

Hodnocení vedoucího diplomové práce

Autor práce: **Bc. Lukáš VLADAŘ**

Název práce: **Analýza velikosti dat pro neurální syntézu řeči**

Jazyková a grafická úprava

Nadprůměrné

Samostatnost zpracování tématu

Nadprůměrné

Vhodnost použitých metod

Nadprůměrné

Způsob zpracování a vyhodnocení

Nadprůměrné

Správnost získaných výsledků

Nadprůměrné

Vlastní přínos

Nadprůměrné

Doplnění hodnocení, připomínky:

Diplomová práce se zabývá problematikou neurální syntézy řeči, konkrétně se zaměřuje na vliv množství zdrojových dat (řečových nahrávek daného hlasu profesionálního nebo amatérského řečníka) a vliv strategie trénování řečových modelů na kvalitu výsledné syntetické řeči. Jde o vysoce aktuální úlohu, neboť volba strategie trénování (např. tvorba “single-speaker” modelu vs. využití předtrénovaného “single” nebo “multi-speaker” modelu) bývá jedním z prvních kroků, kterým se musí tvůrce systému syntézy řeči věnovat. Často je přitom snahou minimalizovat počet zdrojových nahrávek daného hlasu.

Diplomant se svého úkolu zhostil výborně. Podrobně se seznámil s aktuálním vývojem v oblasti neurální syntézy řeči a osvojil si práci s frameworkem pro trénování modelů syntézy řeči Coqui-ai/TTS. Dále navrhl a provedl sadu experimentů pro testování vlivu výše zmíněných aspektů na kvalitu výsledné syntetické řeči. Pro vlastní syntézu řeči byl vybrán state-of-the-art end-to-end model VITS. Zvláštní pozornost si zaslouží samotné vyhodnocení kvality syntetické řeči – zde diplomant navrhl, zorganizoval a pečlivě vyhodnotil velké množství poslechových testů typu MUSHRA pro subjektivní hodnocení kvality a navrhl vlastní implementaci výpočtu melovské kepstrální vzdálenosti (MCD) s využitím dynamického programování (DTW) pro objektivní hodnocení kvality. Hlavním výsledkem je návod, resp. sada doporučení, jakou strategii trénování pro daný počet zdrojových dat použít.

Také po obsahové stránce není práci co vytknout. Práce je psána velmi čitelnou formou, jasně a srozumitelně, s velkým počtem obrázků a tabulek, prakticky bez překlepů. Je dobře členěná, s vyváženou kompozicí teoretické a experimentální části. Oceňuji též přehledný souhrn state-of-the-art metod a modelů neurální syntézy řeči.

K diplomové práci a k přístupu Lukáše Vladaře nemám žádné připomínky. Naopak, velice oceňuji aktivní přístup a samostatnost po celou dobu řešení práce. Diplomovou práci celkově hodnotím jako velmi zdařilou.

Splnění bodů zadání

úplně

Doporučení k obhajobě

ANO

V _____ dne _____

Doc. Ing. Jindřich Matoušek, Ph.D.