|  |  |
| --- | --- |
|  |  |

# CMPT 742 Practices on Visual Computing I – Project Proposal

## The Voice: High-Quality Identity-Specific Voice Conversion

### A 90s computer singing into a microphone, white background, cartoon manga styleDescription

Voice conversion methods aim to transfer speaker identity from one audio sample to another. While recent methods have significantly advanced the state-of-the-art in voice conversion, they mostly focus on a many-to-many scenario with a wide range of identities and work on short and non-emotive audio samples.

In this project, we aim to elevate the quality of voice conversion methods by concentrating on a single target identity with abundant training data. In this many-to-one scenario, our goal is to transform any given audio source to match our target identity. The students will try and show that given this more constrained scenario one can create highly realistic results that surpass the current state-of-the-art.

The project will be composed of three different parts:

* **Benchmarking** Students will start by evaluating existing state-of-the-art voice conversion methods on a selected target identity to establish a baseline solution.
* **Refinement** Building on our benchmark, we will then try to improve the results by following the original training regime of the baseline, but with fine-tuning specifically to our target identity.
* **Innovation** Finally, we will explore new solutions for further improving the quality of our results in our many-to-one context.

This project will be executed in PyTorch with the help of existing open-source solutions to provide the baselines. We will specifically start with the ConsistencyVC and compare it against knn-VC and Urhythmic. Students are expected to document their progress throughout the project and analyze their performance both quantitatively and qualitatively on a chosen set of validation audio tracks. The final project is expected to also include a full project repository with a clear readme file, as well as a presentation of the project.

### Contact Person

John Smith?

### Contributor of the Project Idea

?

Please send the filled template to Ali Mahdavi-Amiri at [**amahdavi@sfu.ca**](mailto:amahdavi@sfu.ca).