

## Posudek oponenta bakalářské práce

**Student:** Horázný František  
**Téma:** Akustická detekce pozice řečníka pomocí mikrofonního pole (id 22685)  
**Oponent:** Mošner Ladislav, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **průměrně obtížné zadání**  
Samotná metoda, která je jádrem aplikace, staví na dvou základech: vzájemné korelaci a hyperbolách. Proto přístup samotný není příliš komplexní. Student ovšem musel zprovoznit zapůjčenou zvukovou kartu a dokázat pro ni naprogramovat (a kvůli menšímu výkonu upravit) aplikaci pro odhad pozice řečníka.
- 2. Splnění požadavků zadání** **zadání téměř splněno**  
Bod 3 zadání požaduje výpočet polárních souřadnic řečníka vůči libovolnému bodu. V textu práce o tom však není zmínka a všechny uváděné pozice jsou počítány v kartézských souřadnicích vzhledem k fixnímu bodu. Osobně však nepovažuji tento nedostatek za závažný. Přepočtení do polárních souřadnic není komplikované a možnost spočítání pozice vůči libovolnému bodu je spíše otázkou zpracování výstupů hlavního programu.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**  
Rozsah zprávy je vyhovující. Uvítal bych však nahrazení některého vzdáleně souvisejícího textu za relevantnější (např. zmínka o zvířecím sluchu).
- 4. Prezentací úroveň předložené práce** **85 b. (B)**  
Technická zpráva je logicky strukturovaná (až na prolnutí návrhu a implementace). Text je většinou plynulý. Kapitoly a podkapitoly by si v některých případech zasloužily popisnější názvy. Zejména část o testování je slabá z pohledu popisu. Z textu tak často není zcela zřejmé, co student zkoušel a jak postupoval.
- 5. Formální úprava technické zprávy** **80 b. (B)**  
Text je z jazykového hlediska dobrý. Často se ale objevují problémy s pojmy (např. microphone field místo microphone array, analogický signál místo analogového, harmonický signál místo periodického, síla signálu místo výkonu signálu, křížová korelace místo vzájemné, instruktážní sada místo instrukční). Z typografického hlediska je práce na dobré úrovni. Mohla by být zlepšena úpravou následujících aspektů. Za rovnicemi bývá v textu zprávy zpravidla odsazení (začátek nového odstavce), i když popis (například proměnných v rovnici) pokračuje. U grafů chybí popisky os a písmo v obrázcích je někdy špatně čitelné. Kvalita grafů by šla zlepšit i vektorovým formátem místo rastrového.
- 6. Práce s literaturou** **85 b. (D)**  
Po stránce práce s literaturou je zpráva podprůměrná. Pozitivní je, že student poctivě citoval převzaté obrázky. Domnívám se, že důležité literární prameny chybí. Zejména v kapitole 3, která se zabývá možnostmi odhadu pozice řečníka i popisem implementované metody, chybí citace relevantních zdrojů. Zvolená metoda není nová, a i přes to se čtenář nedozví o související literatuře. Na několika místech nelze jednoznačně rozlišit autorovy myšlenky od převzatých (například u definice použitých filtrů). O relevantnosti a věrohodnosti některých elektronických zdrojů by se dalo pochybovat (například blog [4], elektronický článek o sluchu zvířat [9]). Některé bibliografické citace jsou neúplné.
- 7. Realizační výstup** **65 b. (D)**  
Studentovi se podařilo splnit požadavek schopnosti běhu výsledného programu v reálném čase na zvukové kartě. Systém se však nezdá být přesným a robustním, jelikož výsledky experimentů jsou často uváděny jen pro několik pozic zdroje, kde jednotlivé přesnosti se značně liší. Kladně hodnotím to, že student provedl analýzu teoretických limitů určení pozice řečníka v závislosti na vzorkovací frekvenci (Sekce 4.2) a snažil se přijít s řešením, které bere v potaz charakteristiku hledaného zdroje - hlasu (implementace detekce hlasu v nahrávce). Oceňuji snahu o zjištění vlivu komponent programu na chování. Ne všechny ale shledávám dostatečnými. Například testování výpočtu prahu využívaného pro detekci hlasu v nahrávce bylo provedeno porovnáním vizualizace a poslechem. Filtrování signálů pásmovou propustí není dobře motivováno (mezí frekvence určeny na základě "frekvenčního rozsahu potřebného pro dorozumívání"). Testování přístupů ke korelaci by zasloužilo jiné zpracování (vzájemná korelace byla vybrána jako lepší, protože odhadnuté posunutí bylo v akceptovaných mezích v 48,8 % případů oproti druhému přístupu s přesností 44,9 %). V rámci testování celého systému bylo odhaleno, že velký problém působí ozvěny a po jejich potlačení se přesnost zlepšila. Výsledky dalších experimentů je obtížné interpretovat v důsledku tvrzení ze Sekce 6.2.4 (při měření mohlo dojít k nezanedbatelným nepřesnostem) a faktu, že ve většině případů byl zdroj zvuku nahrazen telefonem přehrávajícím hlas.

### 8. Využitelnost výsledků

Práce je založena na implementaci známého přístupu pro lokalizaci řečníků (zdrojů zvuku). Pozornost si zaslouží fakt, že je program schopen běhu na sestavě s ARM/SHARC s omezeným výkonem. Implementace je motivována hned v úvodu práce. Výsledek úsilí by byl využitelný pro směřování záře reflektorů či kamery na řečníka. Podle uvedených výsledků je ovšem praktické nasazení ještě vzdáleným cílem.

### 9. Otázky k obhajobě

- V čem se liší Vaše bakalářská práce od té, kterou citujete (Akustická detekce pozice řečníka pomocí mikrofonního pole).

### 10. Souhrnné hodnocení

**70 b. dobře (C)**

Problémem práce je to, že není zřejmé, co je přínosem studenta a které myšlenky byly převzaté, což je zapříčiněno nedůsledným citováním. V důsledku zmiňovaných problémů v experimentální části lze těžko posoudit přesnost výsledného lokalizátoru. V textu se objevují faktické chyby. Schéma na Obrázku 4.5 ne zcela odpovídá implementovanému systému. V odvození výpočtu souřadnic jsou chyby. Nedostatky se vyskytují i v definici normalizované vzájemné korelace.

Na druhou stranu kladně hodnotím to, že student samostatně odvodil výpočet souřadnic, který je založený na průniku hyperbol. Dále jej byl schopen naprogramovat pro výpočetně omezenou platformu. Student nabídl osobní prezentaci bakalářské práce.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 23. června 2020

Mošner Ladislav, Ing.  
oponent