnału (w przypadku obrazu – sygnału dwuwymiarowego) w ciąg sygnałów (subobrazów) o stopniowo zmniejszającej się rozdzielczości. Dekompozycja polega na generowaniu subobrazów na drodze przemiennych filtracji górno i dolno-przepustowych, przy jednoczesnym zmniejszaniu rozdzielczości z interwałem dwa. Realizacja dyskretnego przekształcenia falkowego (DWT – Discrete Wavelet Transform) może odbywać się różnymi drogami (Saha, 1999), za najskuteczniejszą uważa się zaproponowaną przez Shapiro metodę EZW – Embedded Zerotree Wavelet (Shapiro, 1993). Podobnie jak w metodzie JPEG kompresja następuje na etapie kwantyzacji, kiedy rekonstruując obraz dokonuje się pewnych uproszczeń zwłaszcza w zakresie wysokich częstotliwości (Vetterli, 1995).

Obecnie dominują dwa rozwiązania technologiczne oparte o metodę falek: JPEG2000 (standard ISO) i open standard ECW (Enhanced Compressed Wavelet). Dla potrzeb wizualizacji ekranowej dopuszcza się kompresowanie obrazów wielotonalnych chromatycznych (barwnych) nawet w stosunku 50:1 czyli dwukrotnie więcej niż przy standardowej kompresji JPEG.

Przesyłanie obrazu w porcjach „ekranowych”

Nawet kompresja w stosunku 50:1 nie rozwiązuje problemu zasilania on-line przeglądarek publikacji kartograficznych (w uproszczeniu – przeglądarek GIS). Rozwiązuje go dopiero dodanie „ekranowej” strategii przesyłania, polegającej na wysyłaniu tylko takiej porcji obrazu która jest potrzebna do wypełnienia okna przeglądarki. Okno to jest z reguły mniejsze od pełnej matrycy ekranowej, zajmuje tylko pewną część ekranu. Załóżmy, że okno przeglądarki ma rozmiary 800x600, zatem jego „zamalowanie” wymaga przesłania danych opisujących jasności 800x600 pikseli co dla obrazu achromatycznego daje 480 tys. bajtów a dla obrazu barwnego RGB trzy razy tyle czyli ok. 1,4 MB (bez kompresji). Jeśli dysponujemy dobrym łączem internetowym, które zapewnia transmisję z prędkością np. 30 KB/s, to taka nieskompresowana porcja barwnego obrazu ładuje się prawie minutę (47 sek.), ale po kompresji w stosunku 50:1 czas skraca się już do jednej sekundy. Czyli przesyłanie skompresowanych porcji spełnia wymagania przeglądarek on-line w zakresie ładowania rastrowej mapy podkładowej o formule fotomapy.

Zmiana skali obrazu

Kolejnym krytycznym problemem wyświetlania map podkładowych w przeglądarkach GIS jest przeskalowanie obrazu a przypadkiem skrajnym jest wyświetlenie „dużego” obrazu na „małym” ekranie. Załóżmy jak poprzednio, że ekran przeglądarki ma rozmiary 800x600 a rozmiar obrazu liczony w pikselach jest 8000x8000. Aby ten obraz zmieścić na tak małej matrycy ekranu musi być on kilkunastokrotnie przeskalowany (w tym przypadku aż trzynastokrotnie co wynika ze stosunku 8000/600). Tak duże przeskalowanie może być realizowane na drodze splotu obrazu z macierzą kwadratową o rozmiarze równym współczynnikowi przeskalowania (np. macierz jednostkowa, macierz aproksymująca funkcję Gaussa) jednakże jest to związane z dużym wydatkiem obliczeniowym.

Zmiana skali obrazu z wykorzystaniem piramidy obrazowej

Cyfrowe przetwarzanie obrazów od dawna stosuje technikę piramid obrazowych (Burt, 1983). Idea piramidy obrazowej jest bardzo prosta, polega na zbudowaniu serii pomniejszych obrazów, z których każde następne jest – najczęściej – dwukrotnie (liniowo) mniejsze od poprzedniego (Kropatsch, 2000). Piramidy początkowo wykorzystywane do innych zadań (np. automatyczne dopasowanie obrazów – matching) zostały z powodzeniem zaadaptowane do ładowania obrazów w przeglądarkach internetowych. Jedno z pierwszych tego typu zastosowań powstało w 1995 r. na amerykańskim uniwersytecie MIT (Digital Orthophoto Browser). W

zależności od zapotrzebowania zgłoszonego przez przeglądarkę, serwer wybiera to pomniejszenie które najlepiej przybliży rozmiar okna przeglądarki. To dzięki temu internauta ma wrażenie, że w krótkim czasie załadował obraz o dużych rozmiarach, rzędu kilkuset MB. Tymczasem okno przeglądarki zostało wypełnione tylko uprzednio przygotowanym pomniejszeniem obrazu, które po kompresji ma nie więcej niż kilkadziesiąt KB. Sukcesywne powiększanie obrazu w przeglądarce zmusza serwer do wysyłania stosownej porcji obrazu pobieranej z innego poziomu skalowego. Zaletą tej metody jest możliwość dobrania takich filtracji przy generowaniu kolejnych poziomów, aby wartość informacyjna pomniejszeń była jak najwyższa (analogia do generalizacji map).

Zmiana skali obrazu z wykorzystaniem kompresji falkowej

Obraz skompresowany metoda falkową jest w istocie ciągiem sub-obrazów o rozmiarach zmieniających się ze stosunkiem dwa. Dla przykładowego obrazu 8000x8000 ciąg ten będzie zawierał subobrazy: 4000x4000, 2000x2000, 1000x1000, 500x500, 250x250, 125x125, 62x62, czyli razem osiem poziomów skalowych. Rozmiar najmniejszego sub-obrazu nie jest ograniczony, ale ze względów praktycznych ustala się go w pobliżu wielkości 64x64. Jest to optymalny dla zastosowań internetowych rozmiar bloków odświeżania obrazu i może być traktowany jako najmniejsza wspierana przez układ kompresor/dekompresor skala wizualizacji. Pierwsza, inicjalna wizualizacja obrazu w przeglądarce z reguły rozpoczyna się od małej skali, do rekonstrukcji obrazu wystarczą wtedy najmniejsze subobrazy. Dwukrotne powiększanie skali zmusza algorytm rekonstruujący obraz do wykorzystania kolejnego subobrazu, aż do wyczerpania pełnego zbioru (w analizowanym przykładzie ośmio elementowego). Po drodze wystąpi wielokrotnie potrzeba interpolacji pomiędzy rozmiarami poziomu skalowego a rozmiarami okna (np. z 500x500 do 600x600). Koszt tej operacji nie jest duży, gdyż współczynnik przeskalowania nigdy nie jest większy niż dwa. Ten zabieg odbywa się w locie. Opisana własność kompresji falkowej jest określana jako „wieloskalowość” lub „wielorozdzielczość” (Wu, 2000).

Opisana cecha „wieloskalowości” nie jest oczywiście receptą na pokonanie klasycznej rafy kartograficznej jaką jest generalizacja mapy (w tym wypadku fotograficznej), ale jest środkiem łagodzącym ten problem.

Podsumowanie

W ostatnim czasie powiększa się liczba portali dostarczających w trybie on-line porcje obrazowe ekstrahowane z obrazów o rozmiarach rzędu kilku a nawet kilkuset GB! Umożliwia to kompresja falkowa w implementacji Enhanced Compressed Wavelet. Rozwiązania te wyraźnie górują pod względem sprawności nad technikami opartymi o piramidy obrazowe i kompresję standardową JPEG. Ale nie można nie zauważyć dość poważnej wady która uwidacznia się tym bardziej, im większy jest rozmiar obrazu pierwotnego. Wada ta polega na tym, że najmniejsze sub-obrazy powstałe z gigantycznego obrazu pierwotnego tracą zupełnie wartość informacyjną stając się niekomunikatywnymi „chmurami”. Kompresja falkowa buduje subobrazy w ustalony i niezmienny sposób. W tym przypadku elastyczniejsza jest koncepcja piramid, w której każdy poziom skalowy może powstać niezależnie, co pozwala dobrać opty-

malne metody filtracji i interpolacji towarzyszącej powtórnemu próbkowaniu. Ale wada ta jest relatywnie łatwa do usunięcia na drodze fuzji ze strukturą piramid. Rozwiązanie polega na ograniczeniu dekompozycji do ustalonego poziomu – N (np. 10). Poniżej poziomu granicznego, czyli na umownym poziomie $N+1$ musi być wprowadzony nowy obraz bazowy w stosunku do którego prowadzona jest niezależna dekompozycja. Taki obraz można wygenerować stosując operatory morfologiczne albo funkcję Gaussa. Alternatywą jest zastąpienie obrazu pierwotnego zupełnie innym obrazem, zarejestrowanym przy znacznie mniejszej rozdzielczości (np. ortofotomapa lotnicza może być zastąpiona fotomapą satelitarną).

Przedstawiona dyskusja technik wyświetlania obrazów wielotonalnych w przeglądarkach GIS pracujących w trybie on-line dowodzi, że mają one specyfikę wynikającą z uwarunkowań stawianych przez internet. Specyfika ta jest na tyle charakterystyczna, że można mówić o „internetyzacji obrazów” jako działaniach ukierunkowanych na optymalizację wizualizacji on-line.

Literatura

- Burt P., Adelson E., 1983: The Laplacian Pyramid as a Compact Image Code. IEEE, Vol.Com31, No4.
- Kropatsch W., Bischof H. (editors), 2000: Digital image analysis: selected techniques and applications, New York. Springer.
- Mikrut S. 2003: Wpływ skanowania i kompresji według standardu JPEG na wykrywanie obiektów liniowych i punktowych na obrazach cyfrowych. Praca doktorska - AGH.
- Oppenheim A., (red.), 1982: Sygnały cyfrowe. PWN Warszawa.
- Saha S., Vemuri R., 1999: Adaptive Wavelet Coding of Multimedia Images, Proc. ACM Multimedia Conference.
- Shapiro J. M., 1993: Embedded Image Coding Using Zerotrees of Wavelet Coefficients, IEEE Trans. SP, vol. 41, no. 12.
- Vetterli M., Kovacevic J., 1995: Wavelets and Subband Coding. Prentice Hall, Signal Processing Series, Englewood Cliffs, NJ.
- Wallace Gregory K., 1991: The JPEG Still Picture Compression Standard. Communications of the ACM, vol. 34 no. 4.
- Wu J., Amaratunga K, Lui T., 2000: Design of an Online GIS Viewer by Wavelet Technology. Proceedings of the 8th International Conference on Computing in Civil and Building Engineering, Stanford, CA.

Summary

The Internet users of today expect online interactive GIS to provide high quality maps. Large image data sets with a relatively low speed network are the principal barrier for online GIS-viewers, especially of base maps. This paper analyses usefulness of image compression to design hybrid GIS browsers. It is compared some known compression method with wavelet technology. Advantage and disadvantage of this new technology are discussed and an integration of image pyramid system and wavelet technology is suggested as a solution.

Krystian Pyka
Akademia Górniczo-Hutnicza w Krakowie
Wydział Geodezji Górniczej i Inżynierii Środowiska
30-059 Kraków, al. Mickiewicza 30
mail: krisfoto@interia.pl;
kpyk@malopolska.mw.gov.pl