ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1998
Seria: INFORMATYKA z. 34 Nr kol. 1381

Jarosław FLAK
Politechnika Śląska, Instytut Informatyki

KOMPRESJA DANYCH W KOMPUTEROWYM SYSTEMIE MULTIMEDIALNYM[1]

Streszczenie. W artykule opisano zagadnienia występujące przy stosowaniu kompresji danych w systemie multimedialnym i dokonano wprowadzenia w tę tematykę. Przedstawiono powody stosowania kompresji i wymagania, jakie stawia się jej ze względu na występujące usługi i rodzaje danych. Zanalizowano cele kompresji i czynniki wpływające na dobór metod kompresji. Określono miejsce i zadania kompresji w komputerowym systemie rozproszonym o zadanych wymaganiach.

DATA COMPRESSION IN MULTIMEDIA COMPUTER SYSTEM

Summary. In this paper the dependences that appear while applying data compression in multimedia computer system were described. An introduction to the problems has been made. The reasons of applying compression and requirements concerning existing services and data types were presented. The purposes of compression and the factors that influence the choice of compression methods were studied. Application areas and goals of compression in distributed computer system were addressed.

1. Wprowadzenie

Dynamiczny rozwój techniki komputerowej, dokonujący się w ostatnich latach, pozwolił na operowanie dużymi ilościami informacji. W komputerowych systemach multimedialnych są przetwarzane, przesyłane i składowane różnorodne dane. Mają one przeważnie znaczną objętość, zwłaszcza obraz i dźwięk. Każdy system komputerowy posiada jednak swoje ograniczenia dotyczące mocy obliczeniowej komputerów, przepustowości kanałów transmisji

[1] Praca wykonana w ramach Projektu Celowego nr 8T11C 009 96C/2995

oraz ilości miejsca dla danych w pamięciach systemu. Na skutek tych ograniczeń szybkość działania pewnych usług w systemie może być niewystarczająca, a czasem ich poprawne działanie może być nawet niemożliwe. Rozwiązaniem tego problemu może być zastosowanie kompresji [3], która zmniejszy objętość danych. Umożliwi to efektywniejsze przetwarzanie i składowanie danych oraz zwiększy prędkość transmisji przy ich przesyłaniu.

Zapewnienie dużej przepustowości i ilości pamięci w systemie zależy przeważnie od możliwości poniesienia wydatków przez inwestora, ponieważ można wtedy zainstalować bardzo szybką sieć i dużą ilość dysków. Efektywność przesyłu zależy także od działającego oprogramowania, na przykład protokołów, również multimedialnych, pracujących w sieci. W dużym stopniu można poprawić komunikację stosując efektywną kompresję danych, bowiem przepustowość informacyjna kanału wzrośnie o tyle, o ile zmniejszy się objętość danych poddanych procesowi kompresji.

Parametry, konstrukcja i sposób działania systemu multimedialnego są określone przez wymagania stawiane odnośnie do ilości i jakości świadczonych usług oraz rodzajów występujących danych. W artykule przedstawiono problemy wynikające z potrzeby przesyłu i przechowywania dużej ilości informacji z punktu widzenia kompresji danych, a także naszkicowano jej miejsce i zadania w systemie multimedialnym o zadanych parametrach i wymaganiach.

2. Akwizycja i reprezentacja danych

Projektowanie i stosowanie kompresji wymaga wiedzy między innymi o usługach, jakie będą wykonywane w systemie komputerowym oraz rodzajach danych, jakie będą tam występowały. Każdy rodzaj usługi, działającej na określonych danych, generuje strumień informacji o różnej wielkości. Rodzaje danych, występujące w systemie, określają więc sposoby kompresji i ich miejsce w elementach systemu.

Do poszczególnych elementów systemu będą wprowadzane różnorodne informacje. Założono, że będą to:

- papierowe dokumenty tekstowe;
- papierowe dokumenty graficzne: prosta grafika, zdjęcia, mapy;
- informacje alfanumeryczne; wprowadzane do baz danych różnych działów administracji;
- informacje graficzne, mapy wektorowe;
- dźwięk, głos;
- informacje graficzne i tekstowe w postaci elektronicznej wprowadzane z zewnątrz.

Akwizycja *dokumentów papierowych* odbywać się będzie za pomocą skanera. Dalszy ich obieg będzie następował w sposób elektroniczny. Na wyjściu skanera otrzymuje się obraz o zadanej rozdzielczości i liczbie kolorów. Oczywiście im wyższa rozdzielczość i większa liczba kolorów, tym większą objętość zajmuje ta informacja. Dla obrazów są to szczególnie duże objętości, więc z punktu widzenia kompresji tym bardziej należy zwrócić na nie uwagę.

Tekstowe dokumenty papierowe mogą być reprezentowane w różny sposób. Najprostszym sposobem jest postać obrazu barwnego, szczególnie dla dokumentów zawierających kolory. Dla dokumentów czarno-białych można przyjąć reprezentację obrazu w odcieniach szarości lub czarno-białą. Bierze się pod uwagę możliwość poddania dokumentów procesowi rozpoznawania pisma, tzw. OCR i ich reprezentacji w postaci tekstowej.

Graficzne dokumenty papierowe są reprezentowane w postaci obrazów, podobnie jak dokumenty tekstowe, przy czym będą to w większości obrazy o większej liczbie kolorów, szczególnie dla map i zdjęć. Dla map przewiduje się poddanie ich procesowi wektoryzacji.

Problemem, który występuje na etapie akwizycji dokumentów w postaci obrazów, jest ustalenie optymalnej rozdzielczości i liczby barw dla każdego skanowanego typu dokumentu, tak aby zajmował on jak najmniejszą objętość, w zależności od jego treści i znaczenia.

Szeroko pojęte *informacje alfanumeryczne*, jak: elektroniczne dokumenty, arkusze kalkulacyjne, informacje dla baz danych mogą być wprowadzane jako teksty za pomocą klawiatur na poszczególnych stanowiskach. Ich reprezentacja i przetwarzanie zależy od rodzaju aplikacji.

Informacje graficzne będą wprowadzane za pomocą stacji graficznych z odpowiednim oprogramowaniem. Będą to proste grafiki, dokumenty, mapy wektorowe i zdjęcia, reprezentowane jako obrazy. Nie będą one skanowane, lecz edytowane.

Dźwięk i głos wprowadzane będą do systemu za pomocą sieci telefonicznej lub mikrofonów umieszczonych przy określonych stanowiskach. Dźwięk będzie digitalizowany i zapisywany w pamięci dyskowej komputera w postaci cyfrowej. Należy dobrać odpowiednią częstotliwość i rozdzielczość próbkowania w zależności od rodzaju informacji dźwiękowej.

Informacje graficzne i tekstowe wprowadzane *z zewnątrz* w postaci elektronicznej (np. poczty, informacji z innych urzędów) będą musiały być klasyfikowane i w zależności od potrzeb poddawane odpowiedniemu rodzajowi kompresji.

3. Kompresja danych

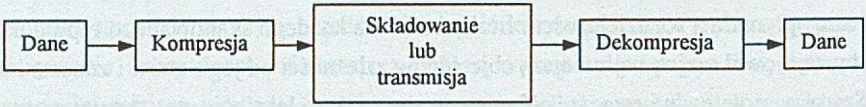
Kompresja danych cyfrowych [1] jest to proces transformacji naturalnej reprezentacji danych w inną reprezentację o mniejszej objętości. Przez objętość danych będziemy rozumieli liczbę jednostek informacji, na przykład bajtów, koniecznych do zapisania tych

danych. Niech $|x|$ oznacza objętość danych x, I oznacza pewną informację, Kn - naturalną reprezentację informacji, K - reprezentację po kompresji. Przez stopień kompresji rozumiemy wartość:

$$r(I) = \frac{|Kn(I)|}{|K(I)|} \tag{1}$$

Ponieważ dane skompresowane zajmują z założenia mniejszą objętość, więc podczas transmisji, na przykład w sieci komputerowej, można je będzie przesłać szybciej. Jeśli podczas przesyłu danych kompresja dokonywana jest na bieżąco, prędkość transmisji zależy również od szybkości kompresji, czyli ilości bajtów, które można skompresować w jednostce czasu.

Schemat przetwarzania danych w procesie kompresji przedstawiono na rys. 1. Dane w postaci skompresowanej zajmują mniejszą objętość, ale nie można bezpośrednio odczytać, co jest w nich zawarte. Oryginalną postać danych przywraca proces dekompresji.



Rys. 1. Przetwarzanie danych w procesie kompresji
Fig. 1. Data processing in compression

Można powiedzieć w uproszczeniu, że techniki kompresji zmniejszają rozmiary danych usuwając z nich redundantne (nadmiarowe) informacje, np. powtarzające się ciągi takich samych wartości. Bardziej złożone metody mogą dokonywać selekcji, wybierając informacje ważniejsze, a usuwając mniej ważne.

Istnieje wiele metod kompresji danych [1] [3] [4]. Można je podzielić na dwie główne rodziny:

- kompresja bezstratna;
- kompresja stratna.

W przypadku *kompresji bezstratnej* po dekompresji otrzymujemy dane dokładnie takie, jakie zostały poddane kompresji. Otrzymany po dekompresji na przykład obraz zgadza się z oryginałem z dokładnością co do piksela.

Stopnie kompresji, uzyskiwane za pomocą tych metod, zależą, w dużym stopniu, od rodzaju danych, np. dla bogatego w szczegóły kolorowego obrazu uzyskamy mniejszy stopień kompresji niż dla prostej, czarno-białej grafiki. Przeciętnie uzyskuje się stopnie kompresji rzędu 2:1 do 10:1, zależnie od typu danych. Daną metodę kompresji bezstratnej można stosować w zasadzie do każdego rodzaju danych, ponieważ, w odróżnieniu od metod kompresji stratnej, metody te z reguły nie uwzględniają semantyki danych.

Kompresja stratna dopuszcza utratę części informacji. Po dekompresji nie otrzymuje się więc dokładnie takich danych, jakie były kompresowane. Metody te stosuje się głównie do kompresji danych uzyskanych z sygnałów analogowych, np. obrazu lub dźwięku. Straty próbuje utrzymać się na jak najniższym poziomie, aby nie były dostrzegalne lub słyszalne. W metodach tych uwzględnia się fizjologiczne i psychologiczne aspekty ludzkich zmysłów wzroku i słuchu oraz fakt, że w zapisie dźwięku lub obrazu znajdują się informacje, których przeciętny człowiek i tak nie zauważa odsłuchując lub oglądając zapisane dane. Informacje takie można więc usunąć bez wyraźnego pogorszenia jakości. Dla człowieka tak przetworzone dane są odbierane jako prawie takie same, a różnice są prawie niezauważalne.

Dla różnych rodzajów danych stosuje się różne metody kompresji stratnej. Algorytmy kompresji np. dźwięku nie nadają się do kompresji obrazu, chociaż mogą posiadać pewne wspólne elementy, jak np. wykorzystanie transformaty Fouriera.

Używając stratnych metod kompresji można uzyskać znacznie większy stopień kompresji niż dla metod bezstratnych. Metody te oferują możliwości dostosowywania, do żądanych wymagań, stopnia kompresji lub współczynnika jakości odtworzonych danych.

3.1. Objętość danych a ilość informacji

Problemem, który stara się rozwiązać kompresja w systemie multimedialnym, jest zbyt duża objętość danych. Drastycznie wtedy spada szybkość działania takiego systemu lub wzrasta jego cena. Komputery przetwarzają, dyski składują, a sieci przesyłają *zbyt wiele* danych.

Zadaniem kompresji jest zmniejszenie *objętości danych* zachowując jednocześnie taką samą (kompresja bezstratna) albo „prawie taką samą" (kompresja stratna) *ilość informacji*. Zakładamy tutaj, że *objętość* danych jest równa ilości bajtów, które zajmują w pamięci. Zadanie kompresji polega więc na usunięciu jak największej liczby bajtów, tak aby usunąć jak najmniej informacji. Ilość *informacji* zawartej w danych jest raczej trudna do precyzyjnego określenia. Zależy ona od semantyki danych i od ich zastosowania. Kompresja stratna usuwa część informacji, mającą mniejsze znaczenie. Ilość informacji w procesie kompresji bezstratnej pozostaje oczywiście niezmienna.

4. Kompresja w systemie obiegu informacji

Z punktu widzenia kompresji w systemie multimedialnym określone są pewne wymagania, które można sformułować następująco:

W systemie występują pewne usługi generujące strumienie danych o różnej wielkości. Występują różne rodzaje informacji określonych typów, objętości i zastosowań. Podlegają one operacjom:

- akwizycji;
- składowania;
- przetwarzania;
- przesyłania.

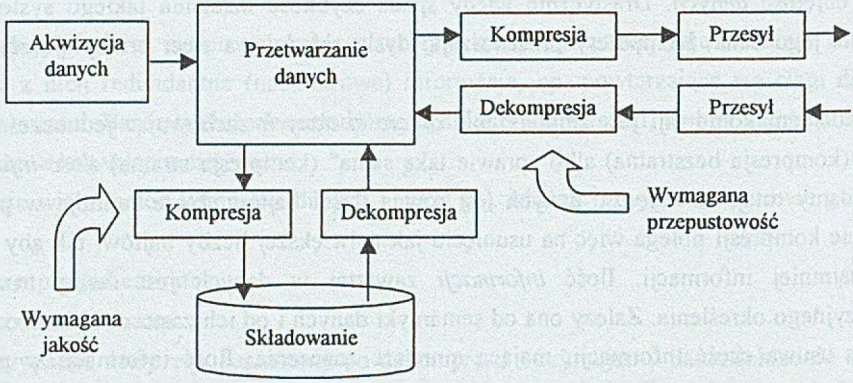
Jednocześnie w systemie komputerowym występują ograniczenia:

- mocy przetwarzania komputerów;
- pojemności pamięci dyskowej komputerów;
- przepustowości sieci komputerowych.

Dodatkowym czynnikiem, jaki należy uwzględnić, jest:

- wzrost ceny systemu w przypadku wymiany istniejącego sprzętu na inny, o lepszych parametrach.

Na podstawie powyższych przesłanek widać, że usługi w systemie mają swoje wymagania, a sprzęt ma swoje ograniczenia. Uzasadnia to potrzebę zastosowania kompresji danych, która w tani sposób polepszy wymienione wyżej parametry systemu i usprawni, a niekiedy wręcz umożliwi, obieg informacji w systemie.



Rys. 2. Cel kompresji danych w zależności od wymagań
Fig. 2. Purpose of data compression depending of demandings

Kompresja danych może odbywać się z uwzględnieniem różnych celów, w zależności od narzuconych z góry wymagań. Na rys. 2 przedstawiono zastosowania kompresji przy przesyle i składowaniu danych oraz różne cele kompresji. Przy przesyle danych kanałem transmisji w systemie multimedialnym można rozpatrzyć dwa przypadki:

1. Jakość. Czynnikiem o najważniejszym znaczeniu dla kompresji jest wymagana *jakość* informacji po dekompresji. Zakładamy, że kanał transmisji posiada wystarczającą przepustowość do przesłania strumienia informacji.
2. Objętość danych. Przepustowość kanału jest niewystarczająca do przesłania strumienia informacji. Najważniejszym czynnikiem dla kompresora jest wymagana *objętość* skompresowanych danych.

W pierwszym przypadku czas dostarczenia danych do odbiorcy nie jest ściśle zdeterminowany. Narzuconym z góry wymaganiem przy kompresji jest jakość, z jaką dane (na przykład obraz lub dźwięk) mają być odtworzone po dekompresji. Objętość, jaką zajmą dane po kompresji, jest sprawą drugorzędną. Kompresja w tym wypadku ma za zadanie zwiększenie efektywności przesyłu lub, częściej, składowania danych.

W drugim przypadku mówimy najczęściej o kompresji w czasie rzeczywistym, np. wideokonferencji. Czynnikiem krytycznym jest objętość przesyłanych danych, a z niej dopiero wynika możliwa do osiągnięcia jakość. Stopień kompresji musi być dopasowany do narzuconej z góry przepustowości, aby zagwarantować ciągłość strumienia danych, nawet kosztem pogorszenia jakości informacji. Kompresja ma za zadanie umożliwienie przesyłu informacji, który bez kompresji byłby niemożliwy na zadanym poziomie jakości.

4.1. Rodzaje danych a kompresja

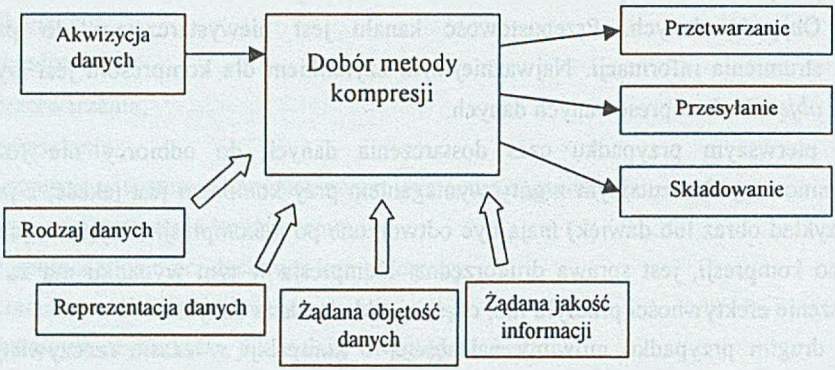
Założymy, że narzuconym kryterium kompresji jest wymagana jakość informacji po dekompresji, a objętość danych jest mniej ważna. Musi być oczywiście osiągnięty pewien kompromis pomiędzy jakością a objętością, ponieważ z reguły żądamy, aby dane zajmowały jak najmniejszą objętość i żeby to osiągnąć jesteśmy skłonni zrezygnować z bardzo dobrej jakości.

Na dobór metod kompresji mają wpływ następujące czynniki:

- *Rodzaj* danych poddawanych kompresji, na przykład obraz dokumentu tekstowego, mapa, zdjęcie, głos, tekst, dane numeryczne. Rodzaje danych określone są przez usługi dostępne w systemie.
- *Reprezentacja* danych. Dane jednego rodzaju mogą być reprezentowane w różny sposób, na przykład dokument tekstowy może być przedstawiony jako tekst, obraz kolorowy, obraz czarno-biały. Reprezentacja danych wynika z *przeznaczenia*, jakie będą miały poszczególne informacje. Z reprezentacją danych ściśle związany jest także *format* danych, w jakim zostają one zapisane.

Od rodzaju i wyboru reprezentacji informacji zależy objętość danych. Dla każdego rodzaju danych stosuje się odpowiednie metody kompresji w zależności od reprezentacji i przeznaczenia danych. Zilustrowano to na rys. 3. Z zależności tych wynika w pierwszym

rzędzie bezstratność lub stratność dobranych metod kompresji. Dla metod stratnych możliwa jest modyfikacja parametrów dotyczących jakości informacji po dekompresji.



Rys. 3. Czynniki wpływające na dobór metody kompresji
Fig. 3. Factors for selection of compression method

W systemach multimedialnych dominującymi rodzajami danych są zdigitalizowane sygnały analogowe, do których stosuje się najczęściej stratne metody kompresji. Jednak w każdym systemie komputerowym zachodzi potrzeba składowania lub transmisji danych, które nie mogą ulegać modyfikacji. Metody bezstratne wykorzystuje się do kompresji obrazów binarnych, baz danych, programów, informacji alfanumerycznych.

Problemem, który trzeba rozwiązać, już na etapie akwizycji informacji z postaci analogowych, jest ustalenie odpowiednich częstotliwości próbkowania i rozdzielczości dla obrazów i dźwięku, w ścisłym powiązaniu z *rodzajem* danych, który podlega akwizycji i jego późniejszym *przeznaczeniem*. Przy akwizycji danych analogowych następuje utrata części informacji i można powiedzieć, że zachodzi już tutaj kompresja stratna, która dokonuje selekcji informacji ważniejszej, w zależności od przeznaczenia. Na przykład dla map będzie przyjęta inna rozdzielczość i liczba barw niż dla zdjęć. Zmniejsza to, w znaczny sposób, objętość danych zachowując w dalszym ciągu istotę tej informacji.

Skanowane tekstowe i graficzne dokumenty papierowe, reprezentowane w postaci obrazów, mogą być poddane kompresji stratnej w zależności od typu obrazu (kolorowy, w odcieniach szarości, czarno-biały). Dla dokumentów kolorowych można zastosować metody zmniejszające ilość kolorów i optymalizujące ich jakość odtworzenia dla konkretnego dokumentu, np. tablicę kolorów LUT. Dla każdego rodzaju danych należy dobrać odpowiedni współczynnik jakości. W szczególnych przypadkach można zastosować kompresję bezstratną.

Dane alfanumeryczne różnych rodzajów, jak: mapy po procesie wektoryzacji, bazy danych, arkusze kalkulacyjne, teksty, należy kompresować bezstratnie. Można zastosować tutaj metody zaprojektowane specjalnie dla bezstratnej kompresji określonych typów danych.

Dźwięk kompresuje się stratnie w zależności od jego rodzaju. Dla głosu ludzkiego można stosować dedykowane metody kompresji.

Dla wszystkich rodzajów danych, podlegających kompresji stratnej, występuje problem: ile informacji usunąć, a ile pozostawić. Rozwiązać go można szczegółowo analizując i klasyfikując poszczególne rodzaje danych występujące w systemie, w zależności od ich przeznaczenia.

5. Przepływ informacji w rozproszonym systemie komputerowym

Sprawne funkcjonowanie systemu rozproszonego jest możliwe jedynie przy prawidłowo funkcjonującym przepływie informacji. Podstawową usługą, jaką należy zapewnić w systemie, jest szybkie przesyłanie informacji i dostęp do przechowywanych zasobów, np. do różnego rodzaju baz danych.

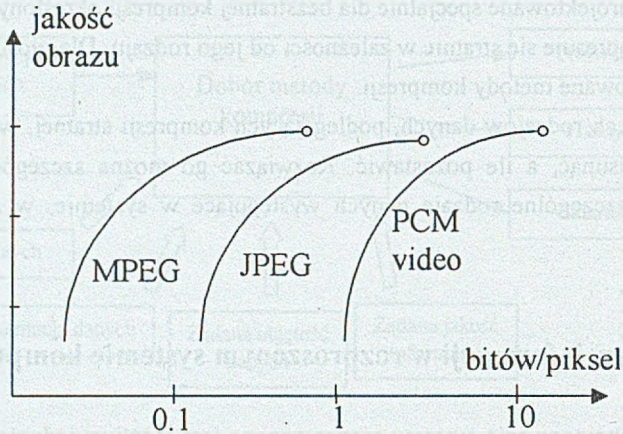
Komunikację między komputerami zapewnia *kanał transmisji* sieci komputerowej. Jest on opisywany przez wiele parametrów. Dla celów kompresji najważniejszym parametrem jest *przepustowość* kanału, czyli ilość bajtów, jaka może zostać przesłana tym kanałem w jednostce czasu. Zakładamy, że w parametrze „przepustowość kanału" zawarte są już informacje charakterystyczne dla danego systemu transmisji, jak na przykład: czas dostępu do łącza, protokoły, stopa błędów i ewentualne retransmisje. Niemożliwe jest więc fizyczne przesłanie przez kanał większej ilości danych w jednostce czasu, niż określa to przepustowość kanału.

5.1. Zwiększanie przepustowości kanału transmisji

Stosując kompresję danych, można przesłać większą *ilość informacji* przez kanał transmisyjny, niż wynikałoby to z jego fizycznej przepustowości. Ilość informacji po kompresji (w przypadku kompresji bezstratnej) pozostaje taka sama, jednak objętość danych, w której się ona zawiera, jest mniejsza. Można więc przesłać w tym samym czasie więcej informacji. Jeśli tylko procedury kompresji i dekompresji nadążają przetwarzać dane, to wzrasta przepustowość „informacyjna" kanału (przepustowość w sensie miary ilości bajtów przesłanych w jednostce czasu pozostaje oczywiście taka sama).

Zwiększanie przepustowości informacyjnej jest możliwe do ok. 2 - 4 razy. Takie stopnie kompresji osiągają bezstratne metody kompresji. Zależność między współczynnikiem

kompresji a przepustowością jest oczywiście liniowa, to znaczy przepustowość kanału wzrasta tyle razy, ile wynosi stopień kompresji.

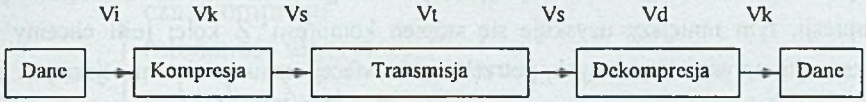


Rys. 4. Zależność jakości obrazu od stopnia kompresji dla różnych algorytmów kompresji
Fig. 4. Picture quality as a function of compression ratio for various compression algorithms

Jeśli ilość danych po kompresji bezstratnej jest nadal większa niż przepustowość kanału, nie można jej przesłać w czasie krótszym, niż określa to przepustowość kanału. Można w takiej sytuacji dokonać kompresji stratnej, która usunie pewną część informacji z danych, ale zmniejszy znacznie rozmiar danych. Ilość usuwanych informacji można zwiększać, uzyskując większy stopień kompresji. Schemat przybliżonej zależności stopnia kompresji od współczynnika strat jest przedstawiony na rys. 4 (według [2]). Stopień kompresji jest wyrażony jako ilość bitów przypadającą na jeden piksel obrazu, współczynnik strat jest odzwierciedlony w postaci jakości obrazu po dekompresji. Współczynnik strat jest trudny do ustalenia i różny dla każdego rodzaju danych. Przeważnie jest to zestaw czynników mierzących straty w określonych dziedzinach.

5.2. Kompresja w czasie rzeczywistym

Kompresja informacji może dokonywać się w czasie niezdeterminowanym lub w czasie rzeczywistym. W pierwszym przypadku nie narzuca się z góry *czasu*, w jakim dane mają zostać skompresowane. Pozwala to na zastosowanie złożonych obliczeniowo i czasowo metod kompresji i osiągnięcie lepszego stopnia kompresji. W przypadku kompresji „na bieżąco" korzysta się z prostych i szybkich algorytmów. Z reguły, dodatkowo narzuca się stopień kompresji, jaki powinny mieć skompresowane dane. Aby w takich warunkach możliwa była kompresja, muszą być spełnione pewne zależności. Rozpatrzmy system pokazany na rys. 5.



Rys. 5. Model kompresji w czasie rzeczywistym
Fig. 5. Real time compression model

Na wejście podajemy dane strumieniem o wielkości Vi, które przyjmuje kompresor z określoną przepustowością Vk. Przez przepustowość kompresora rozumiemy ilość bajtów, jakie potrafi skompresować w jednostce czasu. Na wyjściu kompresora, a jednocześnie na wejściu kanału płynie strumień danych Vs. Kompresor koduje dane z pewnym stopniem kompresji R (R > 1), więc:

$$Vs = \frac{Vi}{R} . \tag{2}$$

Jako jednostkę przepustowości przyjmiemy ilość bajtów na sekundę. Kanał ma przepustowość Vt. Po przesłaniu kanałem informacja płynąca z szybkością Vs trafia do dekompresora, który potrafi dekodować dane z szybkością Vd. Na wyjściu dekompresora płynie strumień danych o przepustowości Vk, jaki również był wprowadzany do systemu.

Można tu zauważyć pewne oczywiste zależności wynikające z warunku efektywności transmisji danych.

Wielkość strumienia danych nie może być większa niż przepustowość kompresora:

$$Vk \geq Vi . \tag{3}$$

Aby zachodziła kompresja w czasie rzeczywistym, przepustowość kompresora nie może być mniejsza niż przepustowość kanału transmisji:

$$Vk \geq Vt . \tag{4}$$

Strumień danych wychodzący z kompresora nie może być większy niż przepustowość kanału transmisji:

$$Vs \leq Vt ; \tag{5}$$

$$\frac{Vi}{R} \leq Vt . \tag{6}$$

Przepustowość dekompresora powinna być co najmniej równa

$$Vs = \frac{Vi}{R} , \tag{7}$$

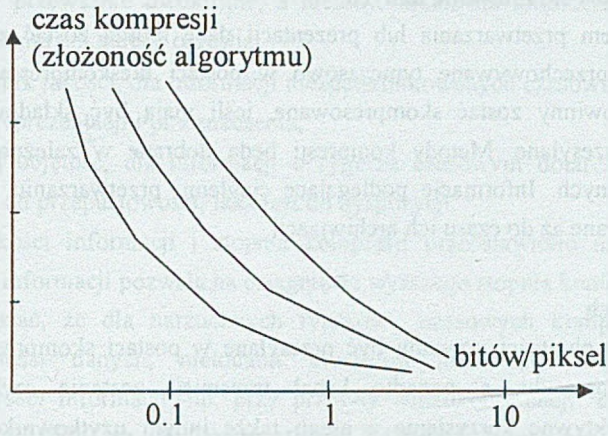
a generalnie powinna być równa Vt.

Przy projektowaniu systemów kompresji czasu rzeczywistego trzeba wziąć pod uwagę powyższe zależności. Jasne jest, że im większy stopień i prędkość kompresji, tym więcej danych można przesłać, przekraczając nawet granicę fizycznej przepustowości łącza. Dla

kompresorów w rzeczywistości występuje pewna ogólna zależność: im większa prędkość kompresji, tym mniejszy uzyskuje się stopień kompresji. Z kolei jeśli chcemy uzyskać większe skompresowanie danych, potrzeba na to więcej czasu, a więc mniejsza jest prędkość kompresji. Zależności te pokazano na rys. 6 (według [2]). W praktyce wypośrodkowuje się te parametry lub dostosowuje do wymaganej przepustowości łącza transmisji.

Trzeba jednak zdawać sobie sprawę, że aplikacje multimedialne operują tak ogromnymi ilościami danych, że dla bardzo ograniczonych przepustowości nawet najlepsze metody kompresji stratnej nie zapewnią odbioru danych wysokiej jakości. Można podnosić poziom stratności, a tym samym stopień kompresji, jednak istnieje wymóg minimalnej przepustowości, jaką musi zapewnić kanał transmisji dla danej usługi, aby odbiór danych multimedialnych był płynny i akceptowalny dla odbiorcy.

Jeśli dane są przesyłane przez sieć komputerową, to w praktyce przepustowość sieci powinna być większa, niż wynika to z wyżej podanych zależności. Dane multimedialne nie są jedynymi danymi, które będą przez sieć przesyłane. Zostaną również rozszerzone o odpowiednie adresy i inne konieczne informacje. Powyższe szacunki zakładają także, że czas dotarcia informacji w sieci od nadawcy do adresata wynika jedynie z prędkości przesyłu danych w sieci. Należy jednak uwzględnić obciążenie sieci i sytuacje, w których, na przykład, kilku użytkowników jednocześnie nadaje strumienie informacji przez jeden węzeł sieciowy lub wspólną magistralę. Problem stanowi również protokół dostępu do łącza, gdyż w sieciach, w których nie jest gwarantowany czas dotarcia informacji do adresata, np. działających w oparciu o zasadę wykrywania konfliktów, nie można zagwarantować również ciągłości odtwarzania na bieżąco i mogą wystąpić opóźnienia w tym procesie. Z powyższych rozważań wynika, że wobec sieci, w której ma być realizowane na przykład odtwarzanie dźwięku na bieżąco, na podstawie przesyłanych danych, trzeba postawić wymagania przewyższające oszacowania tylko szerokości pasma dźwięku. W praktyce można jednak próbować przesyły takie realizować w sieciach nie spełniających wymagania zdeterminowanego czasu przesyłu informacji. Jednak wtedy prędkość przesyłu danych w sieci powinna być co najmniej kilkakrotnie wyższa od wcześniej oszacowanych, aby w ten sposób, w miarę możliwości, zamortyzować wpływ obciążenia sieci.



Rys. 6. Zależność złożoności algorytmów kompresujących od osiąganego stopnia kompresji
Fig. 6. Complexity of compression algorithms as a function of compression ratio

Kompresja danych wymaga czasem dosyć dużych mocy obliczeniowych [5]. Jeśli proces kodowania wymaga większej ilości czasu lub mocy obliczeniowej niż proces dekodowania, to mówimy o kompresji asymetrycznej. Nadaje się ona do systemów, które nie zakładają kompresji w czasie rzeczywistym, a mają oferować dobrą jakość ruchomego obrazu wideo czy dźwięku. W przypadku gdy w procesie dekodowania przeprowadza się dokładnie te same operacje jak przy kodowaniu, tylko w odwrotnej kolejności, mamy do czynienia z kompresją symetryczną. Wówczas potrzebne moce obliczeniowe oraz czasy kompresji i dekompresji są takie same. Algorytmy asymetryczne sprawdzają się w systemach, w których z jednego źródła raz skompresowanych dobrze, a więc z dużym narzutem czasowym, danych korzysta wielu użytkowników stosując szybką dekompresję.

5.3. Umiejscowienie kompresji w systemie

System multimedialny powinien posiadać możliwość dołączenia i posługiwania się informacjami pochodzącymi z innych systemów, a więc skompresowanych w różny sposób. Powinny więc w tym celu być dostępne różnorodne dekompresory danych. Zakładamy, że oprogramowanie powinno rozpoznawać format kompresji i dokonywać automatycznej lub półautomatycznej dekompresji. W systemie komputerowym kompresja danych powinna występować w miejscach, w których następuje przetwarzanie, składowanie, przesyłanie i archiwizacja danych.

Przetwarzanie i składowanie danych

Przed procesem przetwarzania lub prezentacji dane muszą zostać zdekompresowane. Mogą być one przechowywane tymczasowo w postaci nieskompresowanej, jednak po przetworzeniu powinny zostać skompresowane, jeśli mają być składowane na dyskach lokalnych lub przesyłane. Metody kompresji będą dobrane w zależności od rodzaju i przeznaczenia danych. Informacje podlegające ciągłemu przetwarzaniu mogą nie być w ogóle kompresowane aż do czasu ich archiwizacji.

Przesył danych

Dane o dużej objętości powinny być przesyłane w postaci skompresowanej. Zostanie skrócony czas przesyłu, a ponadto kanał transmisji zostanie szybciej zwolniony, umożliwiając efektywne korzystanie z niego także innym użytkownikom sieci. Można założyć, że wysyłane dane będą już skompresowane w procesie kompresji po przetwarzaniu danych, gdy są składowane na dyskach lokalnych. Metody kompresji przy przesyle, np. w modemach, należy stosować tylko dla danych nieskompresowanych. Przy stosowaniu kompresji w czasie rzeczywistym ograniczeniem może być czas kompresji i dekompresji, który dla połączeń typu punkt-punkt ogranicza zastosowanie takiej kompresji do stosunkowo wolnych łączy.

Archiwizacja danych

Dla celów archiwizacji można zastosować metody kompresji o większej złożoności obliczeniowej, kompresujące dłużej, ale uzyskujące większy stopień kompresji. Można też dokonywać dla niektórych rodzajów danych dalszej kompresji stratnej, aby zmniejszyć rozmiar archiwum, pozostawiając jednak istotną informację, na przykład dla obrazów dokumentów tekstowych.

6. Podsumowanie

Kompresja danych w systemie multimedialnym, poprzez redukcję objętości danych, pozwala na zwiększenie efektywności przetwarzania danych, zredukowanie wielkości pamięci dyskowej i skrócenie czasu transmisji danych. Zwalnia jednocześnie od konieczności wymiany sprzętu w systemie na inny, o lepszych parametrach.

Właściwa kompresja danych wymaga podjęcia decyzji o doborze odpowiednich metod kompresji. Należy uwzględnić ich szybkość, efektywność oraz stratność lub bezstratność. Wpływ na to mają rodzaje usług i danych występujące w systemie, a także wymagania odnośnie do jakości usług. Metody bezstratne pozwalają na kilkukrotne zmniejszenie

rozmiaru danych, przeważnie dwukrotne, a metody stratne na kilkudziesięciokrotne. Dla stratnych metod kompresji należy określić:

- współczynnik jakości, dla informacji niezdeterminowanych czasowo, w zależności od wybranej reprezentacji i przeznaczenia;
- wymaganą objętość, dla informacji o rygorze czasowym dotarcia do adresata w zależności od przepustowości, jaka jest do dyspozycji.

Zależności jakości informacji i stopnia kompresji przedstawiono na rys. 4. Niższa wymagana jakość informacji pozwala na osiągnięcie wyższego stopnia kompresji.

Może się okazać, że dla narzuconych rygorów czasowych kompresji, a zarazem wymaganej objętości danych, metodami kompresji niemożliwe będzie osiągnięcie akceptowalnej *jakości* informacji, np. przy przesyle wideokonferencji. Istnieje zależność, pokazana na rys. 6, mówiąca, że lepszy stopień kompresji osiąga się stosując bardziej złożone metody, a więc o dłuższym czasie kompresji. Metody szybkiej kompresji mogą więc nie osiągać żądanego stopnia kompresji dla wymaganej jakości danych.

Na dobór metod kompresji istotny wpływ ma rodzaj danych i wymagana reprezentacja danych. Związane jest z tym przeznaczenie danych i ich format.

W systemie komputerowym kompresja powinna być umiejscowiona w punktach: przetwarzania, reprezentacji i składowania danych, transmisji danych (skrócenie czasu transmisji danych) i archiwizacji („dokładniejsza" kompresja). W zależności od miejsca będzie to kompresja różniąca się celem i doborem metody.

LITERATURA

[1] Skarbek W.: Metody reprezentacji obrazów cyfrowych. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1993.

[2] Haskell B. G., Puri A., Netravali A. N.: Digital video: an introduction to MPEG-2. Chapman & Hall, New York 1997.

[3] Bookstein A., Storer J. A.: Data Compression. Information Processing & Management, 1992, Vol. 28, No. 6, s. 675-680.

[4] Krupiczka A.: Komunikacja multimedialna. Standardy i trendy. ZN Pol. Śl. s. Informatyka z. 30, Gliwice 1996.

[5] Steinbrink B.: Multimedia u progu technologii XXI wieku. Robomatic, Wrocław 1993.

[6] Posnak E. J., Lavender R. G., Vin H. M.: An adaptive framework for developing multimedia software components. Communications of the ACM, October 1997, Vol. 40, No. 10.

[7] Netravali A. N., Haskell B. G.: Digital Pictures - Representation, Compression and Standards, 2nd edit. Plenum Press, New York 1995.

Recenzent: Dr hab. inż. Adam Mrózek, Prof. Politechniki Śląskiej

Wpłynęło do Redakcji 30 grudnia 1997 r.

Abstract

In multimedia computer systems operations on much miscellaneous data types are performed. Every computer system is limited by the amount of memory and throughput of transmission channels. To improve data storage and transmission, especially of high volume, and even make it possible, the data compression is employed.

Depending on imposed requirements, the main aim of data compression is either to achieve the required quality or volume of information. The increasing of the compression ratio degrades the information quality (fig. 4). Data types, their representation and predestination influence the choice of compression method. This implies the usage of lossless or lossy compression methods. Required quality or required amount of data after compression can be forced in lossy methods.

The compression makes it possible to increase the throughput of transmission channel. Dependencies which are essential for the real time compression were defined in chapter 5.2 (equations 3 - 7). Better compression ratio can be achieved by applying high complexity algorithms (fig. 6)

In the computer system data compression should be used in areas connected with data processing, storing, transmission and archiving.