PAPER 1 :A Comparison of Singing Evaluation Algorithms

Author :Partha Lal

Centre for Speech Technology Research,

University of Edinburgh, UK

1 )comparison techniques/Methods :

1)Simple Pitch Comparison :

This was the first approach attempted. Each recording was represented

by taking an evenly spaced selection of 50 pitch samples

and linearly interpolating across unvoiced regions. Since the

clips are now essentially vectors, the Euclidean distance between

a caller’s clip and the reference clip was defined to be the caller’s

score. Various parameters of that algorithm were estimated using

the sung corpus collected in [6], which is described in Section 4.1

2) Dynamically-aligned Pitch Comparison

One of the failings of the simple approach was that it was overly

sensitive to the caller falling slightly out of time compared to the

reference clip. This approach compensated for that by performing

a Viterbi alignment of the caller’s pitch curve and the reference

pitch curve. The algorithm was modified so as to penalise substitution

errors by an amount proportionate to the magnitude of the

pitch difference.

3) Penalising humming

The system scores users who hum just as highly as those who

sing, since the system is based only on pitch estimates. One solution

to this problem that was tested was to combine the existing

pitch-based score, , with the confidence score returned by a

speech recogniser. The two values would then be treated as separate

dimensions and the distance from the origin would represent

the combined score.

1. Algorithms used for the approaches :

There were two main designs considered, both of which involved

comparing pitch estimates of the recordings taken at various intervals.

Pitch estimates were made using the “Robust Algorithm

for Pitch Tracking.

The score returned by each system is a non-negative number

such that a higher score means a greater difference between the

clips. Identical clips score zero.

Another feature common to both approaches is that the reference

clip was manually cleaned up to remove non-vocal audio.

1. Conclusion :

The most effective method considered involved making estimates of frequency and

performing a Viterbi alignment of the pitch curves of the clips to

be compared. Measuring the total pitch difference of the aligned

curves led to a distance metric between the caller and the gold

standard.

1. Datasets :

No dataset

PAPER 2 :The Evaluation of Singing Voice Accuracy: A Comparison Between Subjective and Objective Methods .

Authors :Pauline Larrouy-Maestri, †Yohana Lev eque, ^ ‡Daniele Schon, € †Antoine Giovanni, and \*Dominique Morsomme, \*Liege, Belgium, yzMarseille, France

Algoritms Used : Vocal accuracy can also be objectively quantified by measuring the fundamental frequency (f0) variations along the performance. Since the SINGAD (SINGing Assessment and Development) system of Howard and Welch,6 these acoustic methods have been developed7 and presented as a more reliable solution to evaluate vocal accuracy, as they avoid the natural limits of a subjective judgment

Acoustic analysis consists in segmenting the auditory signal and extracting the f0 of sung vowels. Indeed, vowels carry the maximum of voicing and stable pitch information9 and mark the onsets of musical tones.10 In pitch-matching tasks, vocal accuracy is directly represented by the difference between the produced pitch and the model.

2)Methods used :

1) One hundred sixty-six untrained singers were recruited among the Belgian population. The majority (64%) reported listening to music less than 1 hour a day, only 6% declared going to the concerts more than once a month. None of them were professional musicians, 3% had studied a musical instrument in a conservatory of music and 24% reported a basic musical education in a local music school.

2) The singing voice experts were four speech therapists specialized in singing voice treatment.

3) they performed individually the popular French song ‘‘Happy Birthday’’ a cappella. Participants were instructed to sing ‘‘naturally, while imagining a festive and friendly context.’’ No particular starting note was given to let the participant choose his/her comfortable range.

4) Acoustic analysis. The popular song ‘‘Happy Birthday’’ comprised four phrases. It has 25 notes (each note corresponding to a syllable) and 21 when one ignores the rhythmically concurrent repeated notes. The measurements are based on the melodic intervals. The calculations are grounded on the equal temperament which is a compromise tuning scheme used in Western music for each production, three criteria of vocal accuracy were quantified : pitch interval deviation, number of contour errors, and tonality modulations.

5)Pitch interval deviation : Calculated the difference in cents between each performed interval and the theoretical one. We considered the absolute value of the differences and computed the average score across the entire melody. A small deviation reflects a high precision of intervals.

6)Contour errors :We counted each time that the produced interval direction deviated from the direction of the musical score.

7)Tonality modulation. We computed the number of modulations, defined as an interval error larger than 100 cents not followed by a corrective interval of at least 100 cents in the reverse direction. Thus, modulations indicated that tonality changed during at least three notes.

Conclusion :

We computed a correlation matrix using Spearman coefficient.

Dataset : No dataset was used, live audience were considered.