

ZÁPADOČESKÁ UNIVERZITA V PLZNI
FAKULTA PEDAGOGICKÁ
KATEDRA VÝPOČETNÍ A DIDAKTICKÉ TECHNIKY

**KOMPONENTY PRO VÝUKOVÝ ELEKTRONICKÝ MATERIÁL -
KOMPRESSE V OBLASTI POČÍTAČŮ**
BAKALÁŘSKÁ PRÁCE

Lukáš Smutný

Přírodovědná studia, Informatika se zaměřením na vzdělávání

Vedoucí práce: Mgr. Zbyněk Filipi

Plzeň, 2015

Prohlašuji, že jsem diplomovou práci vypracoval samostatně s použitím uvedené literatury a zdrojů informací.

V Plzni, 30. června 2015

.....
vlastnoruční podpis

ZDE SE MŮŽE, UZNÁ-LI AUTOR ZA VHODNÉ, NACHÁZET
PODĚKOVÁNÍ.

ZDE SE NACHÁZÍ ORIGINAL ZADÁNÍ KVALIFIKAČNÍ PRÁCE.

OBSAH

SEZNAM ZKRATEK	2
ÚVOD	3
1 ALGORITMY BEZZTRÁTOVÉ KOMPRESY DAT	4
1.1 METODA POTLAČENÍ NUL	4
1.2 RLE	4
1.3 LZ77	4
1.4 LZ78	4
1.5 LZW	5
1.6 SHANNON-FANOVO KÓDOVÁNÍ	5
1.7 HUFFMANOVO KÓDOVÁNÍ	5
1.8 ARITMETICKÉ KÓDOVÁNÍ	5
1.9 LZMA	5
1.10 PNG	6
2 ALGORITMY ZTRÁTOVÉ KOMPRESY	7
2.1 JPEG	7
2.2 JPEG 2000	7
2.3 MPEG	7
2.4 FRAKTÁLOVÁ KOMPRESY	8
2.5 MP3	8
2.6 AAC	8
ZÁVĚR	9
RESUMÉ	10
SEZNAM LITERATURY	11

SEZNAM ZKRATEK

AAC	Advanced Audio Coding
LZ77	Lempel-Ziv 77
LZ78	Lempel-Ziv 78
LZMA	Lempel-Ziv-Markov-Chain Algorithm
LZW	Lempel-Ziv-Welch
MP3	MPEG (Moving Picture Experts Group) Audio Layer III
MPEG	Moving Picture Experts Group, přeneseně jejich video formáty
JPEG	Joint Photographic Experts Group, přeneseně jejich obrázkový formát
PNG	Portable Network Graphics
RLE	Run Length Encoding

Úvod

Díky míře elektronizace současného světa jsme bez přehánění obklopeni daty. Protože kapacita elektronických zařízení a přenosových cest není neomezená, je žádoucí, aby tato data byla komprimována. Toho lze dosáhnout tím, že se buď v těchto datech sníží, nebo omezí redundance (bezztrátová komprese – první kapitola) nebo se některá nepodstatná data vynechají (ztrátová komprese – druhá kapitola).

1 ALGORITMY BEZZTRÁTOVÉ KOMPRESY DAT

Bezztrátová komprese se aplikuje tam, kde je potřeba zachovat přesnou informační hodnotu dat, především text, ale lze pochopitelně použít i na obraz či zvuk.

1.1 METODA POTLAČENÍ NUL

Jednoduchá metoda, která nahrazuje často se opakující znak (nulu) zápisem obsahujícím speciální znak a počet nahrazených nul.

Příklad: A00B000C0000D00 -> A#1B#2C#3D#1

1.2 RLE

Metoda proudového kódování se vyplatí při nahrazování velkého počtu opakovaných znaků, například u obrazu s velkými stejnobarevnými plochami. Protože se k zápisu nahrazení používají tři znaky (vodící znak, opakovaný znak a počet opakování), nemá smysl nahrazovat znaky, pokud se opakují méně než čtyřikrát.

Příklad: AAAABBCCDDDDDD -> #A4BBCCC#D5

1.3 LZ77

Tato metoda původem od autorů Lempel a Ziv hledá opakující se posloupnosti znaků a na místa jejich opakování vkládá odkaz na místo jejich předchozího výskytu. Dlouho nepoužité posloupnosti uvolňují místa nově použitým ve vyhledávacím bufferu.

1.4 LZ78

Na rozdíl od LZ77 není tato metoda schopná „zapomínat“ dlouho nepoužité posloupnosti, každou novou přidává do slovníku, který pak může značně narůst. Tato metoda je výhodná na kompresi málo proměnlivých dat. Naopak může být značně neefektivní, pokud se struktura dat v přenášených signálech významně mění během času.

1.5 LZW

Tato vylepšená metoda původem o autorů Lempel a Ziv hledá opakující se posloupnosti znaků, z nichž pak tvoří slovník, kdy každé uložené posloupnosti odpovídá určitý kód ve slovníku. Metoda je adaptivní, obsah slovníku se mění v závislosti na kódovaných datech.

1.6 SHANNON-FANOVO KÓDOVÁNÍ

Tuto metodu vymysleli oba uvedení v názvu nezávisle na sobě ve stejnou dobu. Prefixový kód vzniká podle statisticky určeného počtu výskytu znaku v kódovaném řetězci. Výsledný kód je proměnlivé délky, protože závisí na počtu různých znaků v řetězci.

1.7 HUFFMANOVO KÓDOVÁNÍ

Jedná se o statistickou metodu, která nejprve zjistí četnost výskytu jednotlivých znaků, na jejímž základě sestaví binární strom. Strom je tvořený tak, aby nejčastější znak měl nejkratší, jednobitový, kód. A naopak, nejméně častý znak nejdelší z užívaných kódů.

1.8 ARITMETICKÉ KÓDOVÁNÍ

Při použití této metody je celý vstupní řetězec zakódován jako část intervalu od nuly včetně do jedné. Čím delší je řetězec, tím více se k sobě blíží horní a dolní mez výsledného intervalu. Změna mezí je určena pravděpodobností výskytu posledního zaznamenaného znaku. Výsledný interval je potřeba doplnit buď o počet znaků, který má představovat, nebo řídicí znak označující konec řetězce.

1.9 LZMA

Kombinuje LZ77 s Markovovými řetězci a „range“ kódováním, které, podobně jako například aritmetické kódování, kóduje řetězec do podoby jednoho čísla.

1.10 PNG

Je kombinace LZW a prokládání a kombinací s různými předzpracováními pixelů k dosažení optimálního kompresního poměru. Předzpracování může spočívat v přímých informacích o pixelu, v rozdílu oproti pixelu o jeden vlevo, v rozdílu oproti pixelu o jeden nahoře, průměr z pixelu a pixelů o jeden vlevo a nahoře nebo hodnoty z pixelu a jeho tři sousedů (první o jeden vlevo, druhý o jeden nahoře a třetí o jeden vlevo i nahoře).

2 ALGORITMY ZTRÁTOVÉ KOMPRESSE

Pokud není nutné při dekompresi získat totožná data, jaká vstupovala do komprese, dá se pomocí ztrátových kompresních metod docílit velmi vysokého kompresního poměru na úkor kvality zpětného převodu. Nedílnou součástí ztrátových kompresních metod jsou různé druhy transformací, které data uspořádají například podle důležitosti, a pokud je algoritmus postavený tak, že méně podstatná data zanedbává nebo nahrazuje, dojde ve výsledku k efektivnější kompresi.

2.1 JPEG

Je metoda s volitelnou mírou komprese, kdy dokáže na výstupu docílit poměru 25:1 nebo i 100:1 (záleží mimo jiné na „vhodnosti“ obrázku). Při kompresi je původní obraz rozdělen na čtverce obsahující 64 pixelů, na které je posléze aplikována diskretní kosinová transformace.

2.2 JPEG 2000

Vylepšený JPEG, kombinuje vlnkovou transformaci se skalárním kvantováním a entropickým kódováním.

2.3 MPEG

Pro pohyblivé obrázky, filmy, bylo potřeba vyvinout vlastní kompresní metodu, protože ty používané na statické obrázky nedostačovaly, byť měly relativně velký kompresní poměr. U filmů se navíc využívá kódování pomocí rozdílů oproti předchozím snímkům, protože při frekvenci 25 snímků za sekundu se vychází z předpokladu, že po sobě jdoucí snímky se většinou příliš neliší.

2.4 FRAKTÁLOVÁ KOMPRESSE

Rekurzivně hledá sobě podobné obrazy, což je časově náročné a tak je výsledná kvalita obrazu dána především určeným časovým limitem na kompresi.

2.5 MP3

Jedná se o metodu komprese zvuku využívající frekvenční a časové maskování. Je vhodná na kompresi hudební produkce a nepříliš vhodná na kompresi mluveného slova. Hlavním parametrem je tzv. bitrate, který má značný vliv na kvalitu výsledného zvuku. Bitrate nemusí být pevně daný, může být proměnlivý a kombinovat tak výhody vysokého bitrate (kvalitní zvuk) s výhodou nízkého bitrate (velký kompresní poměr).

2.6 AAC

Vylepšený kompresní audio formát, který oproti mp3 získává především využitím dlouhodobé predikce a zapojením postprocessingu.

ZÁVĚR

Pokud je potřeba po dekompresi dostat stejná data jako před kompresí, je třeba zvolit některou bezztrátovou kompresní metodu. Výběr je tak podmíněn především výsledným kompresním poměrem, případně časovou náročností metody. Tam kde stačí zachovat jen podobná data, lze sáhnout po některé ze ztrátových kompresních metod. Tyto metody výměnou za neúplnost dat po dekompresi nabízí až řádově lepší kompresní poměr.

RESUMÉ

If same data after decompression are needed, you have to choose any of the lossless compression methods. The selection is then conditioned over compression ratio or average time consumption. If it's sufficient to maintain just a similar data you can pick any of the lossy compression methods. These methods in return for incomplete data after decompression offer a better compression ratio.

SEZNAM LITERATURY

ČAPEK, Jan a Peter FABIÁN. *Komprimace dat: principy a praxe*. Vyd. 1. Praha: Computer Press, 2000, viii, 173 s. Internet. ISBN 80-722-6231-9.

JIROUŠEK, Radim a Peter FABIÁN. *Principy digitální komunikace: principy a praxe*. Vyd. 1. Voznice: Leda, 2006, 309 s. Utility. ISBN 80-733-5084-X.

MURRAY, James D. *Encyklopedie grafických formátů*. 2. vyd. Praha: Computer Press, 1997, 922 s. ISBN 80-722-6033-2.

PECINOVSKÝ, Josef a Peter FABIÁN. *Archivace a komprimace dat: principy a praxe*. Vyd. 1. Praha: Grada, 2003, 116 s. Utility. ISBN 80-247-0

