

Vec: Posudok dizertačnej práce.

Autor práce: Ing. Václav Mach

Názov práce: Inpainting of missing audio signal samples

Školiace pracovisko: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií

Odbor: Teleinformatika

Rozsah: 122 strán vrátane zoznamov obrázkov, tabuliek a literatúry plus prílohy na CD.

K aktuálnosti témy

Maskovanie chýb je dlhodobým záujmom výskumu pri videu a statickom obraze. Jednak to bolo spôsobené tým, že pri distribučne prenášanom videu nie je možné použiť ARQ metódy, ale aj preto, že vo videokonferenčnom prenose ide o komunikáciu v reálnom čase. Preto sa prvé vážne úvahy o maskovaní chýb objavili už v štandarde H.263. Pri videu v reálnom čase ide však o metódy vychádzajúce len z rôznej aplikácie váženého priemerovania. Ďalšia spoločenská objednávka prišla z oblasti fotografie, kde úlohou takýchto metód je najmä zbaviť sa nežiadúceho objektu na fotografii a nahradiť ho čo najrealistickejšou reprezentáciou okolia. Špeciálnym prípadom je syntéza textúr. V prípade statického obrazu ide o sofistikované metódy s veľmi uspokojivými výsledkami, ale nesmiernou časovou náročnosťou. Vo zvuku sa podobné metódy objavili neskôr. Zrejme nebola taká silná spoločenská objednávka. V súčasnosti sa však maskovaniu chýb v audio nedá vyhnúť. Súčasné spôsoby prenosu priamo počítajú so stratou informácie. Ďalšia nemenej dôležitá objednávka je spojená s digitalizáciou a záchranou kultúrneho dedičstva.

Metódy maskovania zvuku musia spĺňať prísnejšie kritériá, než sú kladené na obraz. Vyplýva to z rôznej citlivosti a spôsobu vnímania porúch v obraze, videu a zvuku. S metódami maskovania chýb vo zvuku sú veľmi malé skúsenosti, a najmä pri sofistikovaných metódach nie sú ani skúsenosťami odporúčané parametre rôznych metód, s ktorými sú už v obraze skúsenosti viac ako 10 rokov.

Z týchto dôvodov považujem prácu za veľmi aktuálnu a prínosnú nielen z pohľadu vedy, ale najmä praxe.

K cieľom

Ciele sú dizertabilné. Autor si ich zvolil vhodne a splnil ich v plnom rozsahu až na zmenu metriky, čo je však z pohľadu dizertabilnosti a plnenia cieľov nepodstatné. Skôr by som očakával prísnejšie ciele, napríklad zásadnejšiu modifikáciu niektorej zo skúmaných metód.

K použitým metódam

Autor použil pre riešenie štandardné metódy vedeckej práce a vhodne zvolil aj spôsob realizácie experimentálneho overenia. Tu sa vrátim k predchádzajúcemu odseku o možnosti modifikácie existujúcich metód. Autor používa na realizáciu jednej z metód DCT IV. Dá sa dokázať, že táto transformácia nedokáže interpolovať ani signál prestriedaný nulami. Je preto predpoklad, že nedokáže uspokojivo extrapolovať vypadnuté časti signálu. Ak má byť dôvodom použitia DCT IV to, že je implementovaná v kodéroch, tak potom je na zváženie, či by DCT III, čo je jej inverzná verzia (použitá napr. pre dekódovanie), nedávala lepšie výsledky. Táto transformácia interpolačné vlastnosti má.

Možno by si zaslúžila pozornosť aj metóda, pôvodne určená pre vypadané obrazy realizovaná pomocou KLT (Hotelling, PCA), i keď musím priznať, že SVD použitý v práci je akýmsi

reprezentantom KLT. Z jednoduchých metód by si pozornosť zaslúžili napr. Legendreove polynómy.

Pôvodné prínosy

Za pôvodné prínosy považujem:

1. Komplexná analýza algoritmov maskovania chýb v audio, tu pod názvom inpainting, predovšetkým však metód vychádzajúcich z princípu postupnej aproximácie.
2. Implementácia a experimentálne overenie relaxačných algoritmov pre audio inpainting.

K záverom pre ďalší rozvoj

Zo záveru je zrejmé, že autor si je vedomý niektorých nedostatkov riešenia. Škoda, že nedáva čitateľovi ďalšie návrhy možností riešenia častí tohoto zložitého procesu.

K publikovaniu jadra práce

Autor v bibliografii uvádza 9 prác, ktorých bol spoluautorom. Z nich podľa názvov 8 sa priamo týka riešených problémov dizertačnej práce, z toho jeden príspevok bol publikovaný v zborníku z medzinárodnej vedeckej konferencie citovanej v databáze SCOPUS a jeden v neimpaktovanom vedeckom časopise klasifikovanom v odbore Telekomunikácie. V zozname absentujú cennejšie publikácie, napr. v zborníkoch vydaných organizáciami: IEEE, IET, SPIE, ACM, IFIP, EURASIP alebo vydané renomovanými vydavateľstvami: Springer, Elsevier, Willey, prípadne publikácie v impaktovaných vedeckých časopisoch podľa kategorizácie Thompson Reuters. Rozsah publikácií preto považujem za "akceptovateľný s výhradami".

K vedeckej činnosti

Autor uvádza, že pracoval ako riešiteľ 6 výskumných projektov, dva krát bol na výskumnom pobyte na zahraničnej univerzite. Jeho výskumnú činnosť považujem za nadpriemernú.

Pripomienky a otázky

Okrem vyššie uvedených pripomienok mám ešte nasledujúcu:

Tvrdenie v kap. 7.7 o tom, že PSMt je vyhovujúcejšia metóda pre evaluáciu výsledkov, lebo má podobné výsledky ako SNR, odporuje zaužívaným zvyklostiam, aj odporúčaniam ITU. Rozhodujúce pre hodnotenie kvality (ak je to možné realizovať) je subjektívne hodnotenie, teda niektorá perceptuálna metrika. Aby sa nemuselo (alebo ak sa nedá) robiť subjektívne hodnotenie, hľadá sa metrika, ktorá by čo najlepšie zodpovedala subjektívnemu hodnoteniu. A na tieto účely ITU odporúča medzi perceptuálnou metrikou a objektívnou metrikou vypočítať Pearsonovu koreláciu. Ak by sa autor neobmedzoval len na SNR, ale by použil aj iné metriky na báze porovnávania hodnôt vzoriek, histogramov, spektier, korelácie a podobne zo skupiny viac ako 20 rôznych používaných jednoduchých vzdialeností pre hodnotenie kvality, možno by sa takémuto tvrdeniu vyhol. Keďže sa mi nepodarilo v texte nájsť podmienky subjektívneho hodnotenia kvality rekonštrukcie signálu, navrhujem, aby autor vysvetlil, ako bolo robené toto hodnotenie: jednostimulovo alebo dvojstimulovo, koľko hodnotiteľov, interval spoľahlivosti a podobne.

Iné pripomienky a otázky nemám.

Dizertačná práca spĺňa podmienky stanovené pre úroveň dizertačných prác, odporúčam jej obhajobu.

Po úspešnej obhajobe **navrhujem udeliť akademický titul „philosophiae doctor“.**

V Bratislave 26.5.2016.

