

Wstęp

Początek XXI wieku to przede wszystkim era informacji, którą charakteryzuje intensywny rozwój społeczeństwa informacyjnego, przekształcającego się stopniowo w społeczeństwo szeroko dostępnej wiedzy (przetwarzanie i tworzenie wiedzy staje się podstawą jego rozwoju). Wobec dużej i stale rosnącej zwiększającej się roli Internetu można obecnie mówić również o zjawisku społeczeństwa sieciowego.

Cena chwili

Znakiem czasu jest nadmiar informacji. Jesteśmy zalewani informacjami różnej jakości, które dotyczą każdej, niemal najdrobniejszej dziedziny ludzkiego życia). Coraz bardziej pożądane stają się narzędzia nie tyle zwiększające dostęp, ile raczej ułatwiające skuteczną selekcję informacji. Szybki dostęp do istotnych informacji, efektywne przetwarzanie informacji i wyszukiwanie wiarygodnych źródeł wiedzy staje się składnikiem decydującym o sukcesie wszelkiej działalności biznesowo-gospodarczej, organizacyjnej, społecznej, „gorącym” towarem na sprzedaż.

Powszechnym sposobem na życie współczesnego człowieka, żyjącego w ciągłym pośpiechu, staje się oszczędzanie każdej chwili (zobacz pozycję T. H. Eriksen, „*Tyrania chwili*”, PIW, 2003), mające znaczenie zwłaszcza ekonomiczne oraz eksploatacyjne (użytkowe). 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 można uzyskać przede wszystkim poprzez dbałość o jakość informacji, tj. selekcję jedynie tych danych, które są istotne dla użytkownika, oraz poprzez zapewnienie minimalnej długości reprezentacji przekazywanych danych.

Rodzą się więc pytania o naturę informacji, jak ją opisać ilościowo, jak lepiej rozumieć jej sens, wybierać, przetwarzać, porównywać, klasyfikować. Od strony technicznej wymaga to konstrukcji narzędzi i systemów zorientowanych na cyfrową reprezentację informacji, najlepiej tworzoną w sposób upakowany, oszczędny, który ułatwia jej analizę, przechowanie i przekazanie. Celem kompresji danych jest uzyskanie optymalnego sposobu reprezentowania informacji. Kryteria optymalizacji zależą od zastosowań.

Kompresja oszczędza czas

Kompresja zmniejsza przesyłanych zbiorów i strumieni danych, daje szybszy dostęp do informacji w bazach danych, a także porządkuje i wpływa na jakość dostarczanej informacji. Rośnie ilość informacji zapisywanej na nośniku o danej pojemności, można łatwiej (szybciej) tę informację przetwarzać, porównywać czy indeksować. Rośnie ilość informacji dostarczonej w jednostce czasu, jest ona bardziej użytecznie reprezentowana, bez zbędnej, zabierającej czas nadmiarowości. Efektywna kompresja dociera do istoty informacji, także w sensie znaczeniowym, jest więc narzędziem niezwykle przydatnym, niekiedy wręcz koniecznym, jest jednym ze sposobów na zwiększenie „prędkości świata” (lub też lepsze wykorzystanie uciekającego czasu).

Rozwój technologiczny w zakresie doskonalenia jakości i zwiększania pojemności nośników informacji oraz zwiększania przepustowości sieci komputerowych (lokalnych i globalnych) nie nadąża za lawinowym wzrostem liczby danych cyfrowych wymagających przetworzenia, analizy, archiwizacji i przesłania w danym środowisku sieciowym. Typowe skryningowe badanie mammograficzne to dwa obrazy każdej z piersi wyko-

nane w projekcjach kranio-kaudalnej i skośnej. Każdy z obrazów to średnio około 40 MB (megabajtów) danych, co daje $4 \times 40 \text{ MB} = 160 \text{ MB}$ danych z jednego badania. Płyta CD o pojemności 700 MB może pomieścić dane cyfrowe jedynie z 4 takich badań. Czas przesłania pojedynczego mammogramu łączem o przepustowości 1 Mbit to 320 sekund, czyli 5 minut i 20 sekund, a całego badania – 21 minut i 20 sekund. Jeśli dziennie wykonuje się typowo około 30 badań (czas całego badania mammografem cyfrowym wynosi najwyżej 20 minut), to istnieje potrzeba archiwizacji $30 \times 160 \text{ MB} = 4,8 \text{ GB}$ danych dziennie z jednego pracującego mammografu (cała płyta DVD-5, jednostronna, jednowarstwowa). Dane te trzeba przechowywać przez okres przynajmniej 10 lat.

Pascal pisząc list do przyjaciela wspominał, że „nie ma czasu, aby napisać krócej” – pisze więc jak leci. Przepraszał, że zabrakło mu czasu, by zaoszczędzić czas adresata. Opracowanie oszczędzających czas metod kompresji jest niełatwe, wymaga rozwagi, cierpliwej analizy właściwości danych, wielu godzin poszukiwań efektywnych modeli źródeł informacji czy algorytmicznych metod formowania zakodowanej w ciąg bitów postaci informacji. Przez ostatnie 10 lat, od czasu opracowania efektywnych koderów CALIC i JPEG-LS, nie nastąpił żaden przełom, jakościowy postęp w dziedzinie odwracalnej kompresji obrazów. Setki badań, dziesiątki artykułów, konferencji, tysiące godzin mozolnej pracy wielu badaczy nie przełożyło się na wzrost efektywności koderów obrazów.

Cechy kompresji danych

Kompresja staje się zagadnieniem coraz bardziej powszechnym, a stosowane rozwiązania wpływają na kształt współczesnych systemów teleinformatycznych. Skuteczne metody kompresji wychodzą poza proste algorytmy kodowania i redukcji nadmiarowości. Muszą uwzględniać specyfikę zastosowań, chociażby poprzez elastyczne dostosowanie do różnicowanych wymagań użytkownika, warunków transmisji (odporność na błędy), cech systemu informatycznego. Zawierają zabezpieczenia przed kradzieżą danych, dają możliwość ukrycia informacji przed osobą niepowołaną, czy też różnego uporządkowania przekazywanej (odtwarzanej) informacji poprzez dostosowanie do interaktywnych wskazówek odbiorcy, zmiennego kryterium optymalizacji postaci kodu czy też poprzez uwzględnienie wymagań strumieniowania.

Zasadniczą cechą współczesnych metod kompresji jest więc elastyczność, zdolność doboru ilości, jakości i postaci informacji wynikowej (wyjściowej) w zależności od definiowanych potrzeb odbiorcy. Algorytm tworzenia zakodowanej reprezentacji powinien realizować takie zadania przy możliwie skromnych kosztach (obliczeniowych i sprzętowych). Opracowanie skutecznego algorytmu kodowania wymaga więc zwykle optymalizacji wielokryterialnej, wykorzystania mechanizmów adaptacji do lokalnych cech sygnału (zbioru danych), a czasami nawet interakcji zmieniających procedurę kompresji w czasie rzeczywistym.

Algorytmy kompresji wykorzystują oprócz teoretycznych podstaw informacji, modeli źródeł informacji także kody, czyli metody tworzenia efektywnej reprezentacji danych. Przedstawiając to obrazowo: jeśli system wymiany (przekazywania) informacji jest nerwem, krwioobiegim współczesnego społeczeństwa, to kompresja jest sposobem na koncentrację i wzmocnienie przekazywanych bodźców, a wskutek tego – na lepsze odżywianie całego organizmu. Jest jak technologiczne lekarstwo nie tylko dające szansę normalnego życia, ale stwarzające również możliwości dalszego rozwoju.

Kompresja obrazów

Technologiczne doskonalenie systemów rozpowszechniania informacji wymaga stosowania efektywnych metod kompresji danych, w tym koderów obrazów. Obecnie niebagatelną rolę odgrywa przekazywanie informacji za pomocą obrazu – przykładem niech będzie chociażby dominacja systemów operacyjnych z graficznym interfejsem użytkownika typu *Windows*. Nowe technologie akwizycji obrazów, poprawa zdolności rozdzielczej systemów obrazowania, redukcja szumu i zwiększanie dynamiki sygnału użytecznego wymuszają 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 medycznego i biologicznego itd., wymaga użycia efektywnych metod gromadzenia, indeksowania, przeglądania i wymiany tej informacji.

Bezstratna kompresja obrazów polega na tworzeniu możliwie oszczędnej ich reprezentacji, dającej się odtworzyć do postaci oryginalnej (wejściowej dla kodera obrazów) z dokładnością do pojedynczego bitu. Jest więc bezpieczna, niezwiązana z obawą utraty części informacji, co jest szczególnie istotne, np. w zastosowaniach medycznych. 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 medycznych po kompresji, brakuje jasnych reguł prawnych. Sformułowanie obiektywnych kryteriów optymalizacji algorytmów kompresji stratnej jest nadal sprawą otwartą, choć jest przedmiotem licznych badań od wielu lat. Wymaga szerszego spojrzenia na pojęcie informacji przede wszystkim w warstwie znaczeniowej. Kompresja bezstratna pozostaje więc bezpiecznym i najbardziej powszechnym sposobem reprezentowania obrazów medycznych. Podobne dążenia do zachowania źródłowej postaci obrazów w procesie kompresji występują w zastosowaniach cyfrowej telewizji, kina domowego, grafice komputerowej, przy wykorzystaniu niektórych zdjęć satelitarnych itd.

Podstawowym elementem wszystkich użytecznych metod kompresji są efektywne algorytmy kodowania binarnego, wykorzystujące wiarygodne modele zależności danych przeznaczonych do kodowania, tj. danych obrazowych lub danych reprezentacji pośredniej, uzyskanych po wstępnym przetworzeniu danych źródłowych (bardziej podatnych na kodowanie).

Książka

Opracowanie to jest próbą całościowego ujęcia zagadnienia kompresji obrazów, z wykorzystaniem fundamentalnych koncepcji z teorii informacji rozbudowanych o wybrane aspekty semantyczne, zależne od rodzaju informacji, jak i od metody wiarygodnego modelowania danych obrazowych. Ponadto książka daje teoretyczne podstawy metod binarnego kodowania, które są szeroko wykorzystywane we współczesnych standardach kompresji, stosowanych formatach zapisu, przesyłania i prezentacji obrazów. Dla poszczególnych metod podana jest ogólna zasada kompresji, czyli określona reguła modelowania danych i konstruowania kodu z wykorzystaniem prostych zależności matematycznych. Z reguły tej wynika konkretny algorytm umożliwiający opracowanie w praktyce koderów obrazów. Podane algorytmy wzbogacono niekiedy przykładami gotowych realizacji w określonym języku programowania. Istotna jest też analiza właściwości poszczególnych koderów, ich efektywności dla różnego

rodzaju danych. W celu lepszego zrozumienia bardziej złożonych pojęć i metod zamieszczono formalne opisy (wybrane twierdzenia, dowody, własności) oraz przykłady, które ukazują efektywność i użyteczność omawianych rozwiązań, jak też wyjaśniają ich działanie.

Obserwuje się tendencję modelowania źródeł informacji na niższym poziomie, bez pośrednich modeli zależności wartości sąsiednich pikseli źródłowych. Uproszczenia prowadzą do zmian w analizie informacji: zamiast obiektów piksele, zamiast pikseli – rozkłady map bitowych itp. Stosuje się modelowanie lokalnych kontekstów wybranych przestrzeni bitowych niekoniecznie skorelowanych z obrazową interpretacją danych. Szacowanie zależności danych traktowanych bardziej elementarnie (z uproszczonym alfabetem źródła) daje bardziej uniwersalne rozwiązania. Innymi słowy w koderach odwracalnych coraz większą rolę odgrywają uniwersalne metody kodowania danych, wykorzystujące przede wszystkim statystyczne metody modelowania kontekstu w kodowaniu binarnym (zwykle arytmetycznym). Potwierdzają to przytoczone w ostatnim rozdziale rezultaty eksperymentów własnych i innych. Dlatego też uznano za celowe wnikliwe przedstawienie najbardziej użytecznych metod kompresji danych, które mają znaczenie uniwersalne i są wykorzystywane w koderach obrazów. Pokazano to w podsumowującym rozdziale siódmym. Wydaje się, że sukces najbardziej użytecznych dziś rozwiązań zapewniła doskonalenie statystycznych, adaptacyjnych metod modelowania danych w szybkich koderach odpowiednio formowanych strumieni bitowych.

Pierwszy rozdział jest wprowadzeniem w zagadnienie kompresji, podkreśla dużą rolę, jaką odgrywają efektywne metody kodowania we współczesnych systemach teleinformatycznych, definiuje podstawowe pojęcia i zwraca uwagę na potrzebę rzetelnej charakterystyki kodowanych danych. Opisano tam podstawowe cechy obrazów, zwrócono uwagę na zróżnicowanie danych obrazowych, a co za tym idzie, metod ich efektywnej kompresji. Przedstawiono także paradygmat kompresji obrazów w wersji bezstratnej dający wzorzec elastycznego kodera obrazów naturalnych, biomedycznych i innych, przy czym zwrócono uwagę na problem kompresji stron dokumentów czy reklam zawierających obraz i tekst (tzw. obrazy mieszane). Dokonano syntetycznej charakterystyki metod kompresji obrazów, rozwiniętej w dalszej części opracowania. Krótki rys historyczny ukazuje intensywny rozwój różnorodnych koncepcji kompresji danych na przestrzeni ostatnich dziesięcioleci i duże zainteresowanie tą tematyką badań na całym świecie. Rozdział ten zawiera także kilka praktycznych wskazówek, jak tworzyć własne kodery, jak zrealizować prosty algorytm, gdzie znaleźć w Internecie ogólnodostępne biblioteki oprogramowania, bogate w użyteczne procedury kompresji.

W kolejnych pięciu rozdziałach przedstawiono najbardziej użyteczne metody binarnego kodowania, składające się na paradygmat bezstratnej kompresji obrazów z rozdziału pierwszego.

W **drugim rozdziale** przedstawiono teoretyczne podstawy kodowania wynikające z klasycznej teorii informacji, uzupełnionej o wskazanie praktycznych ograniczeń oraz użytecznych modyfikacji, które mogą zwiększyć efektywność kompresji. Ważne jest tutaj przede wszystkim omówienie pojęcia informacji w ujęciu stochastycznym i semantycznym oraz nadmiarowości reprezentacji danych i właściwości stosowanych modeli źródeł informacji. Oprócz ogólnych zasad kodowania według Shannona podano także schematy blokowe bezstratnych metod kompresji oraz adaptacyjnego sposobu realizacji algorytmów kodowania. Zamieszczono przykłady bardziej szczegółowych rozwiązań metod modelowania i binarnego kodowania aby dokładniej wyjaśnić zagadnienia teoretyczne. Treści tego rozdziału umożliwiają lepsze zrozumienie kon-

cepcji zaawansowanych metod kodowania, określają także warsztat niezbędny do projektowania własnych rozwiązań.

Rozdział trzeci dotyczy optymalnej metody kodowania symboli znanej jako metoda Huffmana. Idea kodów symboli przypisujących różne słowa kodowe poszczególnym symbolom danych kodowanych przyrostowo znalazła optymalną realizację już wiele lat temu. David A. Huffman, jeden z największych autorytetów teorii informacji zaproponował w swojej pracy magisterskiej z roku 1952 metodę wyznaczania słów kodowych na podstawie wag poszczególnych symboli (tj. szacowanych prawdopodobieństw ich występowania w kodowanym strumieniu danych). Jest to wstępująca metoda łączenia węzłów o najniższej wadze w nowe węzły z sumą ich wag, w odróżnieniu od metody zstępującej Shannona-Fano polegającej na dzieleniu grupy węzłów na dwie podgrupy o możliwie zbliżonej sumie wag. W rozdziale scharakteryzowano obie te metody, podano algorytmy, wykazano optymalność kodu Huffmana. Ponadto wskazano statyczne i dynamiczne (adaptacyjne) sposoby realizacji algorytmu Huffmana, zwracając uwagę na wiele problemów związanych z konstrukcją i modyfikacją binarnego drzewa Huffmana, struktury przydatnej w wyznaczaniu kodu optymalnego. Szczególne przypadki źródeł o tzw. geometrycznych rozkładach prawdopodobieństw wystąpienia symboli można kodować za pomocą uproszczonych wersji kodu Huffmana, tj. kodów unarnego, Golomba, Rice'a i innych kodów przedziałowych. Kody te, powszechnie wykorzystywane we współczesnych standardach i uznanych formatach kompresji danych (w tym obrazowych), opisano wraz z przykładami zastosowań, zwracając uwagę na łatwość ich realizacji i uzyskiwaną szybkość algorytmów kodowania. Rozdział zakończono syntetycznym zestawieniem omówionych kodów symboli.

Czwarty rozdział dotyczy strumieniowych metod kodowania binarnego według koncepcji kodowania arytmetycznego. Jedno słowo kodowe jest w tym przypadku przypisane całemu zbiorowi kodowanych danych jako liczba ułamkowa z przedziału $[0, 1)$. Podano algorytm kodowania arytmetycznego określający sposób obliczania tej liczby, a także sposób realizacji tej koncepcji w arytmetyce całkowitoliczbowej. Ponadto poruszono dwa istotne zagadnienia: statystyczne modelowanie kontekstu służące określeniu w sposób adaptacyjny aktualnych prawdopodobieństw wystąpienia poszczególnych symboli źródła opisującego dane źródłowe (m.in. metody PPM i CTW) oraz binarny algorytm kodowania (dla alfabetu składającego się z 0 i 1), który umożliwia szybką i efektywną realizację koderu arytmetycznego. Podkreślono duże znaczenie adaptacyjnych koderów binarnych, obecnych we wszystkich nowoczesnych standardach kompresji obrazów. Udoskonalone realizacje binarnych koderów arytmetycznych wykorzystano w takich standardach jak: JBIG, JBIG2, JPEG2000, MPEG-4, zaś uproszczone implementacje przybliżeń koncepcji kodowania arytmetycznego o mniejszej złożoności czasowej i pamięciowej to m.in. koder-Z (ze standardu DjVu), koder ELS (z formatu Pegasus PIC) czy nawet adaptacyjny koder Golomba (ze standardu JPEG-LS). Omówiono sposoby redukcji złożoności algorytmu kodowania przy zachowaniu jego efektywności, zwracając przy tym uwagę na powiązanie koncepcji kodowania arytmetycznego i metody kodowania długości serii (sekwencji).

W **piątym rozdziale** opisano inną grupę metod, tj. algorytmy kodowania słownikowego. Poszczególne algorytmy różnią się sposobem budowania słownika, wykorzystywanego do wyszukania powtarzających się identycznych (z dokładnością do bitu) sekwencji symboli, stanowiących frazy słownika. Do oszczędnego zapisu indeksu kodowanej frazy wykorzystuje się często metody entropijne (np. algorytm *deflate* z pakietu ZLIB). Kodowanie słownikowe dotyczy więc przede wszystkim sposobu modelowania danych wejściowych metodami deterministycznymi. Stosowane w takich formatach lub standardach kompresji obrazów jak GIF, PNG, czy TIFF, są

także podstawą tak znanych narzędzi (ogólnego przeznaczenia) do archiwizacji danych, jak: *ARJ*, *LHarc*, *RAR*, cała rodzina koderów *ZIP* itp.

Rozdział szósty dotyczy metod wstępnego przekształcenia danych w celu ich skuteczniejszej kompresji, czyli modelowania z wykorzystaniem liniowych i nieliniowych modeli predykcji (często zależności funkcyjnych). Aktualnie kodowany symbol (wartość z alfabetu) jest przewidywany na podstawie symboli sąsiednich, przy czym sąsiedztwo określa przyczynowy kontekst (na który składają się pozycje danych wcześniejszej zakodowanych). Przeprowadzone badania w wielu przypadkach potwierdzają, że większe znaczenie od odpowiednio dobranej funkcyjnej zależności przewidywania ma sposób definiowania kontekstu, najlepiej w sposób adaptacyjny (w przód i wstecz), iteracyjny, wielorozdzielczy (dla obrazów). Teoretyczne poszukiwania optymalnych metod predykcji zakładają określony poziom korelacji danych w najbliższym sąsiedztwie. Klasykne kryterium optymalizacji, tj. minimalny błąd średniokwadratowy, nie przekłada się zwykle na minimalny rozmiar skompresowanego pliku. Dobrze jest bowiem uwzględnić w kryterium również minimalizację entropii, której obliczenie wymaga estymacji prawdopodobieństw symboli. Niestety, dla sygnału niestacjonarnego o wyraźnie zróżnicowanej statystyce lokalnej prawdopodobieństwa te nie mogą być szacowane wiarygodnie. Modele nieliniowe można zwykle lepiej dopasować do lokalnych zmian, jednak zmiany o charakterze odmiennym od założonego mogą doprowadzić do gwałtownego pogorszenia efektywności predykcji i całego algorytmu kompresji. Nie ma więc prostych rozwiązań, które umożliwiają uzyskanie efektywności uniwersalnej, tj. na „przyzwoitym” poziomie dla różnorodnych danych wejściowych (optymalność globalną choćby w sensie przybliżonym). Poszukiwania użytecznych modeli predykcji, w tym liniowej DPCM i rozwiązań adaptacyjnych, stanowią najistotniejszą część tego rozdziału. Ponadto podano wiele przykładów modeli predykcji wykorzystanych w efektywnych koderach obrazów.

W **siódmym rozdziale** omówiono wybrane metody łączenia efektywnych sposobów modelowania oraz binarnego kodowania w algorytmach bezstratnej kompresji obrazów. Przedstawiono, wykorzystując wstępną charakterystykę obrazów, sposoby sekwencyjnego porządkowania danych w jednowymiarowy strumień wejściowy koder binarnego, a także hierarchiczny opis zależności danych w oryginalnej przestrzeni dwuwymiarowej, który pozwala zwiększyć efektywność kompresji. Podano rozwiązania standardowe, uznane schematy koderów, wyniki eksperymentów weryfikujących ich skuteczność. W prezentowanych algorytmach wykorzystywane są scharakteryzowane wcześniej metody kodowania i modelowania danych, dopasowane do standardowych zbiorów testowych zawierających różne rodzaje obrazów. Dodano także krótką charakterystykę metod falkowych wykorzystywanych często do zastosowań o opcjonalnym wyborze rodzaju kompresji (od stratnej do numerycznie bezstratnej). Rozdział ten stanowi podsumowanie zagadnień związanych z bezstratną kompresją obrazów. Zwrócono w nim uwagę na rozwiązania praktyczne, na narzędzia znane jako najbardziej efektywne i przydatne w wielu zastosowaniach. Poszukiwano także odpowiedzi na pytanie o możliwość i sposoby zwiększenia wydajności używanych obecnie koderów obrazów. Sygnalizowane są kierunki dalszego rozwoju tej dziedziny wobec znanych ograniczeń.