

GENEROVANIE TEXTÚROVEJ BÁZY ĽUDSKEJ TVÁRE

Ján Mihalík

Laboratórium číslicového spracovania obrazov a videokomunikácií

Katedra elektroniky a multimediálnych telekomunikácií, Fakulta elektrotechniky a informatiky TU Košice

jan.mihalik@tuke.sk

Abstrakt

Článok pojednáva o generovaní textúrovej bázy ľudskej tváre. Navrhujeme spôsob jej generovania na základe predspracovania tréningového súboru ľudských tvárí a následnej PCA (Principal component analysis). Tento spôsob ju umožňuje generovať pre ľudské tváre s pohybom alebo bez pohybu, nezávislé od ich geometrického tvaru ako aj podmienok osvetlenia. Nakoniec poukazujeme na možnosti jej aplikácie v algoritmoch číslicového spracovania vizuálneho objektu ľudskej hlavy pre štandardy MPEG-4 a MPEG-7.

Kľúčové slová: textúrová báza, ľudská tvár, predspracovanie, 3R model, analýza relevantných textúr, MPEG-4, MPEG-7

Abstract

This article deals about generating texture basis of the human face. We propose a way of its generating based on preprocessing training set of human faces and following PCA (Principal component analysis). This method allows it to generate for human face with or without motion, independent from their geometry and light conditions. Finally, we point to the possibility of its application in algorithms of digital processing of visual object - human head – for the standard MPEG-4 and MPEG-7.

Key words: texture basis, human face, preprocessing, 3D model, relevant texture analysis, MPEG-4, MPEG-7

1 Úvod

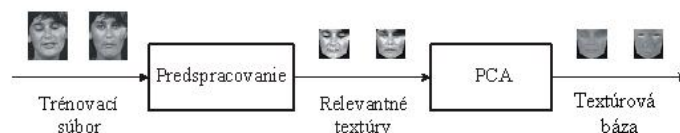
V súčasnosti moderné informačné a komunikačné technológie veľmi často využívajú vizuálne informácie v digitálnej forme. Štandardné videokodeky umožňujú získavať digitálnu reprezentáciu vizuálnych scén. Pritom videokodek MPEG-4 [1] na rozdiel od predchádzajúcich štandardov MPEG-1 a MPEG-2 umožňuje kódovanie jednotlivých videoobjektov namiesto celých vizuálnych scén a preto podporuje rozvoj aplikácií s obsahovou (objektovou) interakciou. Pritom efektívna reprezentácia videoobjektov, manipulácia s nimi, editovanie v ich bitových tokoch ako aj možnosť ich škálovania dávajú úplne nové druhy tejto interakcie. Tým vznikajú nové oblasti aplikácií ako sú: video na Internete, mobilné multimediálne komunikácie a služby, obsahový (objektový) prístup do vizuálnych databáz, interaktívna TV ako aj hry a kiná, atď.

SNHC (Synthetic Natural Hybrid Coding) [2] je podskupinou MPEG-4, ktorá sa špecializuje na kódovanie grafických modelov reálnych alebo virtuálnych 3R objektov. Štandardizácia tohto kódovania rozširuje rozsah pôvodných aplikácií MPEG-4, pretože umožňuje kombinovať reálne a syntetické objekty vo virtuálnom prostredí. Takýmto 3R objektom je aj reálna ľudská hlava pre ktorú SNHC zavádza nové algoritmy kódovania založené na jej klonovaní pomocou geometrických modelov, animácie ako aj textúrovania [3].

Na druhej strane MPEG-7 [4] pozostáva zo štandardných nástrojov pre popisovanie (deskripciu) multimediálneho obsahu za účelom jeho rýchleho a efektívneho vyhľadávania, indexovania, filtrácie ako aj sprístupnenia. Pomocou algoritmov MPEG-7 možno vytvoriť taký popis vizuálneho obsahu, ktorý bude dávať informácie o jeho vytvorení, ale aj podmienkach (právkach) jeho použitia. Tiež štruktúrálnu informáciu o jeho priestorových, časových a priestorovo-časových segmentoch. Potom aj informácie o takých atribútoch akými sú farba, textúry, pohyb, geometria

vizuálnych objektov, včítane ich kompozičnej informácie so vzájomnou interakciou.

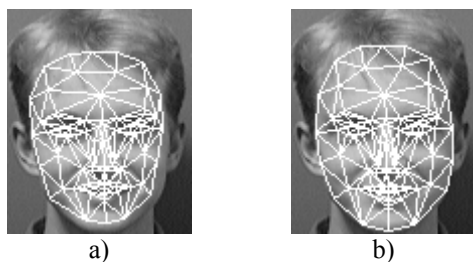
Mnoho algoritmov týchto štandardov pri číslicovom spracovaní vizuálneho objektu ľudskej hlavy je založených na použití textúrovej bázy jej tváre. Navrhujeme spôsob jej generovania na základe predspracovania tréningového súboru ľudských tvárí a následnej PCA (Principal component analysis), ako to vidno z **obr. 1**. Tento spôsob ju umožňuje generovať pre ľudské tváre s pohybom alebo bez pohybu, nezávislé od ich geometrického tvaru ako aj podmienok osvetlenia.



Obr. 1. Princíp generovania textúrovej bázy ľudskej tváre.

2 Predspracovanie ľudskej tváre

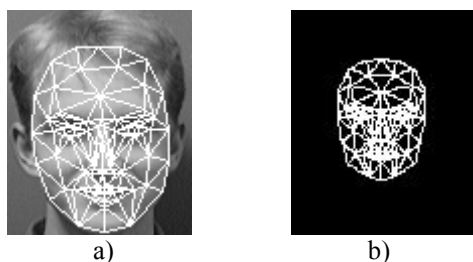
Trénovací súbor ľudských tvárí obsahuje ľudské tváre rôznych ľudí vo všeobecnosti získané pri inom osvetlení, ktoré môžu byť navyše v pohybe. V procese predspracovania sa najprv adaptuje 3R model [5] na ľudskú tvár v každom obraze z tréningového súboru pomocou poloaautomatickej metódy. Táto metóda najprv pomocou afinnej kalibrácie [6] upraví veľkosť a umiestni tento model (**obr. 2a**), a potom sa ručne koriguje pozícia len niektorých jeho vrcholov (**obr. 2b**).



Obr. 2. a) Kalibrácia 3R modelu, b) korekcia 3R modelu.

2.1 Geometrické normovanie ľudskej tváre

Následne sa vykoná v procese predspracovania geometrické normovanie ľudskej tváre, aby sa odstránili textúrové zmeny spôsobené globálnym a lokálnym pohybom, ako aj geometrickými rozdielmi medzi jednotlivcami. Geometrické normovanie je vlastne nelineárna transformácia, ktorá transformuje zdrojový obraz ľudskej tváre, na ktorej je už umiestnený prispôbený 3R model ľudskej hlavy, ako to vidno z obr. 3a na štandardný tvar daný priamym pohľadom 3R modelu s určitou mierkou (obr. 3b).

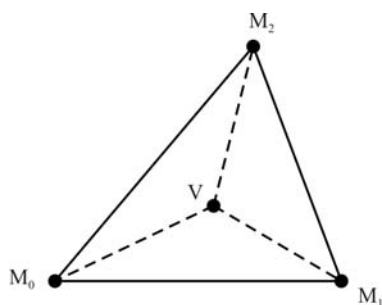


Obr. 3. a) Originálny obraz s prispôbeným 3R modelom, b) štandardný tvar 3R modelu.

Po geometrickom normovaní ľudskej tváre nasleduje proces textúrovania, pri ktorom budeme používať barycentrické súradnice, ako aj algoritmus riadkového rozkladu (scan-line algoritmus).

2.2 Definícia a výpočet barycentrických súradníc

Uvažujme jeden trojuholník $T = [M_0 M_1 M_2]$ na obr. 4 štandardného tvaru 3R modelu ľudskej hlavy. Potom polohu ľubovoľného bodu V v tomto trojuholníku možno vyjadriť



Obr. 4. Trojuholník $T = [M_0 M_1 M_2]$ štandardného tvaru 3R modelu spolu s jeho bodom V a vnútornými subtrojholníkmi.

pomocou jeho barycentrických súradníc $\lambda_0(V)$, $\lambda_1(V)$, $\lambda_2(V)$ takto

$$V = \sum_{i=0}^2 \lambda_i(V) M_i, \text{ pričom } \sum_{i=0}^2 \lambda_i(V) = 1 \quad (1)$$

a M_0 , M_1 , M_2 sú vrcholy tohto trojuholníka. Barycentrické súradnice bodu V tohto trojuholníka sú dané pomerom plochy jednotlivých subtrojholníkov (obr. 4) k ploche celého trojuholníka, a preto ich vypočítame takto

$$\begin{aligned} \lambda_0(V) &= \frac{d(V, M_1, M_2)}{d(M_0, M_1, M_2)} \\ \lambda_1(V) &= \frac{d(M_0, V, M_2)}{d(M_0, M_1, M_2)} \\ \lambda_2(V) &= \frac{d(M_0, M_1, V)}{d(M_0, M_1, M_2)} \end{aligned} \quad (2)$$

kde determinant

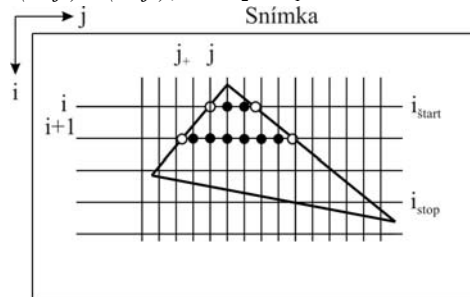
$$d(M_0, M_1, M_2) = \det \begin{pmatrix} 1 & 1 & 1 \\ u_0 & u_1 & u_2 \\ v_0 & v_1 & v_2 \end{pmatrix} \quad (3)$$

a podobne sa vypočítajú aj ostatné determinanty v čitateľoch, kde vymeníme iba jednotlivé vrcholy trojuholníka za bod $V = (u, v)^T$. Potom aj vrcholy trojuholníka sú vyjadrené stĺpcovými vektormi $M_0 = (u_0, v_0)^T$, $M_1 = (u_1, v_1)^T$, $M_2 = (u_2, v_2)^T$. Barycentrické súradnice sú použité k lokalizácii bodu V v trojuholníku $T = [M_0 M_1 M_2]$ nezávisle od súradnicového systému, v ktorom je tento trojuholník popísaný, t.j. nemenia sa v závislosti od zmien súradníc jeho vrcholov. Svoje hodnoty nemenia ani pri spoločnom pohybe alebo rotácii bodu V a trojuholníka v rovnakom smere.

2.3 Algoritmus riadkového rozkladu

Pomocou algoritmu riadkového rozkladu je možné získať súradnice vnútorných bodov štandardného tvaru 3R modelu, ak sú známe súradnice premietnutých vrcholov. Pre tieto vnútorné body zároveň musí platiť, že sú bodmi ortogonálneho rastra, na ktorom potrebujeme zobraziť textúry postupne pre op.

Na obr. 5 je znázornená snímka s ortogonálnym rastrom a premietnutým trojuholníkom 3R modelu, vo vnútri ktorého sú op vyznačené čiernymi bodmi. Každá strana trojuholníka v tejto snímke je v podstate úsečka, ktorá je daná svojimi vrcholmi (i_1, j_1) a (i_2, j_2) , kde $i_2 > i_1$.



Obr. 5. Snímka s ortogonálnym rastrom a premietnutým trojuholníkom.

Potom rovnica priamky, na ktorej táto úsečka leží bude mať tento tvar

$$i = i_1 + \frac{1}{k}(j - j_1) \quad (4)$$

kde (i, j) sú súradnice priesečníka (biely bod) tejto priamky s i -tým riadkom rastra snímky a pre k platí

$$k = \frac{j_2 - j_1}{i_2 - i_1} \quad (5)$$

Potom pre nasledujúci riadok $i + 1$ ortogonálneho rastra bude

$$i + 1 = i_1 + \frac{1}{k}(j_+ - j_1) \quad (6)$$

kde j_+ je súradnica stĺpca priesečníka v riadku $i + 1$. Odčítaním rov. (1) od rov. (3) a po úprave dostaneme iteratívny vzťah

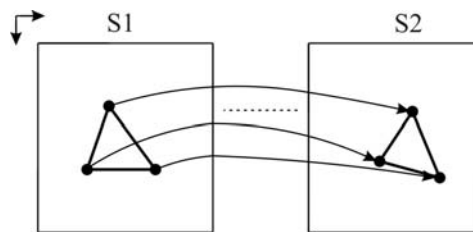
$$j_+ = j + k \quad (7)$$

pomocou ktorého môžeme postupne určiť j_+ súradnicu priesečníka každého nasledujúceho riadka ortogonálneho rastra zväčšeného o jednotku ($i + 1$) a príslušnej strany trojuholníka. Obdobným spôsobom sa získajú súradnice priesečníkov na protiahlnej strane trojuholníka. Je nutné poznamenať, že priesečníky ako aj vrcholy premietnutého trojuholníka nemusia ležať na celočíselnom ortogonálnom rasti.

Algoritmus riadkového rozkladu pre jeden trojuholník je potom vyjadrený pomocou rov. (2) až (4) a začína riadkom i_{start} a končí riadkom i_{stop} . Najprv sme vypočítali pre každý riadok konštantu k z rov. (2). V prvom riadku i_{start} sme vypočítali pomocou rov. (3) súradnice j oboch priesečníkov príslušných protiahlých strán trojuholníka. Potom sme pomocou iteratívneho vzťahu (4) vypočítali priesečníky postupne pre každý nasledujúci riadok. Všetky body v príslušnom riadku s celočíselnou súradnicou j medzi priesečníkmi sú op ležiace vo vnútri tohto trojuholníka. Týmto spôsobom získame súradnice všetkých op ležiacich vo vnútri premietnutého trojuholníka zo znalosti súradníc jeho troch vrcholov. Rovnako algoritmus riadkového rozkladu sa aplikuje na ľubovoľný trojuholník premietnutého štandardného tvaru 3R modelu ľudskej hlavy.

2.4 Textúrovanie štandardného geometrického tvaru 3R modelu

Textúrovanie štandardného geometrického tvaru 3R modelu predstavuje proces, v ktorom sa textúrové hodnoty z originálneho obrazu ($S1$) prenášajú na štandardný tvar ($S2$) 3R modelu. Medzi vrcholmi trojuholníkov štandardného tvaru a vrcholmi príslušných trojuholníkov z originálneho obrazu existuje vzťah, ako je to znázornené na **obr. 6**. Takto sa textúry z vnútra každého trojuholníka v $S1$ pretransformujú do vnútra príslušného trojuholníka v $S2$.



Obr. 6. Vzťah príslušných trojuholníkov a prenos textúr z $S1$ do $S2$.

Ak sú známe priemety 3R modelu v $S1$ a $S2$, z ktorých vyplývajú hodnoty súradníc vrcholov ich korešpondujúcich trojuholníkov navrhnutý algoritmus textúrovania pozostáva z týchto štyroch krokov

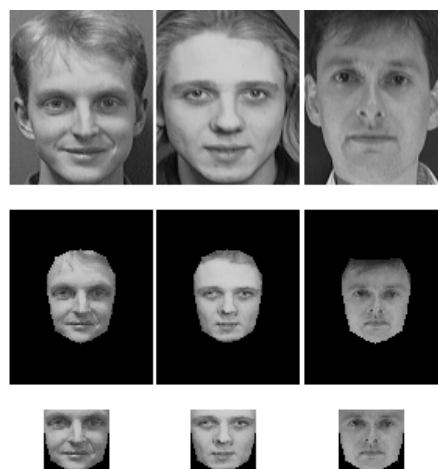
1) *Pomocou rozkladového algoritmu určíme body ortogonálneho rastra vo vnútri jednotlivých trojuholníkov priemetu 3R modelu v $S2$.*

2) *Pre tieto body na základe známych súradníc vrcholov príslušných trojuholníkov vypočítame ich barycentrické súradnice.*

3) *Z týchto vypočítaných barycentrických súradníc a známych súradníc vrcholov trojuholníkov priemetu 3R modelu v $S1$, ktoré zodpovedajú jednotlivo trojuholníkom v $S2$, sa určí poloha korešpondujúcich vnútorných bodov v $S1$. Pritom reálne hodnoty ich kartézskych súradníc sa zaokrúhľujú na celé čísla, aby ležali na ortogonálnom rasti, tak ako v $S2$.*

4) *Prenos textúrových hodnôt op vo vypočítaných bodoch v $S1$ do korešpondujúcich bodoch v $S2$.*

Geometricky normované textúry (tváre) sú viac podobné ako originálne obrazy, pretože zodpovedajú modelu s tým istým geometrickým tvarom. Výsledky textúrovania sú zobrazené na **obr.7** (v strede). Po získaní geometricky normovanej tváre (textúry) je dôležité vybrať jej relevantnú textúru na **Obr. 6** (dole), ktorá obsahuje najdôležitejšie časti ľudskej tváre (oči, ústa, nos...).



Obr. 7. (hore) Originálne obrazy ľudských tvárí, (v strede) geometricky normované textúry, (dole) relevantné textúry.

2.5 Centrovanie a energetické normovanie relevantných textúr

Tým, že relevantné textúry môžu byť získané z obrazov ľudských tvárí pri rôznom osvetlení, je potrebné minimalizovať zmeny tohto osvetlenia ešte pred aplikovaním PCA. Preto za účelom tejto minimalizácie je relevantná textúra na výstupe bloku predspracovania ešte centrovaná a energeticky normovaná. Potom hodnoty relevantnej textúry \mathbf{X} po úplnom predspracovaní budú

$$\mathbf{X} = \frac{\mathbf{Y} - \alpha}{\sigma} \quad (8)$$

kde \mathbf{Y} sú hodnoty necentrovanej relevantnej textúry bez energetického normovania. Konštanta α je stredná hodnota a σ smerodajná odchýlka textúry. Tieto konštanty sa vypočítajú takto

$$\alpha = \frac{1}{n} \sum_{i=1}^n Y_i \quad (9)$$

$$\sigma = \sqrt{\sum_{i=1}^n (Y_i - \alpha)^2} \quad (10)$$

kde n je počet prvkov (hodnôt) Y_i relevantnej textúry. Všetky relevantné textúry po úplnom predspracovaní sa vyznačujú tým, že majú nulový strednú hodnotu a jednotkovú disperziu.

3 Analýza relevantných textúr

Analýza relevantných textúr ľudských tvárí z tréningového súboru za účelom získania textúrovej bázy (vlastné komponenty) korešponduje s Hotellingovou transformáciou (HT) alebo Karhunenovou-Loèveovou transformáciou (KLT). Je známa ako Principal component analysis – PCA [7], výsledkom ktorej je báza vlastných komponentov, ktorá predstavuje textúrovú bázu. Teoreticky sme rozpracovali PCA a následne sme aplikovali odvodené výsledky na získanie textúrovej bázy ľudskej tváre bez pohybu a s pohybom.

Vo všeobecnosti relevantné textúry sú vyjadrené v tvare matice \mathbf{X} , ale pre potreby PCA je ich výhodnejšie preusporiadať po riadkoch (len nenulové hodnoty) do tvaru stĺpcového vektora $\bar{\mathbf{x}}$ o rozmere M . Potom celý tréningový súbor relevantných textúr $\bar{\mathbf{x}}_j$, $j = 1, 2, \dots, N$ možno zapísať pomocou matice $\mathbf{F} = [\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_N]$ o rozmere $M \times N$, t.j. obsahuje N M -rozmerných stĺpcových vektorov relevantných textúr. Následne jej korelačná matica \mathbf{C} sa vypočíta takto

$$\mathbf{C} = \frac{1}{N} \mathbf{F} \mathbf{F}^T \quad (11)$$

Táto matica \mathbf{C} o rozmere $M \times M$ je symetrická a má tvar

$$\mathbf{C} = \begin{bmatrix} c_{11}^2 & c_{12} & \cdots & c_{1M} \\ c_{21} & c_{22}^2 & \cdots & c_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M1} & c_{M2} & \cdots & c_{MM}^2 \end{bmatrix} \quad (12)$$

kde c_{ii}^2 je vlastný korelačný moment i -tej zložky a c_{ij} je korelačný moment medzi i -tou a j -tou zložkou stĺpcových vektorov relevantných textúr, pričom platí $c_{ij} = c_{ji}$.

Vlastné vektory $\bar{\mathbf{u}}_i$, $i = 1, 2, \dots, M$ korelačnej matice \mathbf{C} vypočítame z nasledovnej rovnice

$$\mathbf{C} \bar{\mathbf{u}}_i = \beta_i \bar{\mathbf{u}}_i \quad (13)$$

kde β_i sú vlastné hodnoty prislúchajúce vlastným vektorom $\bar{\mathbf{u}}_i$. Potom tieto vektory budú stĺpcami transformačnej matice $\mathbf{U} = [\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \dots, \bar{\mathbf{u}}_M]$, ktorá predstavuje textúrovú bázu ľudskej tváre. Z rov. (13) vidno, že výpočet vlastných vektorov vyžaduje maticové operácie, čo môže spôsobiť zložitý výpočet pre veľké hodnoty M . Ak je počet N stĺpcových vektorov relevantných textúr podstatne menší ako ich rozmer M , je možné znížiť túto výpočtovú náročnosť aplikáciou SVD (Singular Value Decomposition) [8]. SVD umožní vyjadriť vlastné vektory matice \mathbf{C} ako lineárnu kombináciu vlastných

vektorov $\bar{\mathbf{g}}_i$, $i = 1, 2, \dots, N$ matice $\mathbf{H} = \frac{1}{N} \mathbf{F}^T \mathbf{F}$, ktoré dostaneme z tejto rovnice

$$\mathbf{H} \bar{\mathbf{g}}_i = \lambda_i \bar{\mathbf{g}}_i \quad (14)$$

kde λ_i sú vlastné hodnoty prislúchajúce vlastným vektorom $\bar{\mathbf{g}}_i$, ktoré sú stĺpcami matice $\mathbf{G} = [\bar{\mathbf{g}}_1, \bar{\mathbf{g}}_2, \dots, \bar{\mathbf{g}}_N]$. Potom vlastné vektory $\bar{\mathbf{u}}_i$ budú

$$\bar{\mathbf{u}}_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{F} \bar{\mathbf{g}}_i \quad (15)$$

Matica \mathbf{U} má M vlastných vektorov a hodnôt, ale ak je vypočítaná pomocou SVD, tak má iba N nenulových vlastných hodnôt a vektorov, ostatné $(M - N)$ sú nulové. Výpočet vlastných vektorov matice \mathbf{C} o rozmere $M \times M$ sa zredukuje na výpočet vlastných vektorov matice \mathbf{H} o rozmere $N \times N$, pričom sa predpokladá $N < M$.

Rozklad relevantnej textúry $\bar{\mathbf{x}}$ ľubovoľnej ľudskej tváre, vo všeobecnosti aj mimo tréningového súboru, pomocou takto získanej textúrovej bázy je daný pomocou priamej a spätnej KLT s transformačnou maticou \mathbf{U}

$$\begin{aligned} \bar{\mathbf{y}} &= \mathbf{U}^T \bar{\mathbf{x}} \\ \bar{\mathbf{s}} &= \mathbf{U} \bar{\mathbf{y}} \end{aligned} \quad (16)$$

kde $\bar{\mathbf{y}}$ je vektor spektrálnych koeficientov a $\bar{\mathbf{s}}$ je stĺpcový vektor aproximovanej relevantnej textúry.

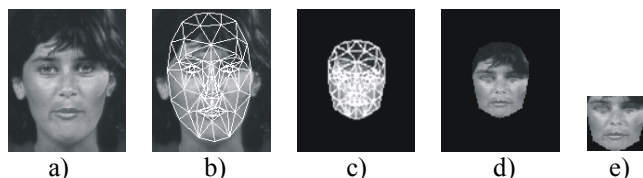
Pritom rozdiel medzi relevantnou textúrou a jej aproximáciou bude minimálny v zmysle kvadratickej chyby

$$SSE = \|\bar{s} - \bar{x}\|^2 \quad (17)$$

čo predstavuje kvadratickú hodnotu Euklidovskej normy rozdielového vektora $\bar{s} - \bar{x}$.

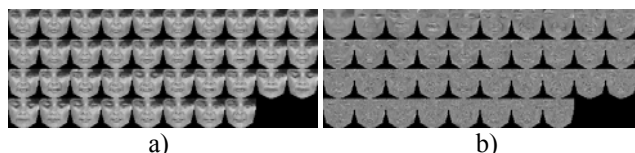
4 Textúrová báza ľudskej tváre s pohybom

Je získaná z obrazov ľudských tvárí v tréningovom súbore, ktoré zachytávajú čo najviac pohybových zmien. V našich experimentoch sme použili tréningový súbor, ktorý bol získaný decimáciou snímok s faktorom 4 videosekvencie „Miss America“, čo je 38 obrazov ľudských tvárí ($N = 38$). Na obrazy s translačným a rotačným pohybom v tomto tréningovom súbore bol adaptovaný 3R model ľudskej hlavy (**obr. 8b**) pomocou poloautomatickej metódy. Následne bolo realizované geometrické normovanie tohto adaptovaného modelu na štandardný tvar daný priamym pohľadom 3R modelu ľudskej hlavy so zatvorenými očami a ústami s určitou mierkou (**obr. 8c**). Takto normovanému modelu zodpovedá geometricky normovaná textúra na **obr. 8d**, ktorej relevantná časť bola centrovaná a energeticky normovaná. Po úplnom predspracovaní relevantná textúra na **obr. 8e** má veľkosť 40×42 obrazových prvkov (op). Takto získaná relevantná textúra bola z maticovej formy preusporiadaná do stĺpcového vektora \bar{x}_j , ktorý obsahuje $M = 1441$ zložiek. Nakoniec pomocou PCA tréningového súboru relevantných textúr bola získaná textúrová báza ľudskej tváre v pohybe.



Obr. 8. a) Originálny obraz ľudskej tváre, b) originálny obraz s adaptovaným modelom, c) štandardný tvar modelu s textúrou, d) geometricky normovaná textúra, e) relevantná textúra.

Obr. 9 ukazuje 38 relevantných textúr získaných z obrazov v tréningovom súbore a tiež textúrovú bázu 38 vlastných ľudských tvárí (vlastné vektory).



Obr. 9. a) Súbor 38 relevantných textúr, b) textúrová báza 38 vlastných ľudských tvárí.

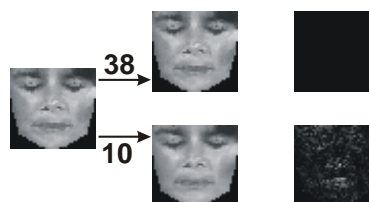
Dostatočne presnú aproximáciu relevantnej textúry ľudskej tváre s pohybom možno dostať aj s obmedzeným počtom vlastných ľudských tvárí z vygenerovanej úplnej textúrovej bázy. Touto redukciou možno zmenšiť výpočtovú zložitosť

tejto aproximácie. V **tab. 1** sú uvedené hodnoty SSE aproximovaných relevantných textúr získané pomocou priamej a spätnej transformácie z rov. (13) s úplnou alebo redukovanými textúrovými bázami s obmedzeným počtom vlastných ľudských tvárí. Pritom z úplnej textúrovej bázy 38 vlastných ľudských tvárí z **obr. 9b** sme redukované bázy získali s počtom 20, 10 a 5 vlastných ľudských tvárí, prislúchajúce k najväčším vlastným hodnotám. Výpočty SSE boli vykonané pre relevantné textúry získané z obrazov v tréningovom súbore, ale aj mimo neho z videosekvencie „Claire“.

Tab. 1. SSE aproximovaných relevantných textúr s použitím úplnej textúrovej bázy 38 vlastných ľudských tvárí a jej redukovanými bázami s ich počtom 20, 10 a 5.

Počet vlastných ľudských tvárí	38	20	10	5
Textúra 1 z „Claire“	521	551	606	656
Textúra 2 z „Claire“	725	772	815	854
Textúra 3 z „Miss America“	0	6	70	82

Z **tab. 1** vyplýva, že ak textúra 3 je získaná z obrazu v tréningovom súbore, tak SSE jej aproximovanej relevantnej textúry pomocou úplnej textúrovej bázy 38 vlastných ľudských tvárí z **obr. 9b** sa rovná nule. Naopak pri použití tejto úplnej textúrovej bázy táto hodnota SSE nie je nulová pre textúry 1 a 2 získané z obrazov videosekvencie „Claire“, pretože tieto neboli získané z obrazov tréningového súboru. Ako je tiež možné vidieť z **tab. 1**, čím väčší počet vlastných ľudských tvárí sa použije, tým je SSE aproximovanej relevantnej textúry menšie, ale naopak tým zložitejší je jej výpočet. Ako vhodný kompromis sú redukované textúrové bázy napr. 10 vlastných ľudských tvárí z úplnej textúrovej bázy. Pre ilustráciu **obr. 10** ukazuje presnosť aproximácie relevantnej textúry 3 z „Miss America“, ktorá bola získaná najprv pomocou úplnej a potom redukovanej textúrovej bázy 10 vlastných ľudských tvárí. Pritom pre úplnú textúrovú bázu s 38 vlastnými ľudskými tvármi jej presnosť je exaktne dokonalá s hodnotou $SSE=0$. Pre danú redukovanú textúrovú bázu zostáva dostatočne presná s hodnotou $SSE = 70$ a veľmi dobrou subjektívnou kvalitou ako to vidno z **obr. 10**.

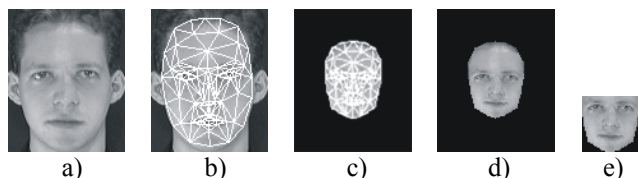


Obr. 10. (zľava) Relevantná textúra, aproximovaná relevantná textúra, rozdielové obrazy (5 krát zosilnené).

5 Textúrová báza ľudskej tváre bez pohybu

Textúrová báza ľudskej tváre bez pohybu sa získava z obrazov ľudských tvárí s čelným pohľadom a bez pohybu v tréningovom súbore, ktoré zachytávajú čo najviac zmien jej

geometrického tvaru. V našich experimentoch sme generovali takúto textúrovú bázu z tréningového súboru 25 rôznych ľudských tvári s čelným pohľadom a bez pohybu. Na tieto boli najprv 3R model ľudskej hlavy adaptovaný (**obr. 11b**) pomocou poloaautomatickej metódy. Následne proces ich predspracovania prebiehal analogicky ako to bolo uvedené v predchádzajúcej časti 4 a výsledky z neho pre jednu ľudskú tvár z tohto súboru sú uvedené v ďalších položkách toho istého **obr. 11**. Nakoniec pomocou PCA relevantných textúr po predspracovaní tréningového súboru bola získaná textúrová báza ľudskej tváre bez pohybu.



Obr. 11. a) Originálny obraz ľudskej tváre, b) originálny obraz s adaptovaným modelom, c) štandardný tvar modelu s textúrou, d) geometricky normovaná textúra, e) relevantná textúra.

Obr. 12 ukazuje 25 relevantných textúr získaných z obrazov v uvažovanom tréningovom súbore a tiež úplnú textúrovú bázu 25 vlastných ľudských tvári. Aj v tomto prípade platia podobné úvahy o dosahovanej presnosti aproximácie ľudskej tváre bez pohybu pomocou tejto generovanej textúrovej bázy, aké boli urobené v predchádzajúcej časti 4. Zároveň sa zachováva analogický vplyv redukcie uvažovanej textúrovej bázy na túto presnosť.



Obr. 12. a) Súbor 25 relevantných textúr, b) textúrová báza 25 vlastných ľudských tvári.

6 Záver

Videokodek MPEG-4 podporuje kódovanie na báze analýzy a syntézy ľudskej hlavy. Využitie tohto kódovania pozostáva z jej detekcie a následnej adaptácie 3R modelu ľudskej hlavy, ktorý sa nachádza ako v kodéri, tak aj v dekodéri. Po detekcii ľudskej hlavy v prvej snímke nasleduje technika tvarovania, ktorá prispôbuje geometrický tvar 3R modelu k objektu ľudskej hlavy vo vstupnej videosekvencii. Tvarovanie geometrie 3R modelu v prvej snímke sa môže komplexne vykonať pomocou algoritmu estimácie jeho tvarovacích parametrov [9], ktorý k svojej činnosti potrebuje textúrovú bázu ľudskej tváre bez pohybu. Jednoduchšie (klasické) algoritmy tvarovania 3R modelu sú založené na extrakcii príznakových bodov na ľudskej tvári v prvej snímke [6], ako sú napr. kútiky úst, špička a koreň nosa, stredy očí atď., z ktorých sa určia parametre upravujúce geometrický tvar 3R modelu. K tejto extrakcii je možno využiť obdobné textúrové bázy, ale vygenerované len pre malé časti ľudskej tváre, ktoré obsahujú tieto príznakové body. Pritom spôsob ich generovania zostáva analogický ako pre celú

ľudskú tvár, pričom ich získavame z tréningových súborov týchto jednotlivých častí. Pohyb takto tvarovaného 3R modelu je potom daný jeho globálnymi ako aj animačnými parametrami, ktoré možno estimovať pomocou algoritmov na základe textúrových báz ľudskej tváre s pohybom [10].

Vizuálne deskriptory [11] pre rozpoznávanie ľudských tvári v štandarde MPEG-7 využívajú algoritmus KLT s generovanou textúrovou bázou vlastných ľudských tvári bez pohybu. Potom tento sa stáva univerzálny a veľmi účinný pre rozpoznávanie ľudských tvári nie len s rôznou textúrovou podobou, ale aj geometrickým tvarom, pričom je málo citlivý na zmenu ich osvetlenia. Dostatočná účinnosť ich rozpoznávania sa v ňom zabezpečuje už textúrovou bázou s malým počtom vlastných ľudských tvári, resp. príznakovým vektorom spektrálnych koeficientov KLT s malým rozmerom. Jeho univerzálnosť je aj v rozpoznávaní ľudských tvári s rôznym geometrickým tvarom ako aj veľkosťou (mierkou), čo je dané ich geometrickým normovaním. Toto normovanie približuje geometrickú podobnosť rôznych ľudských tvári, čím ich rozpoznávanie je dané najmä textúrovou rôznorodosťou.

Literatúra

- [1] Mihalík, J. Štandardný videokodek MPEG-4. *Slaboproudý obzor*, 2003, roč. 60, č.2, s.7-11.
- [2] The special issue of the *IEEE Trans. on Circuits and Systems for Video Technology* on MPEG-4 SNHC, July 2004.
- [3] Strintzis, M., Sarris, N. (Ed.) *3D modeling and animation: Synthesis and Analysis Techniques for the Human Body*. IRM Press, Hershey, PA, July 2004.
- [4] Sikora, T. The MPEG-7 Visual Standard for Content Description – An Overview. *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.11, No.6, 2001, p.696-702.
- [5] Ahlberg, J. *Candide-3: An Updated Parameterised Face*. Rep. No. LiTH-ISY-R-2326, January 2001.
- [6] Mihalík, J., Michalčin, V. 3D Motion Estimation and Texturing of Human Head Model. *Radioengineering*, Vol.13, No.1, 2004, p.26-31.
- [7] Cootes, T. F., Taylor, C. J. *Statistical Models of Appearance for Computer Vision*, Technical report, University of Manchester, 2004.
- [8] Dornaika, F., Ahlberg, J. Fast and Reliable Active Appearance Model Search for 3D Face Tracking, *Proceedings of Model-based Imaging, Rendering, Image Analysis, and Graphical Special Effects (Mirage)*, 2003, p. 113–122.
- [9] Mihalík, J., Kasár, M. Shaping of Geometry of 3D Human Head Model, *Proc. 17th International Conference "Radioelektronika 2007", Brno, Czech Republic*, 2007, pp. 483-486.
- [10] Mihalík, J., Kasár, M. Human Face and Facial Feature Tracking by Using Geometric and Texture Models. *Journal of Electrical Engineering*, Vol.59, No.5, 2008, p.266-271.
- [11] Salambier, P. Overview of the MPEG-7 Standard and of Future Challenges for Visual Information Analysis. *EURASIP Journal on Applied Signal Processing*, No.4, 2002, p.1-11.