

Diplomová práce : **Návrh hashovacího algoritmu pro biometrický podpis uživatele**

Autor práce: **Bc. František Pártl**

Studijní program/obor: N3902 – Inženýrská informatika / Medicínská informatika

Posudek oponenta diplomové práce

Student Bc. František Pártl se ve své diplomové práci zabývá analýzou vlastností, implementací a testováním vybraných technik pro získávání a využití otisků osob, které jsou konstruovány nad zvukovou stopou audiovizuální nahrávky. Práce je řešena v rámci projektu, který je zpracováván na Katedře informatiky a výpočetní techniky ZČU v Plzni, a jedná se o multimodální biometrický podpis adresátů při přebírání poštovní zásilky. Podpis adresátů má tři komponenty, promluva, resp. analýza zvukové stopy, je jednou z nich a je standardizována předpisem formátu. Je nutné hned v úvodu poznamenat, že předepsaným formátem má promluva výrazně limitovanou délku a poměr počtu slov standardního textu a textu variabilního, který obsahuje jméno a příjmení adresáta, je 4 : 2. Tím je manévrovací prostor autora ve výběru příslušných technik a jejich *úspěšnosti* z pohledu volby biometrických charakteristik částečně omezen.

Předložená diplomová práce obsahuje 60 stran vlastního *čistého* textu, 10 stran tištěných příloh a DVD-R s elektronickou přílohou. Práce je členěna do dvanácti kapitol, včetně úvodu a závěru.

Tematicky lze předložený text práce podle mého názoru rozdělit do částí tří. Do první části můžeme zařadit obecný popis biometrie, stručnou část týkající se rozpoznávání obličejů, zúžený přehled nástrojů analýzy řečového signálu a představení některých používaných a současně také použitých technik. Jako druhou část lze chápat část implementační včetně návrhu a realizace algoritmů a tvorbu rozsáhlé knihovny *libpe*. Třetí část pak vidím ve vytvoření a naplnění *datasetu*, testování navržených a realizovaných algoritmů a vyhodnocení dosažených výsledků.

K první tematické části, str. 13 – 31, mám několik detailnějších poznámek a dotazů, protože tuto část považuji z hlediska dalšího řešení diplomové práce za poměrně důležitou. V části popisu fyziologie hlasu na str. 15 autor uvádí rozsah základní hlasivkové frekvence $F_0 \in (60; 400)$ Hz. Z kontextu lze samozřejmě usoudit, že se jedná o tzv. mluvní rozsah, viz dále rozdělení řečového signálu do segmentů a vlastní zadání úlohy. Autor zde ale také hovoří o F_0 u sopranistek, a proto se nabízí otázka dalšího členění podle tzv. mluvního a zpěvního projevu, popř. podle hlasových rejstříků. Co je ale zajímavější z hlediska řešené problematiky je to, že jsem nenašel zmínku o chrapotu, ať už je jeho příčina jakákoliv, a jeho vlivu na charakteristiky, které jsou dále z promluvy extrahovány a použity pro identifikaci řečníka. Je možné, že případný chrapot řečníka ovlivní jeho identifikaci? Kapitola 4. *Analýza řečového signálu* by měla obsahovat přehled a popis metod pro získávání charakteristik k určení řečníka. Je metoda MFCC skutečně jedinou vhodnou metodou, jak by se podle obsahu této kapitoly mohlo jevit? Které další metody lze v této souvislosti aplikovat, resp. kombinovat jejich charakteristiky? Byly jiné takové metody uvažovány, popř. testovány? V kapitole 5. *Použité techniky* postrádám po formální stránce úvodní specifikaci obsahu a cíle této kapitoly. Cíl lze sice vytyčit, ale není zde např. zdůvodnění, proč je právě použita uvedená skladba metod DTW a metoda shlukové analýzy (K-means), zda je to skladba technik, které jsou pro tento účel jediné použitelné a proč. Odpověď na tyto otázky lze částečně dovodit až z dalších částí DP. Obecně, podle mého mínění, by si tato část diplomové práce zasloužila detailnější rešeršní náplň.

K druhé, implementační části, str. 32 – 40, nemám žádných zásadních připomínek, tuto část práce považuji za výbornou, která svědčí o pracovním nasazení, zájmu a schopnostech autora.

Třetí část práce je popsána na str. 41 – 67 a koresponduje také s předchozími částmi textu DP. Můžeme sem zařadit modul pro pořizování dat – *dataset*, který je velmi pečlivě propracován a reprezentován knihovnou *SignatureDataLib*, viz kap. 8. Pro účely testování obsahoval *dataset* celkem 153 videozáznamů pořizovaných od 34 jedinců (15 žen a 19 mužů). Dále tato část obsahuje velmi zajímavý a užitečný nástroj, viz kapitola 9. Aplikace *Hash Visualisation*, který slouží pro realizaci tzv. vizualizačních experimentů. Vizualizační experimenty umožňují, zjednodušeně řečeno, sledovat chování a vyhodnocovat výsledky jednotlivých metod v různých konfiguracích pro extrakci charakteristik otisku osoby nad daty z *datasetu*. Z hlediska struktury textu DP lze v této části nalézt odpovědi a vysvětlení na některé otázky, které vyplynuly v kapitolách 4 a 5. Na základě výsledků

vizualizačních experimentů je pak navržen soubor algoritmů, které jsou popsány v kapitole 10. *Navržené algoritmy*, str. 54 – 56. Tento soubor sedmi algoritmů byl dále systematicky testován, viz kapitola 11. *Aplikace VoiceHashExps*, kde kromě dosažených výsledků, viz kap. 11.4. a 11.5, velmi vysoce hodnotím celkovou metodiku testování, viz kap. 11.1. Opět zjednodušeně řečeno, autor vytvořil soubor nástrojů a metodiku testování tak, že lze modifikovat celý systém pro jiný formát promluvy (skladbu slov, délka promluvy, jazyk řečníka, ...).

Pokud se zastavím u testování a testovaných vzorků, nabízí se dvě otázky. Byly testovány promluvy jiného formátu, než je předepsaný vzor, viz str. 9? Byly jednotlivé promluvy řečníků v *datasetu* nějakým způsobem modifikovány při jejich pořizování ve smyslu zkreslení? Zde se odvolávám na zmínku na str. 41. Jestliže ano, jakým způsobem byly modifikace prováděny, čeho se přesně týkaly a jakých výsledků bylo v této souvislosti dosaženo? A opět se vracím k již zmíněnému možnému vlivu chrapotu.

Diplomová práce je po formální stránce na velmi vysoké úrovni. V textu lze najít drobné nedostatky, které by si zasloužily upřesnění, např. způsob používání odkazů na obrázky do *vzdálených* částí práce, viz str. 24, kde je „odkaz“ na graf 9.4, podobně, viz str. 30, „odkaz“ na graf 3.2. Jednak se jedná o obrázky, ale hlavně, bylo by užitečné z hlediska přehlednosti a rychlé orientace v textu doplnit v takových případech i stránku, na které se obrázek nachází. Také vzniká otázka, podle jakého klíče byly/nebyly autorem práce číslovány matematické vztahy.

Konstatuji, že použitá literatura plně odpovídá tématu a autor se vždy důsledně odkazuje na použité zdroje.

Potvrzuji, že pan Bc. František Pártl splnil zadání v plném rozsahu. Prokázal schopnost samostatné tvůrčí práce, dokázal komplexně a nadstandardně vyřešit zadaný úkol a dospět k vlastní rozsáhlé a velmi propracované plně funkční aplikaci.

Diplomovou práci proto doporučuji k obhajobě a i přes drobné připomínky k první tematické části navrhuji jednoznačně klasifikaci známkou

“ **v ý b o r n ě** ”.

V Plzni 4. 6. 2020

Pavel Nový

Katedra informatiky a výpočetní techniky

FAV – ZČU Plzeň

tlf.: +420 377 632 411

e-mail: novy@kiv.zcu.cz