

Videokompresia

Kováč Ondrej · Informačné technológie

16.03.2015



V príspevku sú popísané základné techniky dekorelácie vstupného obrazu pri jeho kompresii. Kompresia obrazu sa okrem techník zdrojového kódovania dosahuje aj dekoreláciou vstupného obrazu. Prvými technikami dekorelácie boli predikčné algoritmy, ktoré na základe postupnosti predchádzajúcich hodnôt jasu obrazových prvkov predpovedali hodnotu aktuálneho prvku a kódoval sa rozdiel medzi predikčnou a skutočnou hodnotou. Neskôr sa začali používať transformácie. Podstatou transformačnej dekorelácie je prevod obrazu do spektrálnej oblasti. Spektrálne koeficienty sa potom kvantujú a kódujú. Kombináciou predikčných a transformačných techník vznikli hybridné systémy.

Štandard MPEG-4 ako moderný videokodek umožňuje použitie uvedených techník a navyše je možné k obrazu pristupovať ako k hierarchicky usporiadanej koláži videoobjektov, ktoré je možné kódovať zvlášť. Najmodernejším prístupom je kompresia na základe modelového kódovania videoobjektov. Táto technika využíva model objektu, ktorý sa v čase mení. Tieto zmeny je možné zachytiť pomocou zmien polohy bodov modelu. Zmeny modelu je potom možné vysielat na prijímaciu stranu. Výhodou je to, že textúru modelu stačí poslať iba raz, napríklad v prvej snímke, a potom sa posielajú len pohybové parametre modelu. Takýmto spôsobom sa docieli vysoká úroveň kompresie.

1. Úvod do videokompresie

Cieľom kompresie je zmenšenie objemu. V prípade informačných technológií ide o proces zmenšovania objemu dát. Dôvodom na kompresiu je zvyčajne snaha o efektívnejšie využitie pamäťových médií, alebo zvýšenie prenosovej kapacity informačného kanála. Kompresia odstraňuje nadbytočné a nepodstatné časti údajov, aby zostal zachovaný ich obsah, alebo sa zmenil len v prijateľnej miere. Za počiatok kompresie dát sa dá považovať Morseova abeceda z roku 1838. Tá na zmenšenie objemu prenášaných údajov využíva rôznu dĺžku kódov pre jednotlivé znaky abecedy podľa toho, ako často sa vyskytujú v bežnom anglickom texte. V prípade kompresie obrazov alebo videa hovoríme o videokompresii. Metódy videokompresie je možné rozdeliť na stratové a bezstratové[1].

Pri bezstratovej videokompresii je požiadavka taká, že po spätnej rekonštrukcii obrazovej informácie musí byť táto identická s pôvodnou. Toto je docielené odstránením redundancie z pôvodného informačného obsahu. Redundancia je

nadbytočná informácia, ktorá sa dá obnoviť. Pri stratovej videokompresii dochádza okrem potlačenia redundancie aj k strate informácie. Typickým príkladom stratového odstránenia redundancie je kvantovanie. Využíva sa tu takzvaná prijateľná úroveň straty informácie ktorá je vyhodnocovaná subjektívnymi technikami vyhodnocovania kvality. Vo všeobecnosti stratovými technikami je dosiahnutá omnoho väčšia kompresia ako pri bezstratových technikách.

Kompresia sa dosahuje dvoma spôsobmi a v mnohých prípadoch aj ich synergiou. Prvým spôsobom je takzvaná dekorelácia vstupných údajov. Je zrejmé, že vysielané údaje obsahujú veľa redundancie vyplývajúce zo vzájomnej korelovanosti jednotlivých obrazových prvkov (op). Preto je vhodné túto koreláciu odstrániť. Dekoreláciu je možné uskutočniť pomocou diskretných ortogonálnych transformácií (DOT), predikčnými technikami alebo ich hybridnou kombináciou. Druhý spôsob kompresie spočíva v návrhu takého kódu, ktorý využíva štatistickú závislosť vstupnej postupnosti op. Takéto kódy sa nazývajú kódy s premenlivou dĺžkou kódového slova. Do tejto skupiny kódov patrí napríklad Huffmanov, Shano-Fannov kód alebo aritmetický kód. Problematike kódovania sa v tomto príspevku nebudeme venovať.

2. Kódovacie systémy

Pod pojmom kódovací systém (KS) budeme rozumieť celok pozostávajúci z algoritmov dekorelácie, kvantovania a kódovania vstupného obrazu. Základná bloková schéma je na obr.1.



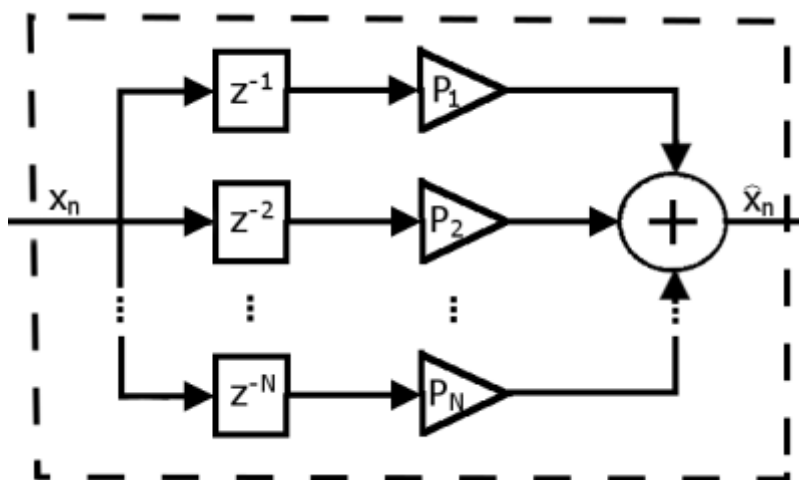
Obr.1 Všeobecná bloková schéma kódovacieho systému.

Z obr.1 je zrejmé, že každý KS obsahuje blok dekorelácie pozostávajúci z algoritmov predikcie alebo transformácie. Tieto algoritmy prevedú vstupný vysoko korelovaný obraz do priestoru predikčných chýb alebo spektrálnych koeficientov, ktoré vykazujú omnoho nižší stupeň korelácie. Kvantovanie vždy predstavuje stratovú operáciu. Je však nevyhnutné, pretože nie je možné kódovať nekonečne veľký rozsah vstupných symbolov. Treba si uvedomiť, že DOT ako aj predikcia v svojej základnej forme majú na výstupe číselné hodnoty nekonečného intervalu reálnych čísel. Kvantovaním je docielené to, že výstupná postupnosť je vo všeobecnosti z konečného intervalu celých čísel.

Z uvedeného teda vyplýva, že kvantovaním dochádza k strate informácie. Pokiaľ hodnota vstupnej premennej prevyšuje rozsah kvantizátora, tak na jeho výstupe bude maximálna hodnota jeho rozsahu. Takto dochádza ku vzniku kvantizačného šumu z preťaženia. Ostatné odchýlky vstupných hodnôt od hodnôt príslušných kvantizačných úrovní spôsobujú granulačný šum[2]. Veľkosť týchto šumov je možné chápať ako mieru straty informácie. Problematika kvantizátorov a kvantizačných šumov je komplexná a nebudeme sa jej ďalej venovať. Posledným blokom je zdrojové kódovanie[3]. Toto kódovanie môže taktiež prispievať k výslednej kompresii ale taktiež môže byť použitý rovnomerný kód, ktorý nijako neredukuje objem dát.

2.1. Predikčné kódovacie systémy

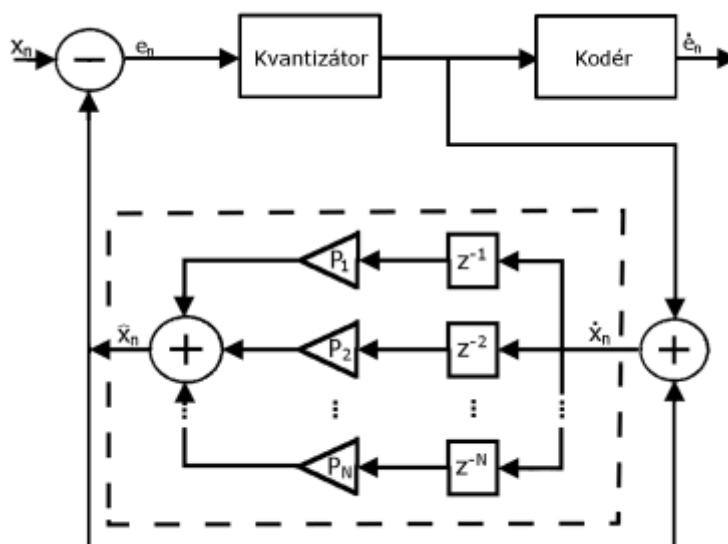
Predikčné kódovacie systémy (PKS) sú založené na predikčnej dekorelácii vstupnej postupnosti. V princípe ide o predpoveď hodnoty op na základe predchádzajúcich hodnôt op. Najjednoduchším prediktorom je jednorozmerný prediktor prvého rádu. Rozmer prediktora určuje to, aké predchádzajúce hodnoty postupnosti op sa budú uplatňovať pri predikcii nasledujúcich. Jednorozmerný (1R) prediktor bude v obrazovom priestore uvažovať iba op z toho istého riadka alebo stĺpca obrazu. Dvojrozmerný (2R) bude k predikcii používať op aj z minulých riadkov. Obdobne 3R prediktor bude k predikcii využívať aj predchádzajúce snímky ak sa jedná o videosekvenciu. Pod pojmom rád prediktora rozumieme počet predchádzajúcich op uplatnených v predikcii. Na obr.2 je bloková schéma prediktora N-tého rádu.



Obr.2 Bloková schéma prediktora N-tého rádu.

Z obr.2 je zrejmé, že predchádzajúce op sa získajú oneskorením postupnosti. Toto oneskorenie je možné dosiahnuť napríklad pomocou FIFO vyrovnávacej pamäte. Takto oneskorené op sú potom násobené predikčnými koeficientami, ktoré sa získavajú procesom optimalizácie tak, aby stredná kvadratická chyba bola minimálna.

Zapojením prediktora do spätnej väzby s odčítačkou dostávame predikčný filter, ktorý bude mať na výstupe postupnosť e_n , ktorá predstavuje postupnosť predikčných chýb. Z uvedeného vyplýva, že z vysoko korelovanej postupnosti dostávame postupnosť zmien medzi op. Je zrejmé, že tieto zmeny budú vykazovať nižší stupeň korelácie. Tieto predikčné chyby sa následne kvantujú a kódujú. Kódovanie môže byť ako už bolo uvedené s premenlivou alebo rovnakou dĺžkou kódového slova. V prípade, že je použitý kód s premenlivou dĺžkou je na výstup pripojený buffer ktorý zabezpečí konštantnú bitovú rýchlosť na výstupe. Bloková schéma všeobecného PKS je na obr.3.



Obr.3 Bloková schéma predikčného systému.

Uvedené techniky kompresie statického obrazu rozšírením o snímkovú pamäť, v ktorej sa uchovávajú predchádzajúce snímky, možno taktiež použiť na 3R predikciu. Taká predikcia sa nazýva medzisnímková predikcia. Pri tejto technike sa však musí uvažovať s tým, že objekty v scéne sa pohybujú a teda nie je možné predikovať hodnotu op na základe jeho súradnice z aktuálnej snímky v minulej snímke. Tento problém je vyriešený pohybovou kompenzáciou. Nevýhodou PKS je, že v prípade poruchy v prenosovom kanáli sa táto porucha rozšíri do celej postupnosti od op ktorý bol prijatý chybné. Obraz LENA po jednorozmernej predikčnej filtrácii prvého rádu, teda obraz predikčných chýb je zobrazený na obr.4. Pre lepšiu ilustráciu je na obr.4 zobrazený jeho negatív.

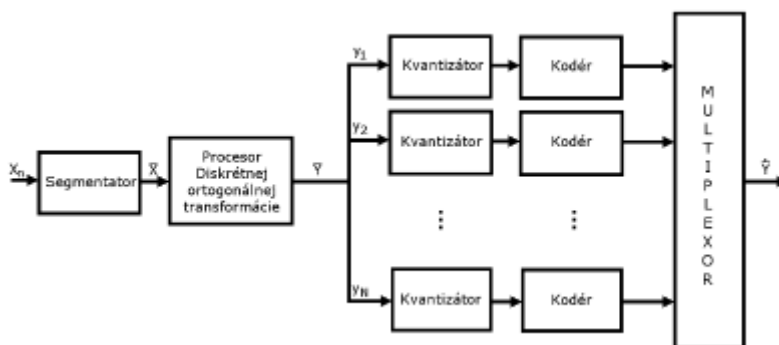




Obr.4 Pôvodný obraz LENA a jeho reprezentácia v priestore predikčných chýb.

2.2. Transformačné kódovacie systémy

Podstata kompresie transformačných kódovacích systémov (TKS) tak ako u PKS spočíva v dekorelácii vstupnej postupnosti op. Dekorelácia sa vykonáva na základe transformácie vstupnej postupnosti z časovej (priestorovej) do frekvenčnej oblasti. K tomu slúži niektorá z diskretných ortogonálnych transformácií. V štandarde JPEG je to napríklad DCT transformácia a v JPEG2000 je použitá waveletová transformácia[4]. Bloková schéma TKS je na obr. 5.

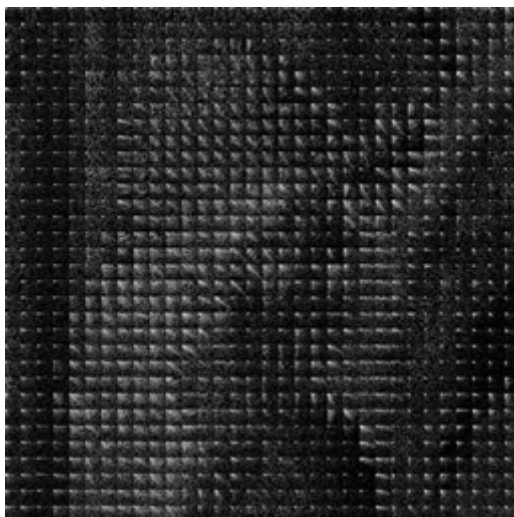


Obr.5 Bloková schéma TKS.

Ako na obr.5 vidieť, vstupná postupnosť X_n sa rozloží na bloky s rozmerom $N \times N$ op. Takto rozložený obraz je potom transformovaný niektorou z transformácií. Následne sa vo všeobecnosti každý spektrálny koeficient kvantuje a kóduje zvlášť. Paralelná reprezentácia kódových slov jednotlivých spektrálnych koeficientov sa potom multiplexorom prevedie na sériový tok. V praxi však nie je nutné kódovať všetky hodnoty spektrálnych koeficientov. Spektrálne koeficienty s vysokou frekvenciou sú zvyčajne veľmi malé a preto ich možno zanedbať tak, že sa prenáša len ľavá horná časť bloku, kde ako je známe je stlačená veľká časť energie a zvyšné koeficienty sa nahradia nulou. Tento postup sa nazýva zonálna filtrácia.

Iným spôsobom je prahová filtrácia. Tá sa vykonáva tak, že spektrálne koeficienty ktoré neprevyšujú zvolený prah sú nahradené nulou. Výhodou tohto systému je pomerne vysoký stupeň kompresie dosiahnutý filtráciou spektrálnych koeficientov pri pomerne vysokej hodnote odstupu signálu od šumu. Ďalšou výhodou oproti predikčnému systému je vysoká odolnosť voči chybám v prenosovom kanáli. Chyby as

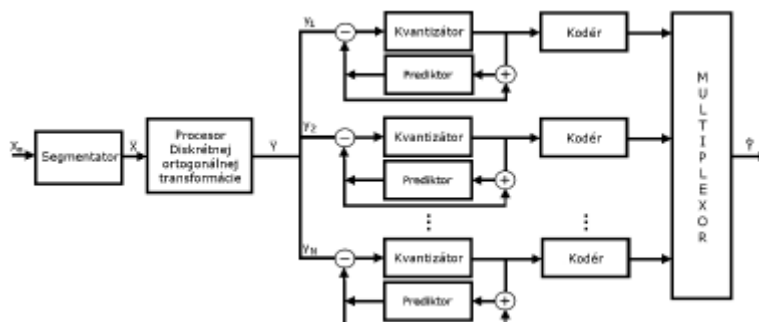
totiž prejavujú iba v bloku, v ktorom nastali. Na obr.6 je typická reprezentácia obrazu v priestore DCT spektrálnych koeficientov s rozmerom bloku 8×8 op.



Obr.6 Obraz LENA v priestore DCT spektrálnych koeficientov.

2.3. Hybridné kódovacie systémy

Hybridné kódovacie systémy (HKS) na vnútrostránkovú dekoreláciu op využívajú ako transformačné tak aj predikčné techniky. Odstraňujú nedostatky ako napríklad nízku odolnosť predikčného systému voči poruchám v prenosovom kanáli a taktiež umožňujú odstrániť medziblokovú koreláciu TKS. Bloková schéma jednoduchého HKS je na obr.7.



Obr.7 Bloková schéma hybridného kódovacieho systému.

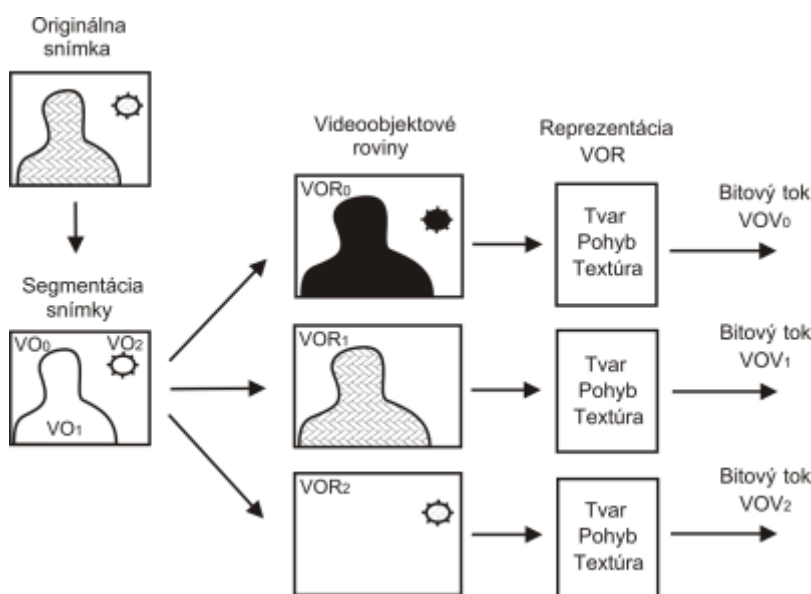
Ako je z obr.7 zrejmé, vstupná postupnosť je tak ako pri TKS rozložená na bloky ktoré sú následne transformované do blokov spektrálnych koeficientov. Potom na rozdiel od TKS sa nekódujú jednotlivé spektrálne koeficienty, ale ich predikčné chyby. Kódové slová sú následne multiplexorom prevedené na sériový tok dát.

2.4. Štandard MPEG-4

MPEG-4 je video štandard, pôvodne určený na kódovanie videa pre veľmi nízke bitové rýchlosti, neskôr však modifikovaný na kódovanie audiovizuálnych objektov[5]. Jednou z jeho základných črt je interaktívnosť na báze obsahu, ktorá prináša možnosť pristupovať k jednotlivým objektom v scéne individuálne. To umožňuje napr. kombináciu reálnych objektov so synteticky vytvorenými. V závislosti od aplikácie ponúka štandard dve skupiny prenosových rýchlostí[3]. Jadro VLBV (Very Low Bitrate Video core) poskytuje prostriedky určené pre aplikácie v rozsahu prenosových

rýchlostí 5 - 64 kb/s. Toto jadro využívajú multimediálne aplikácie v reálnom čase. Jadro HBV (High Bitrate Video core) ponúka vyššie prenosové rýchlosti v rozsahu 64 kb/s - 4 Mb/s. Takéto kódovanie je aplikované v televíznych normách, ako je napr. DVB-T2. Kódovanie reálneho videa využíva tzv. koncepciu videoobjektových rovín (VOR), ktoré predstavujú určitú časť snímky, alebo konkrétny fyzický objekt v scéne. Sekvencia VOR patriacich k tomu istému objektu sa nazýva videoobjekt (VO).

Každý VO, resp. jeho VOR je jednoznačne určený tvarom, pohybom a textúrou. Jednotlivým VOR zodpovedá príslušný bitový tok - videoobjektová vrstva (VOV), ktorý možno separátne kódovať a prenášať. To dáva možnosť flexibilnej manipulácie s jednotlivými VO[6]. Segmentácia snímky na jednotlivé VOR je zobrazená na obr.8. VOR je reprezentovaná informáciou o tvare, pohybe a textúre. V štandarde MPEG-4 je informácia o tvare VOR reprezentovaná maskami, ktoré môžu byť binárne, alebo s rozlíšením 8 bitov (použitie pre hybridné scény). Kódovanie pohybu vychádza z princípu medzismíčkovej predikcie s kompenzáciou pohybu. Tieto princípy sú definované pre VOR typu IVOR, PVOR a BVOR, pričom videoobjektové roviny IVOR a PVOR sú referenčné a BVOR sa získa obojsmernou predikciou.



Obr.8 Segmentácia snímky na VOR.

Kódovanie textúry VOR je kódovanie vnútornej časti VO. V MPEG-4 sa kódovanie textúry vykonáva pomocou transformácie DCT alebo diskkrétnej waveletovej transformácie (DWT), resp. pomocou ich modifikácií SA-DCT (Shape-Adaptive DCT) alebo SA-DWT (Shape-Adaptive DWT)[7]. Na obr. 9a je zobrazený obraz LENA po kompresii s využitím kódovania VOR. Kompresia bola vykonaná tak, že VOR s postavou Leny bol kódovaný bezstratovo a pozadie stratovou DCT s kompresiou 90%. Pre lepšiu ilustráciu na obr.9b je zobrazený obraz Leny po stratovej kompresii s približne rovnakou hodnotou odstupu signálu od šumu (S/Š) ako má obraz komprimovaný pomocou VOR. Z obr.9 je zrejmé, že subjektívna kvalita je pri kódovaní VOR vyššia, hoci kvantitatívna miera šumu je porovnateľná.



(a)



(b)

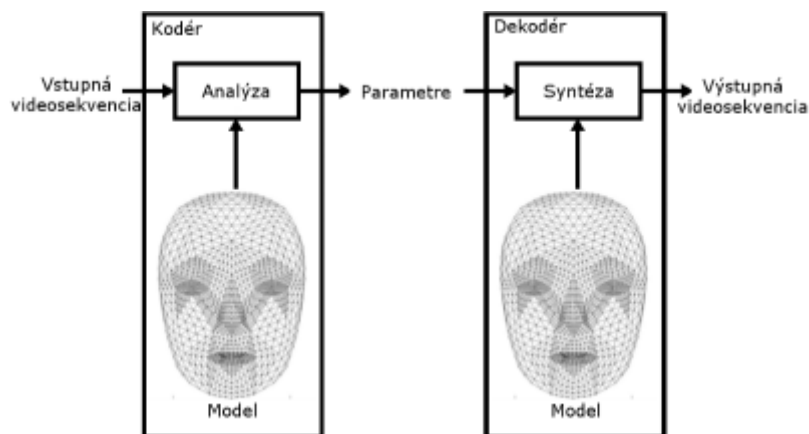
Obr.9 Porovnanie subjektívnej kvality komprimovaného obrazu a) pomocou kódovania VOR v štandarde MPEG4 a b) klasickej kompresie pomocou DCT.

Ďalším prístupom ku kódovaniu obrazu je takzvané modelové kódovanie, ktoré je popísané v štandarde MPEG-4/SNHC (Natural Hybrid Coding). Hlavnou úlohou SNHC je definovať malý počet parametrov ktoré je možné efektívne využiť pri reprezentácii ľudského výzoru a správania. Pre ľudskú hlavu sú základnými parametrami FDP (Facial Definition Parameters) a FAP (Facial Animation Parameters). FDP parametrov môže byť maximálne 84 príznakových bodov, ktoré určujú geometrický tvar modelu ľudskej hlavy. Maximálny počet FAP parametrov je 68 z čoho 3 sú parametre globálneho pohybu a tieto určujú zmenu polohy vrcholov animačných jednotiek alebo komplexného 3R modelu ľudskej hlavy[8]. Modelové kódovanie v štandarde MPEG-4/SNHC zabezpečuje vysokú úroveň kompresie a vysielanie v reálnom čase pri nízkych prenosových rýchlostiach.

2.5. Modelové kódovanie

Klasický prístup ku kompresii obrazových údajov ako už bolo uvedené spočíva v odstránení redundancie, ktorá je dôsledkom vzájomnej korelovanosti op. Táto korelácia môže byť vnútrostránková, kedy sú vzájomne korelované iba op danej snímky alebo medzisnímková, kde dochádza ku korelácii op daných snímok v čase. Účinnosť kódovania je možné zvýšiť pomocou modelovania objektov[9 -11]. Tento proces

pozostáva z identifikácie objektu v obraze a následnej analýzy, tak ako je zobrazené na obr.10.



Obr.10 Bloková schéma modelového kódovania a dekódovania obrazu.

Modely vstupujúce do analýzy môžu byť špecifické alebo všeobecné. Pod pojmom špecifický model budeme rozumieť taký model, ktorý má jasne určenú štruktúru. Táto štruktúra je daná v priestore pomocou vzájomne prepojených vrcholov. Toto prepojenie pozostáva najmenej z troch vrcholov. Takto vzniknutú mriežku budeme nazývať triangulárna sústava. Výhodou popisovania modelu pomocou trojuholníkov je to, že trojuholník v priestore vždy predstavuje rovinu. Je zrejme, že textúry je jednoduchšie prenášať na rovinu ako na iný druh plochy, ktorá vznikne napr. pri použití štyroch vrcholov. Výsledkom analýzy potom budú parametre, ktoré budú reprezentovať škálu modelu, jeho rotáciu, pohyb a podobu.

Tieto parametre budú privedené na vstup dekodéra, ktorý na základe týchto parametrov vykoná syntézu objektov snímky. Výhodou tohto systému je, že obrazové údaje v tvare textúr stačí poslať iba raz, napríklad v prvej snímke. Potom sa už posielajú iba parametre modelov. Výsledná prenosová rýchlosť je potom omnoho menšia ako pri posielaní kompletnej videosekvencie. Hlavným rozdielom medzi PKS, TKS, HKS a modelovým kódovaním je teda to, že k videosekvencii neprístupuje ako k sledu snímok pozostávajúcich z op, ktoré je treba dekorelovať a kódovať, ale snímky sa na základe tvarovaných modelov, pohybových parametrov modelov a textúr generujú na strane prijímača.

3. Záver

Kompresia predstavuje proces zmenšenia objemu dát určený pre efektívne využitie prenosového kanála a kapacity pamäťových médií. Ako bolo uvedené, odstránenie redundancie sa dosahuje dekoreláciou vstupných dát a kompresný pomer ďalej zvyšuje použitím kódu s premennou dĺžkou kódového slova. Použitie kvantovania má za následok určitú stratu informačného obsahu, keďže rozsah výstupných hodnôt je v tomto prípade konečný. Dekorelácia môže prebiehať v dvoch oblastiach - časovej (priestorovej) alebo transformačnej oblasti. V časovej sa využívajú predikčné kódovacie systémy PKS, pri návrhu ktorých sa kladie dôraz na takú optimalizáciu predikčných koeficientov prediktora, pri ktorých rozdielový obraz bude obsahovať najmenšiu energiu.

Na druhej strane boli uvedené transformačné kódovacie systémy. TKS na dekoreláciu

používajú diskkrétne ortogonálne transformácie, ktorých koeficienty predstavujú rozloženie energie jednotlivých frekvenčných zložiek vstupného signálu. Keďže TKS spracováva bloky, v prípade poruchy v prenosovom kanáli sa táto porucha nerozšíri do celej postupnosti, ako je to u PKS. Špeciálny prípad predstavujú hybridné kódovacie systémy HKS, ktoré využívajú ako transformačné, tak aj predikčné techniky. Ich výhodou je odstránenie nízkej odolnosti PKS voči poruchám v prenosovom kanáli a medziblokovej korelácie v TKS.

Opisovaným video štandardom, ktorý využíva HKS je MPEG-4, ktorý vďaka koncepcii videoobjektových rovín VOR dosahuje vyššiu subjektívnu kvalitu pri porovnateľnej miere šumu s predchádzajúcimi MPEG video štandardami. Na záver bolo uvedené, že najvyšší kompresný pomer pri kódovaní videa možno dosiahnuť modelovým kódovaním definovaným v štandarde MPEG-4/SNHC, kde sa prenášajú len parametre daného modelu, ktoré reprezentujú jeho deformácie a pohyb.

Literatúra

1. J. MIHALÍK: Kódovanie obrazu vo videokomunikáciach, LČSOV FEI TU Košice, 2001.
2. B. WIDROW, I. KOLLÁR: Quantization Noise Round off Error in Digital Computation, Signal Processing, Control, and Communications, Englewood Cliffs, NJ: Prentice-Hall, 2008.
3. D. LEVICKÝ: Multimediálne telekomunikácie: multimédia, technológie a vodoznaky, Košice, Elfa, 2002.
4. KOVÁČ, O., VALISKA, J., RUMAN, K.: Application binary state arithmetic encoding in JPEG, Proceedings from conference : 12th Scientific Conference of Young Researchers : May 15th, 2012, Herľany, Slovakia. - Košice : TU, p. 295-298, 2012.
5. MIHALÍK, J.: Štandardný videokodek MPEG-4. Slaboproudý obzor, roč. 60, č.2, s.7-11, 2003.
6. RICHARDSON, Lain E. G.: H.264 and MPEG-4 video compression. Aberdeen: The Robert Gordon University, 2003.
7. SCHAAR, M., TURAGA, D.S., STOCKHAMMER, T.: MPEG-4 beyond conventional video coding. Morgan&Claypool, 2006.
8. JIALIN ZHONG: Flexible face animation using MPEG-4/SNHC parameter streams, Image Processing, 1998. ICIIP 98. Proceedings. 1998 International Conference on , vol.2, no., p.924-928. 1998.
9. MIHALÍK, J., MICHALČIN, V.: 3D Motion Estimation and Texturing of Human Head Model. Radioengineering, Vol.13, No.1, p.26- 31, 2004.
10. KOVÁČ, O., MIHALÍK, J.: Estimácia priestorových súradníc modelu ľudskej hlavy na báze dvoch ortogonálnych pohľadov. In: Electrical Engineering and Informatics 4, Proc. of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. FEI TU, s. 401- 405, 2013.
11. MIHALÍK, J., KOVÁČ, O.: Číslkové spracovanie textúr ľudskej hlavy. LDIPV, FEI TU Košice, 2014.

Spoluautormi článku sú Ing. Vladimír Hajduk, Laboratórium progresívnych komunikačných technológií a Prof. Ing. Ján Mihalík, CSc. Laboratórium číslicového spracovania obrazov a videokomunikácií, Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach.

