



COMPRESIA datelor

Claudiu Soroiu

În cadrul acestui ultim episod al serialului dedicat compresiilor de date vă vom prezenta modul de prelucrare și comprimare a sunetelor și imaginilor.

Dacă în numerele anterioare ale revistei am prezentat tehnici de compresie a datelor fără pierdere de informație, în cele ce urmează vă vom prezenta câteva tehnici de compresie cu pierdere de date. Compresia de date cu pierdere de informație se aplică în special pentru date multimedia cum ar fi sunete, imagini și secvențe video.

Prelucrarea sunetelor

Există multe persoane care și-au pus întrebarea "*Cum sunt stocate sunetele în cadrul unui fișier?*". În continuare vă vom prezenta transformările prin care trece un sunet de la generarea lui de către un dispozitiv sau o persoană până în momentul scrierii lui într-un fișier.

Captarea și stocarea sunetelor

Sunetele generate de dispozitive sau persoane sunt captate de dispozitive de captare a sunetelor.

Cel mai bun exemplu de dispozitiv de captare a sunetelor este microfonul. Principiul de funcționare al acestuia este similar modului de funcționare a urechii umane, prin urmare, sunetele care ajung la microfon interacționează cu membrana acestuia care generează semnale electrice care reprezintă amplitudinea sunetului în funcție de timp.

În continuare, semnalele electrice generate de microfon trebuie transformate în șiruri de numere. Acest lucru se realizează de către o unitate *ADC* (*Analog Digital Converter*) care primește la intrare o tensiune electrică și generează la ieșire un număr reprezentat binar. Datorită faptului că orice tensiune electrică se reprezintă numeric printr-un șir finit de biți, apare o eroare numită *eroare de cuantificare*.

Pentru reprezentarea unui semnal în formă digitală, acesta trebuie eșantionat. Prin eșantionare, dintr-un semnal se extrag cadre la intervale de ΔT secunde.

După eșantionarea unui semnal, acesta poate fi folosit de aplicații multimedia pentru a prelucra și/sau stoca șirurile de biți generate de unitățile *ADC*.

Calitatea sunetului digital care se dorește a fi obținută depinde de numărul de biți folosiți pentru a reprezenta un

cadru și de intervalul de timp dintre cadre. Calitatea sunetului este mai bună cu cât numărul de biți este mai mare și cu cât intervalul de timp dintre două cadre este mai mic.

De exemplu, pentru sistemul telefonic, cadrele sunt reprezentate folosindu-se 7 sau 8 biți, intervalul dintre cadre este 1/8000 secunde (8000 cadre/sec) și se pierd toate sunetele cu frecvențe mai mari de 4 kHz.

Un alt exemplu îl constituie *CD*-urile audio. Pe un *CD*-audio sunetele sunt înregistrate folosindu-se 16 biți pentru a reprezenta un cadru și intervalul dintre cadre este 1/44.100 (44.100 cadre/sec).

Compresia sunetelor digitale

Datorită faptului că pentru a transmite la distanță sunete cu o calitate foarte bună (de calitate *CD*-urilor audio) sunt necesari 1,411 Mbps pentru semnale audio stereo și 705,6 Kbps pentru semnale audio mono. Aceste numere sunt foarte mari și s-a pus problema compresiei sunetelor.

Cei mai importanți algoritmi utilizați în procesul de compresie a sunetelor sunt **codificarea formei de undă** și **codificarea perceptivă**.

Prin **codificarea formei de undă**, semnalul este transformat cu ajutorul transformatei *Fourier* în componentele sale în frecvență (reprezentarea amplitudinilor în funcție de timp). După aplicarea transformatei, amplitudinea fiecărei componente este codificată minimal, folosind cât mai puțini biți. Transformata *Fourier* duce la pierdere de informație.

Șirul de biți obținut se comprimă folosind unul dintre algoritmi prezentați în episoadele anterioare ale compresiilor de date. Scopul este de a reproduce exact forma de undă, la celălalt capăt folosind cât mai puțini biți.

Codificarea perceptivă constă în a transforma sunetul digital, astfel încât un ascultător uman să nu poată face diferența între un sunet transformat și unul netransformat. La baza acestui tip de codificare stă *psihoaustica*.

Psihoaustica este știința care se ocupă de modul în care sunetele sunt percepute de către om.

Proprietatea de bază a **codificării perceptivă** este constituită de faptul că unele sunete pot masca alte sunete, ast-



fel încât acestea din urmă nu se vor mai auzi atâta timp cât primele continuă să fie prezente. Pe baza acestei proprietăți, sunetele care nu vor putea fi auzite de ascultătorii umani vor fi eliminate.

În realitate, dacă generăm un sunet cu frecvența de 1 KHz la un volum de 60 dB și un alt sunet cu frecvența de 1.1 KHz și un volum de 40 dB, nu vom auzi cel de-al doilea sunet. Acest fenomen poartă numele de *mascare a frecvenței*. Sunetele care se aud poartă numele de *sunete mască*.

Un alt fenomen care are loc în realitate este acela că, dacă un sunet maschează un alt sunet, urechea umană este adaptată să audă numai sunetele mască. În cazul în care sunetele mască de la un moment dat dispar, celelalte sunete nu vor putea fi auzite decât după un interval de timp, care reprezintă timpul în care urechea umană se reglează pentru a le putea percepe. Acest fenomen poartă numele de *mascare temporară*.

Prin urmare, o aplicație de compresie a sunetelor care se bazează pe *psihoaustică* va elimina pe baza unor relații sunetele care nu vor putea fi auzite de către om.

După eliminarea sunetelor care nu pot fi percepute asupra rezultatului se aplică unul dintre algoritmi de compresie fără pierdere de informație.

Se observă în acest caz că pierderea de informație este dată de eliminarea sunetelor care nu pot fi auzite de către om.

Raportul *calitate sunet/dimensiune fișier de ieșire* este mai mare dacă se aplică aceste reguli decât în cazul în care se încearcă compresia sunetelor fără a elimina sunetele care nu se aud.

Compresia imaginilor

Pentru a comprima imagini există mai multe metode. Unele aplicații comprimă imaginile fără a pierde informație.

De obicei, pentru a comprima o imagine, aceasta se împarte în mai multe imagini mai mici de formă pătrată sau dreptunghiulară.

În continuare vom prezenta modul de compresie a imaginilor folosit pentru a obține fișiere cu formatul *JPEG* (*Joint Photographic Expert Group*).

Primul pas în compresia imaginii constă în a descompune fiecare punct din imagine în componente *YIQ* (*Y* - luminanță; *I*, *Q* - crominanțe) sau *YUV* (*Y* - luminanță; *U*, *V* - crominanțe). Diferența între *YIQ* și *YUV* este dată de modul de calcul.

Componentele *Y*, *I*, *Q* se calculează pentru fiecare punct al imaginii astfel:

$$Y = 0,30 \cdot R + 0,59 \cdot G + 0,11 \cdot B$$

$$I = 0,60 \cdot R - 0,28 \cdot G - 0,32 \cdot B$$

$$Q = 0,21 \cdot R - 0,52 \cdot G + 0,31 \cdot B,$$

unde *R*, *G*, *B* sunt valorile pentru culorile roșu, verde și albastru.

Valorile componentelor *Y*, *I* și *Q* se păstrează în trei matrice distincte.

Al doilea pas constă în descompunerea fiecărei matrice în pătrate cu latura de 8 puncte.

Asupra fiecărui pătrat se aplică transformarea *DCT* bidimensională (*Discrete Cosine Transform* - Transformarea *Cosinus Discretă*).

După aplicarea transformării, pentru fiecare pătrat se obține o matrice *F* cu 8 linii și 8 coloane. Valorile din matricea *F* trebuie să fie numere întregi.

Valorile matricei *F* se cuantifică prin împărțirea întreagă la un număr întreg. Cuantificarea unei matrice *F* are ca scop reducerea numărului de biți necesari pentru a stoca informații despre imaginea care se comprimă.

Principala sursă a pierderii de informație în cazul formatului *JPEG* o constituie cuantificarea matricelor *F*.

Cuantificarea se poate face cu o constantă pentru toate valorile din matricea *F* (*cuantificare uniformă*). Constanta este diferită în cazul în care este vorba de o matrice *F* obținută prin aplicarea transformării *DCT* asupra unui pătrat obținut din matricea de luminanță și asupra unui pătrat obținut din una dintre matricele de crominanță. O altă modalitate de a cuantifica valorile matricei *F* este aceea de a utiliza două matrice de constante (*cuantificare neuniformă*): una pentru funcțiile *F* obținute din matricea de luminanță și alta pentru funcțiile *F* obținute din matricele de crominanță.

După cuantificarea fiecărei matrice *F* obținute prin împărțirea întreagă a fiecărui element din matricea *F* la elementul corespunzător din matricea constantă, rezultatul se depune într-un tablou unidimensional cu 64 de elemente în zig-zag, după cum indică figura alăturată.

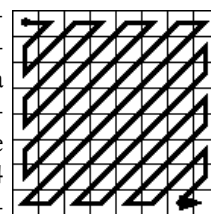


Figura 1

În continuare, tablourile obținute din matricea de luminanță se codifică folosind algoritmul *DPCM* (*Differential Pulse Code Modulation* - transformarea delta este un caz particular al *DPCM*), iar tablourile obținute din matricele de crominanță se codifică folosind algoritmul *RLE*.

Rezultatele obținute la pasul anterior se comprimă folosind algoritmul lui *Huffman* varianta statică sau semi-statică.

În final, la ieșire se transmit constantele de cuantificare, codurile *Huffman* utilizate dacă acestea nu sunt standard (în cazul în care nu s-a folosit varianta statică cu coduri predefinite standard) și datele comprimate.

Bibliografie

1. T. H. Cormen, C. E. Leiserson, R. R. Rivest, *Introducere în algoritmi*, Computer Libris Agora, Cluj-Napoca, 2000
2. A. S. Tanenbaum, *Rețele de calculatoare*, ediția a patra, Byblos, București 2003
3. J. A. Storer, *An Introduction to Data Structures and Algorithms*, Springer, 2002
4. <http://www.datacompression.info/>
5. <http://www.researchindex.com/>

Claudiu Soroiu este redactor GInfo și poate fi contactat prin e-mail la adresa csoroiu@yahoo.com.