

KOMPRIMACE

aneb jak se to tam všechno vejde?

768

576

Položme si hned na začátku zdánlivě nepodstatnou otázku: Kolik místa zabere dvouhodinový film na CD nebo DVD? Uvažujme následující příklad: rozlišení jednoho snímku filmu je **768 x 576 obrazkových bodů (pixelů)**. Kolik místa by tedy zabral jeden snímek?

U barevného filmu na jeden pixel potřebujeme **3 B** (pro každou barvu jeden – červená, zelená a modrá). Při rozlišení 768 x 576 to dohromady dělá **442 368 pixelů**
Jedna snímek tedy zabere $3 \times 442\,369 \text{ B} = 1\,327\,104 \text{ B} \approx \mathbf{1,33 \text{ MB}}$
Na **1 vteřinu záznamu** potřebujeme 25 snímků (televize) $25 \times 1,33 \text{ MB} = \mathbf{33 \text{ MB}}$
Na **1 minutu filmu** bychom potřebovali $60 \times 33 \text{ MB} = \mathbf{1\,980 \text{ MB}}$
Na **1 hodinu filmu** by to dělalo už $60 \times 1\,980 \text{ MB} = 118\,800 \text{ MB} \approx \mathbf{120 \text{ GB}}$
A tedy na **dvouhodinový film** bychom potřebovali $2 \times 120 \text{ GB} = \mathbf{240 \text{ GB}}$
Pokud bychom jako paměťové médium použili **700 MB CD-ROM**, potřebovali bychom jich na dvouhodinovou filmovou produkci rovných **343**, tj. každých **11 vteřin** bychom museli CD vyměnit. Pokud bychom použili **4,6 GB DVD**, potřebovali bychom jich „jen“ **53**. To znamená, že bychom je museli vyměňovat zhruba každé **2 minuty**. Za těchto podmínek bychom si asi film na video CD i DVD klidně odpustili.

Komprimace

To, co činí promítání filmů z CD-ROMů realizovatelným, je **komprimace** neboli laicky řečeno „zhuštění“, „stlačení“, „sbalení“ či „spakování“ dat na mnohem menší velikost. Ve skutečnosti se jedná o vypuštění nadbytečné informace takovým způsobem, aby bylo možno data zpět zrekonstruovat pokud možno bez viditelné či slyšitelné ztráty kvality.

Typickým příkladem komprimace dat je např. *Morseova abeceda*, kdy nejčastěji používané znaky mají nejkratší kód (e = •, t = –, a = •–, n = –• apod.).

Proces, kterým se data převádějí do původního formátu, se nazývá **dekomprimace**. Například u obrázku uloženého v komprimovaném formátu se obnoví původní data automaticky, jakmile jej na svém počítači otevřeme.

Komprimaci lze podle možnosti zpětné rekonstrukce dat rozdělit na bezztrátovou a ztrátovou.

Bezeztrátová komprimace

Bezeztrátová komprimace je taková, při které lze rekonstruovat původní data v plném rozsahu bez ztráty jakékoli informace. Používá se pro komprimaci obrázků, textu a také ji používají programy pro komprimaci pevného disku i různé komprimační programy pro vytváření archivů souborů (WinZip, WinRAR, PowerArchiver apod.).

Např. v textu se často opakující slova nahrazují kratšími symboly (speciálními značkami), jejichž seznam i s odpovídajícími částmi textu se přidá na konec souboru. V obrázku se řádka stejně barevných bodů místo informace o každém bodu jednotlivě nahradí informací o barvě a počtu bodů této barvy bezprostředně za sebou jdoucích.



Bezeztrátový typ komprimace se používá například pro komprimaci disků; dále ho používají formáty BMP, TGA, PCX, TIFF apod., které využívají algoritmu RLE (Run-length Encoding neboli proudové kódování dat) nebo programy ARJ, RAR a formáty GIF, TIFF, které jsou založeny na algoritmech LZ či LZW (Lempel-Ziv-Welchova komprese), jehož princip spočívá v tom, že nahrazuje často se opakující sekvence bitů kratšími kódy (čím častější výskyt, tím kratší kód se snaží zvolit).

Metoda RLE pracuje na principu redukce opakovaných řetězců znaků (proudů). V nekomprimované podobě by 15 znaků „A“ za sebou vypadalo asi takto:

AAAAAAAAAAAAAAAAA

Stejný řetězec po RLE kódování vypadá takto:

15A

Podívejme se, jak se změní 15znakový řetězec, pokud bude obsahovat čtyři různé proudy znaků:

Psssssst

Po RLE kódování:

1P8s1t

Všimněme si, že při kódování jednoho řetězce (znaku „t“) je výsledný kód delší než původní (1t).

A ještě jeden příklad – pokud by nekódovaný řetězec vypadal takto:

Nenapodobitelný

Výsledný řetězec bude mít přesně dvojnásobnou délku:

1N1e1n1a1p1o1d1o1b1i1t1e1l1n1ý

RLE schémata jsou rychlá a účinná, ale jejich kompresní účinnost závisí do značné míry na typu dat. Nehodí se například na běžný text v ASCII (textový dokument).

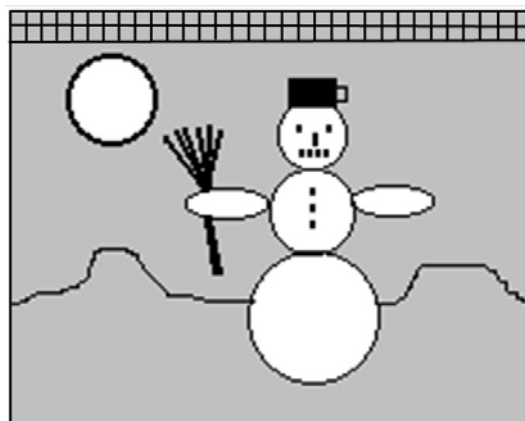
Metoda LZW je jedna z nejrozšířenějších metod používaných v počítačové grafice. LZW je substituční algoritmus, nebo také algoritmus založený na slovníkovém kódování. Algoritmus buduje datový slovník (nebo překládací tabulku) dat, která se vyskytují v nekomprimovaném toku dat. Např. v textu se často opakující se slova nahrazují kratšími symboly (speciálními značkami), jejichž seznam i s odpovídajícími částmi textu se přidá na konec souboru.

Podle typu dat lze bezeztrátovou komprimací dosáhnout zhruba 1/2 až 1/10 původní velikosti (srovnejme například níže uvedenou fotografii skupiny lidí a fotografii draka na obloze).

Ztrátová komprimace

Ztrátová komprimace je taková, při které se původní data dají rekonstruovat jen s větší či menší přesností (s určitou ztrátou některé informace). Ztrátová komprese se používá proto, aby se dosáhlo lepšího kompresního poměru než u neztrátových kompresí. Pomocí ztrátové komprimace lze dosáhnout až 1/25 původní velikosti, ale obvykle na úkor kvality rekonstruovaných dat. Ztrátová komprimace využívá nedokonalosti lidského zraku (popřípadě sluchu), kdy např. dva podobné barevné odstíny vnímá lidské oko jako jednu barvu. Používá se pro komprimaci obrázků, zvuku i videa.

Typickým představitelem ztrátové komprimace je **formát JPEG** (Joint Photographic Experts Group), který se používá pro bitmapové obrázky. Plocha obrázku je rozdělena na čtverce, ve kterých se intenzita dané barvy mění jen nepatrně a všechny odstíny barev každého bodu v tomto čtverci jsou pak nahrazeny barvou jedinou. Čím více odstínů nahradíme barvou jedinou, tím větší komprimace můžeme dosáhnout (ovšem samozřejmě na úkor kvality zrekonstruovaných dat).





Původní velikost tohoto obrázku v nekomprimovaném formátu BMP (při rozlišení 1 024 x 636 pixelů) je 1 860 kB.

Při použití bezztrátové komprimace ve formátu TIFF se soubor obsahující tento obrázek zmenšil na 1 800 kB, což je téměř původní velikost.

Při použití ztrátové komprimace ve formátu JPG (Q faktor 10) se velikost souboru zmenšila na 230 kB, což je téměř 1/8 původní velikosti.



Původní velikost tohoto obrázku v nekomprimovaném formátu BMP (při rozlišení 1 024 x 636 pixelů) je 1 860 kB.

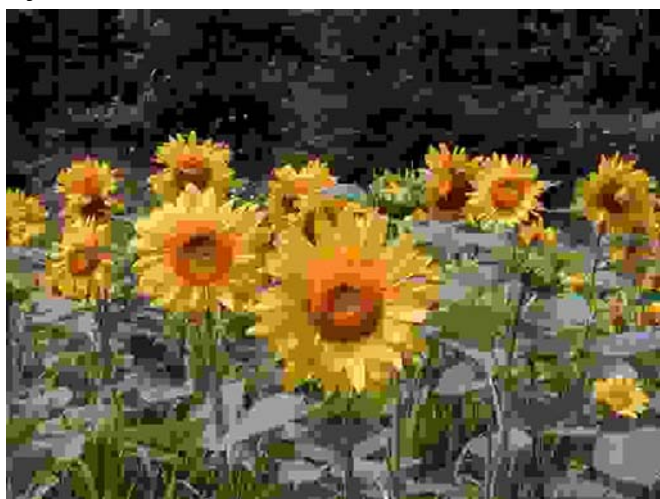
Při použití bezztrátové komprimace ve formátu TIFF se velikost souboru opět téměř nezmění.

Při použití ztrátové komprimace ve formátu JPG (Q faktor 10) se velikost souboru zmenšila na 76 kB, což je 1/24 původní velikosti.

Při používání ztrátové komprimace je potřeba zvolit rozumný kompromis mezi velikostí výsledného zkomprimovaného souboru a kvalitou dat, která obsahuje.



Původní obrázek (při rozlišení 800 x 600 pixelů) – velikost 504 kB.



Špatně zvolený faktor komprimace – velikost výsledného souboru je sice pouhých 11 kB. Ale daň v podobě velmi špatné kvality obrázku je příliš vysoká (příliš mnoho odstínů dané barvy bylo nahrazeno odstínem jediným).

Z principu formátu JPEG vychází i **formát MPEG** (Motion Picture Experts Group), který se používá pro video, animaci a zvuk. Standardní rychlost u videa je 25 nebo 30 snímků za vteřinu (podle použité normy PAL v Evropě či NTSC v USA a Francii). U filmu totiž není nutné komprimovat každý snímek zvlášť. Díky tomu, že se většinou bezprostředně následující snímky od sebe příliš neliší, spočívá princip komprimace v tom, že se zkomprimuje 1 celý (tzv. klíčový) snímek každou půlvteřinu a u zbývajících 14 v ostatním čase se zaznamenávají pouze rozdíly mezi dvěma snímky. Pomocí této komprimace lze dosáhnout 1/200 až 1/400 původní velikosti.

Další způsob, kterým se dá snížit velikost souboru digitálního videa, je, že namísto nahrávání všech 30 snímků za vteřinu, nahrávají někteří výrobci pouze každý druhý snímek (15 snímků za vteřinu), nebo dokonce jen každý třetí snímek (10 snímků za vteřinu).

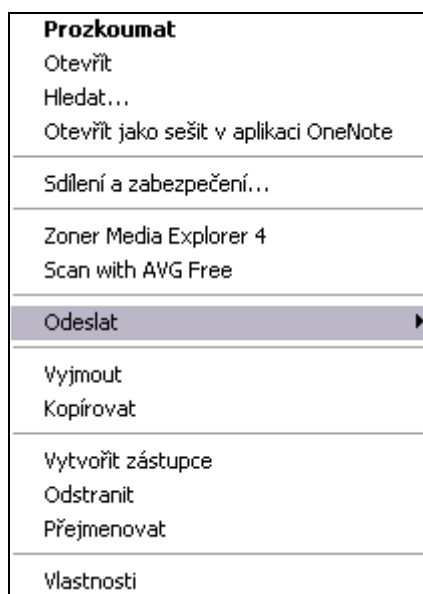
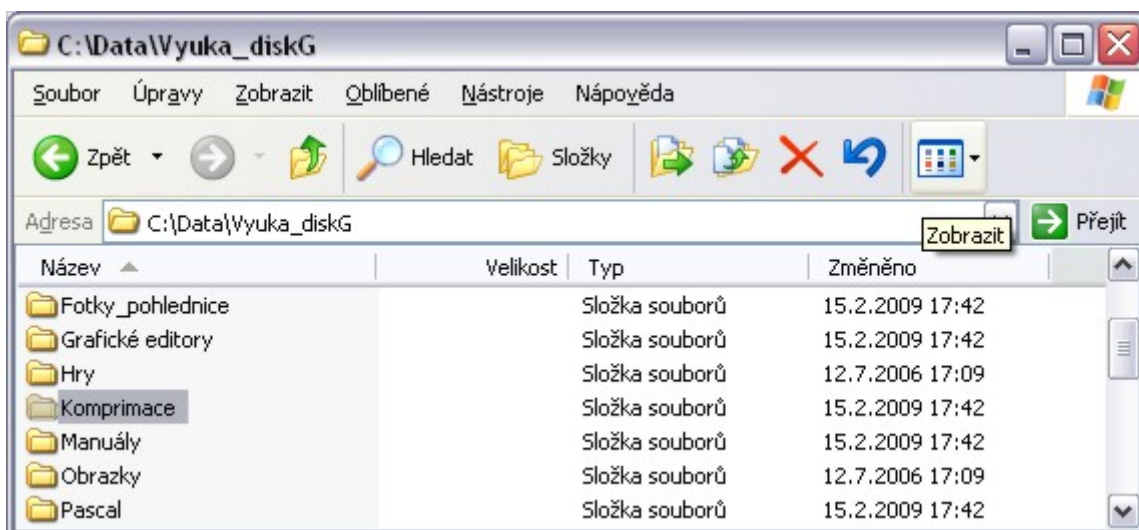
Jiným způsobem pro značné snížení velikosti souboru s digitálním filmem je zmenšení jeho velikosti na obrazovce – standardní televizní obrázky (u evropské normy PAL je to 768 x 576 pixelů, u americké NTSC 640 x 480 pixelů) jsou zmenšeny do formátu 352 x 240 pixelů.

Komprimace se také využívá k zmenšení velikosti souborů, které budou přenášeny přes internet. Pokud posíláme soubory jako přílohu elektronické pošty, měli bychom mít na paměti „přiměřenou“ velikost přenášovaných souborů. Ne každý disponuje neomezeně velkou poštovní schránkou či superrychlým připojením k internetu. Přenos zkomprimovaných souborů je podstatně rychlejší (ve svém důsledku i levnější).

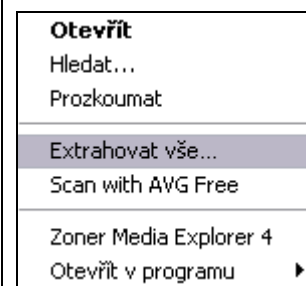
Možnosti komprimace

Kromě výše zmíněných programů pro komprimaci, jako je WinZip, WinRAR, PowerArchiver apod., můžeme také využít souborové manažery, které v sobě možnost komprimace obsahují. Soubory, které chceme zkomprimovat, si obvykle připravíme do samostatné složky. Dále si ukážeme dva velmi jednoduché způsoby jak připravenou složku „zabalit“. (Komprimovanou složku i archiv jsem nazvala Komprimace.)


V **Průzkumníku Windows** nebo v aplikaci **Tento počítač** stačí na připravenou složku ťuknout pravým tlačítkem myši a z nabídky vybrat příkaz **Odeslat** a poté **Komprimovaná složka**. Vytvoří se soubor, takzvaný **archiv**, s původním názvem složky a příponou **ZIP**. Dekomprimace se provede obdobně – ťukneme pravým tlačítkem myši na komprimovaný soubor (archiv) a zvolíme **Extrahovat**. V obou případech se archiv i rozbalený archiv vytvoří v aktuální složce (pokud neurčíme jinak).

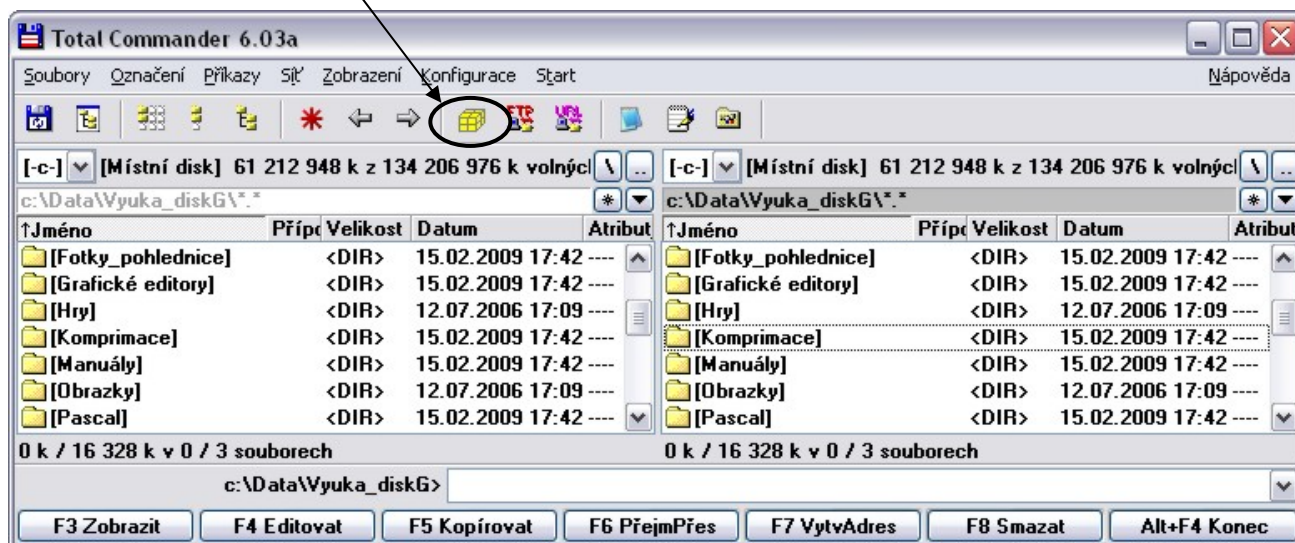


Komprese

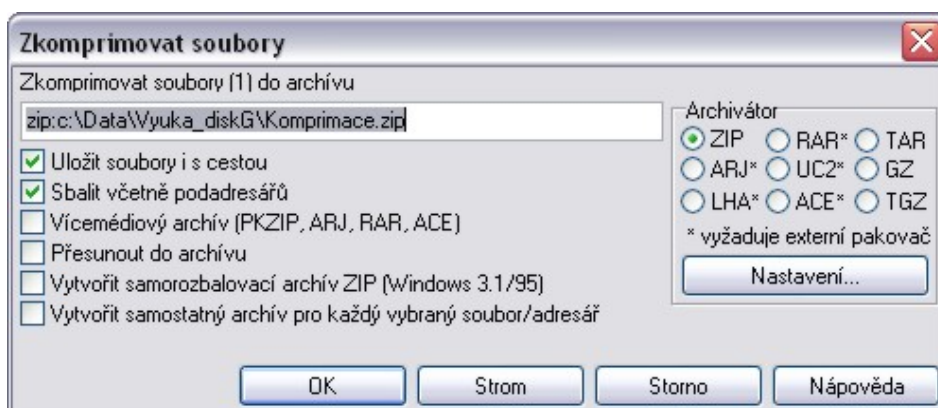


Dekomprese

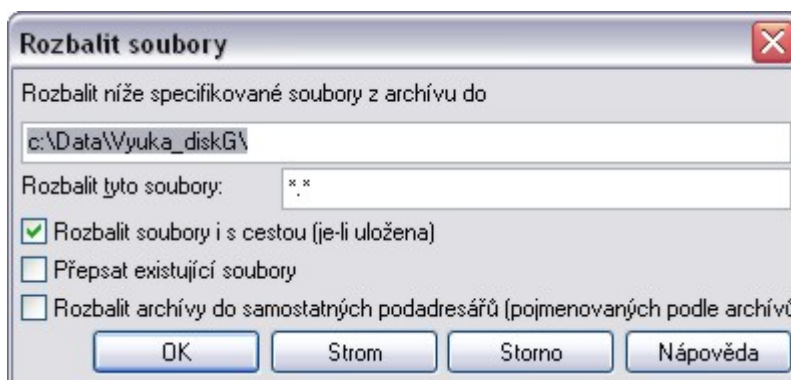
Používáme-li program **Total Commander**, je postup velmi podobný. Opět si nejprve připravíme složku se soubory, které chceme zkomprimovat, a poté zvolíme nabídku **Soubor** → **Komprese...** nebo ťukneme na ikonu  v panelu nástrojů.



V následujícím okně můžeme upřesnit kam, pod jakým jménem a jakou metodou se má archiv vytvořit.



Dekomprimace se opět provede obdobně – zvolíme nabídku **Soubor** → **Dekomprese...**



Literatura a zdroje informací

James D. Murray, William van Ryper: **Encyklopedie grafických formátů**, Computer Press, Brno 1995

Libor Kříž: **Komprimační a archivační programy**, Computer Press, Brno 2002

Erik Holsinger: **Jak pracují multimédia**, UNIS, Brno 1995

Zpracovala: Ing. Simona Martínková

© update srpen 2009