Multimédia tömörítési módszerek

Jellemzően tömörítetlen formában a multimédia adatok igen nagy tárterületet igényelnek. A multimédia tömörítési eljárások általában a veszteséges eljárások csoportjába tartoznak, azaz nem állítják vissza az eredeti adatot, hanem úgy csalnak, hogy az emberi érzékszervek ne érzékeljék.

Példa

pl. Mikrofonnal 1 másodperc sztereó felvétel CD minőségben:

```
44100 * 16 bit * 2 ~ 1,4 [MegaBit]
```

És ez csak 1 másodperc volt!

Példa

Full HD képkocka (1080p): 1920-1080 mérete (elméleti mérete)

```
1920 * 1080 * 24 bit ~ 6 MByte
```

Vegyük figyelembe, hogy ez csak egyetlen képkocka, a modern multimédia rendszerek másodpercenként 30-60 (újabban akár 100) képkockát szoktak megjeleníteni.

Veszteséges tömörítés alapelve: Olyan módon kell átalakítani a bitfolyamokat, hogy az emberi szem és fül számára fontos információ megmaradjon. A tapasztalat szerint a szemünk érzéketlenebb a nagyobb térbeli frekvenciájú képösszetevőkre, azaz egy halvány sűrű mintázatot nehezebben veszünk észre, mint egy ritkább mintázatot. Mozgó kép esetén is az alacsony frekvenciákra érzékeny a szemünk.

A kétdimenziós Fourier transzformáció abban segít, hogy a kép pixeleit átalakítja frekvencia értékekre.

megjegyzés: adott egy mikrofonfelvétel, ami mintavételezett időpillanatokban az érzékelt hangerőértékeket (amplitúdókat) ábrázolja. Tegyük fel, hogy zongorán valaki a 'Boci-boci tarka' c. dalt vette fel. A Fourier transzformációt úgy kell elképzelni, hogy a transzformáció meg tudja állapítani, hogy egy adott időintervallumban pl. egy **C** szólt, máshol egy **E**. Ez olyan, mintha a kottát próbálnánk előállítani hallás után. Ha csak a kottafejeket kódolnánk és megadnánk hogy zongorát használtunk, akkor nagyon nagy tömörítést kapnánk. A valóságban az összes felhangot is megkapjuk csökkenő intenzitással.

JPEG tömörítés elve

A világon a legismertebb veszteséges tömörítéses módszer a JPG vagy JPEG.

Nem célunk a precíz matematikáját bemutatni. Viszont megmutatjuk, hogy milyen lépésekből áll: JPEG eljárás

1.) A képet 8×8-as mátrixokra osztjuk és egyesével haladunk mátrixonként

2.) Elvégezzük az un. diszkrét koszinusz transzformációt a mátrixon, ez alakítja át frekvencia tartományba az adatot, ez teszi lehetővé, hogy úgy tudjunk elhagyni értékeket (információt), hogy ez észrevehető lenne szemmel.

3.) Kvantáljuk az eredmény mátrixot előre meghatározott szabályok alapján. Ez a lépés jelenti az adatvesztést. Ezt úgy lehet elképzelni, mint egy kerekítést. Tegyük fel, hogy az eredmény 16.345 volt egy adott mátix pozícióban, de e helyett 16-ot írunk.

4.) Mátrixot cikk-cakk elrendezésben tekintjük és alkalmazzuk az értékekre az RLE kódolást.

5.) Az eredményt Huffmann kódolással tovább tömörítjük.

A tömörítés paramétere (1-100 érték amit minden jpeg tömörítő mentéskor megkérdez) a kvantálás finomabbá, illetve durvábbá tétele lesz. Általában 80-as értéken szokták alapértelmezetten tartani.

A dekódolás szakasza fordítva halad végig a pontokon és minden pontban inverz műveletet hajt végre.

MPEG tömörítés elve

Az emberi fül a **20 Hz** és **20 kHz**-es tartományban érzékeli a hangot. A legérzékenyebb tartomány **2-4kHz** között van (megjegyzés: a hangvilla **440 Hz** frekvencián rezeg). A hallás jellemző tulajdonsága, hogy egy időben hangzó két közeli frekvencián rezgő hang közül a nagyobb intenzitású elfedi az alacsonyabb intenzitásút. Ez elfedés akkor is érvényes, ha az alacsony intenzitású hang **2ms**-al korábban szólal meg és kb. **15ms**-al később fejeződik be, mint a nagyobb intenzitású közeli frekvencián szóló másik hang.

A kvantálást úgy kell elvégezni, hogy figyelembe vesszük, hogy a **2-4kHz** közötti frekvenciákat nagyobb finomsággal kezeljük, mint a pl. a **10-20Khz** közöttieket. (Emlékezzünk a kvantált jelre korábban az anyagban, az y tengelyt nem egyenletesen osztjuk fel, az alja finomabb a teteje durvább.)

Az MPEG tömörítés kép és hang esetére is sokféle módszert kínál. A képtömörítésnél figyelembe veszi az egyes képkockák közötti változásokat is. Ennek megfelelően **key-frame**-eket (kulcs frémeket) definiál (**I-frame**), amelyek teljes képkockát tömörítenek a korábban megismert **jpeg** módszerrel, valamint a köztes képkockákban felismerik a változásokat (**P-frame**, **B-frame**). Ennek jelentősége a gyorstekerésnél is lényeges, mert egy **P-framehez** tartozó teljes képkocka, csak az őt megelőző utolsó I-frame-től kezdődően készíthető el, mivel nem tartalmazza a teljes képi információt, csak a változásokat.

Az **mpeg** esetén fontos, hogy a kép és a hang időben szinkronban legyen, ezért a hang és kép időbélyegeket is tartalmaz.

Az mpeg mint szabvány a tárlót is definiálja, a jól ismert mp3 és mp4 ennek a szabványnak a változatai.

Mp4

A mobiltelefonunkról ismert mp4 képet és hangot is egyaránt tárol, de a kép és hang sokféleképpen kódolható benne. A telefonunk a h264 vagy (h265) elnevezésű tömörítést használja képkockák tömörítéséhez, ennek sokféle hardveresen támogatott és nem támogatott eljárása van.

From:
https://edu.iit.uni-miskolc.hu/ - **Institute of Information Science - University of Miskolc**

Permanent link:
https://edu.iit.uni-miskolc.hu/tanszek:oktatas:infrendalapjai_architekturak:informacio_feldolgozas:multimedia_toemoerites

Last update: **2024/11/12 19:55**