

Posudek oponenta na dizertační práci

Autor dizertační práce: Ing. Jaroslav Moravec, FEI, Univerzita Pardubice
Název dizertační práce: Nástroje pro identifikaci osob s využitím kontury lidské ruky
Autor posudku: Ing. Petr Pošík, Ph.D.
Katedra kybernetiky, FEL ČVUT v Praze
petr.posik@fel.cvut.cz

1 Téma práce

Předložená dizertační práce se zabývá autentizací osob podle kontury lidské ruky. V současnosti existují běžně dostupné, levné a funkční snímače otisků prstů, příp. jiných biometrických znaků, které podléhají menším změnám u téhož člověka, než kontura lidské ruky. Jistě není lehké přijít s metodou pro autentizaci či dokonce identifikaci člověka založenou právě na kontuře lidské ruky tak, aby byla konkurenceschopná ostatním existujícím metodám z hlediska ceny, rychlosti a přesnosti. Téma práce hodnotím jako poměrně náročné, ale nejsem si jistý, zda se konkurenceschopného výsledku podařilo dosáhnout.

2 Rozsah a struktura práce

Text práce zabírá 211 stran. Z mého pohledu je to způsobené nevhodně zvolenou strukturou práce. Jsou mi známy v podstatě dvě formy, které může dizertační práce mít: buď nový ucelený text (většinou vycházející z publikovaných prací), nebo sbírka publikovaných článků opatřená kvalitním shrnujícím textem. Předložená práce není bohužel ani jedno: tváří se jako nově napsaný ucelený text, ale ve skutečnosti je většina jednotlivých podkapitol patrně tvořena publikovanými články autora (ovšem bez dobrého sjednocujícího textu). To následně vede k dalším nepřijemným vlastnostem textu, např. k tomu, že některé části se v práci několikrát opakují (např. úvod do biometrických metod, popis různých evolučních algoritmů), nebo k tomu, že opakující se věci jsou na různých místech popsány poněkud jinak (aniž by k tomu byl nějaký důvod). Odhaduji, že pokud by byl v práci každý koncept dobře popsán vždy jen na jednom místě, dala by se práce zkrátit o čtvrtinu až třetinu.

Samotný text práce obsahuje mnoho detailů, v nichž se čtenář na mnoha místech doslova utápí. Hlavní a důležité myšlenky jednotlivých algoritmů, přístupů, problémů, atd. se v nich ztrácí, místo aby byly zdůrazněny. Celkově práce spíše popisuje, *jak* je něco udělané, místo toho, aby hlavně vysvětlila, *proč* se to dělá.

Po velmi obecném úvodu a stručném popisu relevantních metod autor popisuje konstrukci biometrického skeneru, který v práci navrhl. Následně v kapitole 5 rovnou přistupuje k popisu hlavních typů evolučních algoritmů, které jsou v práci hojně využívány. Očekával bych ale, že před popisem evolučních algoritmů (tedy nástrojů pro řešení nějakých problémů) už čtenář bude znát formální specifikaci problémů, které se v práci vlastně řeší, a nějaká kritéria, jimiž se bude posuzovat kvalita dosažených řešení. Bez této znalosti je těžké soudit, proč jsou vlastně EA zvoleny jako hlavní (a prakticky jediný) optimalizační nástroj použitý v této práci.

Hlavní část práce pak tvoří kapitoly 6 a 7. Kapitola 6 zabírající 70 stran pojednává o metodách korekce soudkovitosti obrazu. Kapitola 7 se pak na cca 80 stranách věnuje metodám pro autentizaci osob. Očekával jsem, že kvalitní korekce soudkovitosti (kap. 6) bude nepostradatelná pro dosažení dobrých výsledků při autentizaci osob v kapitole 7. Bohužel se ukázalo, že v kapitole 7 se korekce soudkovitosti vůbec nepoužívá, resp. že se používají existující databáze obrazů ruky, u nichž není zřejmé, zda korekce soudkovitosti byla provedena, nebo zda lze vůbec provést. Z hlediska tématu dizertační práce (autentizace osob podle kontury ruky) pak kapitola 6 působí nadbytečně. Algoritmy navržené v kapitole 6 také budí poměrně dost nejasností, viz. doplňující dotazy 4-7.

3 Formální stránka práce

Autor má bohužel potíže s využíváním a správným zápisem matematických výrazů a symbolů. Někdy je použito více různých značení pro stejný koncept (viz připomínky v příloze). To dále ztěžuje porozumění celému textu.

Práce s odbornou literaturou se zdá být na první pohled v pořádku. Seznam použité literatury je uveden v práci mnohokrát, na konci jednotlivých sekcí, či podsekcí, patrně podle toho, kde končily články, z nichž je práce poskládána. Toto není sice zcela obvyklé, ale nepovažuji to za vadu. Přesto jsem narazil na několik nejasností týkajících se využití odborné literatury, viz doplňující dotaz 8.

4 Dosažené výsledky

Kapitola 6 prezentuje 3 různé algoritmy pro korekci soudkovitosti a uvádí jimi dosažené výsledky. Chybí ale jejich vzájemné porovnání a porovnání s výsledky dosaženými jinými metodami známými z literatury, a to jak z hlediska přesnosti rekonstrukce původního obrazu, tak z hlediska časových nároků. Chybí také nějaký jednoznačný závěr, která z metod je nejlepší, příp. která se hodí v jaké situaci. Neumím proto posoudit, jak dobré dosažené výsledky jsou. Protože se algoritmy z kapitoly 6 nijak nepoužívají v kapitole 7 v algoritmech pro autentizaci osob, není možné posoudit ani jejich přínos pro tuto úlohu.

Výsledky metod pro autentizaci osob navržených v kapitole 7 jsou z hlediska chybovosti (FAR, FRR, ERR) srovnatelné s výsledky dosaženými jinými metodami. Není ale zřejmé, kolik času potřebují jednotlivé porovnávané metody od okamžiku pořízení snímku ruky po rozhodnutí, zda bude osobě udělen přístup či ne. Je jasné, že doba zpracování závisí na implementaci a prostředí, v němž program běží, nicméně alespoň orientační srovnání by bylo vhodné v práci uvést.

5 Publikované práce

Autor v seznamu vlastních publikací uvádí 4 časopisecké články, 3 konferenční příspěvky a 1 technickou zprávu.

Tři časopisecké články byly publikovány v časopisech s nízkým (Q4) nebo žádným impact factorem, jedna z publikací je zveřejněna v časopise s IF cca 5.5. Tři z článků se věnují problematice korekce distorze obrazu, jediný se věnuje klasifikaci kontury ruky.

Všechny tři konferenční články byly publikovány na téže konferenci CSOC 2020, dva z nich se věnují klasifikaci kontur ruky.

Publikační výsledky považuji za mírně nevyvážené, ale dostačující.

6 Shrnutí hodnocení

Předložená dizertační práce obsahuje mnoho nedostatků, nejasností, možná i chyb. Mám k ní řadu dotazů a výhrad (viz. doplňující dotazy a seznam nedostatků v příloze). Nicméně je zřejmé, že bylo odvedeno velké množství práce.

Práci není jednoduché ohodnotit. Mám-li hodnotit vyvinuté algoritmy a dosažené výsledky, musím konstatovat, že vytvořené algoritmy řeší problémy, pro něž byly vytvořeny. Méně zřejmé je z práce to, jak dobře tyto problémy řeší ve srovnání s konkurencí, ale minimálně z hlediska dosažené přesnosti považuji algoritmy za porovnatelné s konkurencními metodami. Časová náročnost vytvořených metod není zřejmá.

Samotný text práce nepovažuji za kvalitní a jsem si jistý, že jej jde vylepšit v mnoha směrech.

Je také možné, že jsem některým částem práce špatně porozuměl. Chci proto dát autorovi šanci obhájit svou práci přede mnou i před komisí, a proto **doporučuji předloženou práci přijmout k obhajobě.**

7 Doplňující dotazy do diskuse

1. Na str. 25 autor píše: "Biometrika rozhodně nepatří mezi jednoduché vědní obory. Je to dáno tím, že rozmanitost variant, byť jen jediného biometrického ukazatele je v lidské populaci značná." Znamená to, že pokud by variabilita biometrických ukazatelů byla menší, např. nulová, byla by biometrika jednodušší? Měla by pak ale vůbec význam?
2. Nemění se parametry optické soustavy při pořízení každé fotografie ruky? Má vůbec smysl hledat korekci společnou pro všechny fotografie? Co musí být splněno, aby to smysl mělo?
3. Mnohé komerční programy nabízejí nástroje pro korekce vad optické soustavy. Obsahují profily pro různé objektivy a různé fotoaparáty. Používají se v práci tyto známé profily? Jsou dostatečné pro snímání kontury lidské ruky?
4. Algoritmus (6.1).4: Chápu, že tento algoritmus můžu použít pro korekci obrazu metodou direct-mapping. T.j., při znalosti korekčního polynomu $f: r_d \rightarrow r_u$ dokážu pro každý pixel na souřadnicích (i, j) poškozeného obrazu I_D spočítat, že se má namapovat na pixel na souřadnicích (X_S, Y_S) opraveného obrazu I_U , tedy že $I_U(X_S, Y_S) = I_D(i, j)$. Nerozumím ale, jak je možné pro back-mapping použít pouhé prohození souřadnic pixelů v poškozeném a opraveném obraze, tedy vztah $I_U(i, j) = I_D(X_S, Y_S)$? Abych toto mohl provést, potřeboval bych přece inverzní funkci k polynomu f , ale tu nemám. Co mi zde uniká?
5. Naprosto nerozumím, proč je v sekci 6.1.5 a v algoritmu (6.1).6 potřebná vnitřní optimalizační smyčka řešící optimalizační úlohu (6.1.15). Známe-li pozici kalibrační značky v poškozeném obraze I_D a mám-li kandidátský polynom $f: r_d \rightarrow r_u$ navržený vnějším optimalizačním algoritmem, dokážu spočítat polohu kalibrační značky v opraveném obraze pomocí polynomu f a direct-mappingu. To by mělo stačit k výpočtu hodnoty kritéria (6.1.14). Proč je to nutné dělat vámi navrženým způsobem?
6. Polynomiální model radiální distorze obrazu: Známe-li pozice kalibračních (referenčních) značek v nepoškozeném obraze a pozice značek v obraze zatíženém distorzi, lze koeficienty polynomu snadno odhadnout lineární regresí. Proč jste se rozhodl použít k jejich nalezení evoluci?
7. Popis "division modelu" v sekci 6.3.4 je zmatený. Mohl by jej autor vyjasnit?
 - Rovnice (6.3.4) definuje funkci \mathcal{F} jako $\mathcal{F}(r_d) = \frac{r_u}{1+k_1 r_u^2}$. Funkce \mathcal{F} má být funkcí proměnné r_d , ta se ale v definici funkce vůbec nevyskytuje. Předpokládám tedy, že správně má být buď (a) $r_d = \mathcal{F}(r_u) = \frac{r_u}{1+k_1 r_u^2}$ (tedy funkce s inverzním významem vzhledem k polynomům používaným v předchozích sekcích), nebo (b) $r_u = \mathcal{F}(r_d) = \frac{r_d}{1+k_1 r_d^2}$ (tedy funkce se stejným významem vzhledem k polynomům používaným v předchozích sekcích). Co je správné?
 - Neroumím rovnici (6.3.5): $\frac{r_u}{1+k_1 r_u^2} \approx \frac{r_u}{1-k_1 r_u^2}$. Toto rozhodně neplatí univerzálně pro jakékoli hodnoty k_1 a r_u . Prosím vysvětlíte, za jakých podmínek tyto výrazy dávají podobné výsledky a zda jsou tyto podmínky splněny ve vaší práci.
8. Sekce 6.1.7 se snaží odůvodnit využití evolučních algoritmů k řešení úlohy korekce soudkovitosti (osobně bych to očekával mnohem dříve). Našel jsem v této sekci ale dvě vyjádření, která ve mně vzbudila pochybnosti:
 - Věta "Skutečnost, že starší optimalizační postupy tzv. Gradient Descent Algorithms (GDA) poskytují výrazně horší výsledky, je popsána a analyzována již v (Yang & Illingworth, 1994)." Když však citovanou práci otevřeme, zjistíme, že jedině, co autoři říkají o gradientních algoritmech, je toto: "GD methods are not suitable for this task because: (1) they are susceptible to becoming trapped at local minima, (2) the gradient info is not available." Nevidím zde, že by autoři tvrdili, že tyto algoritmy poskytují výrazně horší výsledky, ani to, že by tuto skutečnost nějak analyzovali. Prosím autora o vyjádření.

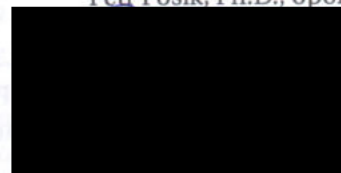
- Věta "(Ji & Zhang, 2001) na základě svého výzkumu popsali, že 'SGA svou výkonností překonává klasické fotogrammetrické metody...'" Autoři v článku skutečně GA používají a samozřejmě argumentují vhodnými vlastnostmi. Ale žádné porovnání s "klasickými fotogrammetrickými metodami" neprovádějí. Uvedenou doslovnou citaci jsem v článku také nenašel. Prosím autora o vyjádření.

9. Konstrukce RDD:

- Pokud se bod P_{RDD} , který se využívá ke konstrukci RDD, nachází ve vzdálenosti $10L_{HAND}$ od P_{WC} , dává smysl počítat vzdálenosti? Nešly by místo toho použít jen x -ové souřadnice bodů kontury ruky?
 - Na str. 126 je uvedeno, že "Z obecného hlediska je výhodnější umístit bod P_{RDD} poblíž bodu P_{WC} ." Proč jej tedy umístit daleko? Vysvětlíte tento rozpor.
10. U některých autentizačních metod je uvedena statistika IR, jejíž vysvětlení jsem v práci nikde nenašel. Co tato zkratka znamená a jak je definována?
 11. Fungují navržené metody autentizace podle kontury ruky i pro osoby s poraněnou rukou (chybějící prst, chybějící článek prstu)? Mám na mysli případ, kdy i vzorové snímky pro uložení databázi jsou pořízeny už s tímto poraněním.
 12. Nakolik je vyvinutá metoda autentizace lidí na základě kontury ruky konkurenceschopná v porovnání s jinými metodami, např. otisky prstů, z hlediska přesnosti a časových nároků? Jak dlouho trvá rozhodnutí, zda člověku má být povolen přístup či nikoli?
 13. Na straně 163 je uvedeno, že "Algoritmus eaICP využívá nejmodernější klasifikační metody, které se běžně využívají v mnoha citlivých lékařských i vojenských oborech." Této větě nerozumím. Jaké klasifikační metody využívá algoritmus eaICP? A proč jsou nejmodernější?
 14. V sekci 7.2.4.1 je uvedeno, že při použití databáze GPDS musely být některé snímky rukou vyřazeny a nahrazeny náhodně vybraným snímkem identické osoby. Pokud se porovnáváte s jinými metodami, provedli stejnou úpravu datové sady, nebo ty "nepoužitelné" snímky v databázi ponechali?
 15. Metody prezentované v sekcích 7.2 a 7.3 nebyly nikde publikované? V seznámech literatury pro tyto sekce jsem nenašel žádnou autorovu relevantní publikaci. Neříkám, že je to špatně, jen se ujišťuji.

30. prosince 2020

Petr Pošík, Ph.D., oponent



A Některé připomínky, nejasnosti, či chyby nalezené v práci

Zde uvádím seznam dalších nejasností a chyb nalezených v předložené práci. Může se zdát, že někde až hnidopišsky trvám na správném formalismu. To ale není samoúčelné. Správně použitý matematický formalismus zjednodušuje čtení práce a poskytuje jednoznačné definice jednotlivých konceptů. Pokud se ale matematický formalismus použije chybně, naopak to čtenáře velmi mate a porozumění práci je mnohem obtížnější. V neposlední řadě může tento seznam dobře posloužit, pokud bude třeba práci přepracovat.

1. Osobně mám problém s použitým značením, konkrétně se symboly P_{op} , D_{im} , D_{om} , G_{en} . Např. P_{op} přece není nějaká varianta symbolu P blíže určená indexem op , je to jen zkratka slova Population; vhodný by tak byl symbol Pop .
2. Na str. 20 a 21 jsem se poprvé setkal s použitými zkratkami RDD a EER. Vysvětlení jejich významu jsem se dočkal až o 100 stran později. Případy, kdy se v práci mluví o konceptu, který je vysvětlen teprve později, jsou z mého pohledu až moc časté.
3. Popis algoritmu CMA-ES na str. 38 a Alg. (5.3).1 jsou pro čtenáře zcela bezcenné a nic mu nepřinesou.
4. Na str. 45 autor píše: „CE (kaskádové estimátory) tvoří vzácnější, řídkce se vyskytující téma. Je to dáno především značnou složitostí algoritmizace.“ O několik řádků níže se mluví o CPR (cascaded pose regression) slovy: „CPR je ... jednoduše implementovatelná metoda“. V těchto tvrzeních vidím rozpor.
5. Na str. 47, vztah (6.1.7) má prohozené funkce \sin a \cos u výpočtu x -ové a y -ové souřadnice.
6. V práci se poměrně často vyskytuje nejednotné značení, např.:
 - V algoritmech (6.1).6 a (6.1).7, pokud to chápu správně, je tentýž bod a jeho souřadnice jednou označeny jako $(X_t.x_1, X_t.x_2)$ a jednou jako (t_x, t_y) .
 - V textu práce se pro střed obrazu používá na str. 47 označení $P_{uc}(x_{uc}, y_{uc})$, později na str. 54 označení $center(x, y)$ a v algoritmu (6.1).7 zase $(center_x, center_y)$.
 - V sekcích 6.1 a 6.2 se používají dva evoluční algoritmy, kdy jeden je zanořený do druhého. V sekci 6.1 se jako primární označuje ten vnitřní a jako sekundární ten vnější. V sekci 6.2 je to naopak.

Pro čtenáře je to neskutečně matoucí.

7. Matematické zápisy na mnoha místech práce obsahují chyby, kvůli kterým jsou některé vzorce a rovnice nesmyslné, např.:
 - Rovnice (6.1.14) $fitness_2 = \|M_F - M_R\|^2$ je nejasná. M_F a M_R jsou matice 2D bodů. Jejich rozdíl je tedy také matice 2D bodů. Jak je ale definována norma matice 2D bodů? A v závislosti na tom, co tato norma je, jak se počítá její druhá mocnina?
 - Rovnice (6.3.4) a (6.3.5). Viz doplňující dotaz 7.
 - Str. 127, rovnice (7.1.11) a (7.1.12). Kromě zmateného názvosloví tam vidím i faktické chyby. V jedné z rovnic definujete \mathcal{R} jako hodnotu kritéria pro nějakou aproximační funkci F , ve druhé definujete tentýž symbol jako *optimální hodnotu* tohoto kritéria.
8. Častým problémem v práci jsou také nesmyslně zapsané optimalizační problémy, např.:
 - Str. 54, (6.1.13): Pokud optimalizace probíhá nad možnými dvojicemi matic (M_F, M_R) , pak výsledkem funkce argopt nemůže být množina optimálních koeficientů.
 - Str. 94, (6.3.9): Pokud optimalizace probíhá nad možnými hodnotami koeficientů k_i , pak výsledkem funkce argopt nemohou být souřadnice (\hat{x}, \hat{y}) .
 - Str. 132, (7.1.22): Pokud množina \mathcal{H} , přes kterou se optimalizuje, obsahuje možné hodnoty parametrů, kterými je ovlivňována kontura \mathcal{S} , pak výsledkem funkce argopt nemůže být skalár \mathcal{E} vyjadřující velikost odchylky obou kontur. Zde by dávalo smysl použít v zápisu optimalizační úlohy jen funkci opt .
 - Str. 137, (7.1.26): Opět, výsledkem funkce argmin zde nemůže být skalár \mathcal{E} . I zde by měla být použita jen funkce min . Naopak na str. 138 v (7.1.27) by místo funkce min měla být použita funkce argmin .

- Str. 176: v rovnici (7.2.1) by opět mělo být jen \min místo $\arg \min$. Naopak v (7.2.2) by mělo být $\arg \min$ místo \min .
 - Str. 179: v (7.2.3) by opět mělo být jen opt , nikoli $\arg \text{opt}$.
9. "Waveletova transformace" není pojmenována podle pana Waveleta, ale podle waveletů (vlnek), které tvoří její báze. Tato transformace má v češtině zažitý název, vlnková transformace.
 10. Str. 134: matice D_{ist} je definována jako matice vzdáleností. V alg. (7.1).6 se ale v kroku 4 do prvků matice přiřazují body P_{min}^M .
 11. Sekce 7.1.3 (a její podsekce) se věnují popisu algoritmu eaICP. Ovšem teprve podsekce 7.1.3.9 obsahuje popis původního algoritmu ICP. Očekával bych to naopak, tedy nejprve popis algoritmu ICP a teprve následně popis úprav využitých v eaICP.
 12. V sekci 7.1.3.10 je velmi nevhodně (chybně) použito slovo "klasifikován" ve větách "V první části je kontura ruky klasifikována, je vytvořen její popis, ..." a "Ve druhé části jsou kontury klasifikovány ve smyslu (7.1.22)...". V první větě není vůbec zřejmé, co je slovem "klasifikována" myšleno, ve druhé větě je tím patrně myšlen výpočet fitness pro danou konturu ruky. Ani v jednom případě ale patrně nejde o klasifikaci ve smyslu rozřazování objektů do několika tříd.
 13. V sekci 7.2.2.2 na str. 175 je uvedeno, že "No Free Lunch theorem ... říká, že v oblasti optimalizačních algoritmů nepřichází nic zcela zdarma." V dizertační práci bych čekal přesnější interpretaci NFL.
 14. Na str. 207 je uvedeno, že "Použité statistické ukazatele spadají do relativně nové metodologie označované jako 'matice záměn' neboli 'confusion matrix'..." Na matici záměn přece není nic nového, je to zcela standardní nástroj. Odhady podmíněných pravděpodobností, jako jsou FPR a FNR slouží jako odhady chyb I. a II. druhu, které jsou ve statistice používány už dlouhá desetiletí.