

## Posudek oponenta bakalářské práce

**Student:** Hodaň David  
**Téma:** Konverze hlasu (id 18959)  
**Oponent:** Novotný Ondřej, Ing., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**  
Práce vyžaduje nastudování rozsáhlou problematiku zpracování a syntézy řeči.
- 2. Splnění požadavků zadání** **zadání splněno s drobnými výhradami**  
Student se podle bodu 2 seznámil s dostupnými toolkity na zpracování řeči, ale opomněl dostupná řešení pro hlasovou konverzi, která mohla být použita pro porovnání výsledků. V práci dále zcela chybí objektivní vyhodnocení navrženého systému z bodu 4.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**  
Zpráva je obvyklého rozsahu.
- 4. Prezentační úroveň předložené práce** **80 b. (B)**  
Práce je logicky strukturována s korektní návazností jednotlivých kapitol od teoretického úvodu do dané problematiky až po návrh a implementaci vlastního řešení. V některých pasážích (především o návrhu) působí práce trochu zmateně.  
V kapitole 3, pracuje se z-transformací, která není nijak uvedena a následně chybí vztah mezi uváděnými vzorci v časové oblasti a v z-rovině.
- 5. Formální úprava technické zprávy** **90 b. (A)**  
Po této stránce je těžké najít něco, co by se práci dalo vytknout.
- 6. Práce s literaturou** **90 b. (A)**  
Student cituje 31 zdrojů. Student správně uvádí citace v textu. Tématem se vybraná literatura váže ke zpracovávanému zadání.
- 7. Realizační výstup** **50 b. (E)**  
Student naimplementoval systém pro konverzi hlasu v prostředí Matlab. Výsledky nejsou ovšem příliš přesvědčivé. Na přiložené nahrávce z DVD sleduji pouze minimální posun hlasu k cílovému mluvčímu. V teoretické části student zmiňuje důležitost buzení/základní tónu a hlasového traktu, ale základní frekvenci v práci nepoužívá. Při samotné konverzi modeluje a transformuje pouze hlasový trakt.  
V experimentální a hodnotící části je zcela vynecháno objektivní testování, ač je to uvedeno i jako jeden bod zadání. Z popisu subjektivních testů vyplývá, že je student prováděl pouze poslechem sám a pravděpodobně na jedině testovací nahrávce, což je pro zhodnocení dosažených výsledků velmi nedostačující.
- 8. Využitelnost výsledků**  
Současný stav práce pravděpodobně neumožňuje její další využití. Hlas se po změně jen velmi vzdáleně blíží k cílovému mluvčímu. Také vzhledem k absenci kvalitního testování nejsou k dispozici žádná čísla, která by se dala použít k porovnání s dalšími systémy.
- 9. Otázky k obhajobě**  
Proč je vhodnější použít Hammingovo okno při výseku části signálu oproti obdelníkovému oknu (při odpovědi se zaměřte především na spektrální oblast signálu)?  
Může se stát, že při odhadu koeficientů LPC filtru bude filtr nestabilní?
- 10. Souhrnné hodnocení** **60 b. uspokojivě (D)**  
Po prezentační a typografické stránce působí práce poměrně dobře. Práce ovšem zřejmě trpí nedostatečným proniknutím do dané problematiky. Student se zaměřil na změnu koeficientů artikulačního traktu, zcela ale opomněl v návrhu a implementaci změny základní frekvenci, což může být jedním z důvodů snížené kvality výsledku.  
Po experimentální stránce zcela chybí kvalitní vyhodnocení výsledku aplikace. Z dostupných informací (přiložené DVD a text zprávy) vyplývá, že systém zhodnotil pouze subjektivně a jak bylo již zmíněno, zřejmě na jedině nahrávce. Tento fakt osobně vnímám jako velké mínus a je hlavním důvodem uděleného celkového hodnocení. Osobně bych doporučil, aby pro obhájení práce s navrženým hodnocením, student dopracoval kvalitní objektivní (spektrální vzdálenost, či jiná metrika) i subjektivní (s alespoň 10 posluchači, viz strana 12 v prezentaci: [http://www.fit.vutbr.cz/study/courses/ZRE/public/pred/06\\_kod/06\\_kod.pdf](http://www.fit.vutbr.cz/study/courses/ZRE/public/pred/06_kod/06_kod.pdf)) hodnocení na adekvátně zvolené datové sadě většího rozsahu (alespoň 5-10 nahrávek) a uvedl své poznatky v prezentaci.

V Brně dne: 1. června 2016

.....  
podpis