Intelligibility of Sung Lyrics: A Pilot Study

Karim M. Ibrahim[1], **David Grunberg**[1], **Kat Agres**[2], **Chitralekha Gupta**[1,3] **and Ye Wang**[1]
`{karim.ibrahim, grunberg, chitralekha, wangye}@comp.nus.edu.sg, kat_agres@ihpc.a-star.edu.sg`
[1]School of Computing, National University of Singapore, [2]Institute of High Performance Computing, A*STAR, Singapore , [3]NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore

1. Introduction

- **Estimating the intelligibility of lyrics in songs is a novel and challenging problem with various applications.**
- **Which of these two songs has more intelligible lyrics?**

Chop Suey?

Imagine?

- **Which one would you recommend to an English learner?**
- **Can a machine also measure lyrics intelligibility?**

2. Problem statement

- **Estimating the level of lyrics intelligibility in a given song.**
- **Elements of Intelligible songs [1]:**
  - Loud/dominant singing voice
  - Clear pronunciation
  - Singing rate is not too fast or too slow
  - The usage of simple language
- **Challenges**
  - Dataset
  - Features and model
  - Evaluation
- **Approach**

Dataset collection → Dataset Labelling → Features Selection → Model Training → Evaluation

3. Behavioral Experiment

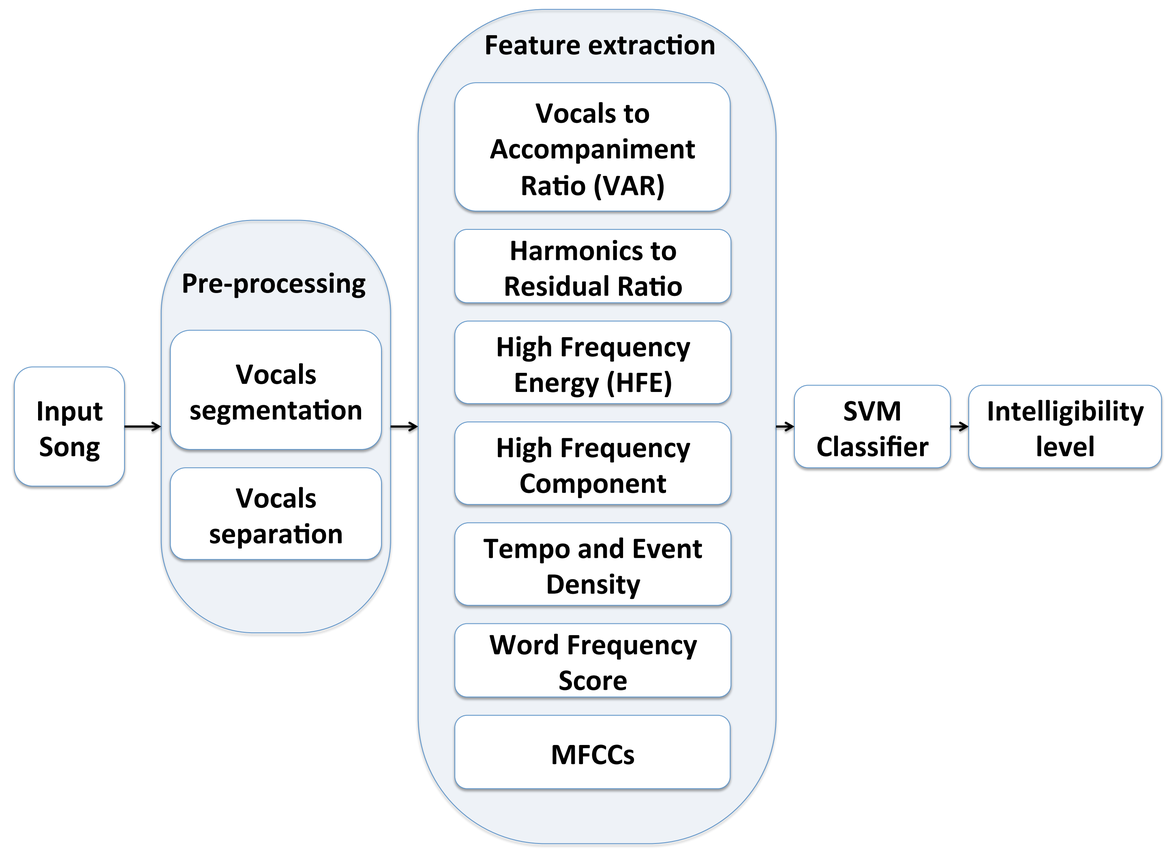
- **Task:** To transcribe the lyrics of an excerpt from a song [2]
- **Purpose:** Transcription accuracy reflected intelligibility
- **Dataset: 100** excerpts from **50** songs, **2** excerpts per song, covering **5** genres, and **10** songs per genre
- **Participants:** 17 participants, 7 females and 10 males
- **Procedure**:
  - Listen to an excerpt twice (average 6.5 seconds)
  - Transcribe the lyrics (10 seconds period)
  - **Intelligibility Score** = avg.(#words correctly transcribed/ total #words)

Sound & Music Computing Lab
http://www.smcnus.org/

4. Computational system

- Extract relevant audio and textual features
- Train a support vector machine classifier to predict level of intelligibility to three classes: High, Moderate and low

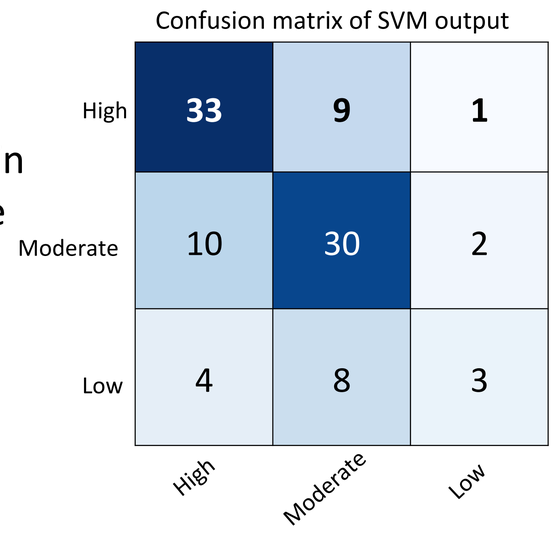
Proposed System



5. Evaluation

- Leave-one-out cross validation.
- Classification accuracy: **66%**
- Model gives higher accuracy in predicting high and moderate intelligibility
- Top Contributing features:
  - MFCC
  - Syllable rate
  - VAR
  - HFE

Confusion matrix of SVM output



	High	Moderate	Low
High	**33**	**9**	**1**
Moderate	10	30	2
Low	4	8	3

6. Conclusion

- The study provides evidence that the proposed system has promising results, and draws attention to the problem of estimating lyrics intelligibility
- Labeling data for intelligibility estimation can be done by a lyrics transcription experiment
- Applications for estimating lyrics Intelligibility include recommending songs for language learners

7. References

[1] Collister, Lauren B., and David Huron. "Comparison of word intelligibility in spoken and sung phrases." (2008).
[2] Condit-Schultz, Nathaniel, and David Huron. "Catching the Lyrics." *Music Perception: An Interdisciplinary Journal* 32.5 (2015): 470-483.