

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta elektrotechniky a komunikačních technologií
Ústav radioelektroniky

Ing. Tomáš Frýza

**KOMPRIMACE OBRAZOVÝCH SIGNÁLŮ POMOCÍ
TRANSFORMACE 3D DCT**

VIDEO SIGNAL COMPRESSION USING 3D DCT TRANSFORM

ZKRÁCENÁ VERZE PH.D. THESIS

Obor: ELEKTRONIKA A SDĚLOVACÍ TECHNIKA
Školitel: Doc. Ing. Stanislav Hanus, CSc.
Oponenti: Ing. Svetozár Ďurovič, DrSc.
Prof. Ing. Dušan Levický, CSc.
Datum obhajoby: 21. března 2006

KLÍČOVÁ SLOVA

Statické snímky, pohyblivé snímky, komprimace obrazových signálů, JPEG, MPEG-1, MPEG-2, DCT, 2D DCT, 3D DCT, kvantování, entropické kódování, MATLAB, C/C++, střih video sekvencí, DSP, TMS320C6000, Code Composer Studio

KEYWORDS

Images, video signals, compression techniques, JPEG, MPEG-1, MPEG-2, DCT, 2D DCT, 3D DCT, quantisation, entropy coding, MATLAB, C/C++, video sequences' cut, DSP, TMS320C6000, Code Composer Studio

DISERTAČNÍ PRÁCE JE ULOŽENA:

Ústav radioelektroniky

Fakulta elektrotechniky a komunikačních technologií

Vysoké učení technické v Brně

Purkyňova 118

612 00 Brno

© Tomáš Frýza, 2007

ISBN 978-80-214-3467-7

ISSN 1213-4198

OBSAH

1 ÚVOD	5
2 TROJROZMĚRNÁ DISKRÉTNÍ KOSINOVÁ TRANSFORMACE	6
2.1 Kvantování 3D DCT koeficientů	7
2.2 Prahování 3D DCT koeficientů	8
2.3 Entropické kódování 3D DCT koeficientů	9
2.4 Dekódování 3D DCT koeficientů	9
2.5 Zlepšení komprimačních vlastností při použití 3D DCT	10
2.6 Efektivní výpočet 3D DCT	15
3 STŘIH VIDEO SEKVENCÍ	17
3.1 Detekce stříhu ve video sekvenci	17
3.2 Předzpracování snímků video sekvence	18
3.3 Modifikace kodéru 3D DCT	19
4 OVĚŘENÍ KOMPRIMAČNÍCH VLASTNOSTÍ 3D DCT	20
4.1 Prostředí Dec-C++	22
4.2 Prostředí Code Composer Studio	23
4.3 Komprimace testovacích video sekvencí	24
4.4 Volba obrazové kvality	25
5 ZÁVĚR	28
LITERATURA	29
PROFESNÍ ŽIVOTOPIS	31
ABSTRAKT	32
ABSTRACT	32

1 ÚVOD

Využití obrazových signálů lze nalézt v mnoha oblastech lidské činnosti. Od bezpečnostních systémů, přes biomedicínské aplikace, až po televizní vysílání s vysokým rozlišením. Jak pro záznam, zpracování, přenos či archivaci video signálů existuje velké množství systémů, které jsou uzpůsobeny pro danou aplikaci.

Neoddělitelnou součástí vývoje digitálních obrazových systémů je také zdokonalování komprimačních metod obrazových informací. Uvažujme nyní případ digitální televize se standardním rozlišením. Obrazová informace je zde soustředěna do snímků o velikosti 720×576 bodů a počet zobrazených snímků za sekundu je roven 25. Parametry libovolného bodu barevného snímku mohou být vyjádřeny pomocí tří dílčích signálů, které jsou kódovány 8 bity. Bitový tok nekompromovaného obrazového signálu tak dosáhne bezmála 250 Mb/s. Systém digitální televize s vysokým rozlišením vyžaduje datový tok ještě několikanásobně vyšší. V oblasti obrazových signálů je proto úkolem komprimačních technik redukovat objem těchto datových toků.

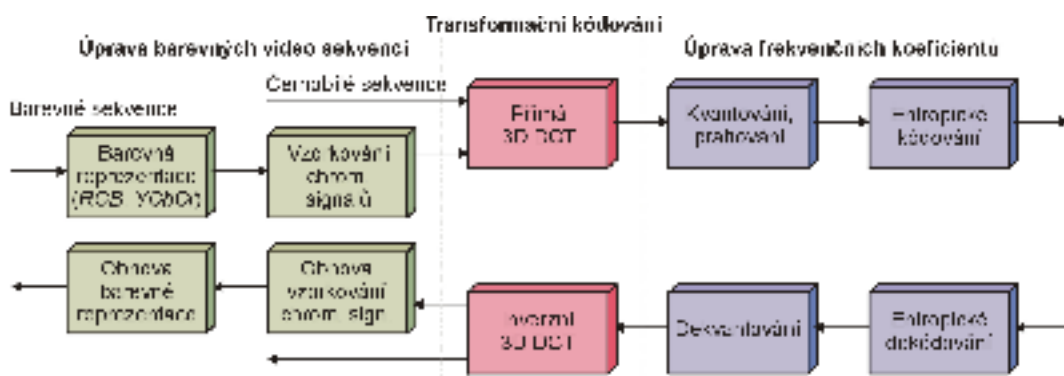
Obecně existují tři základní kroky při komprimaci video signálů. Pro odstranění nadbytečnosti uvnitř snímků je využita podobnost mezi obrazovými body. Tyto kompresní metody se souhrnně nazývají intraframe kódování a jsou základem např. kodéru statických snímků JPEG. Druhým krokem je kódování rozdílu mezi sousedními snímky, tj. redukce nadbytečnosti v čase. V tomto případě hovoříme o mezisnímkovém, tzv. interframe kódování. Mezisnímkové kódování je princip použitý u všech standardních video kodeků, jako jsou H.261, H.263, MPEG-1, MPEG-2, MPEG-4 a také 3D DCT. Posledním krokem k potlačení nadbytečnosti je entropické kódování samotných symbolů s použitím kódování proměnné délky. Trojrozměrná diskrétní kosinová transformace (3D DCT), které je věnována tato disertační práce, slučuje intraframe a interframe kódování do jediného transformačního kódování.

Cílem disertační práce je prostudovat možnosti a využití transformace 3D DCT v oblasti komprimace obrazových dat. Vytyčené cíle lze rozdělit do následujících bodů.

- Prostudování dosavadního vývoje v oblasti komprimace statických i pohyblivých snímků. Důraz bude kladen na metody s diskrétní kosinovou transformací.
- Navržení kódovacího a dekodovacího řetězce, využívajícího 3D DCT. Výchozím bodem bude kodér JPEG, jehož strukturu lze s obměnou aplikovat i na trojrozměrné signály.
- Posouzení faktorů, které ovlivňují vlastnosti 3D DCT. Výběr a naprogramování rychlého algoritmu pro výpočet 3D DCT.
- Rozbor vzniku a potlačení prolínání obsahů video sekvencí při střihu. Prolínání sekvencí je negativní důsledek 3D DCT; budou navrženy komplexní metody jak tento jev odstranit.
- Výše uvedené kroky budou ověřeny experimentálně v prostředí MATLAB a Dev-C++. Konečným krokem bude realizace kodéru 3D DCT s pomocí signálového procesoru.

2 TROJROZMĚRNÁ DISKRÉTNÍ KOSINOVÁ TRANSFORMACE

Tato disertační práce je zaměřena na využití trojrozměrné diskretní kosinové transformace (3D DCT - Three Dimensional Discrete Cosine Transform) v oblasti komprimace pohyblivých snímků. U standardů JPEG a MPEG, které využívají dvourozměrnou verzi DCT, je toto transformační kódování použito výhradně jako intraframe kódování, tj. uvnitř jednotlivých snímků. Přínosem 3D DCT je odstranění nadbytečnosti v čase již samotným transformačním kódováním. Znamená to např. možnost vynechání operací pro hledání pohybových vektorů, které jsou nezbytné u standardu MPEG. Struktura kodéru a dekodéru 3D DCT, jehož blokové schéma je znázorněno na *obrázku 2.1*, vznikla obměnou známé struktury standardu JPEG.



Obrázek 2.1: Blokové schéma kodéru a dekodéru využívající 3D DCT.

Stejně jako u standardu JPEG, ani u 3D DCT se neprovádí kódování najednou pro celé snímky. Ty se nejprve rozdělí na bloky o velikosti $N \times N$ obrazových bodů. Následně jsou odpovídající skupiny bodů od sousedních snímků spojeny do krychlí (tzv. video krychlí) o rozměrech $N \times N \times N$. Tyto video krychle tvoří vstup 3D DCT kodéru a jsou navzájem nezávislé.

Vzhledem k silné korelaci jednotlivých palet barevného prostoru RGB , je výhodné použít prostoru jiného, např. YC_bC_r , který se skládá z jasové složky Y a chrominančních složek C_b, C_r . Výhoda chrominančních složek spočívá především v možnosti redukce počtu jejich vzorků v důsledku snížené citlivosti lidského zraku na jejich změny. Podle požadované barevné věrnosti lze tak chrominanční složky podvzorkovat ve vertikálním (4 : 2 : 2) nebo ve vertikálním i horizontálním směru (4 : 2 : 0).

Pro $N \times N \times N$ obrazových bodů je přímá trojrozměrná diskretní kosinová transformace (3D fDCT - Three Dimensional Forward Discrete Cosine Transform) pro jednu barevnou paletu definovaná rovnicí [14]

$$X^{C^{(II)}}(l, m, n) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sum_{f=0}^{N-1} x(i, j, f) \cdot [C_N^{II}]_{li} \cdot [C_N^{II}]_{mj} \cdot [C_N^{II}]_{nf}, \quad (2.1)$$

kde hodnota $x(i, j, f)$ reprezentuje intenzitu jednoho bodu kódované video sekvence na souřadnicích (i, j, f) sniženou o hodnotu 128, $X^{C(II)}(l, m, n)$ je 3D DCT koeficient na pozici (l, m, n) vypočtený pomocí definičního vztahu DCT-II [14] a čítací indexy nabývají hodnot $l, m, n = 0, 1, \dots, N - 1$.

Počet obdržených frekvenčních koeficientů je stejný jako počet vstupních obrazových bodů, tj. $N \times N \times N$. Účelem transformačního kódování ovšem není redukce počtu koeficientů, ale dekorelace jednotlivých obrazových bodů. Obdobně jako v případě dvourozměrné DCT, i zde reprezentuje koeficient $X^{C(II)}(0, 0, 0)$ stejnosměrnou složku (DC - Direct-current Component) kódované video krychle. Ostatní frekvenční koeficienty označme za střídavé (AC - Alternating-current Component). Přínos komprimační metody tkví v efektivní formě zápisu těchto DC a AC koeficientů. Lze odvodit, že minimální hodnota stejnosměrného koeficientu je po zaokrouhlení rovna -2 897 a maximální hodnota je rovna 2 874. Pro zápis hodnot DC koeficientů je tedy potřeba vyčlenit 13 bitů.

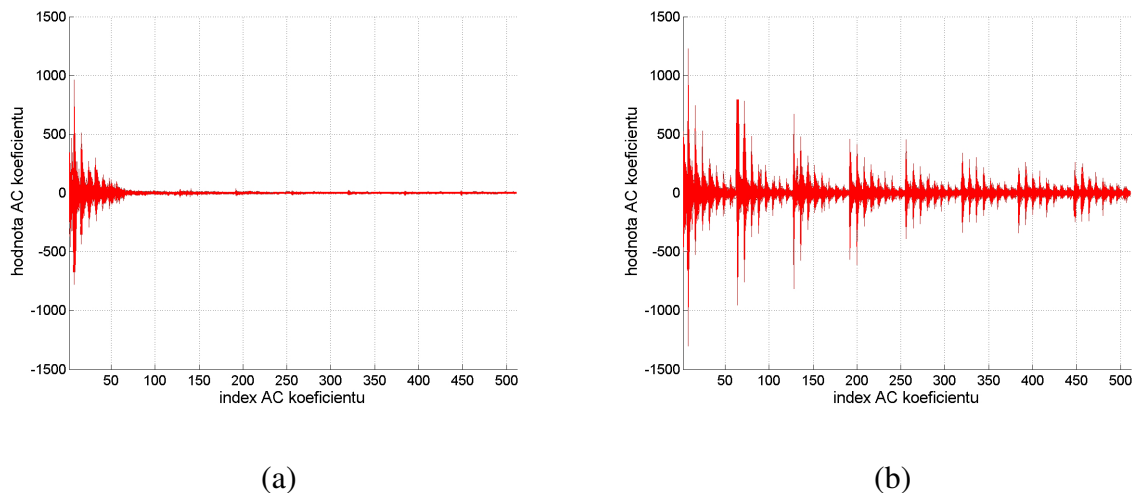
Druhou část frekvenčních koeficientů tvoří střídavé složky. Interval hodnot, ve kterém se AC koeficienty pohybují, lze odvodit z definice přímé diskrétní kosinové transformace. Analýzou funkčních hodnot báze funkce vyplývá, že maximální hodnota AC koeficientů je po zaokrouhlení rovna 2 885 a minimální hodnota rovna -2 885. Reálné sekvence však poukazují na možnost zúžení tohoto intervalu, od -2 048 do 2 047.

Přestože z analyzovaných sekvencí vyplývá, že mezní hodnoty AC koeficientů se přibližně pohybují kolem $\pm 1 600$, většina koeficientů nabývá hodnot podstatně menších. Zpravidla se jedná o hodnoty menší než 100 (v absolutní míře), přičemž většina koeficientů je blízká nebo rovna nule. Také střední hodnota koeficientů je nulová, jak dokládá ukázka na *obrázku 2.2*. Jednoznačná absence vyšších spektrálních složek na prvním obrázku klasifikuje sekvenci jako statickou s malým počtem detailů. Pokud kódovaná sekvence obsahuje dynamické změny, projeví se to ve zvýšené velikosti AC koeficientů a také v jejich pronikání do vyšších frekvenčních oblastí. Tato situace je patrná u sekvence na obrázku druhém.

2.1 KVANTOVÁNÍ 3D DCT KOEFICIENTŮ

Díky transformačnímu kódování lze převést obrazová data na spektrální složky. Velikost a poloha (tj. hodnoty indexů l, m a n) jednotlivých složek charakterizují obsah kódované sekvence a ve své podstatě také důležitost dané informace pro diváka. Frekvenční koeficienty s nižšími hodnotami indexů l, m a n obsahují informaci, jejichž význam je pro diváka podstatný. Koeficienty, reprezentující vyšší prostorové frekvence a tím i detaily v obraze obsahují informaci, na kterou není lidský zrak příliš citlivý. Operace, která ovlivňuje velikost jednotlivých frekvenčních koeficientů podle míry jejich důležitosti se nazývá kvantování [21].

Jelikož velikost výstupního datového toku úzce souvisí s obsahem kódované scény, je z hlediska garantování konstantního datového toku nutné korigovat průběh kódování v 3D DCT kodéru. To se provádí změnou kvantovacích parametrů. Necht'



Obrázek 2.2: Analýza frekvenčních koeficientů. AC koeficienty pro skupinu snímků 0 – 7 sekvencí: a) `cerveny_kostel` a b) `zavod`. Jednotlivé koeficienty jsou načítány po řádcích.

parametr qIt určuje kvalitu dekódované video sekvence a je dán celým číslem z intervalu $\langle 1; 100 \rangle$. Pak lze kvantizační tabulky pro libovolnou hodnotu n upravit podle vztahu [20]

$$Q(l, m, n) = \text{round} \left(\frac{Q_{org}(l, m) \cdot prm + 50}{100} \right), \quad (2.2)$$

kde $Q_{org}(l, m)$ je originální kvantizační tabulka pro dvojrozměrné signály a hodnota prm je definována takto:

$$prm = \begin{cases} \frac{5\,000}{qIt} & \text{pokud } qIt < 50 \\ 2 \cdot (100 - qIt) & \text{pokud } qIt \geq 50. \end{cases} \quad (2.3)$$

Pro správné nastavení kvantizační tabulky je třeba zajistit okrajovou podmínku: pokud platí, že $Q(l, m, n) > 255$, pak je uvažováno $Q(l, m, n) = 255$. Pro nejnižší zvolenou kvalitu dekódované video sekvence ($qIt = 1$), jsou všechny prvky kvantizační tabulky rovny 255 a naopak, pro nejvyšší kvalitu ($qIt = 100$), jsou kvantovací parametry pro všechny frekvenční koeficienty rovny 1.

2.2 PRAHOVÁNÍ 3D DCT KOEFICIENTŮ

Pro zvýšení kompresního poměru je do struktury kodéru 3D DCT zařazen blok prahování. Jak napovídá název, úkolem tohoto bloku je odstranit frekvenční koeficienty, jejichž hodnota je nižší než stanovený práh [3]. Výběr kvantizační tabulky, úroveň kvality (qIt) a míry prahování ($thrs$) jsou hlavními kritérii pro získání dobrých výsledků při komprimaci obrazových dat.

2.3 ENTROPICKÉ KÓDOVÁNÍ 3D DCT KOEFICIENTŮ

Zápis 3D DCT koeficientů je stejně jako u standardu JPEG rozdělen na zápis koeficientů stejnosměrných a střídavých. Hodnota DC koeficientu se kóduje pomocí diferenčního kódování. U AC koeficientů se kódují nenulové koeficienty společně s informací o počtu předcházejících nulových koeficientů (délka běhu). Také pořadí načítání AC koeficientů vychází ze standardu JPEG. Standardní cik-cak načítání však bylo nutné uzpůsobit pro 3D signály. Rozdělíme-li blok frekvenčních koeficientů na matice $X^{C(II)}(l, m, 0)$, $X^{C(II)}(l, m, 1)$ až $X^{C(II)}(l, m, N - 1)$, jsou prvky z těchto matic řazeny způsobem podle standardu JPEG. Postupně je tak načteno všech $N \times N \times N$ frekvenčních koeficientů. Označme tento postup pro $N = 8$ jako “8×cik-cak”.

Z histogramu frekvenčních koeficientů je patrné, že malé hodnoty jsou častější než hodnoty velké. Pro efektivní zápis koeficientů je tedy vhodné použít kratší kódová slova právě pro tyto koeficienty a naopak, hodnoty s menší četností kódovat slovy delšími. Tento způsob zápisu dat se nazývá kódování proměnné délky (VLC - Variable Length Coding). V oblasti kódování video signálů se používají dva typy kódů s proměnnou délkou: Huffmanovo kódování a aritmetické kódování. U kodéru 3D DCT bylo vybráno první z nich.

2.4 DEKÓDOVÁNÍ 3D DCT KOEFICIENTŮ

Obnova původní obrazové informace se provádí na dekódovací straně a je zajištěna posloupností inverzních operací ke straně kódovací. Datový tok, který reprezentuje komprimovanou sekvenci je nejprve podroben entropickému dekódování. Pro správnou činnost tohoto dekodéru je nutná znalost kódových slov Huffmanova kódu. Vzhledem k tomu, že inverzní operace k procesu prahování neexistuje, je dalším blokem dekodéru operace přeskládání frekvenčních koeficientů do správného pořadí a dekvantizace [21].

Dalším krokem je obnova obrazových dat pomocí inverzní trojrozměrné diskrétní kosinové transformace (3D iDCT - Three Dimensional Inverse Discrete Cosine Transform). Ta je, pro jednu barevnou paletu, popsána vztahem [14]

$$\hat{x}(i, j, f) = \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{X}^{C(II)}(l, m, n) \cdot [C_N^{II}]_{li} \cdot [C_N^{II}]_{mj} \cdot [C_N^{II}]_{nf}. \quad (2.4)$$

Pokud byla kódována barevná sekvence a na kódovací straně byla změněna její barevná reprezentace, příp. způsob vzorkování chominančních signálů, obsahuje dekódovací řetězec také obnovu formátu 4 : 4 : 4 a barevné palety *RGB*. Posledním krokem je přesun intervalu hodnot obrazových bodů do původních mezí $\langle 0; 255 \rangle$ přičtením konstanty 128.

2.5 ZLEPŠENÍ KOMPRIMAČNÍCH VLASTNOSTÍ PŘI POUŽITÍ 3D DCT

Snahou při kódování video signálů je dosažení velkého kompresního poměru a žádné, nebo přijatelné odchylky v rekonstruované sekvenci. Zaměříme-li se na kódovací řetězec využívající 3D DCT, lze faktory které ovlivňují kompresní poměr na straně jedné a kvalitu rekonstruované sekvence na straně druhé, přesně identifikovat.

Hlavní ukazatel libovolné komprimační metody je její schopnost redukce datového toku. V kódovacím řetězci 3D DCT dochází k určité míře redukce již při podvzorkování chrominančních signálů (při barevné reprezentaci YC_bC_r), dále při kvantování, prahování a při entropickém kódování. Zbývající operace, tj. výběr typu řádkování, převod barevné sekvence z reprezentace RGB na reprezentaci obsahující jasovou složku a barvonosnou informaci a transformační kódování nelze přímo zařadit mezi operace pro snížení datového toku. Z hlediska počtu vzorků a velikosti jejich bitového vyjádření se totiž nejedná o kompresní postupy. Paradoxně, právě tyto operace se ve svém dopadu podílí na výši kompresního poměru nejvíce.

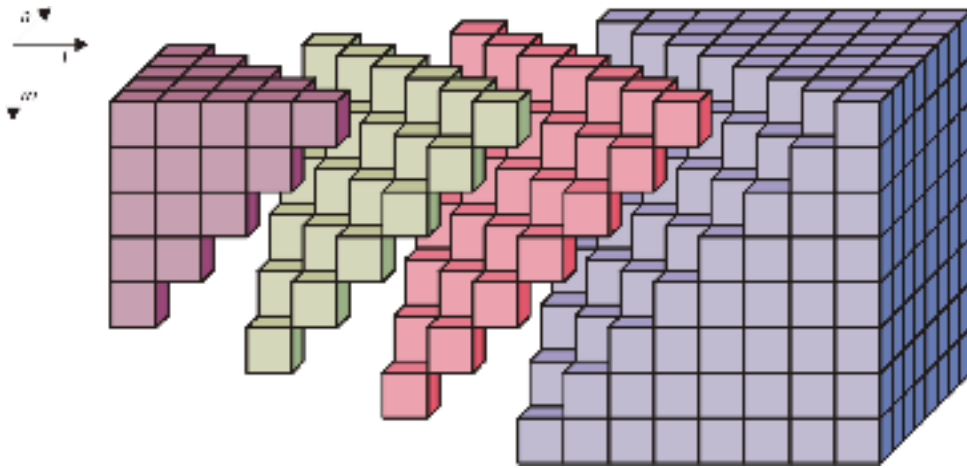
Druhým hlediskem komprimačních metod je velikost chyby, kterou do rekonstruované sekvence vnáší. Některé části kodéru s 3D DCT patří mezi operace ztrátové a zdroje těchto chyb jsou následující. Prvním je podvzorkování chrominančních signálů u formátů $4:2:2$ a $4:2:0$, kdy dochází k vynechání části vzorků nesoucí informaci o barvě ve video sekvenci. Vliv na výslednou kvalitu má i způsob obnovy vynechaných vzorků na dekódovací straně. Z hlediska výrazné degradace kvality rekonstruované sekvence jsou nejdůležitějšími bloky kvantování a prahování. U všech ostatních operací, tj. u převodu mezi barevnými reprezentacemi, u transformačního a entropického kódování nedochází ke vnášení chyb, protože se jedná o bezztrátové operace.

2.5.1 Huffmanův kód pro 3D DCT koeficienty

Vraťme se nyní k pořadí načítání frekvenčních koeficientů cik-cak způsobem, který byl popsán dříve. Koeficienty $X^{C(II)}(l, m, n)$ jsou v rovinách pro $n = 0, 1, \dots, N - 1$ načítány stejně jako u standardu JPEG. Proto byl tento způsob označen $8 \times$ cik-cak.

Pro řazení 3D DCT koeficientů lze však odvodit velké množství odlišných postupů. U standardu JPEG jsou koeficienty postupně načítány v liniích, které splňují podmínku $l = m = 0, 1, \dots, 2(N - 1)$. Rozšířením tohoto způsobu na trojrozměrný prostor lze definovat řazení koeficientů také pro 3D DCT. V tomto případě lze podmínku pro výběr roviny v krychli frekvenčních koeficientů zapsat jako $l = m = n = 0, 1, \dots, 3(N - 1)$, jak ilustruje *obrázek 2.3*. Krychle je zde rozdělena na řezy, kolmé k tělesové úhlopříčce. Proto označme tento způsob řazení jako “ \perp řez”. Takto přerovnané koeficienty je nutné následně kódovat Huffmanovým kódem.

Jak bylo odvozeno výše, pro kódování hodnot DC koeficientů je nezbytné vyčlenit až 13 bitů na symbol a pro AC koeficienty stačí o jeden bit méně. Uvážíme-li, že DC i AC koeficienty mohou nabývat kladných i záporných hodnot, je pro DC koeficienty vyčleněn interval hodnot od -4 095 do 4 095 a pro AC koeficienty interval od -2 047 do 2 047. Velikost celého intervalu je pro potřeby entropického kódování rozdělena



Obrázek 2.3: Pořadí načítání frekvenčních koeficientů v rovinách \perp na tělesovou úhlopříčku krychle frekvenčních koeficientů.

na dílčí intervaly s rostoucím počtem prvků [10]. Každému intervalu je pak přiřazeno jedno kódové slovo podle četnosti výskytu koeficientů. Vzhledem k tomu, že délka běhu je závislá na pořadí načítání frekvenčních koeficientů, byly vyvinuty odlišné slovníky Huffmanova kódu jak pro načítání 8×cik-cak tak i pro načítání uzpůsobené AC koeficientům s řazením \perp řez.

Ze struktury kodéru 3D DCT plyne, že koeficienty zastupující stejnosměrnou složku a složky střídavé jsou kódovány pomocí odlišných Huffmanových tabulek. Pro každý způsob řazení frekvenčních koeficientů byly proto vytvořeny dva slovníky. Počet prvků v tabulkách Huffmanova kódu pro DC koeficienty je neměnný a obsahuje 13 kódových slov (12 intervalů + hodnota nula). Počet prvků ve slovníku Huffmanova kódu pro AC koeficienty závisí na maximální délce běhu. Pro všech 511 AC koeficientů je maximální délka běhu rovna 510, tj. počet prvků ve slovníku je 5 622. Tak obsáhlý slovník klade nezanedbatelné nároky na paměť.

Počty bitů jednotlivých kódových slov pro zápis stejnosměrných koeficientů jsou vypsány v *tabulce 2.1* a počty bitů pro střídavé koeficienty jsou uvedeny v *tabulce 2.2*. První sloupec udává velikost délky běhu. U DC koeficientů není z pochopitelných důvodů tento parametr použit. Interval hodnot identifikuje jednotlivá kódová slova a skládá se ze dvou subintervalů, symetrických kolem nuly. Vyjimku přirozeně tvoří samotná nula. Délka slova popisuje počet bitů navržených kódových slov pro dva způsoby řazení frekvenčních koeficientů a délka následujícího symbolu udává počet bitů, kterými je kódována hodnota samotného frekvenčního koeficientu.

Délky kódových slov pro stejnosměrné koeficienty nabývají stejných hodnot pro oba typy řazení frekvenčních koeficientů. Ty se pohybují od 2 do 8 bitů na symbol. Je to pochopitelné, protože v obou případech jsou DC koeficienty načítány jako první z množiny všech koeficientů a také způsob jejich zpracování, tj. diferenční kódování je totožný. Odlišná kódová slova vykazují slovníky AC koeficientů. Pověsimně si počtu bitů pro zakódování koeficientů s malou hodnotou délky běhu. Nižší

Tabulka 2.1: Délka Huffmanova kódu pro DC koeficienty (nezkrácená verze slovníku)

Délka běhu	Interval hodnot		Délka slova [b]		Délka následujícího symbolu [b]
			8×cik-cak	└řez	
–	0		4	4	0
–	1(-1)	1(-1)	3	3	1
–	2(-3)	3(-2)	2	2	2
–	4(-7)	7(-4)	3	3	3
–	8(-15)	15(-8)	3	3	4
–	16(-31)	31(-16)	3	3	5
–	32(-63)	63(-32)	4	4	6
–	64(-127)	127(-64)	4	4	7
–	128(-255)	255(-128)	5	5	8
–	256(-511)	511(-256)	6	6	9
–	512(-1 023)	1 023(-512)	7	7	10
–	1 024(-2 047)	2 047(-1 024)	8	8	11
–	2 048(-4 095)	4 095(-2 048)	8	8	12

hodnota počtu bitů jednotlivých slov koresponduje s větší pravděpodobností výskytu slov z tohoto intervalu. Z toho plyne, že způsob └řez vykazuje častější koeficienty s malou délkou běhu. V první skupině kódových slov, tj. ve skupině s délkou běhu rovné nule má druhý způsob u dvou slov menší délku a ve druhé skupině je to již u pěti kódových slov. To ukazuje na efektivnější uspořádání koeficientů pro entropické kódování. Ideální situací je způsob, kdy jsou všechny nenulové koeficienty zařazeny bezprostředně za sebou. V tomto směru nedosahuje způsob odvozený od zápisu 2D DCT koeficientů značných úspěchů.

Jak bylo uvedeno, celkový počet kódových slov ve slovníku Huffmanova kódu při načítání všech 511 AC koeficientů je neúnosně dlouhý. Řešením by mohlo být rozdělení těchto koeficientů do menších skupin a ty kódovat odděleně. Necht' je skupina 512 frekvenčních koeficientů rozdělena na osm podskupin o 64 prvcích. První koeficient z každé skupiny zde představuje jakýsi pseudo stejnosměrný koeficient a je kódován pomocí Huffmanových tabulek pro DC koeficienty. Pravý DC koeficient, tedy $X^{C(II)}(0, 0, 0)$ je navíc kódován diferenčním kódováním. Ostatní AC koeficienty jsou kódovány standardním způsobem společně s délkou běhu v definovaném pořadí. Uvažujeme-li maximální délku běhu jen 62, jedná se mimochodem o způsob použitý pro zápis 2D DCT koeficientů, bude slovník AC koeficientů obsahovat slov podstatně méně, $11 \times (62 + 1) + 1 = 694$. Požadavky na velikost paměti pro uložení takového slovníku jsou výrazně nižší.

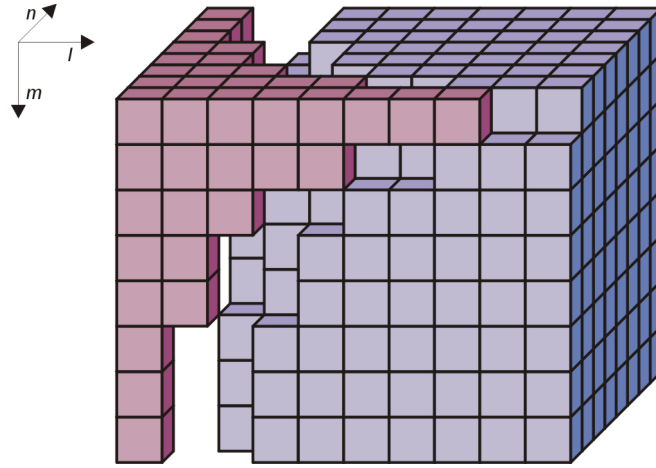
2.5.2 Vnitřní a vnější region 3D DCT koeficientů

Odlišný přístup ke kvantování frekvenčních koeficientů navrhuje Lee [13]. Ten vychází z vlastností koeficientů, resp. z jejich hodnot kolem hlavních os v krychli trans-

Tabulka 2.2: Délka Huffmanova kódu pro AC koeficienty (nezkrácená verze slovníku)

Délka běhu	Interval hodnot	Délka slova [b]		Délka následujícího symbolu [b]
		8×cik-cak	⊥řez	
0	EOB	6	6	0
0	1(-1) 1(-1)	2	2	1
0	2(-3) 3(-2)	3	3	2
0	4(-7) 7(-4)	3	3	3
0	8(-15) 15(-8)	4	4	4
0	16(-31) 31(-16)	6	5	5
0	32(-63) 63(-32)	7	7	6
0	64(-127) 127(-64)	9	8	7
0	128(-255) 255(-128)	11	11	8
0	256(-511) 511(-256)	15	15	9
0	512(-1 023) 1 023(-512)	26	26	10
0	1 024(-2 047) 2 047(-1 024)	27	27	11
1	1(-1) 1(-1)	3	3	1
1	2(-3) 3(-2)	5	5	2
1	4(-7) 7(-4)	7	7	3
1	8(-15) 15(-8)	9	8	4
1	16(-31) 31(-16)	11	10	5
1	32(-63) 63(-32)	13	12	6
1	64(-127) 127(-64)	15	14	7
1	128(-255) 255(-128)	17	17	8
1	256(-511) 511(-256)	21	20	9
1	512(-1 023) 1 023(-512)	28	28	10
1	1 024(-2 047) 2 047(-1 024)	29	29	11
⋮	⋮	⋮	⋮	⋮
510	1(-1) 1(-1)	4 731	5 080	1
510	2(-3) 3(-2)	4 732	5 081	2
510	4(-7) 7(-4)	4 733	5 082	3
510	8(-15) 15(-8)	4 734	5 083	4
510	16(-31) 31(-16)	4 735	5 084	5
510	32(-63) 63(-32)	4 736	5 085	6
510	64(-127) 127(-64)	4 737	5 086	7
510	128(-255) 255(-128)	4 738	5 087	8
510	256(-511) 511(-256)	4 739	5 088	9
510	512(-1 023) 1 023(-512)	4 740	5 089	10
510	1 024(-2 047) 2 047(-1 024)	4 740	5 089	11

formovaných obrazových dat. Necht' oblast méně důležitých koeficientů je soustředěna do hyperboloidu a jeho doplňkem je oblast koeficientů s vyšší důležitostí, jak je naznačeno na *obrázku 2.4*. Tento doplněk je pro kladnou konstantu C definován vztahem



Obrázek 2.4: Rozdělení krychle frekvenčních koeficientů na oblasti vnitřní a vnější podle důležitosti jednotlivých koeficientů.

$$(l + 1)(m + 1)(n + 1) \leq C. \quad (2.5)$$

Koeficienty ležící v oblasti doplňku k hyperboloidu nazvěme jako vnitřní a koeficienty patřící do hyperboloidu jako vnější. Velikosti obou oblastí lze měnit parametrem C .

Pro odvození hodnot kvantizačních parametrů je navržen vztah, který zohledňuje zda se daný koeficient nachází ve vnitřní nebo vnější oblasti. Možností jak tento vztah definovat je více. V [13] je růst kvantizačních parametrů realizován pomocí funkce $\exp(\cdot)$. Tyto vztahy jsou definovány následovně

$$Q(l, m, n) = \text{round} \left\{ \begin{array}{l} A_{inn} \cdot \left\{ 1 - \frac{\exp[-\beta_{inn} \cdot (l + 1)(m + 1)(n + 1)]}{\exp(-\beta_{inn})} \right\} + 1 \\ \quad \text{pokud } (l + 1)(m + 1)(n + 1) \leq C, \\ \\ A_{out} \cdot \{1 - \exp[-\beta_{out} \cdot (l + 1)(m + 1)(n + 1)]\} + 1 \\ \quad \text{pokud } (l + 1)(m + 1)(n + 1) > C. \end{array} \right. \quad (2.6)$$

Pro odvození jednotlivých parametrů kvantizační tabulky je nutné specifikovat hodnoty parametrů $A_{inn,out}$ a $\beta_{inn,out}$. Konstanty s indexy *inn* se vztahují ke koeficientům ležícím v oblasti s důležitými koeficienty a konstanty s indexy *out* se vztahují ke koeficientům vnějším. Maximální hodnoty kvantizačních parametrů pro vnitřní a vnější koeficienty udávají amplitudy $A_{inn,out} \geq 0$. Rychlost růstu hodnot v kvantizační tabulce a jejich přiblížení k maximální hodnotě $A_{inn,out} + 1$ je ovlivněna parametry $\beta_{inn,out} \geq 0$.

Způsob řazení frekvenčních koeficientů vychází z velikostí kvantizačních parametrů a z příslušnosti k vnitřní nebo vnější oblasti. Princip řazení je následující. Jednotlivé kvantizační parametry jsou v obou regionech seřazeny vzestupně podle své velikosti. Řazení frekvenčních koeficientů probíhá nejprve ve vnitřním regionu a první v pořadí je vždy stejnosměrný koeficient. Následují koeficienty s nejnižší hodnotou kvantizačního parametru až po ty s největšími. Teprve poté jsou obdobným způsobem seřazeny koeficienty z vnějšího regionu.

Byly odvozeny celkem tři typy kvantizačních tabulek, které odpovídají různým stupňům komprimace, od nejnižší po nejvyšší. Počet koeficientů ve vnitřní oblasti u všech definovaných kvantizačních tabulek koresponduje s parametry $C = 12$ a $A_{inn} = A_{out} = 254$. Pro kvantizační tabulku, která by umožňovala dosažení nejvyššího kompresního poměru na úkor zhoršené kvality některých typů rekonstruovaných sekvencí, byla pro parametr β_{inn} zvolena hodnota 0,09 a pro β_{out} hodnota 0,011. Kvantizační tabulka pro kvalitativně lepší výsledky splňuje podmínku $\beta_{inn} = 0,05$ a $\beta_{out} = 0,05$. Nejlepší kvality dosahuje komprimace pomocí třetí kvantizační tabulky s nejpomalejším růstem kvantizačních parametrů ve vnitřní oblasti. Parametr β_{inn} byl podle hodnot kompresního poměru a chyby v tomto případě zvolen 0,02 a $\beta_{out} = 0,002$.

2.6 EFEKTIVNÍ VÝPOČET 3D DCT

Definiční vztahy pro výpočet přímé (2.1) a zpětné (2.4) trojrozměrné diskrétní kosinové transformace nejsou i přes dnešní výkonné hardwarové prostředky vhodné pro použití v systémech, pracujících v reálném čase. Jsou totiž příliš početně náročné. K výpočtu jednoho frekvenčního koeficientu je potřeba provést $N^3 - 1$ operací součtů a $3N^3$ operací součinů. Počet frekvenčních koeficientů při transformaci jedné video krychle je přitom $N \times N \times N$. Uvažujeme-li dále sekvenci s rozlišením 720×576 obrazových bodů (systém PAL), snímková frekvence necht' je pro jednoduchost 24 snímků za sekundu a velikost DCT je rovna $N = 8$, pak počet video krychlí nutných pro kódování sekvence v šedotónové paletě s dobou trvání jedné sekundy je 19 440. Z toho vyplývá, že pro provedení transformačního kódování jedné sekundy vstupních dat systému PAL je zapotřebí 5 086 126 080 součtových operací a 15 288 238 080 operací násobení. V případě barevné sekvence je tento počet ještě třikrát vyšší.

Pro ilustraci byl proveden výpočet 3D fDCT, doplněný kvantováním frekvenčních koeficientů. Vstupní sekvence měla rozlišení 720×576 bodů a obsahovala pouze černobílou informaci. Výpočet byl realizován na počítači typu PC s procesorem AMD Athlon XP 2700+ s taktovací frekvencí 2,17 GHz a s velikostí operační paměti 512 MB. V kódovacím řetězci bylo vynecháno entropické kódování a také režijní operace s načítáním vstupních dat a plněním výstupních vyrovnávacích pamětí. Za těchto podmínek, program napsaný v jazyce C, pracoval více než 38 minut!

Praktické využití 3D DCT při komprimaci video signálů je podmíněno vývojem algoritmů pro rychlejší výpočet. Počet operací násobení lze snížit již pouhou modifikací zápisu trojrozměrné transformace následujícím způsobem

$$X^{C(II)}(l, m, n) = \sum_{f=0}^{N-1} [C_N^{II}]_{nf} \cdot \sum_{j=0}^{N-1} [C_N^{II}]_{mj} \cdot \sum_{i=0}^{N-1} x(i, j, f) \cdot [C_N^{II}]_{li}, \quad (2.7)$$

kde vypočítat jeden 3D DCT koeficient znamená realizovat $N^3 + N^2 + N$ různých násobení. Počet součtů zůstal ovšem nezměněn. Úprava výpočtu tímto způsobem je pouhou kosmetickou změnou definičního vztahu a přes dílčí snížení počtu násobení, z hlediska komprimace video signálů nic neřeší. V historii byly proto hledány jiné postupy, jak výpočet 3D DCT zefektivnit. Metoda podle [14] využívá substituci diskrétní Fourierovy transformace (DFT - Discrete Fourier Transform), kdy jsou 3D DCT koeficienty vypočteny pomocí reálných částí Fourierových koeficientů. Transformace 3D DCT je zde rozdělena na posloupnost tří jednorozměrných diskrétních kosinových transformací, čímž se nejen sníží počet nutných operací, ale výrazným způsobem se redukuje také paměťové nároky při ukládání hodnot mezivýpočtů. Vhodnou úpravou soustavy rovnic pro výpočet osmibodové DFT lze počet nutných matematických operací zmenšit na 29 součtů a 5 násobení. Pro výpočet N^3 3D DCT koeficientů (kde $N = 8$) tak stačí $3 \times 29 \times 8 \times 8 = 5\,568$ součtů a $3 \times 5 \times 8 \times 8 = 960$ násobení.

Pro video sekvenci v systému PAL a dobou trvání jedné sekundy (resp. 24 snímků) je nutné provést 108 241 920 operací součtů a 18 662 400 operací součinů za sekundu. Ve srovnání s početní náročností definičních vztahů se u operací součtů jedná o pokles více než třicetinásobný a u násobení se počet operací snížil téměř 820×! Přehledné srovnání počtu operací pro výpočet jednoho frekvenčního koeficientu, pro jednu video krychli a pro transformaci jedné sekundy černobílé video sekvence v systému PAL pro tři různé způsoby výpočtu 3D DCT koeficientů je znázorněno v *tabulce 2.3*. Počty operací pro barevné sekvence mohou v závislosti na formátu chrominančních signálů nabývat až třikrát větších hodnot.

Tabulka 2.3: Počty operací součtů a součinů nutných pro výpočet přímé nebo zpětné 3D DCT pro tři různé algoritmy.

Počet koeficientů	Početní operace	Definice 3D DCT (2.1), (2.4)	Modifikovaná definice (2.7)	Pomocí DFT
Jeden 3D DCT koeficient	+ ×	511 1 536	511 584	– –
Videokrychle (512 koeficientů)	+ ×	261 632 786 432	261 632 299 008	5 568 960
Sekvence o 24 snímcích	+ ×	5 086 126 080 15 288 238 080	5 086 126 080 5 812 715 520	108 241 920 18 662 400

Redukce počtu nutných operací napovídá, že způsob výpočtu 3D DCT koeficientů pomocí DFT bude mnohem efektivnější než při použití definičních vztahů. Proved' me nyní výpočet pomocí navrženého postupu, naprogramovaném v jazyce C a odladěném v prostředí Dev-C++. Výpočet přímé trojrozměrné diskrétní kosinové transformace a operace kvantování frekvenčních koeficientů jednosekundové sekvence v rozlišení PAL s jednou barevnou paletou, trval bez dalších operací (prahování, entropické kódování, apod.) 1,28 s. Připomeňme, že s použitím definičních vztahů přesahovala tato hodnota 38 min, což je bezmála $1\,800\times$ více. Dosažené časy naznačují možnost kódování video signálů pomocí 3D DCT v reálném čase.

3 STŘIH VIDEO SEKVENCÍ

Podstatou využití trojrozměrné diskrétní kosinové transformace je současné kódování skupiny N snímků. Z pohledu kodéru jsou množiny těchto snímků navzájem nezávislé. V případě, že každá skupina je reprezentována tématicky shodnou video sekvencí, komprimační vlastnosti 3D DCT (tj. výše kompresního poměru a výsledné chyby) jsou velmi dobré. Uvažujme nyní situaci, kdy vstupní sekvence je tvořena dvěma nebo i více obsahově odlišnými scénami. Samotný kodér však není schopen tuto skutečnost identifikovat a vstupní data transformuje do frekvenční oblasti bez ohledu na jejich obsah. Dosažitelný kompresní poměr se tak sníží. Při vyšší míře kvantování, příp. prahování frekvenčních koeficientů dojde také v rekonstruované sekvenci k patrnému prolínání obsahů dílčích sekvencí. Ten se nejvíce projeví v okolí stříhu [4].

Odstranění prolínání mezi video sekvencemi nelze provést bez vnějšího zásahu do kódovacího řetězce 3D DCT. Teoreticky existují dva odlišné postupy jak takového odstranění dosáhnout. První vychází z myšlenky ovlivnění vstupních dat, kdy operace kodéru a dekodéru 3D DCT zůstávají neměnné. Druhý způsob spočívá v modifikaci samotného kompresního algoritmu, kdy se snímky dílčích sekvencí transformují odděleně. Oba přístupy spojuje společný předpoklad: ke své činnosti potřebují znát přesnou polohu stříhu v kódované posloupnosti.

3.1 DETEKCE STŘIHU VE VIDEO SEKVENCI

Principiálně lze skokovou změnu obrazové informace detekovat již v časové oblasti, kdy se porovnávají obrazové body sousedních snímků, nebo v oblasti frekvenční, kdy se porovnávají hodnoty 2D DCT nebo Fourierových koeficientů.

Výhoda detekce v časové oblasti tkví v nulových požadavcích na přídatnou paměť pro porovnávané snímky - skupina všech kódovaných snímků je již uložena ve vyrovnávací paměti kodéru. Hlavním přínosem detekce ve frekvenční oblasti je možnost porovnávání pouze malé skupiny frekvenčních koeficientů. Ty totiž obsahují dostatek informace o charakteru celého snímku. Pokud budou pro detekci stříhu použity 2D DFT koeficienty, je nutné včlenit do kódovacího řetězce blok Fourierovy transformace. Zvyšuje se tak složitost kodéru a současně se zvyšují také paměťové nároky pro uchování těchto koeficientů.

V obou případech je dále možné detekovat stříh buď na úrovni celých snímků nebo pro každý blok $N \times N \times N$ obrazových bodů (koeficientů) zvlášť. Druhá možnost přímo vyplývá z nezávislosti bloků při kódování pomocí 3D DCT. Výhoda prvního přístupu je v posouzení celého obsahu snímků, kdy je pozice stříhu již neměnná pro všechny kódované video krychle dané skupiny N snímků. Nevýhodou jsou značné početní a paměťové nároky zejména při detekci ve frekvenční oblasti, kdy je potřeba transformovat celé snímky mnohdy s vysokým rozlišením. Nezávislé posuzování přítomnosti stříhu v jednotlivých kódovaných segmentech může na druhou stranu citlivěji reagovat na změnu intenzity obrazových bodů v malých oblastech snímků. Dopad tohoto postupu se projeví v nutnosti začlenit pozici potencionálního stříhu v záhlaví každého kódovaného bloku dat pro správnou funkci dekodéru.

Všechny uvedené možnosti při detekci stříhu mají jeden společný aspekt. Změny v množinách obrazových bodů či frekvenčních koeficientů jsou vždy kvalifikovány vhodnými parametry. Mezi tyto parametry patří např. střední kvadratická chyba, střední absolutní chyba, či normovaná vzdálenost. Hodnoty těchto parametrů jsou porovnány s prahovými údaji a následně je rozhodnuto o eventuální přítomnosti stříhu. Tato informace je sdělena kodéru 3D DCT. Lze ukázat, že pro spolehlivou detekci stříhu ve video sekvenci stačí analyzovat přibližně čtvrtinové rozlišení kódovaných snímků.

3.2 PŘEDZPRACOVÁNÍ SNÍMKŮ VIDEO SEKVENCE

První metodou jak potlačit prolínání dvou sekvencí je předzpracování vstupních dat bezprostředně před transformačním kódováním. Jednoduchou metodou, jak toho docílit je snížit hodnotu jasu jednotlivých snímků v závislosti na vzdálenosti od stříhu. Obrazové body se váhují konstantou $\nu(f)$, která zůstává konstantní pro obrazové body uvnitř každého snímku. Vhodným nastavením těchto korekčních faktorů lze tak potlačit vliv prolínání mezi snímky uvnitř kódované video krychle [2]. Matematicky lze popsanou operaci vyjádřit následujícím vztahem

$$\mathbf{x}_{pre}(f) = \mathbf{x}(f) \cdot \nu(f), \quad (3.1)$$

kde $\mathbf{x}_{pre}(f)$ představuje hodnoty intenzity obrazových bodů upraveného a $\mathbf{x}(f)$ původního snímku, $f = 0, 1, \dots, N - 1$. Hodnoty vektoru $\nu(f)$ lze popsat libovolnou funkcí, např. harmonickou, exponenciální, lineární apod. Analýzou jednotlivých funkcí nebyla zjištěna významnější závislost komprimačních vlastností kodéru na charakteru váhovacích funkcí [5].

Následující proces transformačního kódování pak probíhá bezezměn, stejně jako zbývající operace kodéru 3D DCT. Obnova původních hodnot obrazových bodů se provádí na dekódovací straně bezprostředně po bloku 3D iDCT. Musí být přítom použity shodné korekční faktory jako na kódovací straně.

Nespornou výhodou této metody je její jednoduchost a snadná implementace do kódovacího řetězce. Na druhé straně její efektivita není velká. Zvláště u vysokých kompresních poměrů a také pokud stříh nenastává uprostřed skupiny N snímků, tato metoda selhává.

3.3 MODIFIKACE KODÉRU 3D DCT

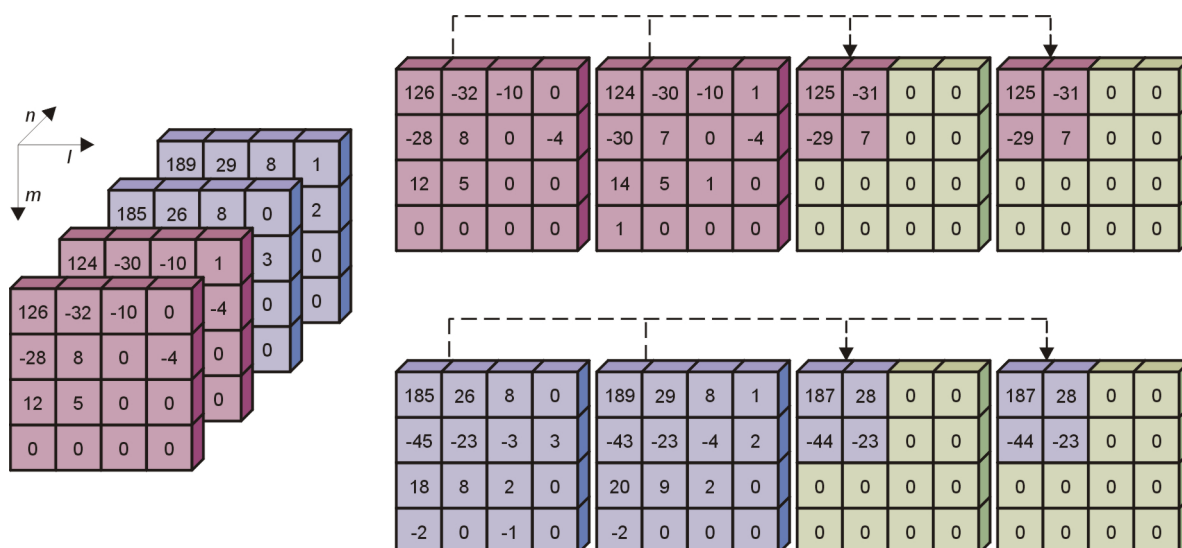
Sofistikovanější metodou jak docílit potlačení nežádoucího prolínání mezi jednotlivými snímky je modifikace samotného kodéru. V principu to znamená, že transformační kódování vstupních dat je prováděno běžným způsobem pouze v případě, kdy skupina N snímků neobsahuje střih. V opačném případě dojde k nepatrně odlišnému způsobu kódování. Modifikace se týká především způsobu výpočtu frekvenčních koeficientů z posloupnosti vstupních dat. Nepatrných odlišností se dočká také blok kvantování a prahování.

Celý princip je založen na separaci dílčích scén v rámci jedné video krychle a v jejich oddělené transformaci. V předešlém textu byla představena myšlenka rozdělení početně složitého výpočtu 3D DCT na posloupnost mnohem jednodušších jednorozměrných transformací. Tohoto přístupu lze s úspěchem využít také pro minimalizaci nárůstu početní složitosti nové verze kodéru. Provádíme-li totiž první dvojici transformací ve směru os i a j , nemůže k žádnému prolínání dojít. Jednotlivé skupiny takto kódovaných matic totiž reprezentují intraframe kódování jednotlivých snímků. Modifikaci kódovacího řetězce vyžaduje pouze poslední z jednorozměrných transformací, tedy ve směru osy f .

Nechť skupina kódovaných snímků obsahuje pouze jeden střih. Skupina N snímků je tak rozdělena na dvě tématické skupiny. Každá z nich bude podrobena zbývajícím třetím 1D DCT odděleně [5]. Vzhledem k situaci, že odvozené početní úkony vyžadují vždy skupinu osmi vstupních vzorků, je nutné 2D DCT koeficienty každé tématické sekvence doplnit daty, jejichž vliv na výslednou kvalitu bude minimální. Navržená metoda vychází z myšlenky realizovat chybějící data aritmetickým průměrem existujících dat určitého počtu koeficientů. Zbývajícím koeficientům budou mít nulovou hodnotu. Všechny dodatečně vytvořené matice koeficientů o velikosti $N \times N$ jsou totožné. Schématické znázornění této operace je uvedeno na *obrázku 3.1*.

Vliv počtu uvažovaných 2D DCT koeficientů, ze kterých se průměrováním odvozují data pro doplnění transformovaných koeficientů byl experimentálně ověřen. Při doplnění kódovaných množin 2D DCT koeficientů pouze nulovými hodnotami dojde ke snížení celkového jasu v rekonstruované sekvenci. Tento důsledek je patrný ve střední části *obrázku 3.2*, kde jsou zobrazeny dva snímky rekonstruované sekvence v nejbližším okolí střihu. V horní části obrázku jsou znázorněny identické snímky bez použití metod pro redukci prolínání. Při použití malého či žádného počtu průměrovaných 2D DCT koeficientů je kromě snižování celkového jasu v sekvenci také patrná přítomnost blokových artefaktů. Výskyt těchto jevů je vázán na malou množinu frekvenčních koeficientů.

Spodní část *obrázku 3.2* zobrazuje rekonstruované snímky při použití průměrování všech 2D DCT koeficientů. Stejně jako v případě s nulovým počtem dodatečných 2D DCT koeficientů, dochází i zde k absolutnímu odstranění prolínání, ale navíc zde nejsou patrné žádné přídavné blokové artefakty, či rušivé kolísání celkového jasu v obraze.



Obrázek 3.1: Princip separace 2D DCT koeficientů odlišných sekvencí uvnitř kódované video krychle. Určitý počet chybějících dat je nahrazen průměrnou hodnotou koeficientů stávajících.

Z uvedených poznatků byla pro strukturu kodéru a dekodéru 3D DCT zvolena varianta s průměrováním množiny 3×3 doplňkových koeficientů. Subjektivním hodnocením všech testovacích sekvencí byla za těchto podmínek zjištěna dostatečná míra potlačení rušivého prolínání kódovaných scén.

4 OVĚŘENÍ KOMPRIMAČNÍCH VLASTNOSTÍ 3D DCT

Cílem disertační práce bylo ověřit komprimační vlastnosti trojrozměrné diskrétní kosinové transformace. Pro testování možností 3D DCT byly postupně vytvořeny softwarové prostředky v prostředí MATLAB, Dev-C++ a Code Composer Studio. Základní popis zpracování signálu byl proveden v prostředí MATLAB. V prostředí Dev-C++ byla navržená struktura kódovacího řetězce přeprogramována pomocí jazyka C a následně ověřena na signálovém procesoru TMS320C6711 v prostředí Code Composer Studio.

Struktura aplikace pro ověření vlastností 3D DCT při komprimaci obrazových dat vychází ze snahy o jednoduchou variabilitu kódovacího řetězce. Byla proto nejprve zvolena struktura v prostředí MATLAB, která není vhodná pro kódování a dekódování dat v reálném čase. Nesplnění požadavku na zpracování v reálném čase není v této fázi chápáno jako nevýhoda, protože stěžejním cílem je nejprve výběr takové struktury, která umožňuje dosáhnout nejlepších výsledků s ohledem na velikost kompresního toku a výsledné chyby. Uvedeným způsobem byla ověřena všechna nastavení kodéru 3D DCT, např. volba typu řádkování, pořadí načítání frekvenčních koeficientů, či volba kvantizačních tabulek.



(a)



(b)

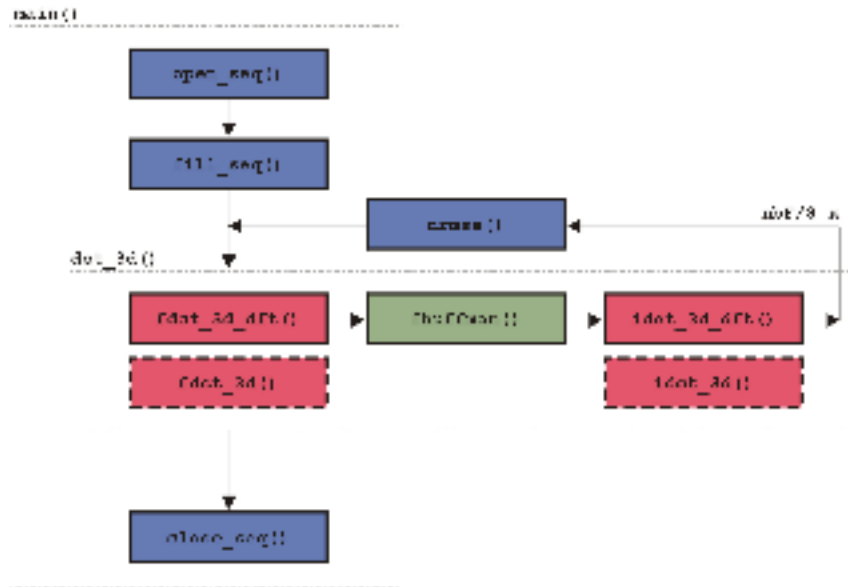


(c)

Obrázek 3.2: Odstranění prolínání video sekvencí pomocí modifikovaného kodéru 3D DCT. Snímky v okolí stříhu při: a) absenci redukce prolínání, b) při použití nulového počtu 2D DCT koeficientů a c) při použití všech $N \times N$ 2D DCT koeficientů pro tvorbu průměrovaných doplňkových koeficientů.

4.1 PROSTŘEDÍ DEC-C++

Jestliže bylo prostředí MATLAB určeno pro testování vlastností 3D DCT a hledání vhodného nastavení všech parametrů, prostředí Dev-C++ bylo využito pro vytvoření finální verze programu kódovacího řetězce 3D DCT pomocí programovacího jazyka C. Oproti původní testovací verzi se program v jazyce C vyznačuje strukturou použitelnou pro zpracování vstupních dat v reálném čase. Posloupnost jednotlivých funkcí je zobrazena na *obrázku 4.1*.



Obrázek 4.1: Bloková struktura kodéru a dekodéru 3D DCT v prostředí Dev-C++, naprogramovaná v jazyce C.

Hlavní důraz byl v tomto případě kladen na rychlost výpočtu v jednotlivých blocích. Hlavní funkce celé aplikace `main` obsahuje dílčí funkce pro alokaci paměťového prostoru pro osm vstupních a výstupních snímků (`open_seq`) a pro jednorázové naplnění vstupních dat obrazovými body reprezentující snímky testovací sekvence v šedotónové paletě (`fill_seq`). Aplikace vytvořená v jazyce C byla koncipována pouze pro černobílé video sekvence, nicméně její rozšíření pro barevné signály je pouze formální. Při tomto rozšíření je potřeba realizovat kódování obrazových dat pro všechny tři barevné složky vstupní sekvence.

Vstupní data jsou po blocích $8 \times 8 \times 8$ podrobena přímé trojrozměrné diskrétní kosinové transformaci, kvantování pomocí funkce `fdct_3d_dft` a entropicky kódována funkcí `fhuffman`. Výstupem této funkce je hodnota kompresního poměru pro aktuální skupinu všech osmi snímků. Kvantované frekvenční koeficienty jsou bezprostředně podrobena zpětné trojrozměrné diskrétní kosinové transformaci (`idct_3d_dft`) a získané hodnoty obrazových bodů jsou porovnány s hodnotami vstupními pomocí funkce `nrms`. Při praktickém využití kodéru 3D DCT nebudou data po dokončení entropického kódování bezprostředně dekodována, ale budou uložena do souboru nebo do výstupní vyrovnávací paměti.

Z hlediska složitosti a početní náročnosti kodéru a dekodéru 3D DCT jsou klíčové funkce pro výpočet přímé a zpětné trojrozměrné diskretní kosinové transformace. Množství početních operací obou funkcí lze považovat za totožné a pro testovací sekvenci v šedotónové paletě s rozlišením $320 \times 240 \times 8$ trvá průběh funkce `fdct_3d_dft` 0,078 s (AMD Athlon XP 2700+). Při následném entropickém kódování (`fhuffman`) trvá celý proces 0,094 s. Uvažujeme-li barevnou sekvenci se vzorkováním chrominancních signálů 4:2:0, lze takto vytvořeným kóděrem v reálném čase zpracovávat až 40 snímků s rozlišením 320×240 za jednu sekundu.

4.2 PROSTŘEDÍ CODE COMPOSER STUDIO

Prostředí Code Composer Studio (CCS) je softwarový nástroj firmy Texas Instruments, umožňující vývoj a ladění algoritmů pro signálové procesory této firmy. Také v tomto prostředí byl realizován kódovací řetězec 3D DCT v jazyce C.

Pro přiblížení časové náročnosti výpočtů v kodéru 3D DCT, byla analyzována funkce realizující výpočet přímé trojrozměrné diskretní kosinové transformace pro jednu video krychli vstupních dat: `fdct_3d_dft`. Spolu s výpočtem inverzní transformace 3D DCT představuje tato funkce nejkritičtější pasáž celého kódovacího řetězce. Časová analýza byla provedena pomocí počtu cyklů, během nichž signálový procesor vykoná celou funkci. Hodnoty počtu cyklů byly získány v prostředí CCS v šesti krocích. První dvě analýzy byly provedeny bez pomoci optimalizačních nástrojů softwarového prostředí. Zbývající čtyři odpovídaly konkrétnímu nastavení optimalizace.

Lze konstatovat, že optimalizační nástroje firmy Texas Instruments dokázaly efektivněji zapsat analyzovanou funkci `fdct_3d_dft`. Nejlepších výsledků bylo dosaženo pomocí automatické optimalizace na lokální úrovni. Počet cyklů, a s tím související také doba výpočtu se od ručně optimalizovaného kódu snížila téměř o 77 %. Přesto trvá výpočet přímé transformace 3D DCT jedné video krychle vstupních dat přibližně 8,5 ms. Počet obdobných video krychlí černobílého vstupního signálu s rozlišením 320×240 za jednu sekundu je více než 3 600. Převod těchto vstupních vzorků do frekvenční oblasti by tedy trval více než 30 sekund. Důvodů, proč je daný algoritmus nedostatečně rychlý je několik. Mezi tzv. hardwarové důvody je možno zařadit malý počet registrů v CPU a především nízká frekvence hodinového signálu. Nedostatky je možné hledat pochopitelně také na softwarové úrovni. Při výpočtu $N \times N \times N$ frekvenčních koeficientů je nutné velké množství mezivýsledků ukládat do paměti. Právě tato komunikace mezi procesorem a pamětí, je z hlediska časové náročnosti nejkritičtější. Připomeňme, že doba uložení nebo načtení 32bitové hodnoty do nebo z paměti trvá procesoru TMS320C6711 pět strojových cyklů. To je více než při sčítání nebo dokonce násobení dvou operandů. Větší efektivnosti signálového procesoru je možné dosáhnout pomocí instrukcí, které provádí násobení nebo sčítání několika 8bitových operandů v rámci dvou 32bitových slov. Využití tohoto přístupu ovšem procesor TMS320C6711 neumožňuje. Realizaci algoritmu 3D DCT na signálovém procesoru lze tedy v rámci disertační práce chápat jen jako úvod do dané problematiky, za kterým musí následovat další vývoj.

4.3 KOMPRIMACE TESTOVACÍCH VIDEO SEKVENCÍ

Konečná podoba kódovacího řetězce 3D DCT byla testována v prostředí MATLAB v několika krocích. První z nich si kladl za cíl určit vhodnou kvantizační tabulku, dále způsob řazení frekvenčních koeficientů při entropickém kódování a konečně výběr vhodných tabulek Huffmanova kódu.

Lze konstatovat, že pro nejlepších komprimační vlastnosti metody 3D DCT je nutné ke kvantování frekvenčních koeficientů jasové složky použít upravené tabulky MPEG-4 (příp. MPEG-2) a pro chrominanční složky modifikované kvantizační tabulky standardu JPEG, určené pro chrominanční složky. Způsob řazení frekvenčních koeficientů je dále vhodné provést pomocí navržené metody \perp řez. Frekvenční koeficienty musí být také rozděleny na DC a AC komponenty a ty kódovány s využitím Huffmanových slovníků s celkovým počtem 5 622 kódových slov (pro AC koeficienty).

Při detailním rozboru parametrů kódovacího řetězce 3D DCT je nutné uvažovat také vliv parametrů, které mají menší dopad na komprimační vlastnosti než mají operace kvantování, či řazení jednotlivých koeficientů. Prvním parametrem je způsob řádkování. Existují dva typy řádkování obrazových dat. V televizní technice se používá řádkování prokládané, kdy jsou odděleně přenášeny a zobrazovány nejprve liché řádky televizního snímku a následně řádky sudé. Při neprokládaném řádkování je přirozené pořadí řádků obrazového signálu zachováno.

Je zřejmé, že odstranění plynulých vertikálních přechodů v zobrazované scéně při prokládaném řádkování, má za důsledek přítomnost většího množství frekvenčních koeficientů ve spektrální oblasti a s tím spojenou menší hodnotu kompresního poměru. Větší míry komprimace lze tedy dosáhnout pomocí neprokládaného řádkování. Ověřená velikost chyby navíc dosahuje nižších hodnot také při použití neprokládaného řádkování. Je tedy evidentní, že při kódování reálných video signálů metodou 3D DCT je s ohledem na komprimační vlastnosti nutné použít právě toto řádkování.

Dalším nástrojem, který ovlivňuje komprimační vlastnosti je změna barevné reprezentace. Částečné dekorelace barevných složek video sekvencí je možné dosáhnout převodem jejich barevné reprezentace z RGB na soubor jasové složky a složek chrominančních YC_bC_r . Tato barevná reprezentace soustřeďuje stěžejní část informace do jasové složky Y a chrominanční složky představují pouhou dodatkovou informaci o barevném složení vstupních dat.

Z porovnání hodnot výsledné chyby plyne, že převod barevných sekvencí do formátu YC_bC_r způsobí nepatrné zvýšení hodnoty chyby v rekonstruovaných sekvencích. Jedná se však o natolik nízký nárůst, že subjektivním hodnocením nebylo možné odhalit žádné viditelné zhoršení kvality. Podstatných rozdílů ovšem dosahují hodnoty kompresních poměrů. Změnou barevné reprezentace bylo docíleno zvýšení kompresního poměru v průměru o 150 %. Počet zpracovávaných vzorků obrazové informace přitom zůstává u obou reprezentací shodný.

Výhoda chrominančních signálů pramení také z nižší citlivosti lidského zraku na její hodnoty. V oblasti zpracování digitálních obrazových dat je proto běžná praxe pod-

vzorkovávat tato data. Dva nejčastější způsoby redukce počtu vzorků chrominančních složek jsou podvzorkování v horizontálním směru (4 : 2 : 2) a v horizontálním i vertikálním směru (4 : 2 : 0). Celkový počet vstupních vzorků se tak v prvním případě sníží o jednu třetinu a ve druhém případě až o jednu polovinu. Dojde proto nejen ke zvýšení kompresního poměru, ale současně i ke snížení počtu operací, které musí kodér i dekodér realizovat.

Při dekódování obrazových dat musí před obnovením barevné reprezentace *RGB* dojít k obnově původního rozlišení ve všech barevných složkách. Vynechané vzorky chrominančních signálů je možné získat opakováním stávajících vzorků či interpolací vždy dvou sousedních hodnot.

Závěry dokazují, že snižováním počtu kódovaných vzorků se výsledná chyba zvyšuje. Subjektivním hodnocením ale opět není možné odhalit tuto degradaci kvality. Výsledná chyba je také větší při obnově vynechaných vzorků pomocí interpolace. Snižování počtu transformovaných vzorků má na druhé straně za důsledek zvýšení kompresního poměru. Hodnoty kompresního poměru pro způsob vzorkování 4 : 2 : 0 dosahují nárůstu minimálně 10 % ve srovnání s úplným rozlišením chrominančních signálů 4 : 4 : 4. Pro statickou sekvenci se však tento nárůst blíží ke 30 %.

4.4 VOLBA OBRAZOVÉ KVALITY

Pro hodnocení komprimačních vlastností kódovacího řetězce 3D DCT byly použity dva hlavní parametry: kompresní poměr a výsledná chyba. Kompresní poměr vyjadřuje snížení počtu bitů, potřebných pro kódování komprimované sekvence nbB_{com} vzhledem k počtu bitů originální sekvence nbB_{org} , jak popisuje rovnice [2]

$$CR = \frac{nbB_{org}}{nbB_{com}}. \quad (4.1)$$

Stupeň kvality rekonstruované video sekvence v závislosti na své originální předloze lze hodnotit subjektivními nebo objektivními kritérii. Subjektivní hodnocení vychází z osobního dojmu každého diváka. Objektivní metody jsou založeny na matematickém porovnání odpovídajících snímků originální a rekonstruované sekvence. Pro posouzení komprimačních vlastností metody 3D DCT byl vybrán objektivní přístup pomocí parametru *NRMSE*, definovaný pro jednu barevnou složku vztahem [2]

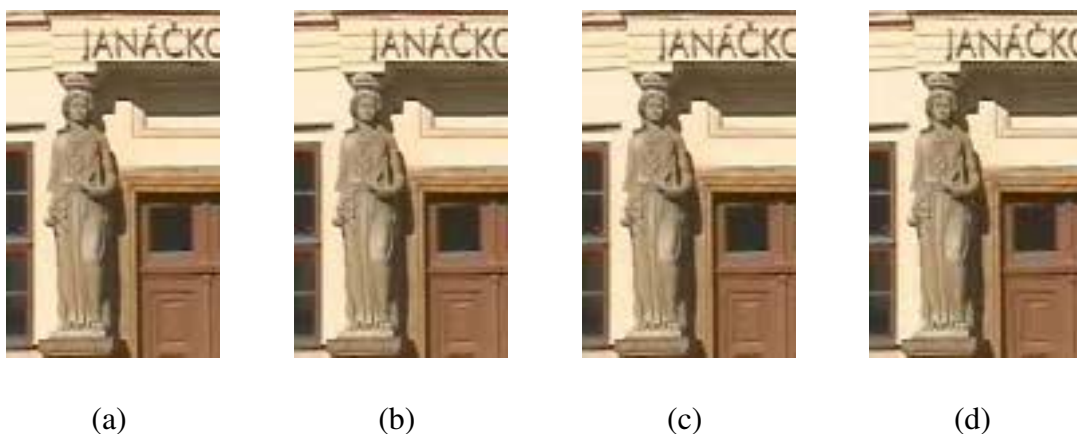
$$NRMSE = \sqrt{\frac{\sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \sum_{f=0}^{nbF-1} [x(i, j, f) - \hat{x}(i, j, f)]^2}{\sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \sum_{f=0}^{nbF-1} x(i, j, f)^2}}, \quad (4.2)$$

kde $x(i, j, f)$ je intenzita obrazového bodu originální a $\hat{x}(i, j, f)$ rekonstruované video sekvence na pozici (i, j, f) , konstanta nbF definuje celkový počet snímků ve video sekvenci a W, H určují rozlišení obou video sekvencí.

Stěžejní vliv na kompresní poměr a výslednou chybu 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 parametru qlt . Ten nabývá hodnot z intervalu $\langle 1; 100 \rangle$, přičemž má charakter kvality; tj. pro $qlt = 100$ je výsledná kvalita nejlepší.

Konečné nastavení kodéru a dekodéru 3D DCT bylo ověřeno na množině sedmi testovacích sekvencí. Postupně byly přepočítány hodnoty kvantovacích parametrů pro devět nastavení parametru qlt . Vypočtené hodnoty kompresního poměru a výsledné chyby jsou znázorněny v tabulce 4.1.

Snižování hodnoty parametru qlt bezprostředně způsobuje zvýšení hodnot všech kvantovacích parametrů. Po operaci kvantování dojde ke snížení počtu nenulových frekvenčních koeficientů a ke zvýšení kompresního poměru. Současně však dochází k větší degradaci kvality rekonstruované sekvence. Dále lze pozorovat, že komprimační vlastnosti 3D DCT závisí na typu kódované video sekvence. Statická sekvence dosahuje pro danou hodnotu kvalitativního parametru qlt nejvyšších hodnot kompresního poměru a nejnižších hodnot chyby. Naopak, video sekvence s rychlým pohybem vykazuje ve všech případech nejnižší hodnotu kompresního poměru, která je současně doprovázena nejvyšší odchylkou originální a rekonstruované sekvence. Tento důsledek je zapříčiněn rozdílným počtem nenulových frekvenčních koeficientů obou sekvencí po přímé trojrozměrné diskrétní kosinové transformaci. Ke snížení tohoto počtu (a současně ke zvýšení kompresního poměru) je zapotřebí aplikovat kvantizační tabulky s velkými kvantizačními parametry, což nutně přináší zhoršení kvality.



Obrázek 4.2: Část rekonstruované sekvence j_{amu} pro: a) $qlt = 100$, b) $qlt = 60$, c) $qlt = 40$ a d) $qlt = 20$.

Příklad statické a dynamické sekvence je uveden na obrázku 4.2 a 4.3. Se zvyšující mírou kvantování lze pozorovat vznik blokových artefaktů v obraze a částečné prolínání obsahu sousedních snímků. To je totožný jev, ke kterému dochází při při-

Tabulka 4.1: Vliv míry kvantování na velikost kompresního poměru a výsledné chyby při komprimaci video signálů pomocí 3D DCT

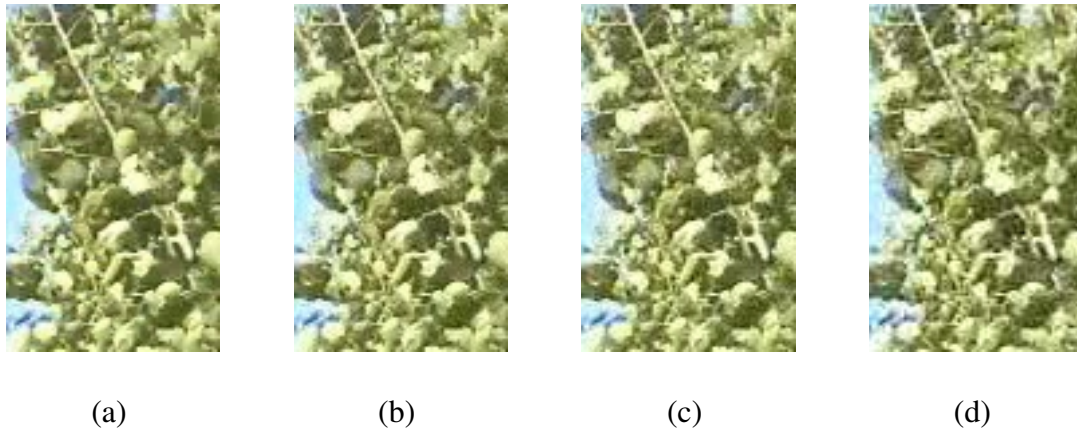
Název sekvence	$qlt = 100$		$qlt = 90$		$qlt = 80$	
	CR	NRMSE	CR	NRMSE	CR	NRMSE
(1) cerveny_kostel	7,67	0,00910	34,76	0,01234	65,44	0,01446
(2) hradecka	4,19	0,01171	11,16	0,01849	17,95	0,02540
(3) jamu	4,83	0,01648	14,26	0,02064	24,18	0,02500
(4) jedovnicka	4,47	0,00997	11,20	0,01491	17,65	0,02062
(5) skok_vysoky	4,03	0,01910	11,17	0,02642	18,46	0,03405
(6) vitr	2,23	0,02567	4,60	0,03440	6,27	0,04482
(7) zavod	3,44	0,02562	9,11	0,03156	14,41	0,03782

Název sekvence	$qlt = 70$		$qlt = 60$		$qlt = 50$	
	CR	NRMSE	CR	NRMSE	CR	NRMSE
(1) cerveny_kostel	93,88	0,01599	115,62	0,01730	137,89	0,01855
(2) hradecka	24,14	0,03071	30,26	0,03547	36,04	0,03944
(3) jamu	33,18	0,02845	41,35	0,03160	48,83	0,03413
(4) jedovnicka	23,59	0,02539	29,49	0,02971	35,02	0,03340
(5) skok_vysoky	25,28	0,03999	32,16	0,04521	38,67	0,04958
(6) vitr	7,69	0,05506	9,08	0,06509	10,43	0,07417
(7) zavod	19,28	0,04308	24,09	0,04797	28,71	0,05216

Název sekvence	$qlt = 40$		$qlt = 30$		$qlt = 20$	
	CR	NRMSE	CR	NRMSE	CR	NRMSE
(1) cerveny_kostel	162,52	0,02003	197,26	0,02233	258,08	0,02674
(2) hradecka	43,42	0,04409	54,83	0,05040	77,71	0,06094
(3) jamu	57,83	0,03723	70,35	0,04182	93,12	0,05042
(4) jedovnicka	41,87	0,03767	51,82	0,04361	73,28	0,05372
(5) skok_vysoky	47,01	0,05457	59,97	0,06134	86,28	0,07273
(6) vitr	12,20	0,08501	15,08	0,10018	21,43	0,12554
(7) zavod	34,70	0,05706	43,97	0,06393	62,84	0,07509

tomnosti stříhu v množině kódovaných snímků. Vznik rušivých blokových artefaktů o velikosti $N \times N$ je u dynamické sekvence patrný již od nízkého stupně kvantování. Na statickém pozadí první sekvence je degradace kvality patrná pouze při nejvyšším stupni komprese a to pouze v malém měřítku.

Hodnocení výsledné kvality není vhodné provádět pouze na základě vypočteného parametru $NRMSE$ či rozborem jednotlivých, statických snímků jak tomu bylo doposud. Nejdůležitějším hodnotícím kritériem je subjektivní dojem diváka na celý průběh dekódované video sekvence. Lze konstatovat, že na všechny typy sekvencí lze aplikovat kvantování s hodnotou parametru $qlt = 40$ a přitom zachovat dobrou kvali-



Obrázek 4.3: Část rekonstruované sekvence `vit_r` pro: a) $q_{lt} = 100$, b) $q_{lt} = 60$, c) $q_{lt} = 40$ a d) $q_{lt} = 20$.

tu obrazu. Při nižších úrovních kvalitativního parametru vykazuje již většina sekvencí přítomnost rušivých artefaktů. Hodnoty kompresního poměru jsou pro toto nastavení vyšší pro sekvence s pomalým pohybem. Z toho plyne ideální použití metody 3D DCT při komprimaci barevných video sekvencí, které obsahují velké plochy a pomalu se měnící obsah.

5 ZÁVĚR

Předložená disertační práce si kladla za cíl posoudit možnosti trojrozměrné diskrétní kosinové transformace při komprimaci video sekvencí. V práci jsou splněny a rozebrány následující cíle:

Prostudování dosavadního vývoje. Úvodní část disertační práce je věnována stručnému popisu komprimačních postupů statických a pohyblivých snímků. Důraz je kladen na metody používající diskrétní kosinovou transformaci: JPEG, MPEG.

Návrh kódovacího a dekódovacího řetězce 3D DCT. Kódovací řetězec JPEG lze s úspěchem využít také pro inspiraci při tvorbě kódovacího řetězce využívající 3D DCT. Změny je pochopitelně nutné provést při ukládání vstupních a výstupních dat, při procesu kvantování a v neposlední řadě také při entropickém kódování.

Rozbor faktorů ovlivňujících komprimační vlastnosti. Komprimační vlastnosti 3D DCT lze ovlivnit různými způsoby. Mezi základní patří volba typu řádkování vstupní video sekvence, míra kvantování frekvenčních koeficientů, způsob řazení koeficientů při entropickém kódování či použití vhodných Huffmanových tabulek.

Odstranit prolínání sekvencí při střihu. Negativní vlastností 3D DCT je vznik prolínání scén při komprimaci odlišných video sekvencí. Příčiny vzniku tohoto pro-

línání, nástroje pro jeho účinné potlačení, včetně možnosti detekce stříhu jsou rozebrány v disertační práci.

Experimentální ověření komprimačních možností 3D DCT. Vlastnosti 3D DCT byly testovány na množině reálných video sekvencí. Z toho důvodu byl realizován kódovací řetězec 3D DCT v prostředí MATLAB, Dev-C++ a Code Composer Studio. Byly doporučeny oblasti použití, kde je možné pomocí 3D DCT dosáhnout nejlepších komprimačních výsledků.

LITERATURA

- [1] AVCIBAŞ, I., MEMON, N., SANKUR, B. Steganalysis Using Image Quality Metrics. *IEEE Transactions on Image Processing*. Únor 2003, roč. 12, č. 2, str. 221–229. ISSN 1057-7149.
- [2] CHROMÝ, I., *Compression of Digital Video Signals*. [Disertační práce], Brno (Česká republika): UREL FEKT VUT v Brně, 1999.
- [3] FRÝZA, T., HANUS, S. Relation Between Character of Real Video Sequences and 3D DCT Compression. In *Proceedings EC-VIP-MC 2003*, vol. 1. Záhřeb (Chorvatsko): FEEC Záhřeb, červenec 2003, str. 107–112. ISBN 95-318-4059-8.
- [4] FRÝZA, T., HANUS, S. Video Signals Transparency in Consequence of 3D DCT Transform. In *Radioelektronika 2003 Conference Proceedings*. Brno (Česká republika): UREL FEKT VUT v Brně, květen 2003, str. 127–130. ISBN 80-214-2383-8.
- [5] FRÝZA, T., HANUS, S. Transparency Reduction Algorithms for 3D DCT Encoders and Decoders. In *Recent Advances in Intelligent Systems and Signal Processing*. Korfu (Řecko): WSEAS, červenec 2003, str. 192–197. ISBN 96-080-5287-4.
- [6] FRÝZA, T., HANUS, S. Algorithms for Fast Computing of the 3D DCT Transform. *Radioengineering*. Duben 2003, roč. 12, č. 1, str. 23–26. ISSN 1210-2512.
- [7] FRÝZA, T., HANUS, S. Image Compression Algorithms Optimized for MATLAB. *Radioengineering*. Prosinec 2003, roč. 12, č. 4, str. 18–20. ISSN 1210-2512.
- [8] FRÝZA, T., HANUS, S. Color Video Signals Compression Based on 3D DCT Transform. *Elektrorevue*. Únor 2003, roč. 2003, č. 9. ISSN 1213-1539. [Online] Available: <http://www.elektrorevue.cz/clanky/03009/english.htm> (březen 2005).
- [9] FRÝZA, T., HANUS, S. Concept of Variation Detector Used in Video Signal Compression Domain. *WSEAS Transactions on Circuits*. Říjen 2004, roč. 3, č. 9, str. 1776–1780. ISSN 1109-2734.

- [10] GHANBARI, M. *Standard Codecs: Image Compression to Advanced Video Coding*. Londýn (Velká Británie): IEE Publishing, první vydání, červen 2003. 430 stran. ISBN 08-529-6710-1.
- [11] JIMENEZ, P., ARROYO, S., BUENO, R., MONTERO, R., PITA, R. Background Behavior Analysis in Video Sequences. In *Recent Advances in Intelligent Systems and Signal Processing*. Korfu (Řecko): WSEAS, červenec 2003, str. 213–216. ISBN 96-080-5287-4.
- [12] JO, W., KIM, J., JUNG, Y., BAE, M. On a Reduction of Computation Time of FFT Cepstrum. In *Recent Advances in Intelligent Systems and Signal Processing*. Korfu (Řecko): WSEAS, červenec 2003, str. 228–232. ISBN 96-080-5287-4.
- [13] LEE, M., CHAN, R., ADJEROH, D. Quantization of 3D DCT Coefficients and Scan Order for Video Compression. *Journal of Visual Communication and Image Representation*. Prosinec 1997, roč. 8, č. 4, str. 405–422. ISSN 1047-3203.
- [14] 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.
- [15] RICHARDSON, I. *Video Codec Design: Developing Image and Video Compression Systems*. Chichester (Velká Británie): John Wiley & Sons, duben 2003. 303 stran. ISBN 04-714-8553-5.
- [16] SAAVEDRA, E., GRAUEL, A., MORTON, D. Combined Methods for Image Compression. In *Recent Advances in Intelligent Systems and Signal Processing*. Korfu (Řecko): WSEAS, červenec 2003, str. 233–235. ISBN 96-080-5287-4.
- [17] SMÉKAL, Z., SYSEL, P. Signálové procesory VLIW firmy Texas Instruments. *Sdělovací technika*. Leden 2003, č. 1, str. 18–20. ISSN 0036-9942.
- [18] STUPÁK, C. Filtering of the Color Images Distorted by Impulse Noise. *Radio-engineering*. Zář 2001, roč. 10, č. 3, str. 21–27. ISSN 1210-2512.
- [19] VÍT, V. *Televizní technika. Přenosové barevné soustavy*. Praha (Česká republika): BEN, Technická literatura, první vydání, leden 1997. 719 stran. ISBN 80-860-5604-X.
- [20] 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.
- [21] WESTWATER, R., FURHT, B. *Real-Time Video Compression*. Boston (USA): Kluwer Academic Publishers, leden 1997. 176 stran. ISBN 07-923-9787-8.
- [22] ŘÍČNÝ, V. *Televizní technika a video technika*. [Skripta], Brno (Česká republika): UREL FEKT VUT v Brně, 1998. 115 stran. ISBN 80-214-1104-X.

PROFESNÍ ŽIVOTOPIS



Ing. Tomáš Frýza

Sekaninova 6, 614 00 Brno

Email fryza@feec.vutbr.cz
Tel. +420 541 149 134
Narozen 16. listopadu 1977 v Novém Jičíně
Národnost česká
Rodinný stav svobodný

Vzdělání

- 2002 – 2005** Tříleté postgraduální studium na Ústavu radioelektroniky FEKT VUT v Brně. Disertační práce z oblasti komprimace číslicových video signálů metodou 3D DCT.
- 2004** Pětiměsíční stáž na ISEP (Institut Supérieure de l'Électronique de Paris) v Paříži, Francie.
- 1996 – 2002** Magisterské studium na FEKT VUT v Brně. Diplomová práce na téma “*Komprimace video signálů pomocí transformace 3D DCT*” byla oceněna cenou děkana.
- 2000 – 2001** Jednoleté studium na ENST (Ecole Nationale Supérieure des Télécommunications) v Paříži, Francie.

Odborná praxe

- 2002** Programování TPU funkcí pro PowerPC Microcontroller 555. Motorola, Rožnov p. Radhoštěm.
- 2003** Tvorba “*beanu*” zapouzdřujících periférie MTS a SIU mikrokontroléru CopperHead. UNIS, Brno.

Informatika

C/C++, UNIX-shells, Delphi, Assembler, MATLAB, MathCad, AutoLISP
Microsoft Office, L^AT_EX
UNIX, Linux, Windows 9x/NT/2000/XP

Cizí jazyky

Čeština rodný jazyk
Angličtina plyně slovem i písmem
Francouzština plyně slovem i písmem

Oblasti vědeckého zájmu

Zpracování číslicových signálů, komprese video signálů, televizní technika, programování v C/C++, mikroprocesorová technika

ABSTRAKT

Předložená disertační práce pojednává o možnosti komprimovat barevné video signály pomocí trojrozměrné diskrétní kosinové transformace (3D DCT). Základní snahou všech komprimačních metod je potlačení nadbytečnosti v jednotlivých snímcích a také v čase, mezi sousedními snímky. Metoda 3D DCT využívá současného kódování skupiny snímků a slučuje tak oba požadavky do jediného transformačního kódování.

Vytvořený kódovací řetězec 3D DCT je odvozen od standardu JPEG, určený pro komprimaci statických snímků. Obměny, ke kterým muselo dojít ve struktuře kodéru a dekodéru 3D DCT jsou v textu podrobně popsány. Jedná se především o použití vícerozměrné diskrétní kosinové transformace, dále o odlišný způsob kvantování frekvenčních koeficientů, a také o vývoj nových Huffmanových tabulek pro entropické kódování 3D DCT koeficientů.

Praktické možnosti komprimační metody byly ověřeny na množině testovacích video sekvencí, které svým obsahem pokrývají širokou škálu aplikací. Bylo zjištěno, že metoda 3D DCT dosahuje nejlepších komprimačních vlastností při kódování scén s pomalým pohybem a velkými plochami shodné barvy. Pro tuto kategorii scén není výjimkou dosažený kompresní poměr o velikosti převyšující hodnotu 100. Prioritní oblastí použití metody 3D DCT jsou tedy video konference a video telefonie.

ABSTRACT

Thesis presents the possibilities of the Three Dimensional Discrete Cosine Transform (3D DCT) in a video compression domain. All video compression methods are focused on removing of any kind of redundance, both in space and temporal dimensions. The 3D DCT combines these principles in a single transform coding.

Proposed structure of the 3D DCT coder is based on the JPEG standard, dedicated for compression of the static pictures. Unavoidable modifications were realized mainly in usage of the three dimensional transform, in quantisation of the frequency coefficients and in the code words dictionary, used in entropy coding.

Practical capabilities of the compression method were tested with the aid of several color video sequences. Each of them represents different type of a visual scene: from the static scenes to the sequences with dynamic changes in temporal dimension. It was discovered the best compression properties of the 3D DCT is obtained when input video sequence contains slow motion accompanied by large areas of the same color. In that case the compression ratio values higher than 100 can be repeatedly reached. Therefore the main domain of using the 3D DCT is in the video conference and video telephony applications.