

LOSSY IMAGE COMPRESSION USING DCT

Radek Liška

Bachelor Degree Programme (4), FIT BUT

E-mail: xliska09@stud.fit.vutbr.cz

Supervised by: David Bařina

E-mail: ibarina@fit.vutbr.cz

Abstract: This paper deals with currently used methods of lossy image compression. Nowadays widely used method is DCT which is basis of JPEG codec. There is a basic description of each compression step used in JPEG. Existing methods are presented to improve compression ratio and/or image quality. Effect of these methods is shown in experiments on image data.

Keywords: JPEG, JPEG-Plus, DCT, image compression

1. ÚVOD

Cílem tohoto příspěvku je popsat možná vylepšení kodeku JPEG, která by mohly zvýšit kvalitu či kompresní poměr obrázků tímto zpracovávaných. V příspěvku je nejprve popsána posloupnost algoritmů, které se dnes běžně používají pro ztrátovou obrazovou kompresi v podobě kodeku JPEG, poté jsou kapitoly popisující jednotlivé části. Poté jsou uvedeny návrhy možných vylepšení a výsledky experimentování s nimi.

2. KOMPRESSE OBRAZU

Ztrátová komprese obrazu s využitím diskrétní kosinové transformace (DCT) využívá převodu do barevného modelu YCbCr. Složka Y je jasová, na kterou je lidský zrak nejvíce citlivý, zatímco Cb a Cr složky jsou barvonosné, na které je zrak citlivý méně než na jasovou složku. Proto je možné z barvonosných složek použít menší datový objem při zachování nebo menším snížení výsledné vnímané kvality obrazu než při obdobné ztrátě informací jasové složky.

Zpracovávaný obraz je dále rozdělen na menší bloky ve velikosti 8×8. Na každý blok je aplikována DCT, výsledné bloky frekvenčních koeficientů jsou po ní kvantizovány. Poté jsou linearizovány a nakonec zakódovány na výstup. Ztráty na kvalitě a vznikající artefakty způsobuje kvantizační krok, který tím zároveň zvyšuje kompresi vstupního obrazu. Při dekódování je pořadí kroků opačné.

2.1. DISKRÉTNÍ KOSINOVÁ TRANSFORMACE

DCT (diskrétní kosinová transformace) [1] převádí signál z prostorové do frekvenční domény. Nejčastěji používaná DCT-II (1), běžně zvaná jen „DCT“, je tvořena následující rovnicí:

$$X_k = \sum_{n=0}^{N-1} x_n \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad k = 0, \dots, N-1 \quad (1)$$

Inverzní transformace k DCT-II je DCT-III (2) násobená 2/N („IDCT“). Samotná DCT-III:

$$X_k = \frac{1}{2} x_0 + \sum_{n=1}^{N-1} x_n \cos \left[\frac{\pi}{N} n \left(k + \frac{1}{2} \right) \right] \quad k = 0, \dots, N-1 \quad (2)$$

Na bloku se provádí dvourozměrná DCT, která se provede jako série jednorozměrných DCT. DCT koncentruje nejvíce energie na nízkých frekvencích obrazu, zejména v prvním elementu tvořícím DC koeficient (Direct Current = stejnosměrný proud; frekvence tohoto koeficientu je vždy nulová). Ostatní elementy jsou AC koeficienty (Alternating Current = střídavý proud).

2.2. KVANTIZACE

DCT je v podstatě bezztrátová transformace, případné ztráty informací se dá vyhnout zvýšením přesnosti reprezentace čísel. Zhoršení kvality ve ztrátové kompresi obrazu způsobuje kvantizace. Kvantizace probíhá vydělením matice DCT koeficientů pomocí kvantizační tabulky. Obvykle existuje jedna kvantizační tabulka pro jasovou složku a jedna pro barvosné složky. Koeficienty kvantizační tabulky jsou zpravidla přímo úměrné výšce frekvence na dané pozici – zvyšují se podél hlavní diagonály. Hodnoty koeficientů kvantizačních tabulek je možné škálovat pomocí uživatelského nastavení kvality. V tabulkách udaných JPEG standardem [2] lze vidět důraz na snížení objemu informace v barvosných složkách a také ve vyšších frekvencích všech tří barevných složek.

2.3. LINEARIZACE

Blok transformovaný pomocí DCT po kvantizaci stále tvoří dvourozměrnou matici. Jelikož většina kodérů zpracovává pouze jednorozměrný proud dat, je třeba je do jednorozměrné podoby převést pomocí linearizace. JPEG kodek používá cik-cak průchod (obrázek 2.a).

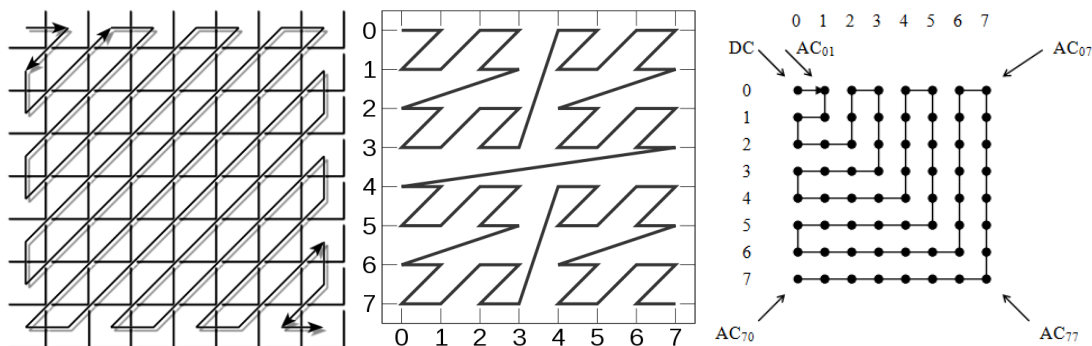
2.4. ENTROPICKÝ KODÉR

Výsledný proud linearizovaných obrazových dat je třeba nějakým způsobem zakódovat do souboru. V JPEG kodeku se využívá Huffmanův kodér [5] pro jeho nízkou náročnost a dobrou výkonnost. Další možnosti popisovanou v JPEG standardu je aritmetický kodér, ale šíření jeho implementace v tomto obrazovém formátu bylo omezeno kvůli patentovému zatížení. Aritmetický kodér poskytuje zlepšení kompresního poměru o 5 až 10 % ve srovnání s Huffmanovým, navýšení času zpracování je ale výraznější a tvoří další důvod problematického prosazování aritmeticky kódovaných obrázků.

2.5. MOŽNOSTI VYLEPŠENÍ

S dnešním výkonnějším hardwarem lze pro vyšší účinnost použít bloky 16×16 namísto současných 8×8. Tímto se sice zvýší hardwarové nároky, ale také se zvýší účinnost komprese díky vyššímu vlivu kvantizace.

Ve standardu JPEG je použit pouze průchod cik-cak. V JPEG-Plus je navržen pyramidový průchod a Mortonův průchod pro JPEG nebyl použit ani navržen; tyto průchody mohou mít vliv na kompresní poměr některých obrázků. Obrázky 2.a a 2.b byly převzaty z Wikipedie, obrázek 2.c je převzat z návrhu JPEG-Plus.



Obrázek 2.a

Obrázek 2.b

Obrázek 2.c

Další možné přístupy k entropickému kódování bloků by mohly být Golombovo kódování či 3R (Recursive Range Reduction). Golombovo kódování je vhodné při geometrické distribuci prvků, když většinu obsahu tvoří několik nízkých čísel, zatímco 3R je pro kódování bloků zajímavou volbou z toho důvodu, že její účinnost na rozdíl od slovníkových metod nezávisí na objemu dat a je ideální pro klesající posloupnosti, které se v jednotlivých kvantizovaných blocích mohou nacházet.

2.6. VÝSLEDKY EXPERIMENTŮ

Všechny experimenty byly prováděny na standardním testovacím obrázku Lena o rozměrech 512×512 pixelů. Experimenty byly provedeny se statickým Huffmanovým kóděrem.

Byly provedeny testy srovnávající různé velikosti bloků a různé linearizační průchody. Vliv rozměrů bloku na velikost výsledného souboru při stejné hodnotě špičkového poměru signálu k šumu (PSNR) je shrnut v tabulce 1. Vliv volby linearizačního průchodu na velikost výsledného souboru při shodné hodnotě indexu strukturální podobnosti SSIM je shrnut v tabulce 2. Velikosti výstupních souborů jsou v bajtech.

PSNR (dB)	Bloky 8×8	Bloky 16×16	Zvolený průchod (SSIM = 0.8634)	Bloky 8×8	Bloky 16×16
33,62	74 231	72 587	Cik-cak	73 716	79 433
35,90	156 131	137 717	Mortonův	74 231	76 739
38,53	281 948	258 952	Pyramidový	74 014	82 945

Tabulka 1

Tabulka 2

3. ZÁVĚR

V příspěvku byly popsány metody ztrátové komprese obrazových dat, které jsou používány ve formátu JPEG/JFIF. Návrh standardu JPEG-Plus, který nebyl schválen ani uveden do praxe, poskytuje další možnosti formátu JPEG/JFIF. Byly shrnuty některé existující metody, které by mohlo být možné využít k vylepšení kompresního poměru či kvality obrazu. Experimenty s testovacím obrázkem bylo zjištěno, že zvětšení velikosti bloku se může pozitivně projevit na výsledné velikosti souboru, zatímco volba linearizačního průchodu neměla na velikost výstupu výrazný vliv s výjimkou pyramidového průchodu u bloků velikosti 16×16.

REFERENCE

- [1] AHMED, N.; NATARAJAN, T.; RAO, K. R. Discrete Cosine Transform. In *IEEE Transactions on Computers*. New York : Institute of Electrical and Electronics Engineers, 1974. s. 90-93.
- [2] ISO/IEC 10918-1:1994. *Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines*. Geneva : International Organization for Standardization, 1994. 182 s.
- [3] MORTON, G. M. *A computer Oriented Geodetic Data Base; and a New Technique in File Sequencing*, 1966. Ottawa : IBM Canada Ltd.
- [4] VOLLBEDING, G. *ITU-T JPEG-Plus Proposal for Extending ITU-T T.81 for Advanced Image Coding* [online]. Geneva : Independent JPEG Group, 2006 [cit. 2011-03-03]. Dostupné z WWW: <http://jpegclub.org/temp/ITU-T-JPEG-Plus-Proposal_R3.doc>.
- [5] HUFFMAN, D. A. A Method for the Construction of Minimum-Redundancy Codes. In *Proceedings of the I.R.E.* New York : The Institute of Radio Engineers, Inc., 1952. s. 1098–1102.