

# NOVÉ MOŽNOSTI KOMPRESNÍ VIDEO SIGNÁLŮ V TELEVIZNÍ TECHNICE

T. Frýza

Ústav radioelektroniky, FEKT VUT v Brně

## Abstrakt

Článek je zaměřen na využití trojrozměrné diskrétní kosinové transformace (3D DCT – Three Dimensional Discrete Cosine Transform) v oblasti komprimace pohyblivých snímků. Podstatou komprese jsou vhodně upravené hodnoty transformovaných koeficientů a jejich efektivní řazení do výstupního datového toku. Tyto operace se provádí obdobně jako u standardu JPEG, určeného pro kompresi statických snímků. Jsou to především operace kvantování, prahování a entropického kódování pomocí Huffmanova, příp. aritmetického kódování. V textu je představen celý komprimační řetězec, vytvořený v prostředí MATLAB.

## 1 Základní informace o 3D DCT

Metoda 3D DCT patří mezi tzv. blokově založené kodéry (mimo jiné JPEG, MPEG), tzn. že kompresi obrazových dat provádí vždy po dílčích blocích. Výhodou 3D DCT je současné odstranění nadbytečnosti jak v jednotlivých snímcích, tak i v čase mezi snímky sousedními, což je nezbytnou operací u všech kompresních metod, zaměřených na video signály. Operace snižující redundanci v čase (hledání shodných objektů, vektorů pohybů) jsou obecně velmi časově náročné, proto lze metodou 3D DCT snížit celkový počet nutných operací pro kompresi/dekompresi obrazových dat.

Pro jeden blok kódovaných obrazových dat o velikosti  $N \times N \times N$  bodů je přímá trojrozměrná diskrétní kosinová transformace (3D fDCT – Three Dimensional Forward Discrete Cosine Transform) definována rovnicí [2]

$$X(l, m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} x(i, j, k) \cdot [C_N]_{li} \cdot [C_N]_{mj} \cdot [C_N]_{nk}, \quad (1)$$

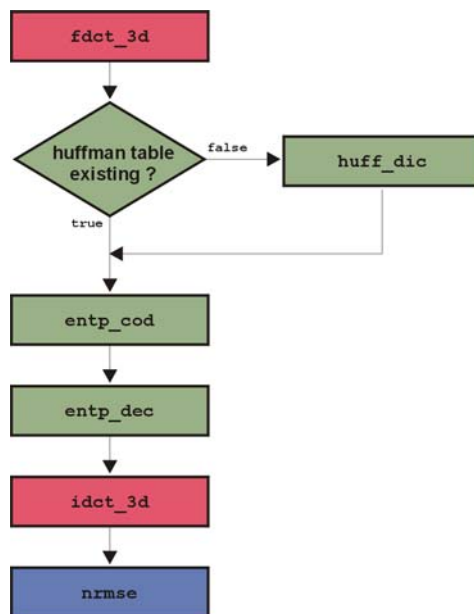
kde hodnota  $x(i, j, k)$  reprezentuje intenzitu jednoho bodu kódované video sekvence na souřadnicích  $(i, j, k)$  sníženou o hodnotu 128,  $X(l, m, n)$  je 3D DCT koeficient na pozici  $(l, m, n)$  vypočtený pomocí definičního vztahu DCT-II [2], básová funkce 3D DCT  $[C_N]_{li} \cdot [C_N]_{mj} \cdot [C_N]_{nk}$  zajišťuje převod hodnot z časové oblasti do spektrální a číselné indexy nabývají hodnot  $l, m, n = 0, 1, \dots, N - 1$ .

Transformační kódování 3D DCT samo o sobě neumožňuje snížení obrovského datového toku kódované video sekvence. Naopak, lze dokázat, že počet nutných bitů pro zakódování frekvenčních koeficientů je vyšší než počet bitů reprezentujících vstupní video sekvenci v časové oblasti. Podstatou komprese jsou vhodné úpravy hodnot a počtu koeficientů a jejich efektivní řazení do výstupního datového toku. Tyto operace se provádí obdobně jako u standardu JPEG, určeného pro kompresi statických snímků. Jsou to především operace kvantování, prahování a entropického kódování pomocí Huffmanova, příp. aritmetického kódování.

Na rozdíl od jednodušší situace u kódování statických snímků, existuje v třírozměrném prostoru (tj. při kódování pohyblivých snímků) mnohonásobně větší počet kombinací, jak u vývoje kvantovacích tabulek, pořadí načítání frekvenčních koeficientů (obdobu cik-cak načítání u JPEG), tak i u vývoje kódových slov pro efektivní zápis Huffmanovým kódem.

## 2 Popis aplikace v prostředí MATLAB

Struktura aplikace pro ověření vlastností 3D DCT při komprimaci obrazových dat vychází ze snahy o jednoduchou variabilitu kódovacího řetězce. Byla proto zvolena struktura, která není vhodná pro kódování a dekódování dat v reálném čase. Nesplnění požadavku na zpracování v reálném čase není chápáno jako nevýhoda, protože stěžejním cílem je výběr takové struktury, která umožňuje dosáhnout nejlepších výsledků s ohledem na velikost kompresního toku a výsledné chyby. Na obrázku 1 je znázorněno zjednodušené blokové uspořádání vytvořeného kodéru v prostředí MATLAB.



Obrázek 1: Bloková struktura kodéru a dekodéru 3D DCT v prostředí MATLAB

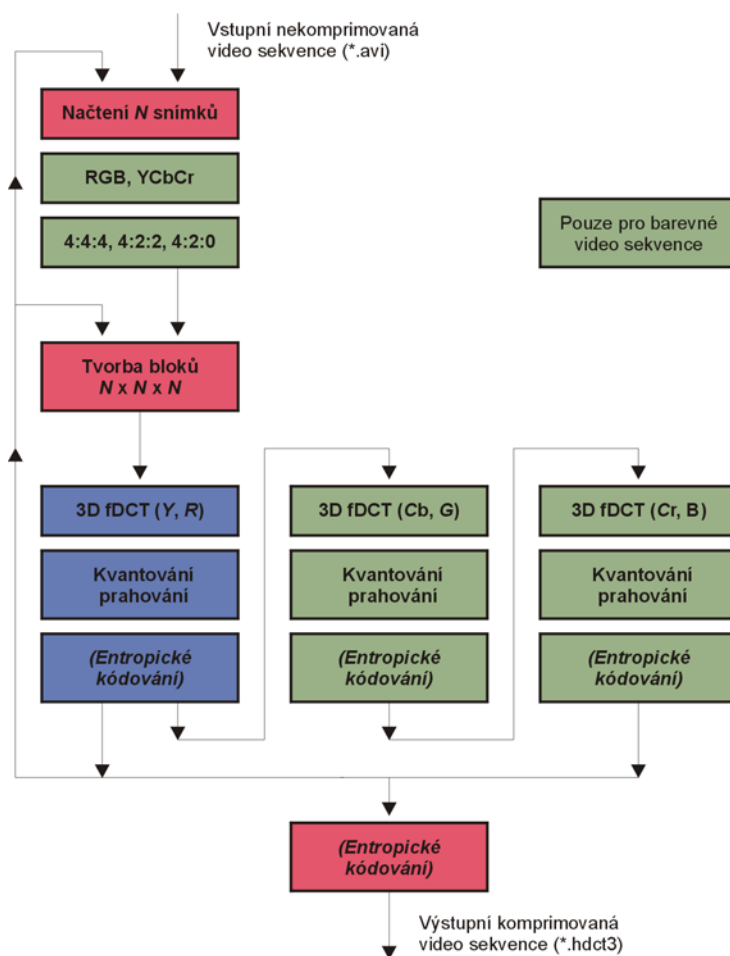
Posloupnost jednotlivých funkcí kodéru a dekodéru 3D DCT odpovídá struktuře, odvozené od standardu JPEG [3]. První blok kodéru, tj. funkce s názvem `fdct_3d` provádějící transformaci vstupních dat do frekvenční oblasti, bude popsán později. Druhým blokem je testování dostupnosti Huffmanových tabulek pro entropické kódování. V případě, že tyto slovníky kódových slov nejsou přístupné, provede kodér analýzu vstupní sekvence a na jejím základě vytvoří slovníky nové, přizpůsobené vstupním datům (funkce `huff_dic`). Pokud Huffmanovy tabulky existují, provede aplikace entropické kódování frekvenčních koeficientů a z velikosti vytvořeného bitového toku vypočte první požadovaný parametr, tj. hodnotu kompresního poměru.

Druhým kritériem pro posouzení komprimační metody je výsledná chyba. Vzhledem k tomu, že cílem aplikace bylo ověření možností 3D DCT, byla komprimovaná data bezprostředně dekódována a pomocí funkce `idct_3d` zpětně transformována na obrazové body. Poslední blok kódovacího řetězce (`nrmse`) provádí objektivní srovnání vstupní a výstupní sekvence pomocí střední normované chyby [1]. Uvedeným způsobem byla ověřena všechna nastavení kodéru 3D DCT, jako např. volba typu řádkování, pořadí načítání frekvenčních koeficientů, či volba kvantizačních tabulek.

Stěžejní operací kodéru je převod obrazových bodů na frekvenční koeficienty a jejich váhování pomocí kvantizačních tabulek. Jak bylo uvedeno, posloupnost těchto operací vykonává funkce `fdct_3d`. Blokové schéma této funkce je znázorněno na *obrázku 2*. Vstupem je barevná, příp. černobílá video sekvence v nekomprimované podobě. Je realizována souborem ve formátu `*.avi`. Kodér postupně načítá skupiny  $N$  snímků této sekvence a ty zpracovává. Pokud je vstupní sekvence barevná, jsou jednotlivé barevné složky převedeny z reprezentace `RGB` na `YCbCr`. Pro testování byla vybrána reprezentace pro standardní rozlišení SDTV a počítačové aplikace [4]. Vypočtené chrominanční signály mohou být dále podzorkovány s ohledem na volbu formátu 4:4:4, 4:2:2, či 4:2:0. Pokud je vstup kodéru tvořen černobílou sekvencí, jsou oba tyto bloky pochopitelně vynechány.

Následně je skupina vstupních snímků rozdělena na video krychle o velikosti  $N$ , které jsou postupně kódovány trojrozměrnou diskrétní kosinovou transformací, algoritmem využívající DFT [2]. Získané frekvenční koeficienty jsou kvantovány, prahovány a mohou být také bezprostředně podrobeny entropickému kódování. V testovací aplikaci byl ale s ohledem na ověření různých způsobů řazení frekvenčních koeficientů a rozdílných Huffmanových tabulek, proces entropického zápisu realizován až dodatečně na všechna komprimovaná data (viz. *obrázek 1*). Z hlediska nároků na paměť při uložení všech frekvenčních koeficientů je zvolený postup nevýhodný, ale ve fázi testování kodéru je naopak vhodnější frekvenční koeficienty nejprve uložit do souboru a následně s nimi pracovat.

Posloupnost operací přímé 3D DCT, kvantování a prahování je pro barevné sekvence realizována celkem třikrát pro každou barevnou složku zvlášť.

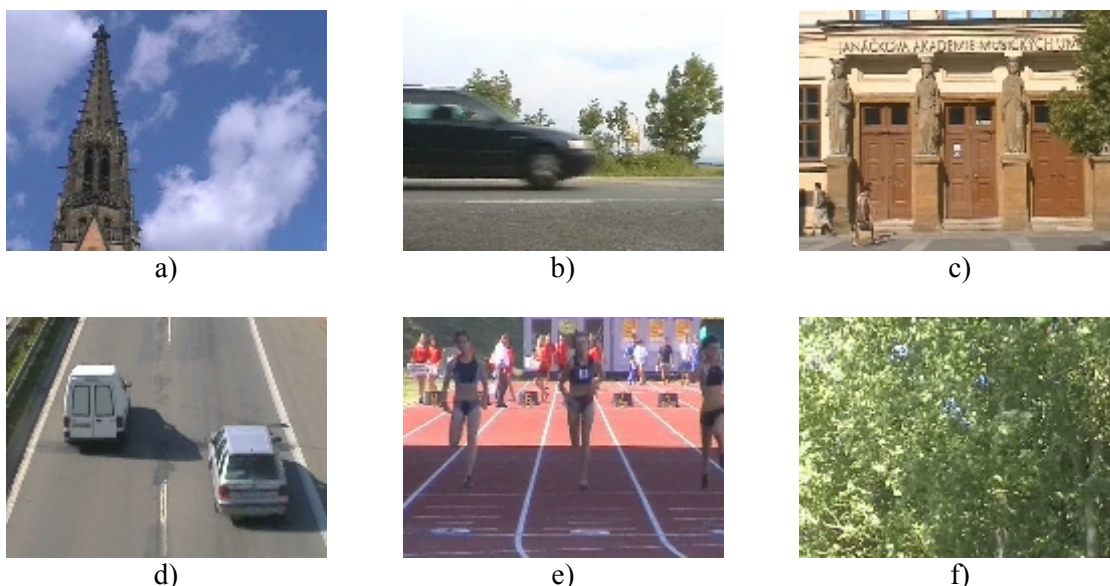


Obrázek 2: Bloková struktura funkce `fdct_3d` v prostředí MATLAB. Bloky se zelenou barvou jsou realizovány pouze pro barevné vstupní sekvence

### 3 Praktické výsledky komprimační metody

Pro potřeby testování vyvinuté aplikace bylo vytvořeno sedm barevných video sekvencí s rozlišením  $320 \times 240$  obrazových bodů a dobou trvání 96 snímků. Množina testovacích sekvencí obsahuje zástupce rozdílných typů. Video sekvence s názvem `cerveny_kostel` a `jamu` obsahují pomalu se měnící děj a velké barevné plochy. Na druhou stranu, sekvence `zavod` a především `vitr` jsou zástupci video sekvencí s rychlým pohybem a množstvím detailů v obraze. Tématická rozmanitost vytvořených sekvencí souvisí s testováním komprimačních vlastností 3D DCT a s empirickým odvozením vhodné aplikace pro kódovací řetězec 3D DCT. Náhledy jednotlivých video sekvencí jsou uvedeny v *obrázku 3*.

Stěžejní vliv na výši kompresního poměru a výsledné chyby má volba kvantizačních tabulek. Pro kvantování transformovaných vzorků jasové složky byla vybrána modifikovaná tabulka standardu MPEG-4. Pro transformované obrazové body chrominančních složek byla přiřazena tabulka uzpůsobena těmto složkám u standardu JPEG. Míra redukce hodnot jednotlivých frekvenčních koeficientů je nastavitelná pomocí tzv. parametru *qlt*. Ten nabývá hodnot z intervalu  $1, 2, \dots, 100$ , přičemž má charakter kvality; tj. pro  $qlt=100$  je výsledná kvalita nejlepší a zmenšuje se při snižování hodnoty tohoto parametru. Také význam parametru *qlt* byl převzat ze standardu JPEG [3].



Obrázek 3: Náhledy testovacích video sekvencí: a) cerveny\_kostel, b) jedovnicka, c) jamu, d) hradecka, e) zavod, f) vitr

## 4 Zhodnocení

Cílem článku, bylo představit aplikaci v prostředí MATLAB, vhodnou pro komprimaci video sekvencí. Pomocí vyvinuté aplikace lze ověřit jednotlivá nastavení kodéru a dekodéru 3D DCT, např. vliv řazení frekvenčních koeficientů při entropickém kódování na velikost kompresního poměru. Lze ukázat, že zásadní vliv na kvalitu rekonstruované video sekvence mají hodnoty prvků použité kvantizační tabulky. Bylo testováno několik typů kvantizačních tabulek. Mezi aspekty, které ovlivňují kompresní poměr a výslednou chybu lze zařadit také výběr použitého řádkování (prokládané/neprokládané), volba barevné reprezentace kódovaných sekvencí, způsob podvzorkování chrominančních signálů a výběr míry kvantování. Všechny tyto parametry lze v prostředí MATLAB libovolně nastavovat.

Pomocí vývojové aplikace v prostředí MATLAB bylo dále ověřeno, že ideální kombinací zmíněných variant je kódovat barevné video sekvence s neprokládaným řádkováním, reprezentované kombinací jasové složky a chrominančních signálů podvzorkovaných způsobem 4:2:0. Nejlepších výsledků dosahovalo kvantování jasové složky pomocí modifikovaných kvantizačních tabulek, odvozených od standardu MPEG-4 a chrominančních složek pomocí tabulek standardu JPEG, určených pro tyto složky.

## Literatura

- [1] Frýza, T., Hanus, S. Optimal Conditions of Video Cube Composition for 3D DCT Encoder and Decoder. In *Proceedings of 9th Conference and Competition Student EEICT 2003*, vol. 2. Brno: FEKT VUT v Brně, duben 2003, str. 52-56. ISBN 80-214-2378-1.
- [2] Rao, K., Yip, P. *Discrete Cosine Transform. Algorithms, Advantages, Applications*. San Diego (USA): Academic Press, Inc., první vydání, leden 1990. 490 stran. ISBN 01-258-0203-X.
- [3] Wallace, G. The JPEG Still Picture Compression Standard. *IEEE Transactions on Consumer Electronics*. Únor 1992, roč. 38, č. 1, str. xviii-xxxiv. ISSN 0098-3063.
- [4] Keith, J. *Video Demystified*. Burlington (USA): Butterworth-Heimann, čtvrté vydání, září 2004. 960 stran. ISBN 07-506-7822-4.