

Metody kompresji obrazów: uwarunkowania dalszego rozwoju

Tomorrow's image compression today

Artur Przelaskowski

Instytut Radioelektroniki, Politechnika Warszawska, Warszawa

Streszczenie: W pracy zwrócono uwagę na kilka istotnych problemów związanych z doskonaleniem metod kompresji obrazów. Krótka charakterystyka współczesnych uwarunkowań pozwala lepiej zrozumieć ograniczenia i perspektywę rozwoju wobec rosnących oczekiwań 'ery informacji'. Aby zobrazować niektóre z ograniczeń przeprowadzono testy wybranych koderów bezstratnych i stratnych. Zwrócono uwagę na konieczność praktycznego wykorzystania semantycznej teorii informacji, szczególnie w kontekście kompresji z selekcją informacji. Podkreślono znaczenie uniwersalnych, bezstratnych koderów danych oraz hierarchicznego uporządkowania informacji w elastycznych koderach z progresją informacji. Prowadzi to do weryfikacji obowiązujących paradygmatów kompresji, szczególnie w zakresie modelowania źródeł informacji. Sugerowane kierunki działań wydają się być użyteczne.

Summary: Analysis and suggestions of image compression improvements were presented. Characteristics of modern conditions of development showed limitations and perspective of development in the following areas: utilization of semantic information theory, selection of useful information by interactive user oriented algorithms, application of universal lossless archivers for image compression, flexibility of ordered hierarchical representation of information in multiscale image coders. Selected lossless and lossy coders were verified experimentally and new compression paradigms were concluded.

Słowa kluczowe: kompresja obrazów, kompresja stratna i bezstratna, teoria informacji, JPEG
Keywords: image compression, lossy and lossless compression, information theory, JPEG

1. Wprowadzenie

Kompresję można sprowadzić do odwracalnego lub nieodwracalnego procesu redukcji długości reprezentacji danych. W opracowaniu nowoczesnych narzędzi kompresji danych, szczególnie obrazów, dużego znaczenia nabiera szersze rozumienie pojęcia kompresji jako wyznaczania efektywnej reprezentacji przesyłanej lub gromadzonej informacji. Wiarygodne definicje informacji, modele źródeł informacji oraz sposoby obliczania ilości informacji i optymalizacji kodów stają się tutaj zagadnieniem kluczowym. Celem pracy było szersze spojrzenie na zagadnienie kompresji obrazów, które pozwoli lepiej zrozumieć istniejące ograniczenia, a także ocenić realne szanse doskonalenia istniejących koderów obrazów.

1.2. Krótka charakterystyka współczesnych uwarunkowań

Jesteśmy świadkami 'ery informacji' o technologicznych podstawach, którą charakteryzuje intensywny rozwój społeczeństwa informacyjnego, przekształcającego się stopniowo w społeczeństwo szeroko dostępnej (tzw. wolnej) wiedzy. Wobec rosnącej roli Internetu można mówić również o zjawisku tworzenia się społeczeństwa sieciowego (prace Castellsa [1]). Coraz większe znaczenie rozwoju technologicznego, zdominowanego przez zagadnienie

efektywnej wymiany informacji oraz świata nauk przyrodniczych i ścisłych widać chociażby w propozycjach tzw. trzeciej kultury [2].

Szybki dostęp do istotnych informacji, efektywne jej przetworzenie oraz wyszukiwanie wiarygodnych źródeł wiedzy staje się decydującym o sukcesie składnikiem nowoczesnej działalności edukacyjnej, biznesowo-gospodarczej, organizacyjnej, społecznej. Wobec nadmiaru pseudoinformacji (tysiące reklam, setki bezużytecznych zachęt zapychających skrzynki pocztowe i kanały telewizyjne) coraz bardziej pożądane stają się narzędzia inteligentnie porządkujące przestrzeń informacyjną.

Sposobem na życie współczesnego człowieka będącego w ciągłym pośpiechu staje się oszczędzanie (wręcz wydzieranie) każdej wolnej chwili [3]. B. Pascal (1623-1662) pisząc list do przyjaciela wspominał, że 'nie ma czasu, aby napisać krócej'. Pisał więc jak leci, przeproszał, że zabrakło mu czasu, by zaoszczędzić czas adresata. Poszanowanie czasu przekłada się na wzrost ilości informacji dostępnej (przekazywanej) w jednostce czasu w warunkach, jakimi dysponuje konkretny odbiorca (określone ramy czasowe, sprzętowe, finansowe). Oszczędność czasu odbiorcy informacji zapewnia przede wszystkim odpowiednia jakość informacji, tj. selekcja danych, które są istotne dla użytkownika, oraz efektywna jej reprezentacja (minimalizacja długości zapisu danych i czasu ich przekazywania, właściwa struktura informacji).

Technologiczne doskonalenie systemów rozpowszechniania informacji wymaga więc stosowania efektywnych metod kompresji danych, w tym koderów obrazów. Obecnie niebagatelną rolę odgrywa przekazywanie informacji za pomocą obrazu - przykładem niech będzie chociażby dominacja obrazkowych systemów operacyjnych typu *Windows*. Kompresja obrazów pozwala szybciej zobaczyć obraz o lepszej jakości, więcej informacji umieścić w danej objętości nośnika, szybciej wyszukać żadaną informację, porównać obszerne zasoby informacji obrazowej etc. W koderach obrazów wykorzystuje się obok elementów teorii informacji także zaawansowane metody numeryczne, teorię sygnałów, analizę funkcjonalną, psychowizualne modele percepcji czy inną wiedzę o człowieku i jego relacji do świata (zależnie od zastosowań).

Doskonalenie metod akwizycji obrazów, poprawa zdolności rozdzielczej systemów obrazowania, redukcja szumu i zwiększanie dynamiki sygnału użytecznego wymusza operowanie coraz większymi zbiorami danych obrazowych. Szybko rosnąca liczba obrazów cyfrowych przenoszących informację w coraz większej gamie zastosowań, takich jak fotografia cyfrowa, kamery cyfrowe na użytek domowy, obrazy satelitarne wykorzystywane w meteorologii, kartografii czy urbanistyce, systemy obrazowania biomedycznego itd., wymaga użycia efektywnych metod gromadzenia, indeksowania, przeglądania i wymiany tej informacji. W opinii wielu odbiorców i nadawców informacji obrazowej wzrost efektywności i użyteczności opracowywanych narzędzi kompresji obrazów jest zbyt wolny wobec wymagań współczesnych zastosowań. Przykładem może być rozwijany obecnie standard JPEG2000, w którym spełnienie żądań nowoczesnego użytkownika okupiono zbyt dużą złożonością obliczeniową algorytmów i trudnościami w praktycznej realizacji założeń standardu.

1.2. Ograniczenia

Głównym ograniczeniem efektywności metod kompresji są zbyt uproszczone modele źródeł informacji wskutek dalece niedoskonałego modelowania rzeczywistości (założenia stacjonarności, gaussowskiej statystyki źródeł). Względność 'optymalnych' dziś rozwiązań wynika także ze sposobu definiowania informacji bez odniesienia do warstwy semantycznej. Często nie sposób przełożyć posiadanej wiedzy *a priori*, dotyczącej analizowanego procesu, na parametry modelu statystycznego. Nie ma jednoznacznych definicji nadmiarowości, która

może występować tak na poziomie pojedynczych danych, jak też obiektywnym i semantycznym. Opisując informację trudno jest stwierdzić, jak zmiana parametrów modelu obiektu decyduje o utracie przez niego 'tożsamości' wpływającej na ilość przesyłanej informacji. Wymaga to definicji pojęcia informacji na wyższym poziomie abstrakcji, a jest to przecież poziom użytkowy typowego odbiorcy.

Interesująca staje się postać reprezentacji danych optymalnej nie w kontekście przyjętych założeń, ale wobec bogatej rzeczywistości form, jakie przyjmuje informacja we współczesnym świecie. Rozważmy prosty przykład. W telewizyjnym teleturnieju uczestnicy zabawy odkrywają kolejno fragmenty obrazu próbując możliwie szybko rozpoznać jego treść. Niekiedy potrzeba bardzo niewiele odsłoniętych elementów, by zidentyfikować znany obraz. Cyfrowy obraz Mona Lisa skompresowany według standardu JPEG z zachowaniem wysokiej jakości rekonstrukcji to 150 kilobajtów danych. Jeśli zaczniemy stopniowo odsłaniać ten obraz rekonstruuując progresywnie kolejne bity skompresowanej reprezentacji, to już po dekompresji 1000 bajtów większość specjalistów potrafi rozpoznać ten obraz. Do rozpoznania obrazu Abrahama Lincolna w podobnym teście przeprowadzonym przez L. Harmona wystarczyło 756 bitów [4].

Problemem jest włączenie wiedzy o świecie dostępnej *a priori* w poszukiwanie optymalnej reprezentacji informacji obrazowej. Zniekształcona, cyfrowa rekonstrukcja Mona Lisy według JPEG jest tylko przybliżeniem, aproksymacją oryginału. O poziomie stratności (nieodwracalności) metody kompresji decyduje poziom aproksymacji danych źródłowych wymagany przez odbiorcę. Niekiedy informację może stanowić jedynie adres internetowy (wskaźnik odpowiedniego obiektu) lub też tekst „Obraz Mona Lisa” powodując reakcję odbiorcy adekwatną do upodobań, posiadanej wiedzy i dostępnych środków (np. wizytę w muzeum Louvre, obejrzenie wysokiej jakości reprodukcji posiadanej w domu czy też książki o dziełach sztuki etc). W innym przypadku odbiorca wymaga reprodukcji najwyższej jakości, co daje kompresja bezstratna cyfrowej postaci obrazu o najwyższej jakości. Występuje tutaj duże podobieństwo z zagadnieniem indeksowania.

1.3. Możliwości udoskonalenia

Możliwe jest bardziej ogólne podejście do zagadnień informacji, kompresji, granic efektywności kodowania, indeksowania. Informacja definiowana jest wtedy jako w dużym stopniu odwołanie do ogólnej wiedzy odbiorcy, pojęcie nadmiarowości konstruowane jest nie jako przewidywalna statystyczna zależność danych w obrębie przekazywanego zbioru (strumienia), ale w znacznie szerszej skali – jako nawiązanie do zasobów wiedzy i środków dostępnych odbiorcy, do realnej semantyki danych. Takie podejście pozwala opracować doskonalsze narzędzia kompresji, czego przykładem są próby optymalizacji kompresji obrazów w ramach standardu JPEG2000.

Zasadniczą cechą współczesnych metod kompresji jest elastyczność, zdolność doboru ilości, jakości i postaci informacji wynikowej (wyjściowej) w zależności od definiowanych potrzeb odbiorcy. Opracowanie skutecznego algorytmu kodowania wymaga zwykle optymalizacji wielokryterialnej, wykorzystania mechanizmów adaptacji do lokalnych cech sygnału (zbioru danych), a czasami nawet interakcji zmieniających procedurę kompresji w czasie rzeczywistym.

Dokładniej te zagadnienia omówiono w kolejnych punktach zwracając uwagę na semantyczną teorię informacji (p.2), dwojaką rolę metod kompresji obrazów (p.3) oraz śledząc nowe paradygmaty kompresji (p.4).

2. Pojęcie informacji

Statystyczna teoria informacji jest znacznie częściej wykorzystywana w metodach kompresji, jednak w nowoczesnych rozwiązaniach coraz większą rolę odgrywa podstawowy, czyli semantyczny aspekt informacji.

2.1. Teoria statystyczna

Prace C. Shannona sprzed ponad 50 lat [5] określiły matematyczne podstawy statystycznej teorii informacji formalizując pojęcia:

- bezstratnej kompresji danych ze źródeł informacji modelowanych procesem losowym o dyskretnym zbiorze wartości (w tym entropia, modele źródeł, twierdzenia o kodowaniu źródeł);
- stratnej kompresji danych ze źródeł informacji modelowanych procesem losowym o ciągłym zbiorze wartości (w tym średnia informacja wzajemna, funkcja zniekształceń źródeł informacji $R(D)$).

Zaproponowany przez Shannona opis informacji znalazł powszechne zastosowanie w różnych dziedzinach nauki, m.in. w biologii, medycynie, filozofii. Występuje w nim stochastyczne rozumienie informacji jako „poziomu niepewności” odbiorcy związanej z przekazywanymi danymi. Wśród transmitowanych danych (symboli źródła informacji) tylko te stanowią informacje, które są nieokreślone (odbiorca nie ma pewności, jaką daną otrzyma). Po otrzymaniu ciągu takich symboli ‘poziom niepewności’ odbiorcy maleje o wartość równą ilości przekazanej informacji.

Informacja występuje w kontekście transmisji (przekazywania) zbioru (strumienia) danych w schemacie: nadawca (źródło informacji), kanał transmisyjny (medium służące przekazaniu sygnału, czyli nośnika informacji) wyposażony w nadajnik (przygotowujący sygnał przenoszący wiadomość do transmisji) i odbiornik (rekonstruujący wiadomość z sygnału), oraz odbiorca (interpretator, użytkownik informacji).

Trudno podważyć ogromne znaczenie teorii Shannona w rozwoju współczesnej nauki. Jednak warto zwrócić uwagę na ograniczenia tej teorii, szczególnie w zakresie uproszczonych, nierealistycznych założeń statystycznych oraz poprzez pominięcie semantyki w definiowaniu źródeł informacji. Ograniczenia te stają się istotne w zastosowaniach do kompresji obrazów.

W statystycznej teorii informacji dominuje składniowy (syntaktyczny) aspekt informacji. Nie prowadzi się rozważań dotyczących prawdziwości czy użyteczności tego, co jest przesyłane. Informacja rozumiana jest wtedy jako zakodowany ciąg symboli źródła informacji nad ustalonym alfabetem z przypisaną wartością semantyczną (znaczeniem) u nadawcy z założenia zgodną z wartością semantyczną u odbiorcy (zobacz rys. 1a)). Reprezentacja R przesyłanych danych powstaje zwykle z kryterium minimalizacji entropii źródła przesyłanej informacji H_S .

[Rys. 1]

2.2. Teoria semantyczna

W niektórych przypadkach semantyka obrazów odgrywa na tyle znaczącą rolę w interpretacji (użytkowaniu, odczytaniu zawartej informacji) przekazywanych danych, że powinna stanowić ważny element w procesie optymalizacji algorytmów kompresji jako ‘uzupełnienie’

statystycznej teorii informacji. Ważnym obszarem zastosowań jest tutaj obrazowanie medyczne.

Dwa zasadnicze cele teorii Shannona to: a) obliczanie ilości informacji dostarczanej przez źródła możliwie wiernie opisujące przesyłane dane (inaczej wyznaczanie teoretycznych wartości granicznych wydajności koderów tych źródeł), b) opracowanie efektywnych kodów dla zdefiniowanych źródeł informacji. Oba te elementy wymagają ustalenia znaczeń pojedynczej danej, kontekstu jej wystąpienia, charakteru reprezentowanej informacji.

Należy tutaj odwołać się do podstaw ogólnej teorii informacji, w której występuje pojęcie semantycznej teorii informacji. Początki tej teorii stanowią prace Bar-Hillela i Carnapa [6] również z lat 50 zeszłego wieku, a istotą jest określanie dodatkowo znaczenia poszczególnych symboli alfabetu źródła informacji. Wykorzystanie ontologicznych i aksjologicznych aspektów rozumienia informacji jest trudne do przełożenia na formalny i algorytmiczny opis kodowania źródeł informacji, jednak wykorzystanie elementów tych teorii wydaje się konieczne w takich zastosowaniach jak np. rekonstrukcja i indeksowanie treści obrazów medycznych w sposób wiarygodny diagnostycznie. Przy przesyłaniu informacji może wtedy odwoływać się jedynie do różnicy semantycznej danych źródłowych u nadawcy i odbiorcy (rys. 1b).

Algorytmiczne wykorzystanie zasad semantycznej teorii informacji może być nieco łatwiejsze poprzez skorzystanie z tych metod matematycznych, które ułatwią semantyczną selekcję informacji w koderach obrazów. Chodzi tu o wykorzystanie analizy funkcjonalnej, które pozwala w większym uolnić się od nierealistycznych założeń statystycznych dobierając bazy przekształceń przybliżających efektywnie lokalne, chwilowe właściwości sygnału. Ułatwia to także dostosowanie metody kompresji do semantyki źródła informacji poprzez większe uporządkowanie informacji (zhierarchizowanie jej opisu) i dokładniejszy jej opis (czasowo-częstotliwościową charakterystykę) w nowej dziedzinie przekształcenia.

W latach 50 zeszłego stulecia powstała 'moskiewska szkoła teorii informacji' z jej najznakomitszym przedstawicielem A.N. Kołmogorowem. Obok probabilistycznych sposobów modelowania źródeł informacji wykorzystano tam także teorię przybliżania (aproksymacji) źródeł z wykorzystaniem metod analizy funkcjonalnej. Stochastyczny proces opisujący źródło informacji zastąpiony jest przez klasę funkcji (sygnałów) f określonych w dziedzinie T . Dowolna funkcja jest aproksymowana i dyskretyzowana przez koder za pomocą sieci aproksymacji. Sieć aproksymacji według koncepcji Kołmogorowa [7] jest zbiorem funkcji możliwie zupełnym i małym (kontrolowanym przez inne pojęcie entropii), aproksymującym istotne cechy sygnałów (funkcji) kodowanych. Kryteria doboru i optymalizacji postaci sieci formułowane jako minimalizacja błędu przybliżenia definiują proces kodowania z dopuszczeniem strat.

Na poziomie ogólności proponowanym przez teorię Kołmogorowa bardzo niewiele można powiedzieć o strukturze optymalnej sieci aproksymacji kształtowanej przez optymalny koder. Choć alternatywna teoria opisu informacji nie prowadzi to dokładnych oszacowań relacji złożoności modelu źródła do ilości reprezentowanej informacji, to jest istotna ze względu na sugestię funkcjonalnego modelowania źródeł. Chodzi o wyznaczenie efektywnych baz przekształceń obrazów w celu lepszego opisu informacji.

Wykorzystanie doświadczeń semantycznej teorii informacji, dodatkowo coraz doskonalszej wiedzy medycznej na temat zasad percepcji psychowizualnej danych obrazowych i obiektywizacji metod ich interpretacji pozwala konstruować doskonalsze sposoby selekcji i porządkowania przesyłanej informacji na podstawie optymalizowanych metod opisu danych (duża rola analizy funkcjonalnej, przede wszystkim analizy harmonicznej, czyli analizy funkcji z wykorzystaniem transformacji o bazie wielu skal dobieranej w przestrzeni czas-częstotliwość).

3. Dwojaka rola kompresji

Kompresja na użytek archiwizacji wymaga niekiedy znacząco różnego podejścia od zastosowań dopuszczających selekcję informacji poprzez jej odpowiednie uporządkowanie czy kwantyzację danych (nazwijmy je selektywnymi). W tradycyjnym podziale odpowiada to kompresji bezstratnej z wiernym zapisem danych (dokładność do pojedynczego bitu) oraz stratnej, gdzie tracona jest możliwość dokładnej rekonstrukcji wszystkich danych źródłowych, jednak niekoniecznie następuje utrata informacji. W tym drugim znaczeniu, kiedy nie sposób do końca określić dopuszczalnych poziomów strat, relacji dane-informacja użyteczna dla odbiorcy, obiektywnych kryteriów i użytecznych zakresów wartości parametrów, algorytmy kompresji mają wiele cech wspólnych z metodami indeksowania.

3.1. Odwracalna kompresja danych

Bitowo bezstratna kompresja obrazów to obecnie przede wszystkim tworzenie możliwie oszczędnej ich reprezentacji z wykorzystaniem zasad statystycznej teorii informacji do celów archiwizacji. Przez ostatnie 10 lat, od czasu opracowania efektywnych koderów CALIC i według standardu JPEG-LS, nie nastąpił żaden przełom, jakościowy postęp w dziedzinie odwracalnej kompresji obrazów. Wspomniane kodery odwracalne na etapie modelowania źródeł informacji wykorzystują 2W (dwuwymiarową) charakterystykę danych w postaci efektywnych modeli predykcji.

Obserwuje się natomiast tendencję modelowania źródeł informacji na niższym poziomie bez pośrednich modeli sąsiednich wartości pikseli źródłowych. Uproszczenia prowadzi do zamiany: zamiast obiektów piksele, zamiast pikseli – rozkłady map bitowych. Stosuje się modelowanie lokalnych kontekstów wybranych przestrzeni bitowych niekoniecznie skorelowanych z obrazową interpretacją danych, szacowanie zależności danych traktowanych bardziej elementarnie (z uproszczonym alfabetem źródła), a co za tym idzie bardziej uniwersalnie.

W koderach odwracalnych kluczową rolę odgrywają metody modelowania kontekstu przy kodowaniu binarnym (zwykle arytmetycznym). W PPM (ang. *prediction with partial string matching*) [8] przy modelowaniu prawdopodobieństw warunkowych adaptacyjnego kodera arytmetycznego wykorzystywane są konteksty różnych rozmiarów. Alfabet linii prawdopodobieństw określonego kontekstu rozszerzany jest dynamicznie, tj. niezerowy podprzedział określonego symbolu a_k pojawia się w linii prawdopodobieństw dopiero wówczas, gdy symbol a_k wystąpi w sekwencji wejściowej poprzedzony tym kontekstem. Ponadto w każdej linii zarezerwowany jest podprzedział symbolu 'PRZEŁĄCZ', który służy do sygnalizacji zmiany rozmiaru kontekstu (dobierany jest możliwie najdłuższy kontekst zaczynając od kontekstu maksymalnego rzędu m). Poszczególne odmiany metody PPM, dostosowane do danych tekstowych, obrazowych, innych, różnią się sposobem określania linii prawdopodobieństw i kontekstów (rozmiarem, kształtem, uproszczeniem statystyki w danym kontekście).

W przypadku, gdy efektywne opisanie jednym kontekstem złożonych lokalnych zależności danych źródłowych nie jest możliwe, stosowane jest niekiedy ważenie kilku modeli prawdopodobieństw bazujących na różnych kontekstach. Jeśli rozkład $P_S^{C_1}$ określono na podstawie kontekstu C_1 , a $P_S^{C_2}$ na podstawie C_2 , wówczas do kodowania można wykorzystać ważony rozkład prawdopodobieństw $(P_S^{C_1} + P_S^{C_2})/2$. Metoda ważenia rozkładów prawdopodobieństw z wykorzystaniem struktury drzewa CTW (ang. *context-tree weighting*) [9] pozwala wykorzystać wszystkie możliwe konteksty i modele kontekstu ograniczone

rozmiarem m . CTW bazuje na binarnym alfabecie źródła informacji bez względu na rodzaj kodowanych danych (obraz, dźwięk, tekst itp.). Wymaga to wstępnej konwersji ciągu symboli nad alfabetem A_S w ciąg bitowy (tj. na wstępnej serializacji, jak w standardzie JBIG i testowanej metodzie BKO), kodowany z wykorzystaniem binarnej struktury drzewa określającej konteksty i prawdopodobieństwa modeli źródła z pamięcią i bez do sterowania kodowaniem arytmetycznym.

Do określenia prawdopodobieństw ciągu symboli wykorzystuje się tzw. prawdopodobieństwo blokowe całego bloku bitów, obliczane na podstawie częstości dotychczasowych (w modelu przyczynowym) wystąpień zer i jedynek zakładając model bez pamięci. Zależności pomiędzy danymi uwzględnione zostały w tzw. prawdopodobieństwach ważonych obliczanych w węzłach tworzonego drzewa binarnego.

Inaczej wyznaczone jest też prawdopodobieństwo symboli, tj. według wyrażenia:

$$P_e(s = a_1) = \frac{n_1 + \frac{1}{2}}{n_1 + n_2 + 1}, \text{ gdzie alfabet źródła } A_S = \{a_1, a_2\}, n_1, n_2 - \text{ liczba wystąpień}$$

odpowiednio a_1 i a_2 w kodowanym dotychczas ciągu źródłowym. $P_e(s = a_2)$ definiowane jest analogicznie.

Coraz większą rolę w praktycznych zastosowaniach odgrywają uniwersalne narzędzia do archiwizacji danych (tzw. archiwizery) wykorzystujące podobne metody statystycznego modelowania źródeł informacji z doбором parametrów zależnie od typu danych. Wykazują one dużą efektywność kompresji skutecznie konkurując z rozwiązaniami dedykowanymi np. kompresji obrazów. Dowodem tego są wyniki licznych eksperymentów (zobacz np. [10]), a także przedstawione poniżej rezultaty przeprowadzonych testów własnych (rys. 2).

[Rys. 2]

Uniwersalny archiwizator WinRK (wykorzystujący odmianę metody PPM do modelowania kontekstu) pozwolił w kilku przypadkach na wyraźne zmniejszenie długości zakodowanych danych obrazowych (średnio o blisko 6%) w stosunku do najefektywniejszego kodera obrazów CALIC. W stosunku do standardów JPEG-LS i JPEG2000 poprawa efektywności sięgnęła odpowiednio 10% i 11%. Nowa wersja WinRK (niestety na obecnym etapie realizacji nieco zbyt złożona obliczeniowo) z algorytmem modelowania PWCM pozwala skrócić skompresowane dane obrazowe o dodatkowe kilka procent.

Dwa uniwersalne kodery map bitowych: według standardu JBIG oraz wykorzystujący nieco bardziej dopasowaną do danych obrazowych metodę serializacji BKO pozwoliły również uzyskać wysoką efektywność kompresji obrazów testowych, przy czym średnia bitowa dla BKO jest na poziomie średniej kodera CALIC.

Okazuje się, że modele informacji obrazowej wyższego poziomu (obiektywne, predykcyjne, transformacji liniowych) nie usprawniają procesu redukcji nadmiarowości wejściowej reprezentacji danych, są za mało elastyczne w dopasowaniu modeli źródeł do lokalnych (chwilowych) cech sygnału. Są one zaś szczególnie istotne przy porządkowaniu, tworzeniu hierarchii informacji uwzględniając niekiedy przesłanki semantyczne w kompresji selektywnej.

3.2. Kompresja selektywna

W kompresji selektywnej poszukiwana jest efektywna reprezentacja semantycznie definiowanej informacji wyrażonej wstępnie danymi w postaci źródłowej, zaś indeksowanie

to tworzenie indeksu (indeksów) obiektów (kolekcji obiektów) danego typu, czyli spisu (zestawu) tych obiektów ze względu na określone ich właściwości (cechy, czyli wartości ustalonego argumentu indeksu – zobacz [16]) mające znaczenie dla użytkownika. Przy wyborze argumentu indeksu (tj. określonych właściwości indeksowanych obiektów) istotna jest rozróżnialność obiektów, a więc wydobycie wyjątkowej treści każdego z obiektów z ogólnego, mało znaczącego tła. Kluczową rolę odgrywa tu dotarcie do realnej informacji związanej z każdym obiektem (w naszym przypadku obrazem), tj. istoty przekazywanej treści. Semantyczne określenie informacji odgrywa tutaj decydującą rolę.

W indeksie występują identyfikatory obiektów wskazujące (pośrednio lub bezpośrednio) ich fizyczną lokalizację. Korzystające z indeksów wyszukiwarki dostarczają na żądanie odbiorcy obiekty o określonych cechach sięgając niekiedy do ich skompresowanej reprezentacji, dekodując i prezentując np. obraz.

Nowe standardy kompresji, np. JPEG2000 przewidują efektywne łączenie kompresji z indeksowaniem, wyznaczanie cech w dziedzinie przekształconych w kompresji danych, co często daje lepszą selektywność wyszukiwania pozwalając porównywać optymalizowane reprezentacje informacji z uwzględnieniem ich znaczenia.

Praca nad standardem kompresji obrazów JPEG2000 jest przykładem szerszego spojrzenia na zagadnienia kompresji selektywnej w ostatnich latach, bogatszego o uwzględnienie wymagań odbiorcy i subiektywnego znaczenia przesyłanych mu danych. Możliwa jest ingerencja w procedury przetwarzania danych oraz organizacji strumienia wyjściowego, a przez to wykorzystanie wiedzy wynikłej z doświadczenia użytkownika, tradycji, kultury itd. Opracowano zagadnienia interakcyjnej transmisji, rekonstrukcji obrazów z ingerencją odbiorcy w jakość i ilość otrzymywanych danych, które stają się realną informacją weryfikowaną przez odbiorcę (jego cele). Wprowadzono elementy zabezpieczenia danych, ochrony własności intelektualnych, ukrywania informacji przed postronnymi odbiorcami. W ostatnich miesiącach trwają intensywne prace nad indeksowaniem obrazów w ramach nowego standardu JPSearch, tworzonego przez członków komitetów JPEG i MPEG jako zintegrowane uzupełnienie standardów JPEG.

Realizacja takich założeń nie jest prosta i na obecnym etapie, niedoskonała. Przeprowadziliśmy szereg testów porównawczych kodera JPEG2000 z realizacją starszego standardu JPEG. Ocena uzyskanych rezultatów nie jest jednoznaczna – zobacz przykład z rys. 3.

[Rys. 3]

O ile w przypadku obrazu testowego Lenna wyraźnie lepsza jakość rekonstrukcji kompensuje kilkukrotnie większą złożoność obliczeniową kodera JPEG2000, to jednak w przypadku obrazu Frog trudno wskazać lepszą wersję w ocenie subiektywnej. Główną zaletą JPEG2000 nie jest więc jego efektywność, ale większa możliwość dostosowania do semantycznej warstwy przekazywanej informacji.

Zastosowania medyczne wymagają jeszcze szerszego spojrzenia na zagadnienie kompresji. Nie bez przyczyny nie ma dotąd powszechnej zgody środowisk lekarskich na stosowanie kompresji stratnej w obrazowaniu medycznym. Obawy budzi wiarygodność diagnostyczna rekonstruowanych obrazów oraz sposób ustalenia optymalnych parametrów procesu kompresji. Nie przekonuje wiele prac dowodzących nawet poprawy jakości obrazów po kompresji [17,18], nie ma jasnych reguł prawnych [19]. Sformułowanie obiektywnych kryteriów optymalizacji algorytmu kompresji tych obrazów jest nadal sprawą otwartą, choć jest przedmiotem licznych badań od wielu lat.

4. Nowe paradygmaty kompresji

Ciągle powracają pytania o naturę informacji, skuteczny jej opis ilościowy i jakościowy, metody ekstrakcji, klasyfikacji, indeksowania informacji zmieniającej swój charakter, zasięg, nośnik, znaczenie. Zmianie ulegają paradygmaty, czyli wzorce, standardy stosowanych z sukcesem rozwiązań. Standardy opracowywane przez kilka lat, w przeciągu kilku następnych tracą swoją aktualność, użyteczność (np. standard kompresji obrazów JPEG, coraz mniej użyteczny w wielu zastosowaniach).

Uzupełnieniem przedstawionych rozważań są zebrane na podstawie licznych eksperymentów (w tym własnych) w zakresie kompresji danych (w tym obrazowych) stwierdzenia i wskazówki sugerujące kierunek doskonalenia obowiązujących dziś paradygmatów kompresji

4.1. Uniwersalny archiwizator

W przypadku odwracalnej kompresji danych istotne okazują się następujące spostrzeżenia:

- świat 'ujęty' w formie rejestracji określonego stanu rzeczywistości w danej chwili nie ma charakteru gaussowskiego; znaczy to, że:
 - dekorelacja danych nie oznacza ich statystycznej niezależności, a model źródła bez pamięci staje się mało skuteczny; wynika stąd, że wykorzystanie modeli z pamięcią w algorytmach kwantyzacji i binarnego kodowania znacząco zwiększa efektywność opisu źródeł informacji;
- obrazy zazwyczaj nie mają charakteru stacjonarnego, ani nawet ergodycznego więc:
 - ważną rolę odgrywa adaptacyjność algorytmów modelowania źródeł, stosowanie metod segmentacji i lokalnej klasyfikacji, zarówno w skali globalnej (na poziomie obrazów) jak i lokalnej (na poziomie fragmentów obrazów);
 - wymagane są wiarygodne sposoby szacowania prawdopodobieństw symboli na podstawie niewielkiej statystyki danych, skuteczne metody upraszczania, kwantyzacji, wyboru kontekstów;
 - inne niż statystyczne metody opisu źródeł informacji okazują się nieskuteczne (predykcyjne, transformacyjnego kodowania, obiektowe).

Najlepszym rozwiązaniem do celów archiwizacji wydaje się uniwersalny koder danych z możliwością doboru modelu źródła zależnie od charakteru danych.

4.2. Elastyczny koder wielu skal

Selektywna kompresja informacji rządzi się innymi prawami. Warto zauważyć, że:

- modelowane obrazy zachowują cechę samopodobieństwa w zmieniającej się skali (po rozciągnięciu obrazu do większej dziedziny i przeskalowaniu zachowane zostają właściwości statystyczne oryginału), a więc:
 - kluczową rolę odgrywają przekształcenia (transformacje) danych ze skalowaniem rozdzielczości, zachowujące informację o położeniu i treści częstotliwościowej współczynników poszczególnych skal;
 - szczególnie ważnym elementem wynikającym z rozkładu zależności danych w przestrzeni ze skalowaniem rozdzielczości jest kodowanie pozycji współczynników transformaty i ich znaczeń (względem ustalonej wartości progowej).

Efektywny koder obrazów powinien zapewnić następujące cechy reprezentowanej informacji:

- hierarchiczność korelującą z semantycznym znaczeniem kolejnych przybliżeń źródła informacji (progresja treści) zapewniająca progresję optymalną w sensie R(D) (tj. od maksymalnego przyrostu informacji na bit transmitowanych danych do minimalnego);
- selektywność rozkładu informacji w poszczególnych obszarach wielu skal dziedziny z możliwością rekonstrukcji informacji zarówno w pełnej, niezakłóconej wersji (nawet w sensie odwracalnej kompresji danych), jak też odpowiednio wybranej, przybliżonej (w możliwie szerokim zakresie) na poziomie pojedynczych pikseli, obiektowym lub semantycznym; elastyczność pozwalającą na realizację porządku kodowania i dekodowania informacji zależnie od wymagań użytkownika (interaktywność procesu kompresji/dekompresji).

Trwają poszukiwania optymalnych przekształceń obrazów w wielu skalach. Klasyczne falkowe przekształcenie obrazów zakłada separowane jądro transformacji (sekwencja przekształceń 1W po wierszach i kolumnach), co nie pozwala dobrze opisać obiektów w obrazach (występuje uwypuklenie informacji w kierunku pionowym i poziomym kosztem innych kierunków). Aproksymacja nieliniowa gładkiej krzywej jest często lepsza w przypadku wykorzystania transformacji wielorozdzielczych z jądrem 2W (*curvelets*), filtrów kierunkowych (*contourlets*), przekształceń geometrycznych po podziale obrazu na bloki o różnej wielkości, co daje lokalny opis informacji w różnej skali (*beamlets*, *wedgelets*, *bandlets* i inne). Zastosowanie oszczędniejszej, bardziej upakowanej i uporządkowanej reprezentacji danych w dziedzinie tych przekształceń pozwala również zwiększyć selektywność rozwiązań uwzględniających znaczenie rozkładów symboli na poszczególnych poziomach tak uzyskanej hierarchii informacji obrazowej. Wykorzystywane są tutaj podstawy teoretyczne kształtowania baz – atomów czas-częstotliwość dostarczane przez analizę harmoniczną i funkcjonalną. Uzyskane już rezultaty poprawy efektywności w stosunku do JPEG2000 są obiecujące [20] i sugerują kierunek doskonalenia metod kompresji sygnałów w ramach planowanego (przez komitet JPEG) na 2010 rok nowego standardu kompresji – AIC (ang. *advanced image coding*).

5. Podsumowanie

Trudno sobie wyobrazić rozwój nowoczesnej telekomunikacji, telemedycyny, e-biznesu, obieg różnego typu dokumentów w Internecie, doskonalenie nowoczesnych technologii sieciowych typu grid, skuteczne wyszukiwarki czy globalne bazy danych bez optymalizowanych algorytmów kompresji, oszczędnych formatów, użytecznych standardów z wieloaspektową charakterystyką strumieni danych przesyłanych i gromadzonych, sprzętowych realizacji kodeków etc.

Kompresja danych jako efektywna i uniwersalna koncepcja reprezentowania informacji poparta efektywnymi realizacjami, przedmiot działalności naukowej i sztuki inżynierskiej jest obiektem zainteresowań w wielu obszarach zastosowań współczesnych technologii teleinformatycznych. Rozumiana jako poszukiwanie, tworzenie i badanie oszczędnej, a zarazem funkcjonalnej reprezentacji danych związanych z informacją także w sensie znaczeniowym pozwala przyspieszyć i wzbogacić przekaz wiadomości, zwiększyć niezawodność wymiany informacji pomiędzy nadawcą i odbiorcą, lepiej archiwizować posiadane zasoby.

W przypadku obrazów szczególnie istotny okazuje się aspekt semantyczny, interpretacyjny, domagający się uwzględnienia charakterystyki (profilu) odbiorcy w procesie wyznaczania użytecznej reprezentacji informacji. Praca nad metodami kompresji danych obrazowych w zakresie określania znaczeń przesyłanych pakietów danych zyskuje przez to

wymiar bardziej humanistyczny i służebny wobec wielu oczekiwań, zamierzeń i ograniczeń współczesnego człowieka.

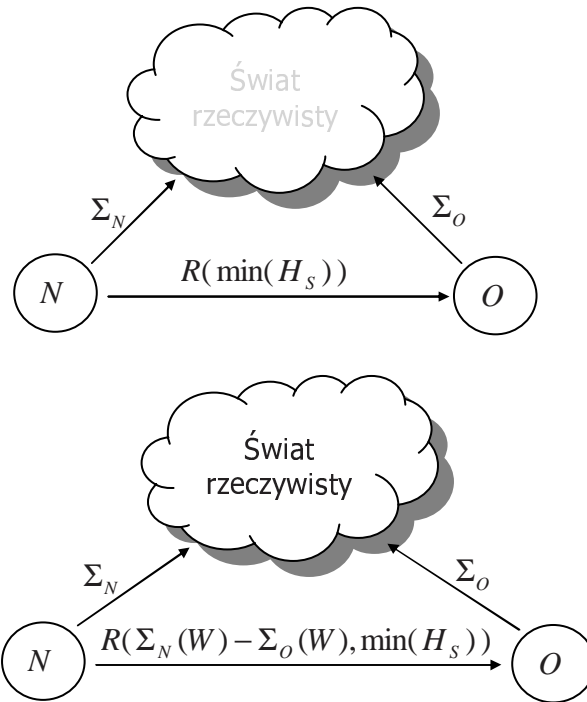
Zwrócono uwagę na wiele otwartych kwestii przede wszystkim w zakresie rozumienia pojęcia informacji, sposobów selekcji informacji zależnie od subiektywnych wymagań odbiorcy, charakterystyki percepcyjnych zdolności odbiorcy, integracji z rozwiązaniami stosowanymi w indeksowaniu i innych metodach przetwarzania informacji. Nieustające, zintensyfikowane w ostatnich latach prace nad multimedialnymi standardami komitetów JPEG i MPEG stanowią dowód dużego zaangażowania badaczy oraz rosnących potrzeb i niesłabnącego zainteresowania rynku. Aby im sprostać konieczna jest silniejsza integracja zarówno w obrębie różnych dyscyplin i specjalności technicznych, jak i obszarów wiedzy związanych z zastosowaniem technik informacyjnych.

Ponieważ dziś coraz cenniejszy staje się czas i selekcja realnej informacji, a każdy dzień przynosi nowe technologie gromadzenia i przekazu informacji, poszukiwanie optymalnych metod reprezentowania informacji staje się palącą potrzebą chwili. Sugerowane kierunki badań wydają się być użyteczne.

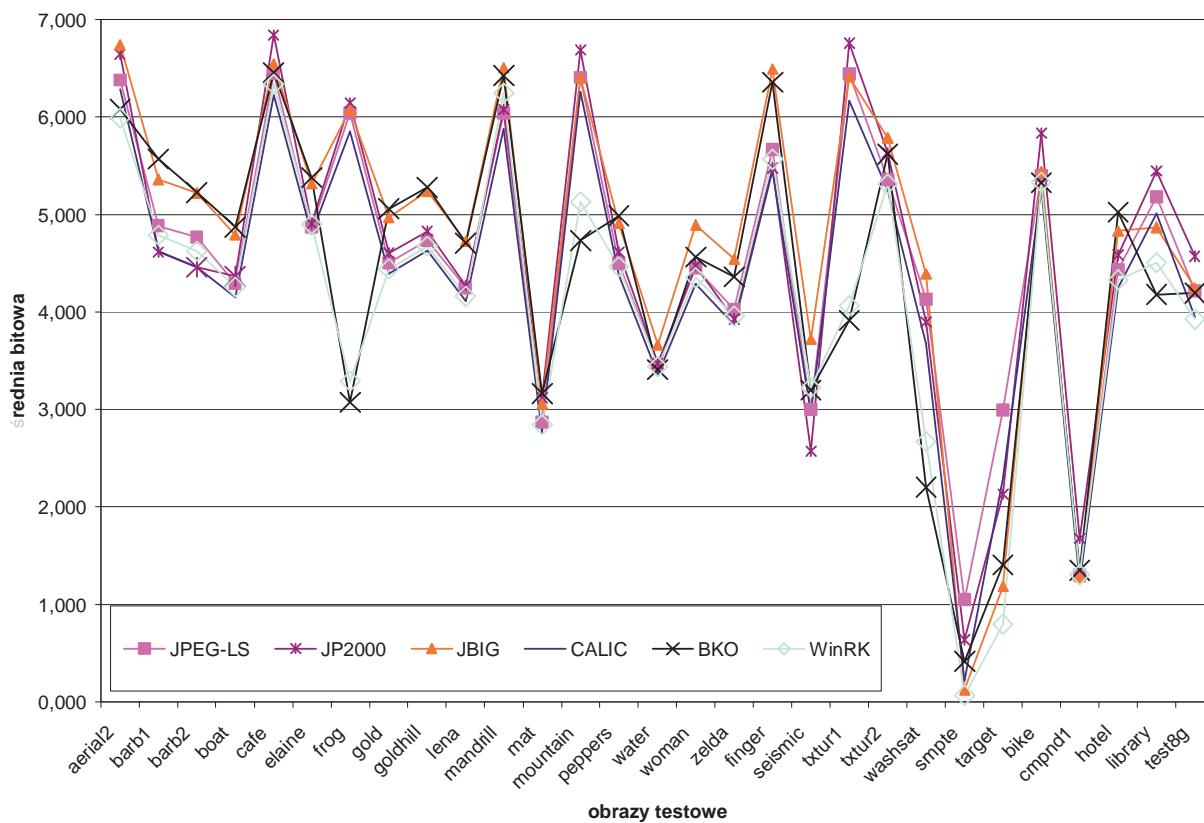
Referencje:

- [1] Manuel Castells on the Network Society: <http://www.tidec.org/geovisions/Castells.html>
- [2] Brockman J.: *The Third Culture: Beyond the Scientific Revolution*. Simon & Schuster, 1995.
- [3] Eriksen T.H.: *Tyrania chwili*. PIW, 2003.
- [4] Donoho D.L., Vetterli M., DeVore R.A., Daubechies I.: *Data compression and harmonic analysis*. Invited paper, IEEE Trans. Inform. Theory, Special Issue, Inform. Theory: 1948-1998 Commemorative Issue, nr 44(6), str. 2435-2476, Oct. 1998.
- [5] Shannon C.E.: *A Mathematical theory of communications*. Bell System Technical Journal, nr 27, str. 379-423 i 623-656, 1948.
- [6] Bar-Hillel Y., Carnap R.: *Semantic information*. British Journal for the Philosophy of Science, nr 4, str. 47-157, 1953.
- [7] Kolmogorov A.N, Tikhomiriv V.M: *ϵ -entropy and ϵ -capacity*. Uspekhi Mat. Nauk, nr 14, str. 3-86 (tłum. ang. Amer. Math. Soc. Transl., Ser 2 Vol 17, str 277-364, 1959.
- [8] Moffat A.: *Implementing the PPM data compression scheme*. IEEE Trans Comm COM-38, nr 11, str 1917-1921, 1990.
- [9] Willems F.M., Shtarkov Y.M., Tjalkens T.J.: *The context-tree weighting method: basic properties*. IEEE Trans. Information Theory, IT-41, str 653-664, 1995.
- [10] Lossless data compression software benchmarks/comparisons: <http://www.maximumcompression.com/>
- [11] SPMG/ JPEG-LS V.2.2 codec, <ftp://spmgece.ubc.ca/pub/jpeg-ls/ver-2.2/>, Signal Processing & Multimedia Group at the University of British Columbia, USA.
- [12] <http://www.cl.cam.ac.uk/~mgk25/jbigkit/>
- [13] X. Wu, CALIC codec, ftp://ftp.csd.uwo.ca/pub/from_wu/v.arith/
- [14] Przelaskowski A.: *Binarny koder obrazów ze skalą szarości*. Prace Naukowe Instytutu Telekomunikacji i Akustyki Politechniki Wrocławskiej, nr 85, seria: Konferencje, nr 29, X Sympozjum Nowości w Technice Audio i Wideo, str. 243-252, Wrocław, 2004.
- [15] http://www.msoftware.co.nz/downloads_page.php
- [16] W. Skarbek: *Podstawy multimediiów*. Podręcznik elektroniczny dla studiów zaocznych OKNO, Politechnika Warszawska, 2004.
- [17] Chang S.G., Yu B., Vetterli M.: *Adaptive wavelet thresholding for image denoising and compression*. IEEE Trans Image Proces. nr 9, str1532-1546, 2000.
- [18] Sarty G.E., Atkins M.S., Pierson R.A.: *Sharper MRI of ovarian and uterine masses via wavelet-based compression*. 20th Annual Canadian Western Society for Reproductive Biology Workshop, 1998.
- [19] Brownrigg P., Bak P.R.G.: *What is the legal risk in using irreversible compression on diagnostic images in medical management? - a legal assessment and review of the regulatory framework in the USA and Canada*. SCAR 2005 Annual Meeting, Orlando, USA.

[20] Wakin M., Romberg J., Choi H., Baraniuk R.: Wavelet-domain approximation and compression of piecewise smooth images. To appear in IEEE Tran Image Proc, 2005.



Rys. 1. Zróźnicowanie pojęcia przesyłanej informacji: a) reprezentacja danych R źródła S o minimalizowanej entropii przesyłanej od nadawcy N do odbiorcy O przy założeniu zgodnych wartościach semantycznych $\Sigma_N(W) \cong \Sigma_o(W)$, gdzie $\Sigma(\cdot)$ jest funkcją semantyczną; b) reprezentacja informacji o minimalizowanej entropii źródła opisującego jedynie różnicę znaczeń u nadawcy i odbiorcy $\Sigma_N(W) - \Sigma_o(W)$ w odniesieniu do przesyłanej wiadomości W .



Rys. 2. Porównanie efektywności bezstratnych koderów obrazów. Wykorzystano 29 obrazów testowych w 8 bitowej skali szarości (z prac standaryzacyjnych komitetu JPEG), o nazwach jak na rysunku. Wartości średnich bitowych zbioru wszystkich obrazów testowych dla poszczególnych koderów wynoszą odpowiednio: 4,55 bnp (JPEG_LS [11]), 4,60 bnp (JPEG2000 VM8.6), 4,75 bnp (JBIG [12]), 4,35 bnp (CALIC [13]), 4,36 (BKO [14]) oraz 4,1 bnp (WinRK v. 2.1.6 [15]). Skrót bnp oznacza ‘bitów na piksel’.

JPEG2000

JPEG

a)



b)



Rys. 3. Porównanie efektywności selektywnej kompresji obrazów według standardów JPEG i JPEG2000: a) obraz testowy Lenna odtworzony po kompresji w stopniu 50:1 (JPEG daje wartość PSNR równą 28,18 dB, a dla JPEG2000 PSNR=30,5 dB); b) obraz testowy Frog zrekonstruowany po kompresji w stopniu 30:1 (JPEG daje PSNR=24,8 dB, a JPEG2000 – PSNR=25,46 dB). W ocenie subiektywnej trudno wskazać wersję Frog o lepszej jakości.