

CYFROWE PRZETWARZANIE DANYCH MEDIALNYCH

Wojciech Zając

**Instytut Informatyki i Elektroniki, Uniwersytet Zielonogórski
65-246 Zielona Góra, ul. Podgórna 50**

e-mail: W.Zajac@iie.uz.zgora.pl

STRESZCZENIE

W artykule przedstawiono stan techniki w zakresie przetwarzania cyfrowych danych multimedialnych, wskazano obszary prowadzonych badań i kierunki dalszych prac.

1. WPROWADZENIE

Tradycyjne techniki telekomunikacji oparte na łączach analogowych tracą swe znaczenie na korzyść metod, wykorzystujących sygnał w postaci cyfrowej. Motywacja zmian ma charakter czysto ekonomiczny – łącza cyfrowe mają istotnie większą pojemność, dane w postaci cyfrowej dają się w szerokim zakresie obrabiać, przetwarzać i stanowią całkowicie nową jakość w dziedzinie telekomunikacji.

Nowe systemy komunikacji, oprócz oferowania oczywistych korzyści, stawiają także nowe wyzwania. Aby zmieścić więcej informacji w danym kanale transmisyjnym, dokonuje się kompresji danych, dążąc do uzyskania możliwie wysokiego jej stopnia. Jednak czym większa kompresja danych, tym większa ich wrażliwość na zakłócenia. Odrębnym, choć nie mniej istotnym problemem jest wykorzystanie kompresji stratnej do transmisji materiałów multimedialnych. Możliwe jest „ściśnięcie” danego materiału do żądanych rozmiarów, jednak trzeba dokonać znacznych wysiłków, by utrata jakości obserwowana przez odbiorcę była możliwie niewielka.

Współczesne badania dotyczą wykorzystania technik falkowych i hybrydowych do konstrukcji koderów, dających lepsze stopnie kompresji przy zachowaniu jakości sygnału oraz zagadnień detekcji, korekcji i maskowania zakłóceń transmisyjnych skompresowanego sygnału.

2. CYFROWE PRZETWARZANIE SYGNAŁÓW

2.1. Przetwarzanie bezstratne i stratne

Wyróżnia się dwa tryby przetwarzania użytecznych sygnałów cyfrowych: bezstratne i stratne. Przetwarzanie bezstratne pozwala na odtworzenie zakodowanego sygnału w dokładnie takiej

postaci, w jakiej został on poddany kodowaniu. Osiągany stopień kompresji jest z reguły niewielki, ponadto idealna postać sygnału rzadko jest istotnie potrzebna, dlatego też metody te używane są w nielicznych zastosowaniach (np. środowiska medyczne słyną z tego, że nie godzą się na utratę części, nawet nieistotnych, informacji).

Przetwarzanie stratne polega na zastosowaniu określonych kryteriów do wyeliminowania pewnej części informacji – odtworzony sygnał jest jedynie podobny do zakodowanego, jednak możliwa jest znaczna kompresja (np. standard MPEG-1 posiada nominalny stopień kompresji 200:1, choć w praktyce osiąga się 20:1). Kryteria eliminacji danych są przedmiotem rozległych badań, dąży się do takiego oszukania zmysłów człowieka, by nie dostrzegł on braku danych.

W dalszej części artykułu omawia się techniki kompresji stratnej.

2.2. Kompresja różnego rodzaju danych

Cyfrowe sygnały audio oraz wizyjne (stałe i ruchome), są reprezentowane w systemie cyfrowego przetwarzania danych w jednej z trzech postaci:

- jednowymiarowej – wektorów (dla próbek dźwięku),
- dwuwymiarowej – macierzy (dla obrazów stałych)
- trójwymiarowej – sekwencyjnego zbioru macierzy (w przypadku obrazu ruchomego).

Techniki kompresji danych wykorzystują specyfikę sygnału, zależną od jego charakteru.

2.3. Kompresja sygnałów audio

Algorytmy kompresujące sygnał dźwiękowy działają dwuetapowo. Pierwszym krokiem jest wykorzystanie złożonych zjawisk, zachodzących w aparacie słuchu człowieka. Należą do nich: zmienny próg słyszalności w zależności od częstotliwości, zjawiska maskowania dźwięków oraz bezwładność w czasie. Dane, które algorytm kodujący uzna za nieistotne dla odbiorcy pomijane. Pozostałe, istotne dane poddane są bezstratnym procesom dekorelacji (najczęściej z wykorzystaniem transformacji) i kompresji (algorytmem Huffmana, arytmetycznym itp.).

2.4. Kompresja obrazów

Przetwarzanie obrazów stałych przebiega także w dwóch etapach. Ze względu na fakt, że dane obrazu mają charakter dwuwymiarowy, krok pierwszy to obejmuje dekorelację danych w dwóch kierunkach (wymiarach). Zazwyczaj ze względu na szybkość operacji oraz jej koszt numeryczny, prowadzi się odrębnie dekorelację w jednym kierunku (np. w poziomie) a następnie powtarza ją w drugim kierunku (np. pionie). Operacja dekorelacji (prowadzonej z użyciem transformaty) uzupełniona jest o moduł kwantyzatora, który dokonuje degradacji rozdzielczości analizy częstotliwościowej na zadanym poziomie. Zdekorelowane dane kompresowane są algorytmem Huffmana lub innym.

2.5. Kompresja sekwencji ruchomych

Sekwencje wideo charakteryzują się złożonością w trzech wymiarach. Prócz dwóch wymiarów obrazów, składających się na poszczególne obrazy sekwencji, występują także zależności między kolejnymi obrazami. Techniki kompresji sekwencji wykorzystują zarówno metody eliminacji danych z pojedynczych ramek (klatek) sekwencji, jak i tzw. kodowanie różnicowe między kolejnymi ramkami. Kodowanie pojedynczych ramek (tzw. kompresja wewnątrzklatkowa, ang. intraframe) eliminuje dane w sposób podobny jak w przypadku kompresji obrazów stałych, jednak kryteria eliminacji danych są nieco inne. Wykorzystuje się fakt mniejszej czułości oka na zmiany barwy niż na zmiany jasności, co pozwala na mniej dokładne kodowanie informacji o barwach. Wykorzystuje się także zjawisko nieostrego widzenia obiektów ruchomych – istotniejsze od dokładnej informacji o konturze czy wyglądzie obiektu są szczegółowe informacje o przemieszczeniu (tzw. motion blur).

Kodowanie różnic między ramkami (tzw. kodowanie międzyramkowe, ang. interframe) wykorzystuje fakt, że zwykle kolejne obrazy sekwencji różnią się od siebie niewielką częścią informacji. Wyznacza się ramki odniesienia i te koduje w sekwencji w sposób pełny, pozostałe zaś nie są kodowane jako dane ramki, lecz jako dane o różnicach względem odpowiedniej ramki odniesienia.

3. JAKOŚĆ DANYCH MEDIALNYCH

3.1. Miary jakości

$$\text{PSNR} = 10 \log_{10} \frac{x_{\max}^2 N^2}{\sum_{\text{obrazu}} (x - x')^2} \text{dB}, \quad (1)$$

gdzie:

| | |
|------------|--|
| x_{\max} | - maksymalna wartość elementu sygnału, |
| N^2 | - liczba elementów sygnału, |
| x | - oryginalny element sygnału, |
| x' | - odtworzony element sygnału. |

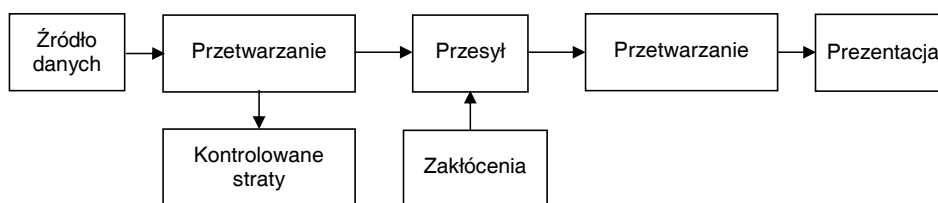
Jakość danych medialnych mierzy się na dwa sposoby. Pierwsza miara, PSNR (ang. Peak Signal-to-Noise Ratio, impulsowy stosunek sygnał-szum, równanie 1) ma charakter matematycznej miary różnicowej (określa odstęp sygnał-szum sygnału mierzonego w odniesieniu do danego sygnału wzorcowego) i, choć jest uznawana i często stosowana w laboratoriach, posiada dwie istotne wady. Po pierwsze, wymaga posiadania sygnału odniesienia, np. obrazu przez zakodowaniem. Jest to łatwe do osiągnięcia w warunkach laboratoryjnych, jednak w pomiarach w rzeczywistych systemach stanowi istotny problem. Druga wada to oparcie wyniku pomiaru na różnicy w wartościach liczbowych poszczególnych elementów składowych sygnału (np. luminancji i chrominancji obrazu), bez

uwzględnienia modelu psychofizycznego aparatu wzroku człowieka. Prowadzi to do sytuacji, gdy ocena dokonana z użyciem PSNR kłóci się z pozytywną oceną subiektywną odbiorcy.

Rozwiązaniem problemu jest wykorzystanie miary subiektywnej. Wymaga ona przeprowadzenia prezentacji ocenianego materiału grupie osób, zebraniu ich opinii i statystycznym opracowaniu wyników. Sposób ten jest kłopotliwy, lecz daje wyniki znacznie cenniejsze niż PSNR.

3.2. Niekonalność danych

Dane medialne przetworzone przez systemu transmisyjne charakteryzuje szereg niekonalności. Część z nich powstaje celowo w skutek stosowania kompresji stratnej, inne pochodzą z zakłóceń procesu transmisji (rys. 1.).



Rys. 1. Schemat systemu transmisji danych cyfrowych

4. POPRAWA JAKOŚCI SYGNAŁU

4.1. Klasyfikacja usterek sygnału

Zwalczanie niekonalności sygnału przyjmuje dwie formy: usuwa się skutki zakłóceń i/lub wykonuje się tzw. post-processing, zabiegi łagodzące celowo wprowadzone przy kodowaniu niekonalności.

Zakłócenia sygnału mogą przyjmować jedną z trzech form:

- błędy bitowe (ang. Random Bit Errors) - w procesie transmisyjnym występuje skończone prawdopodobieństwo niepoprawnej transmisji każdego bitu, określane jako współczynnik błędów bitowych (ang. Bit Error Rate, BER); błędy bitowe powstają na skutek krótkotrwałych zmian warunków transmisji (zakłócenia elektromagnetyczne, termiczne, itp.), częstość ich występowania jest losowa,
- przejściowe zaniki i przerwy w łączności – spowodowane są krótkotrwałymi zmianami warunków transmisji i interferencją; czas ich trwania jest na tyle długi, że zakłócają jednorazowo całe grupy bitów, jednak nie powodują trwałego zerwania łączności,

- trwale zaniki i przerwy w łączności – wywołane dłuższymi zmianami warunków transmisji, zwykle na skutek zmiany położenia urządzeń transmisyjnych (np. pojazd poruszający się w terenie zabudowanym bądź górzystym); ich rezultatem jest czasowe zerwanie łączności.

W praktyce jedynym typem błędów, jakie można zwalczać po odebraniu sygnału są błędy bitowe.

Upośledzenie jakości sygnału, pochodzące od celowo wprowadzonych niedokładności przetwarzania mają zawsze charakter specyficzny dla zastosowanej techniki kodowania. Przykładowo powszechnie stosowana transformata kosinusowa (MPEG, JPEG, H.261, H.263, itd.) powoduje powstawanie artefaktów transformaty (tzw. duszków) oraz uwidocznienia bloków obrazu.

4.2. Zwalczenie zakłóceń

Zakłócenia transmisji zwalczą się na trzy sposoby: przez żądanie ponownego przesłania danych, umieszczenie w sygnale danych dodatkowych (tzw. nadmiarowanie sygnału) lub z użyciem technik maskowania błędów.

Żądanie retransmisji ma bardzo ograniczone zastosowanie – wymagane jest w takiej sytuacji zestawienie i utrzymywanie kanału sprzężenia zwrotnego, na przesłanie żądania oraz ponowną transmisję potrzebna jest określona ilość czasu, co może całkiem tę transmisję pozbawić sensu (np. przy transmisji sekwencji wideo konieczne byłoby jej chwilowe zatrzymanie do czasu uzyskania poprawnych danych). W praktyce rozwiązania takie są rzadkością.

Popularna jest natomiast technika nadmiarowania sygnału (stosowana m.in. w standardach cyfrowych transmisji telewizyjnych DVB, w zapisie danych na płytach DVD, itp.). Polega ona na wykorzystaniu technik takiego kodowania danych, że wystąpienie zakłócenia prowadzi w procesie dekodowania do logicznej niespójności sygnału. Kody autokorygujące mają możliwość detekcji zakłóceń i przywrócenia najbardziej prawdopodobnych (choć niekoniecznie prawdziwych) wartości uszkodzonych słów kodowych.

Najnowsza rodzina technik, zwana maskowaniem błędów (ang. error concealment) jest rozwiązaniem z pogranicza korekcji i post-processingu. Jest to także najmniej inwazyjna z metod podnoszenia jakości sygnału. Idea polega na przyjęciu sygnału (przesłanego bez nadmiarowania) z odbiornika w takiej postaci, w jakiej został odebrany, z dopuszczeniem w nim błędów i podjęciu kroków w celu eliminacji lub ograniczenia wpływu zakłóceń na sygnał. Podejmuje się próby odtworzenia uszkodzonych danych, lecz nie do postaci najbardziej prawdopodobnej, lecz najmniej różniącej się od otoczenia. Autor niniejszej pracy opracował hybrydowy algorytm maskowania zakłóceń pewnej klasy obrazów stałych. Uzyskane wyniki wskazują na duże znaczenie tej techniki, szczególnie wobec faktu, że daje

się ona stosować w odniesieniu do sygnałów transmitowanych w istniejących systemach, bez ingerencji w ich strukturę czy funkcjonalność.

4.3. Post-processing

Podnoszenie jakości sygnału ma zastosowanie niezależne od technik zwalczania zakłóceń i przede wszystkim znajduje zastosowanie w systemach o założonym dużym stopniu utraty jakości na etapie kodowania sygnału.

Zabiegi podnoszenia jakości sygnału po jego odebraniu i odtworzeniu opierają się na założeniu, że kryterium oceny jakości sygnału jest odczucie odbiorcy (miara subiektywna). Podejmuje się zabiegi filtracyjne, wygładzające, normalizujące i inne, prowadzące do redukcji widoczności, zauważalności i dokuczliwości usterek w poprawnie przesłanych danych.

5. PODSUMOWANIE

Trwają intensywne prace nad ulepszeniem istniejących technik kodowania i kompresji danych cyfrowych, szczególnie sekwencji wideo. Zagadnienie to ma istotne znaczenie naukowe i ekonomiczne. Poszukuje się rozwiązań hybrydowych, łączących najkorzystniejsze cechy istniejących metod, wprowadza się nowe techniki, stosuje złożony aparat matematyczny (teoria falkowa, złożone metody analizy matematycznej, itd.). Szczególnie interesującym zagadnieniem wydaje się być maskowanie błędów i post-processing w odniesieniu do sygnałów, kodowanych koderami falkowymi. Planuje się dalsze prace w tym kierunku.

LITERATURA

- [1] W. Zając: *An Error Concealment Algorithm for Digital Image Transmission*. Proceedings of XIV International Symposium on Computer and Information Sciences, Kusadasi Turcja, October 1999
- [2] W. Zając: *Cyfrowe przetwarzane i transmisji obrazów z wysokim współczynnikiem kompresji*, Zeszyty Naukowe WST w Legnicy, Legnica 1999
- [3] W. Zając: *A Hybrid Method of Error Elimination in Digital Image Transmission*, Proceedings of the 5TH International Conference MIXDES'98, Łódź 1998
- [4] W. Zając: *A hybrid error concealment algorithm for digital image transmission*. Proceedings of The Second International Workshop on Multidimensional (nD) Systems. NDS-2000. Czocho Castle, Poland, June 27-30, 2000. Zielona Góra: Technical University Press [2000], p. 263-268
- [5] W. Zając: *Estymacja optymalnych parametrów algorytmu korekcji obrazu w oparciu o odebrany sygnał*. Materiały XVIII Ogólnopolskiej Konferencji „Poliptymalizacja i CAD”, Mielno 2000, s. 387-392