

# ALGORITMY KVANTOVANIA ŠTANDARDNÝCH VIDEOKODEKOV (QUANTIZATION ALGORITHMS OF STANDARD VIDEOCODECS)

Ján MIHALÍK, Radoslav ŠTEFANIŠIN

Laboratórium číslicového spracovania obrazov a videokomunikácií, Katedra elektroniky a multimediálnych telekomunikácií,  
FEI TU Košice, Park Komenského 13, 041 20 Košice, Slovenská republika  
tel.:055/602 2854, 055/602 4309, E-mail: Jan.Mihalik@tuke.sk, Radoslav.Stefanisin@tuke.sk

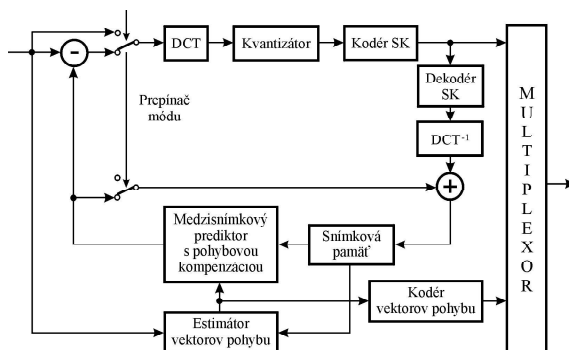
## SUMMARY

The paper gives a review of algorithms of quantization of the standard videocodecs. On the basis of the results of statistical analysis of videosignals in DCT domain different modifications of linear and nonlinear algorithms of quantization are presented in intraframe and interframe modes of the videocodecs. Finally an evaluation and comparison of their performances by archived experimental results have been made.

**Keywords:** standard videocodecs, videosequence, statistical analysis, linear and nonlinear quantization, algorithm intraframe and interframe mode

## 1. ÚVOD

V oblasti metód kompresie statických a dynamických obrazov [1] je charakteristickým znakom štandardizácia, výsledkom ktorej boli mnohé medzinárodné štandardné videokodeky ako H.261[2], H.263 [3] a MPEG-1,2,4 [4,5,6]. Tieto štandardné videokodeky umožňujú jednoduchú výmenu obrazových informácií medzi zariadeniami rôzneho charakteru pomocou pamäťových médií alebo telekomunikačných sietí.



**Obr. 1** Bloková schéma zdrojového kódovania v štandardných videokodekoch

**Fig. 1** The block diagram of the source coding in the standard videocodecs

Kompresia videosignálov vo všetkých druhoch štandardných videokodekov sa dosahuje najmä pomocou medzisnímkového hybridného kódovacieho systému (MHKS) [7] s diskretnou kosínusovou transformáciou (DCT) a medzisnímkovou predikciou s pohybovou kompenzáciou. Tento okrem medzisnímkového módu umožňuje aj vnútro-snímkový mód zodpovedajúci vnútro-snímkovému transformačnému kódovaniu s DCT [8]. Medzisnímková predikcia s pohybovou kompenzáciou sa vykonáva v obrazovom priestore. Pre určenie vektorov pohybu sa používajú metódy estimácie pohybu na princípe blokového

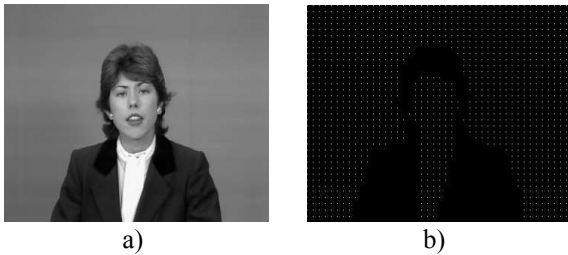
prispôsobovania. Informácia o vektoroch pohybu je prídavnou informáciou, ktorú je potrebné prenášať spolu s hlavnou informáciou o spektrálnych koeficientoch (SK). Na obr.1 je znázornená bloková schéma zdrojového kódovania, ktorá sa používa vo všetkých štandardných videokodekoch.

Tento článok sa zaoberá algoritmi kvantovania štandardných videokodekov, ktoré výraznou mierou ovplyvňujú jeho celkovú účinnosť kódovania. Sú to algoritmy lineárneho a nelineárneho kvantovania vo vnútro-snímkovom a medzisnímkovom móde. Pre dosiahnutie väčšej kompresie údajov sa budú tiež uvažovať ich modifikácie s prázdnu zónou. Na základe vlastných výsledkov simulácie týchto algoritmov sme vykonali porovnanie a hodnotenie ich účinnosti kvantovania.

## 2. ŠTATISTICKÁ ANALÝZA V PRIESTORE DCT PRE VNÚTRO-SNÍMKOVÝ A MEDZISNÍMKOVÝ MÓD

Kvantovanie videosignálu v priestore DCT vyplýva zo zapojenia MHKS v štandardných videokodekoch, kde samotnej operácii kvantovania predchádza priama DCT. Štatistická analýza v priestore DCT bude vykonaná pre vnútro-snímkový a medzisnímkový mód videokodeka. Tieto módy a výsledky štatistickej analýzy je potrebné zohľadniť aj pri samotnom kvantovaní transformovaného videosignálu. Bloky obrazových prvkov (op) vstupného alebo diferenčného videosignálu sa transformujú pomocou DCT na bloky SK s rovnakými rozmermi pri zachovaní energie blokov v obrazovom a transformovanom priestore. DCT dekoreluje tieto bloky op, pričom ich energia v priestore DCT sa koncentruje do SK s nižšími frekvenciami.

Na obr. 2 je uvedená vybraná snímka videosekvencie „Claire“ a jej transformovaný obraz získaný pomocou dvojrozmernej DCT (2RDCT) s rozmermi bloku 8×8. Z transformovanej snímky vidno, že najväčšiu energiu v bloku nadobúda jednosmerný spektrálny koeficient (JSK).

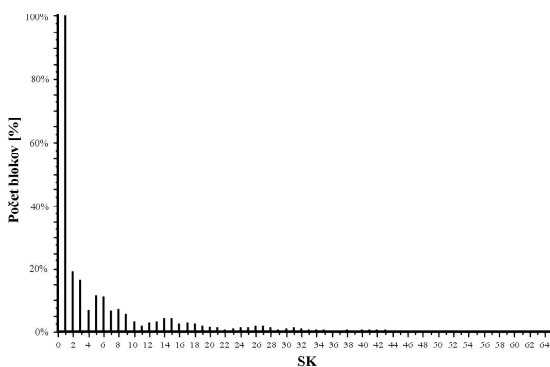


**Obr. 2** a) Vybraná snímka videosekvencie „Claire“  
b) jej transformovaný obraz

**Fig. 2** a) The chosen frame of videosequence „Claire“  
b) its transform image

DCT síce dekoreluje snímky videosekvencie vo vnútri jej blokov, ale môže zachovávať medziblokovú koreláciu, ktorá závisí od veľkosti týchto blokov [9]. Histogramy JSK možno dobre aproximovať pomocou Gaussovho rozdelenia pravdepodobnosti (GRP) a naopak histogramy ostatných SK reprezentujúce striedavé zložky blokov pomocou Laplaceovho rozdelenia pravdepodobnosti (LRP).

Účinnosť kódovania štandardných videokodekov je možné zvýšiť zanedbaním irelevantných SK. Najlepší spôsob odstránenia týchto SK je pomocou prahovej filtrácie. Ako kritérium zvolenia prahu je možné použiť vo vnútrosnímkovom móde pomer signál/šum (S/Š) rekonštruovaného videosekvencie v kodéri získaného aplikovaním spätnej DCT na už filtrované, ale nekvantované SK. Hodnota S/Š takto rekonštruovaného videosekvencie musí zohľadňovať následné zmenšenie svojej hodnoty vplyvom kvantovania. Obr. 3 znázorňuje rozloženie relevantných (nenulových) SK v blokoch transformovanej videosekvencie „Claire“ s počtom 15 snímok vo vnútrosnímkovom móde a s prahovou hodnotou 16.



**Obr. 3** Počet blokov s daným SK v transformovanej videosekvencii „Claire“ pre diagonálne snímanie bloku a vnútrosnímkový mód

**Fig. 3** The number of blocks with given SK in the transformed videosequence „Claire“ for „zig-zag“ scanning of the block and intraframe mode

Z tohto obrázka je vidno, že JSK sa vo vnútrosnímkovom móde vyskytuje v každom bloku, pretože jeho pravdepodobnosť nad prahovou hodnotou je 100%. Pritom pravdepodobnosť výskytu SK s vyššími frekvenciami prudšie klesá končiac SK

s nižšou frekvenciou v porovnaní s obdobným priebehom v medzislímkovom móde (obr. 5).

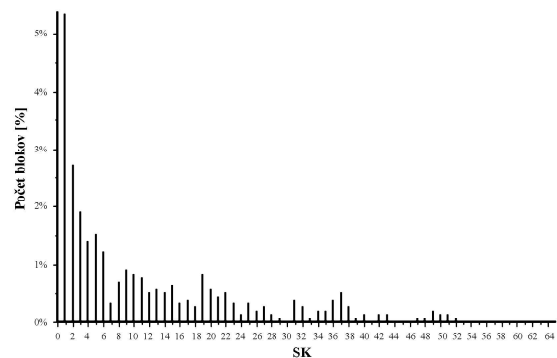
V medzislímkovom móde štandardného videokodeka v ktorom je aplikovaná 2RDCT na diferenčný videosekvencie sa podobne ako vo vnútrosnímkovom móde dekorelujú snímky vo vnútri ich blokov. Na obr. 4 je znázornená vybraná snímka diferenčnej videosekvencie „Claire“ a jej transformovaný obraz získaný pomocou 2RDCT (8×8).



**Obr. 4** a) Vybraná snímka diferenčnej videosekvencie „Claire“ b) jej transformovaný obraz

**Fig. 4** a) The chosen frame of the difference videosequence „Claire“ b) its transform image

Medzislímková korelácia všetkých SK je v tomto móde veľmi slabá a tiež majú menšiu medziblokovú koreláciu. Na obr.5 je znázornené rozloženie relevantných SK v blokoch transformovanej diferenčnej videosekvencie „Claire“ s počtom 14 snímok v medzislímkovom móde a s prahovou hodnotou 16.



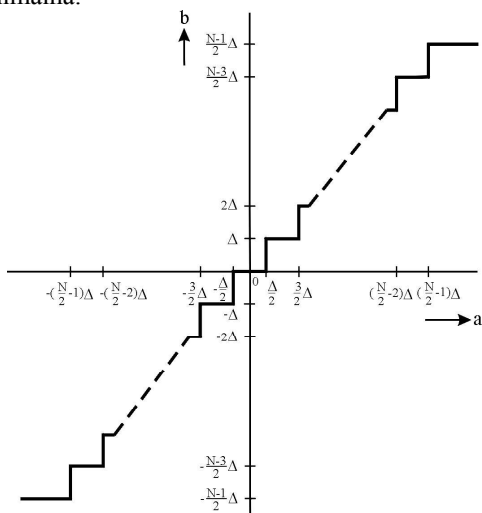
**Obr. 5** Počet blokov s daným SK v transformovanej diferenčnej videosekvencii „Claire“ pre diagonálne snímanie bloku a medzislímkový mód

**Fig. 5** The number of blocks with given SK in the transformed difference videosequence „Claire“ for „zig-zag“ scanning of the block and interframe mode

Ako kritérium zvolenia prahu je možné použiť v tomto móde S/Š rekonštruovaného videosekvencie v kodéri získaného súčtom predikcie a chybového signálu po spätnej DCT filtrovaných, ale nekvantovaných SK. Rozloženie SK na obr. 5 síce naznačuje, že najčastejšie sa vyskytuje JSK, ale poukazuje aj na výskyt SK s vyššími frekvenciami a to aj napriek ich prahovej filtrácii. V tomto móde na rozdiel od vnútrosnímkového módu histogramy všetkých SK možno dobre aproximovať pomocou LRP.

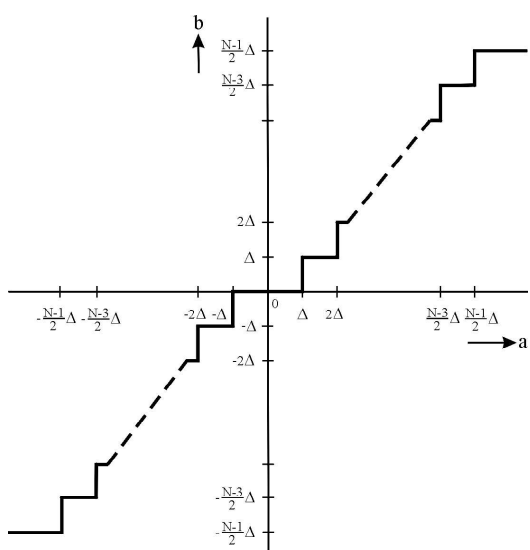
### 3. LINEÁRNE ALGORITMY KVANTOVANIA

Proces kvantovania SK v štandardných videokodekoch nie je presne stanovený, ale vychádza z určitých odporúčaní. V týchto odporúčaníach pre štandardné videokodeky sa pre kvantovanie SK najčastejšie uvažuje lineárny skalárny kvantizátor (LSK) [10]. Typickou vlastnosťou LSK je, že kvantizačné aj rozhodovacie úrovně jeho prevodovej charakteristiky sú rovnomerne odstupňované s kvantizačným krokom  $\Delta$ , ako to vidno z obr.6 Optimalizácia LSK pre zvolené rozdelenie pravdepodobnosti bude znamenať nájdenie takého  $\Delta$ , aby stredná kvadratická hodnota kvantizačného šumu bola minimálna.



Obr. 6 Prevodová charakteristika LSK s nepárnym počtom kvantizačných úrovní

Fig. 6 The transfer characteristic of LSQ with an odd number of quantization levels



Obr. 7 Prevodová charakteristika LSK s prázdnu zónou

Fig. 7 The transfer characteristic of LSQ with dead zone

Odstránením irelevantných SK pomocou vhodnej prahovej filtrácie a následným kvantovaním len relevantných SK je možné podstatne zvýšiť účinnosť štandardných videokodekov. Takýto spôsob kvantovania nazývame „kvantovanie s prázdnu zónou“. Prevodová charakteristika LSK s prázdnu zónou na obr. 7 sa vyznačuje prázdnu zónou v rámci ktorej sa kvantovanie nevykonáva. Hodnoty z tohoto intervalu sú vynulované pomocou prahovej filtrácie. LSK s prázdnu zónou je použitý aj v štandardných videokodekoch ako je H.261 alebo H.263 [11]. Veľkosť prázdnej zóny ako aj celkový tvar prevodovej charakteristiky závisí od algoritmu kvantovania v jednotlivých štandardných videokodekoch.

Pre kvantovanie v H.261 a H.263 sú uvažované dva druhy kvantizátorov [2,3] zodpovedajúce vnútroštruktúrnemu a medzisímkovému módu týchto videokodekov. Algoritmus kvantovania možno popísať nasledujúcimi rovnicami

- vo vnútroštruktúrnem móde

- pre striedavé SK

$$q_i = Y_i / 2QP \quad (1)$$

- pre JSK

$$q_i = Y_i // 8 \quad (2)$$

- v medzisímkovom móde pre všetky SK

$$q_i = (Y_i - QP / 2) / 2QP \quad (3)$$

keď „/“ predstavuje delenie so zaokrúhľením na celé číslo smerom nadol, „//“ predstavuje delenie so zaokrúhľením na najbližšie celé číslo. Potom  $Y_i$  je SK, ktorý sa má kvantovať,  $q_i$  je kvantizačný index a  $QP = \Delta/2$  je kvantizačný parameter, ktorý môže nadobúdať celočíselné hodnoty z rozsahu 1÷31. Tieto hodnoty sa môžu meniť pre každý makroblok alebo skupinu makroblokov, pričom vo vnútri daného makrobloku je použitá rovnaká hodnota QP pre všetky spektrálne koeficienty, okrem JSK vo vnútroštruktúrnem móde. Prázdna zóna kvantizátorov pre ostatné koeficienty je vytváraná operáciou zaokrúhľovania  $q_i$ , pričom jej veľkosť lineárne závisí od samotnej hodnoty QP. Ako vyplýva z rov.(2) JSK vo vnútroštruktúrnem móde sa bude kvantovať pomocou LSK bez prázdnej zóny a s konštantným kvantizačným krokom. Vo všeobecnosti vidno, že výsledkom kvantovania sú kvantizačné indexy  $q_i$  s menším dynamickým rozsahom ako mali pôvodne SK. V dekodovacom systéme je potrebné obnovovať dynamický rozsah týchto SK a to pomocou nasledujúcich rovnic

- JSK vnútroštruktúrnem módu

$$\tilde{Y}_i = 8q_i \quad (4)$$

- všetky ostatné SK

- QP = nepárna hodnota

$$\tilde{Y}_i = QP(2q_i + 1), \quad q_i > 0$$

$$\tilde{Y}_i = QP(2q_i - 1), \quad q_i < 0 \quad (5)$$

-  $QP = \text{párna hodnota}$

$$\begin{aligned}\tilde{Y}_i &= QP(2q_i + 1) - 1, & q_i > 0 \\ \tilde{Y}_i &= QP(2q_i - 1) + 1, & q_i < 0 \\ \tilde{Y}_i &= 0, & q_i = 0\end{aligned}\quad (6)$$

kde  $\tilde{Y}_i$  je rekonštruovaný SK.

Algoritmus kvantovania [4] v štandardnom videokodeku MPEG-1 je podobný kvantovaniu v štandarde JPEG, kde každý koeficient po DCT je delený svojim kvantizačným krokom a zaokrúhlený na celé číslo. Pritom každý kvantizačný krok pre MPEG-1 sa skladá z prvku kvantizačnej matice  $Q$  a mierky kvantovania (MK). Výnimkou je JSK vnútrostránkového módu pre ktorý sa využíva len prvok kvantizačnej matice. Tento algoritmus kvantovania SK sa dá popísať nasledujúcimi rovnicami

- vo vnútrostránkovom móde

- pre striedavé SK

$$q(i, j) = 8Y(i, j) // MKQ(i, j) \quad (7)$$

- pre JSK

$$q(i, j) = Y(i, j) // 8 \quad (8)$$

- v medzisímkovom móde pre všetky SK

$$q(i, j) = 8Y(i, j) / MKQ(i, j) \quad (9)$$

Rov.(9) popisuje LSK s prázdnu zónou, ktorého prevodová charakteristika je znázornená na obr. 7, pričom  $\Delta = \frac{MKQ(i, j)}{8}$ . Kvantovaniu vo

vnútrostránkovom móde zodpovedá prevodová charakteristika na obr.6, kde kvantizačný krok zostáva rovnako definovaný okrem toho pre JSK.

Mierka kvantovania môže nadobúdať hodnoty 1÷31. Táto informácia je prenášaná v hlavičke makrobloku. Pomocou nej je možné meniť kvantizačný krok, ktorý priamo ovplyvňuje veľkosť prázdnej zóny kvantizátora. SK po kvantovaní, vyjadrené pomocou  $q(i, j)$ , sa nachádzajú v intervale  $\langle -255, 255 \rangle$ . Podobne ako pre H.261 a H.263 je potrebná obnova dynamického rozsahu SK. Táto obnova je realizovaná pomocou nasledujúcich rovníc

- JSK vnútrostránkového módu

$$\tilde{Y}(i, j) = 8q(i, j) \quad (10)$$

- všetky ostatné SK

$$\tilde{Y}(i, j) = \frac{q(i, j)MKQ(i, j)}{8} \quad (11)$$

Pre každý z módov štandardného videokodeka sú uvažované vlastné kvantizačné matice pre jasovú a farebné zložky, pričom je možné použiť pri kvantovaní súčasne tú istú kvantizačnú maticu pre obidve zložky. Na obr. 8 sú znázornené štandardné kvantizačné matice pre vnútrostránkový a medzisímkový mód videokodeka MPEG-1. Ako je vidno z tohto obrázka kvantizačná matica pre

medzisímkový mód je monotónna a všetky jej prvky majú rovnakú hodnotu. Medzinárodné odporúčanie pre videokodek MPEG-2 a MPEG-4 predpokladá už nemonotónnu kvantizačnú maticu ako to vidno z obr. 9 pre uvedený mód.

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 8  | 16 | 19 | 22 | 26 | 27 | 29 | 34 |
| 16 | 16 | 22 | 24 | 27 | 29 | 34 | 37 |
| 19 | 22 | 26 | 27 | 29 | 34 | 34 | 38 |
| 22 | 22 | 26 | 27 | 29 | 34 | 37 | 40 |
| 22 | 26 | 27 | 29 | 32 | 35 | 40 | 48 |
| 26 | 27 | 29 | 32 | 35 | 40 | 48 | 58 |
| 26 | 27 | 29 | 34 | 38 | 46 | 56 | 69 |
| 27 | 29 | 35 | 38 | 46 | 56 | 69 | 83 |

a)

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 |

b)

**Obr. 8** Štandardné kvantizačné matice pre MPEG-1

a) vnútrostránkový mód b) medzisímkový mód

**Fig. 8** Standard quantization matrices for MPEG-1

a) intraframe mode b) interframe mode

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 19 | 20 | 21 | 22 | 23 | 24 | 26 | 27 |
| 20 | 21 | 22 | 23 | 25 | 26 | 27 | 28 |
| 21 | 22 | 23 | 24 | 26 | 27 | 28 | 30 |
| 22 | 23 | 24 | 26 | 28 | 28 | 30 | 31 |
| 23 | 24 | 25 | 27 | 30 | 30 | 31 | 33 |

**Obr. 9** Nemonotónna kvantizačná matica pre medzisímkový mód videokodekov MPEG-2 a MPEG-4

**Fig. 9** Non-flat quantization matrix for interframe mode of videocoders MPEG-2 and MPEG-4

Podobne aj v týchto štandardných videokodekoch sa kvantovanie SK vykonáva pomocou kvantizačnej matice a mierky kvantovania [5,6]. Algoritmus kvantovania v MPEG-2 je popísaný nasledujúcimi rovnicami

- vo vnútrostránkovom móde

- pre striedavé SK

$$q'(i, j) = 32Y(i, j) // Q(i, j)$$

$$q(i, j) = [q'(i, j) + \text{sign}(q'(i, j))(pMK) // q] / (2MK) \quad (12)$$

$$\text{pričom} \quad \text{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \quad (13)$$

a konštanty  $p=3, q=4$ .

- pre JSK

$$q(i, j) = Y(i, j) // k, \quad k = 8, 4, 2, 1 \quad (14)$$

- v medzisímkovom móde pre všetky SK

$$q'(i, j) = 32Y(i, j) // Q(i, j)$$

$$q(i, j) = q'(i, j) / (2MK) \quad (15)$$

Kvantizačná matica vnútrostránkového módu je podobná tej pre MPEG-1 s tým rozdielom, že pre JSK môže byť použitý kvantizačný krok

s hodnotami 8,4,2,1 v rov.14. Rekonštrukcia dynamického rozsahu SK je realizovaná v MPEG-2 pomocou nasledujúcich rovníc

- **JSK vnútrošnímkového módu**

$$\tilde{Y}(i, j) = kq(i, j), \quad k = 8,4,2,1 \quad (16)$$

- **všetky ostatné SK**

$$\tilde{Y}(i, j) = ((2q(i, j) + k)Q(i, j)MK)/32 \quad (17)$$

$$\text{kde } k = \begin{cases} 0 & \text{pre vnútrošnímkový mód} \\ \text{sign}(q(i, j)) & \text{pre medzisnímkový mód} \end{cases} \quad (18)$$

Hodnoty rekonštruovaných SK v prípade, že neležia v rozsahu  $\langle -2048, 2047 \rangle$ , sa upravujú do tohto rozsahu pomocou saturácie, ktorá je súčasťou rekonštrukcie dynamického rozsahu SK. Okrem toho MPEG-2 predpokladá úpravu SK  $\tilde{Y}(7,7)$ , ktorá sa vykoná na základe celkovej sumy všetkých SK v danom bloku. MPEG-4 môže prepínať medzi dvoma algoritmi kvantovania SK [6]. Jeden algoritmus zodpovedá kvantovaniu v H.263 a druhý kvantovaniu v MPEG-2, pričom na kvantovanie JSK vnútrošnímkového módu sa predpokladá optimálne nelineárne skalárne kvantovanie.

Všetky kvantizačné matice môžu byť modifikované pre konkrétne vizuálne scény za účelom dosiahnutia čo najvyššej kompresie a najlepšej kvality dekódovanej videosekvencie.

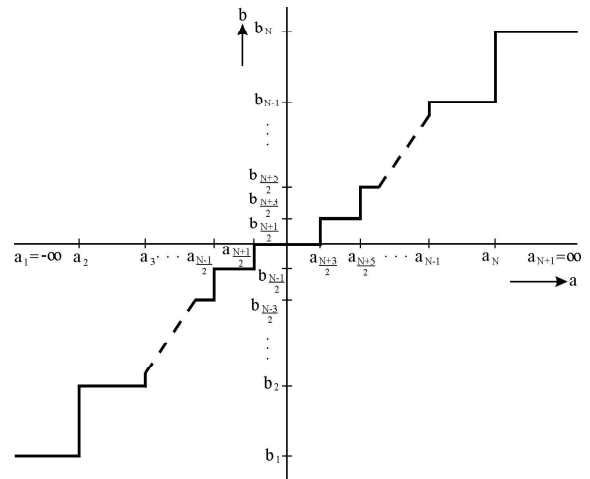
#### 4. NELINEÁRNE ALGORITMY KVANTOVANIA

Nelineárne algoritmy kvantovania sa vyznačujú väčšou účinnosťou v porovnaní s lineárnymi algoritmi. Dosiahne sa pomocou nerovnomerného rozloženia kvantizačných a rozhodovacích úrovní v prevodovej charakteristike nelineárneho skalárneho kvantizátora (NSK) [10], ktorá je tak lepšie prispôbená k rozdeleniu pravdepodobnosti  $f(x)$  vstupnej postupnosti náhodných SK. Podobne ako pre LSK väčšia účinnosť nelineárneho kvantovania sa dosiahne s využitím prázdnej zóny. Potom prevodová charakteristika NSK s prázdnu zónou je na obr. 10.

Optimalizácia NSK znamená nájdenie optimálneho rozloženia kvantizačných ( $b_i$ ) a rozhodovacích ( $a_i$ ) úrovní tak, aby stredná kvadratická hodnota kvantizačného šumu bola minimálna. Jednotlivé úrovne je možné vypočítať pomocou nasledujúcich vzťahov

$$b_i = \frac{\int_{a_{i-1}}^{a_i} x f(x) dx}{\int_{a_{i-1}}^{a_i} f(x) dx}, \quad i = 1, 2, \dots, N \quad (19)$$

$$a_i = \frac{b_i + b_{i-1}}{2}, \quad i = 2, 3, \dots, N \quad (20)$$



Obr. 10 Prevodová charakteristika NSK s prázdnu zónou

Fig. 10 The transfer characteristic of NSQ with dead zone

Všetky štandardné videokodeky používajú lineárne skalárne kvantovanie zodpovedajúce buď kvantizačnému parametru QP (rov.1,2,3) pre H.261 a H.263, alebo kvantizačným maticiam (obr.8,9) pre MPEG-1,2,4. Jedinou výnimkou je kvantovanie JSK pomocou optimálneho NSK v MPEG-4. Uvažovaním nelineárneho skalárneho kvantovania pre všetky SK je možné dosiahnuť väčšiu kompresiu údajov ako s použitím LSK. Použitím NSK v procese kvantovania sa však zvýšia nároky na technickú realizáciu. Tieto sa dajú podstatne znížiť rozdelením bloku SK do určitých zón. Každé zóny by potom zodpovedal jeden NSK. V krajnom prípade by sme mohli uvažovať len jeden NSK pre kvantovanie všetkých SK.

Vo vnútrošnímkovom móde by boli potrebné minimálne dva druhy NSK, jeden pre kvantovanie JSK a ďalší optimalizovaný pre ostatné SK. Na základe štatistickej analýzy SK v medzisnímkovom móde by sme mohli použiť na kvantovanie všetkých SK jeden NSK. V každom z uvedených módoch by bolo vhodné uvažovať nelineárne skalárne kvantovanie s prázdnu zónou z dôvodu väčšej kompresie údajov.

#### 5. IMPLEMENTÁCIA A EXPERIMENTÁLNE VÝSLEDKY VYBRANÝCH ALGORITMOV KVANTOVANIA

Na simuláciu vybraných algoritmov kvantovania štandardných videokodekov vo vnútrošnímkovom a medzisnímkovom móde bola použitá videosekvencia „Claire“ so snímkovou frekvenciou 10Hz a počtom 15 snímkov s rozmermi 288×352 op. DCT bola aplikovaná na bloky 8×8 op a v medzisnímkovom móde bola použitá trojkroková metóda estimácie pohybu bloku op [12].

Ako referenčný algoritmus pre porovnanie ostatných algoritmov kvantovania štandardných videokodekov bol použitý algoritmus kvantovania z MPEG-1. Objektívnou mierou hodnotenia bol priemerný počet bitov (PPB) na op, ktorý je

potrebný na jeho kódovanie a vypočíta sa nasledovne

$$PPB = W \cdot PB \quad (21)$$

$$\text{kde } W = \frac{\text{Počet nenulových } q(i,j) \text{ v blokoch}}{\text{Celkový počet } q(i,j)} \quad (22)$$

a PB je počet bitov potrebný na kódovanie  $q(i,j)$ . Táto miera nezohľadňuje efektívne kódovanie v ktorom sa sleduje počet nulových  $q(i,j)$  medzi dvoma nenulovými  $q(i,j)$  a kódujú sa dvojice (počet nulových  $q(i,j)$ , hodnota nenulového  $q(i,j)$ ).

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| 8 | 8 | 4 | 5 | 5 | 4 | 2 | 2 |
| 6 | 6 | 4 | 2 | 2 | 2 | 2 | 1 |
| 7 | 3 | 3 | 3 | 2 | 2 | 1 | 1 |
| 8 | 5 | 2 | 3 | 1 | 1 | 1 | 1 |
| 7 | 3 | 3 | 2 | 1 | 1 | 1 | 1 |
| 7 | 3 | 4 | 2 | 1 | 1 | 0 | 0 |
| 4 | 2 | 2 | 1 | 1 | 1 | 0 | 1 |
| 2 | 2 | 3 | 1 | 1 | 1 | 0 | 0 |

**Obr. 11** Maximálne absolútne hodnoty  $q(i,j)$  pre MPEG-1 v medzisnímkovom móde a MK=8

**Fig. 11** Maximum absolute values of  $q(i,j)$  for MPEG-1 in interframe mode and MK=8

Obr. 11 znázorňuje výsledky kvantovania pre MPEG-1 v medzisnímkovom móde a MK=8, ktorý dosiahol priemerný špičkový S/Š=41,131 dB. Hodnota MK je pre jednoduchosť uvažovaná rovnaká pre všetky bloky snímok. V štandardoch sa môže adaptívne meniť, tak aby bola čo najlepšie prispôbená pre daný blok.



**Obr. 12** Vybraná rekonštruovaná snímka videosekvencie „Claire“

**Fig. 12** The chosen reconstructed frame of the videosequence „Claire“

Pre maximálnu absolútnu hodnotu 8 z obr. 11 je potrebných včítane znamienkového bitu PB=5 bitov. Tento počet bitov sa použije pre všetky SK. Pri zohľadnení celkového počtu 7121 nenulových  $q(i,j)$  vychádza  $PPB=2,5 \cdot 10^{-2}$ . V tomto móde sa neuvažovala medzisnímková predikcia prvej snímky, pretože tá sa kóduje vo vnútrošnímkovom móde. To znamená, že celkový počet  $q(i,j)$  je určený zo 14 snímok. So zväčšujúcou sa hodnotou MK klesá PPB, ale tým aj S/Š. To je spôsobené

zväčšovaním kvantizačného kroku  $\Delta$ , ktorý priamo ovplyvňuje aj veľkosť prázdnej zóny. Pre ilustráciu dosiahnutého referenčného S/Š je znázornená na obr.12 vybraná rekonštruovaná snímka testovacej videosekvencie „Claire“, ktorej hodnota S/Š je 41,149 dB.

Pre dosiahnutie porovnateľnej hodnoty priemerného špičkového S/Š, ako v MPEG-1, videokodek H.263 bude mať  $PPB=4,13 \cdot 10^{-2}$ . Celkový počet nenulových  $q(i,j)$  je 11738 a PB=5 bitov (max hodnota  $q(i,j)$  je 10). Táto hodnota PPB je dosiahnutá pre hodnotu QP=7. Taká porovnateľná hodnota tohto S/Š sa dosiahne v MPEG-2 s  $PPB=3,28 \cdot 10^{-2}$  (MK=12) pri celkovom počte 9312 nenulových  $q(i,j)$  a PB=5bit (max. hodnota  $q(i,j)$  je 12). Z týchto výsledkov vyplýva, že v medzisnímkovom móde je pre videosekvenciu „Claire“ najúčinnější algoritmus kvantovania z MPEG-1. Videosekvencia „Claire“ je charakteru hlava – plecia, ktorá je typická pre videotelefónne a videokonferenčné systémy, kde sa používajú najmä štandardy H.261 a H.263. Algoritmus kvantovania týchto štandardných videokodekov je podobný tomu v MPEG-1, kedy kvantovanie všetkých SK sa vykonáva v jednom LSK. Zlepšenie PPB v MPEG-1 v porovnaní s H.261 alebo H.263 je dosiahnuté voľbou kvantizačného kroku, ktorého veľkosť je v MPEG-1 lepšie prispôbená vstupnej videosekvencii. Použitie nerovnomernej kvantizačnej matice nezlepšilo PPB pre danú videosekvenciu. Nerovnomerná kvantizačná matica zlepšuje účinnosť kvantovania videosekvencií s veľkým počtom detailov v snímanej scéne. Toto má vplyv aj na zväčšenie počtu nenulových  $q(i,j)$  v MPEG-2 oproti MPEG-1. Na základe štatistickej analýzy v priestore DCT a poznatkov získaných z algoritmov kvantovania štandardných videokodekov vyplýva, že postačuje v medzisnímkovom móde použiť jeden lineárny alebo nelineárny kvantizátor pre všetky SK. Vyššiu účinnosť kvantovania možno dosiahnuť pomocou nelineárneho skalárneho kvantovania s prázdnu zónou. V našom algoritme bola použitá na vytvorenie prázdnej zóny  $\langle -16,16 \rangle$  prahová filtrácia v priestore DCT s prahom 16. Potom na dosiahnutie referenčnej hodnoty S/Š bol potrebný NSK s PB=3bit. Vtedy počet nenulových SK bol 7255 a  $PPB=1,53 \cdot 10^{-2}$ . Zmenšenie PPB potvrdzuje význam náhrady LSK s NSK v štandardných videokodekoch. V tab.1 sú zhrnuté výsledky všetkých použitých algoritmov kvantovania v medzisnímkovom móde štandardných videokodekov.

Obr. 13 znázorňuje výsledky kvantovania v MPEG-1 vo vnútrošnímkovom móde a MK=8, pre ktorý sa dosiahol špičkový S/Š=41,22 dB. Túto hodnotu dosahuje MPEG-1 pre  $PPB=0,349$ , pričom počet nenulových  $q(i,j)$  je 69233 (23760 pre JSK a 45473 pre striedavé SK). Pre kódovanie JSK je potrebných 9 bitov a pre striedavé SK 7 bitov. Pomocou algoritmu kvantovania v H.263 sa porovnateľná hodnota tohto S/Š dosiahne s  $PPB=0,317$  (QP=7), pričom počet nenulových  $q(i,j)$  je 62156. Pre kódovanie 23760 JSK je potrebných 9

bitov (maximálna absolútna hodnota JSK je 235) a pre 38396 striedavých SK je potrebných 7 bitov (maximálna absolútna hodnota 43). Opäť zväčšovaním hodnoty MK klesá hodnota PPB ako aj S/Š.

| ALGORITMUS KVANTOVANIA | S/Š [dB] | Počet nenulových SK | PPB                  |
|------------------------|----------|---------------------|----------------------|
| H.261, H.263           | 41,164   | 11738               | $4,13 \cdot 10^{-2}$ |
| MPEG-1                 | 41,131   | 7121                | $2,5 \cdot 10^{-2}$  |
| MPEG-2                 | 41,163   | 9312                | $3,28 \cdot 10^{-2}$ |
| ONSK (3 bit)           | 41,158   | 7255                | $1,53 \cdot 10^{-2}$ |

**Tab. 1** Výsledky použitých algoritmov kvantovania v medzislímkovom móde štandardných videokodekov

**Tab. 1** The results of applied quantization algorithms in interframe mode of the standard videocodecs

|     |    |    |    |   |   |   |   |
|-----|----|----|----|---|---|---|---|
| 235 | 38 | 16 | 10 | 8 | 4 | 3 | 2 |
| 19  | 13 | 7  | 5  | 3 | 2 | 1 | 1 |
| 11  | 6  | 3  | 4  | 3 | 1 | 1 | 1 |
| 15  | 5  | 3  | 2  | 2 | 1 | 1 | 1 |
| 6   | 4  | 2  | 2  | 1 | 1 | 0 | 0 |
| 6   | 2  | 2  | 1  | 1 | 0 | 0 | 0 |
| 3   | 2  | 2  | 1  | 0 | 0 | 0 | 0 |
| 1   | 1  | 2  | 1  | 0 | 0 | 0 | 0 |

**Obr. 13** Maximálne absolútne hodnoty  $q(i,j)$ , pre MPEG-1 vo vnútroslímkovom móde a MK=8

**Fig. 13** Maximum absolute values  $q(i,j)$  of MPEG-1 in intraframe mode and MK=8

V MPEG-2 je potrebný pre tú istú referenčnú hodnotu priemerného špičkového S/Š PPB=0,325, pričom počet nenulových  $q(i,j)$  je v tomto prípade 63854 (23760 JSK a 40096 striedavých SK). Z maximálnej absolútnej hodnoty JSK 235 vyplýva 9 bitov a z maximálnej absolútnej hodnoty striedavých SK 41 vyplýva 7 bitov potrebných na ich kodovanie. Tento výsledok je dosiahnutý pre MK=15. Treba pripomenúť, že JSK sa kvantuje vo všetkých štandardoch rovnako. Najmenšia hodnota PPB je vo vnútroslímkovom móde dosiahnutá pomocou algoritmu kvantovania používaného v H.261 alebo H.263. V týchto videokodekoch sa používajú dva LSK, jeden pre JSK a druhý pre striedavé SK. Použitie kvantizačnej matice s nerovnomerným rozložením kvantizačných krokov v MPEG-1 neprineslo zlepšenie PPB. Je to spôsobené tým, že takéto nerovnomerné rozloženie kvantizačných krokov nie je najlepšie prispôbené skutočnému rozloženiu striedavých SK v danej testovacej videosekvencii. Zmenšenie PPB oproti MPEG-1 bolo v MPEG-2 aj napriek použitiu rovnakej kvantizačnej matice v dôsledku lepšieho algoritmu kvantovania ako aj zaokrúhľovania jednotlivých  $q(i,j)$ , ktorým sa dá čiastočne priblížiť k skutočnému rozloženiu SK. To spôsobilo aj pokles nenulových  $q(i,j)$  v porovnaní s MPEG-1. Zmenšenie PPB je možné dosiahnuť v MPEG-1 a MPEG-2 prispôbením kvantizačnej matice k

testovacej videosekvencii. Dosiahnutie porovnateľného S/Š v MPEG-2 oproti MPEG-1 je zabezpečené pre omnoho väčšie MK. Je to spôsobené použitím veľkej konštanty (32) v čitateli rov.(12). Použitie tejto konštanty má vplyv na celkový rozsah dosiahnutého špičkového S/Š v rámci možných hodnôt MK. MPEG-1 dosahuje pre maximálnu hodnotu MK omnoho menší S/Š ako MPEG-2. Vysoké hodnoty S/Š v MPEG-2 vyplývajú z toho, že tento štandard je určený pre oveľa väčšie bitové rýchlosti ako MPEG-1. Z celkového hľadiska postačuje použiť vo vnútroslímkovom móde MPEG-1 pre danú testovaciu videosekvenciu dva kvantizátory, jeden pre JSK a druhý pre striedavé SK, podobne ako je to v H.261 a H.263. Zlepšenie účinnosti algoritmov kvantovania v štandardných videokodekoch vo vnútroslímkovom móde je možné dosiahnuť pomocou nelineárneho skalárneho kvantovania s prázdnu zónou uvažovaného pre JSK a striedavé SK. Podobne ako v medzislímkovom móde predpokladá sa vytvorenie prázdnej zóny pomocou prahovej filtrácie v priestore DCT s prahom 16. Na dosiahnutie referenčnej hodnoty S/Š je potrebný NSK pre JSK s počtom 9 bit a pre striedavé SK s počtom 6bit. Pre 23760 nenulových JSK a 33669 nenulových striedavých SK dostaneme PPB= $2,73 \cdot 10^{-1}$ .

| Algoritmus kvantovania |              | S/Š [dB] | Počet nenulových SK | PPB                  |
|------------------------|--------------|----------|---------------------|----------------------|
| JSK                    | Striedavé SK |          |                     |                      |
| H.261,H.263            | H.261,H.263  | 41,521   | 62156               | $3,17 \cdot 10^{-1}$ |
| MPEG-1                 | MPEG-1       | 41,22    | 69233               | $3,49 \cdot 10^{-1}$ |
| MPEG-2                 | MPEG-2       | 41,177   | 63854               | $3,25 \cdot 10^{-1}$ |
| ONSK (9 bit)           | ONSK (6 bit) | 41,254   | 57429               | $2,73 \cdot 10^{-1}$ |

**Tab. 2** Výsledky použitých algoritmov kvantovania vo vnútroslímkovom móde štandardných videokodekov

**Tab. 2** The results of applied quantization algorithms in intraframe mode of the standard videocodecs

V tab. 2 sú zhrnuté výsledky všetkých použitých algoritmov kvantovania vo vnútroslímkovom móde štandardných videokodekov. Podobne ako v medzislímkovom móde tak aj vo vnútroslímkovom móde sa dosiahol zmenšenie PPB pomocou algoritmu nelineárneho skalárneho kvantovania s prázdnu zónou. Použitie dvoch NSK je postačujúce pre dosiahnutie potrebnej kvality kódovanej videosekvencie. Medzi prezentovanými výsledkami algoritmov kvantovania nie sú samostatne uvádzané výsledky pre MPEG-4, pretože ten môže používať algoritmy kvantovania z H.261, H.263 a MPEG-2.

## 6. ZÁVER

Tento článok sa zaoberal algoritmi kvantovania štandardných videokodekov, ktoré významnou mierou ovplyvňujú ich bitovú rýchlosť. Jeho cieľom bolo dať prehľad, včítane zhodnotenia

účinnosti týchto algoritmov a na základe toho navrhnúť efektívnejšie algoritmy kvantovania. Pritom výsledky štatistickej analýzy videosignálu v priestore DCT určujú ich vnútroštruktúrovňové a medzisíťkové modifikácie.

Pre kvantovanie SK bolo použité lineárne skalárne kvantovanie charakteristické pre jednotlivé štandardy. Bol použitý kvantizačný parameter QP zodpovedajúci H.261 a H.263 ako aj predvolené kvantizačné matice spolu s MK zodpovedajúce MPEG-1,2,4. Pre zvýšenie účinnosti kvantovania bolo nahradené lineárne kvantovanie s nelineárnym kvantovaním s prázdnu zónou. V medzisíťkovom móde sa všetky SK kvantovali jedným lineárnym alebo nelineárnym kvantizátorom. Vo vnútroštruktúrovňovom móde boli potrebné dva kvantizátory, jeden pre JSK a druhý pre striedavé SK. Použitie NSK v štandardoch prinieslo zmenšenie PPB, ktorý sa používal ako objektívna miera hodnotenia účinnosti kvantovania. Aplikácia nelineárných algoritmov kvantovania v štandardných videokodekoch zvyšuje ich kompresiu v porovnaní s doposiaľ používanými lineárnymi algoritmi.

**Táto práca bola podporená s GAV MŠ a SAV SR v projekte č. 1/0384/03.**

## LITERATÚRA

- [1] Mihalík, J.: Kódovanie obrazu vo videokomunikáciach. Mercury, ISBN 80-89061-47-8, Košice, 2001.
- [2] ITU-T Recommendation H.261. Video Coding for Audiovisual Services at p×64 kbit/s. Geneva, 1990.
- [3] ITU-T Recommendation H.263. Video Coding for Narrow Telecommunication Channels at <64 kbit/s. Geneva, 1996.
- [4] ISO/IEC 11172-2, MPEG-1. Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s (Part Video). Geneva, 1993.
- [5] ISO/IEC 13818-2, MPEG-2. Generic Coding of Moving Pictures and Associated Audio Information (Part Video). Geneva, 1995.
- [6] ISO/IEC JTC1/SC29/WG11, MPEG98/N1902, MPEG-4 Video Verification Model Version 10.0. San Jose, 1998.
- [7] Mihalík, J.: Adaptívne hybridné kódovanie obrazu. Elektrotech. časopis, 44, č.3, 1993, s.85-89.
- [8] Mihalík, J.: Adaptívne transformačné kódovanie obrazu. Slab. obzor, 52, č.11-12, 1991, s.253-257.
- [9] Chmúrny, J. - Mihalík, J.: Korelačná analýza obrazu v priestore Walshovej –Hadamardovej transformácie. Slab. obzor, 46, č.10, 1985, s.498-501.
- [10] Mihalík, J.: Číslícové spracovanie signálov. Alfa Bratislava, 1987.
- [11] Bhaskaran, V. – Konstantinides, K.: Image and Video Compression Standards. Kluwer Academic Publisher, 1995.
- [12] Mihalík, J. - Štefanišin, R. - Michalčin, V.: Medzisíťkové hybridné kódovanie videosignálov. Zborník vedeckej konferencie s medzinárodnou účasťou "Nové smery v spracovaní signálov VI", Tatranské zrub, 2002, s.223-226.

## BIOGRAPHY

**Ján Mihalík** graduated from Technical University in Bratislava in 1976. Since 1979 he joined Faculty of Electrical Engineering and Informatics of Technical University of Košice, where received his PhD degree in Radioelectronics in 1985. Currently, he is Full Professor of Electronics and Telecommunications and the head of the Laboratory of Digital Image Processing and Videocommunications at the Department of Electronics and Multimedia Telecommunications. His research interests include information theory, image and video coding, digital image and video processing and multimedia videocommunications.

**Radoslav Štefanišin** was born on 1978 in Svidník, Slovak Republic. He received the Ing. (M.Sc.) degree from Technical University of Košice in 2001. At present he is PhD student at the Department of Electronics and Multimedia Telecommunications of Technical University, Košice. His research interest includes video coding with very low bit rate.