

Analysis and Detection of Singing Techniques in Repertoires of J-POP Solo Singers



website



Dataset logo

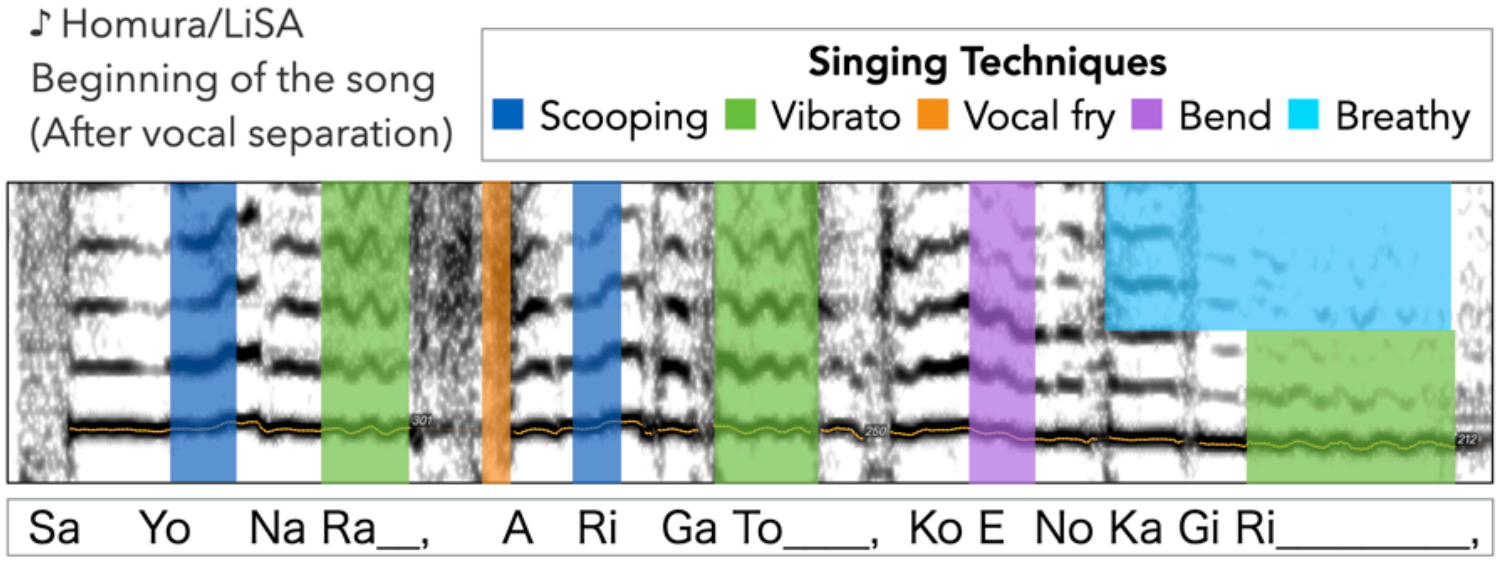
Yuya Yamamoto¹, Juhan Nam², Hiroko Terasawa¹

1: University of Tsukuba, Japan
2: KAIST, South Korea

Abstract

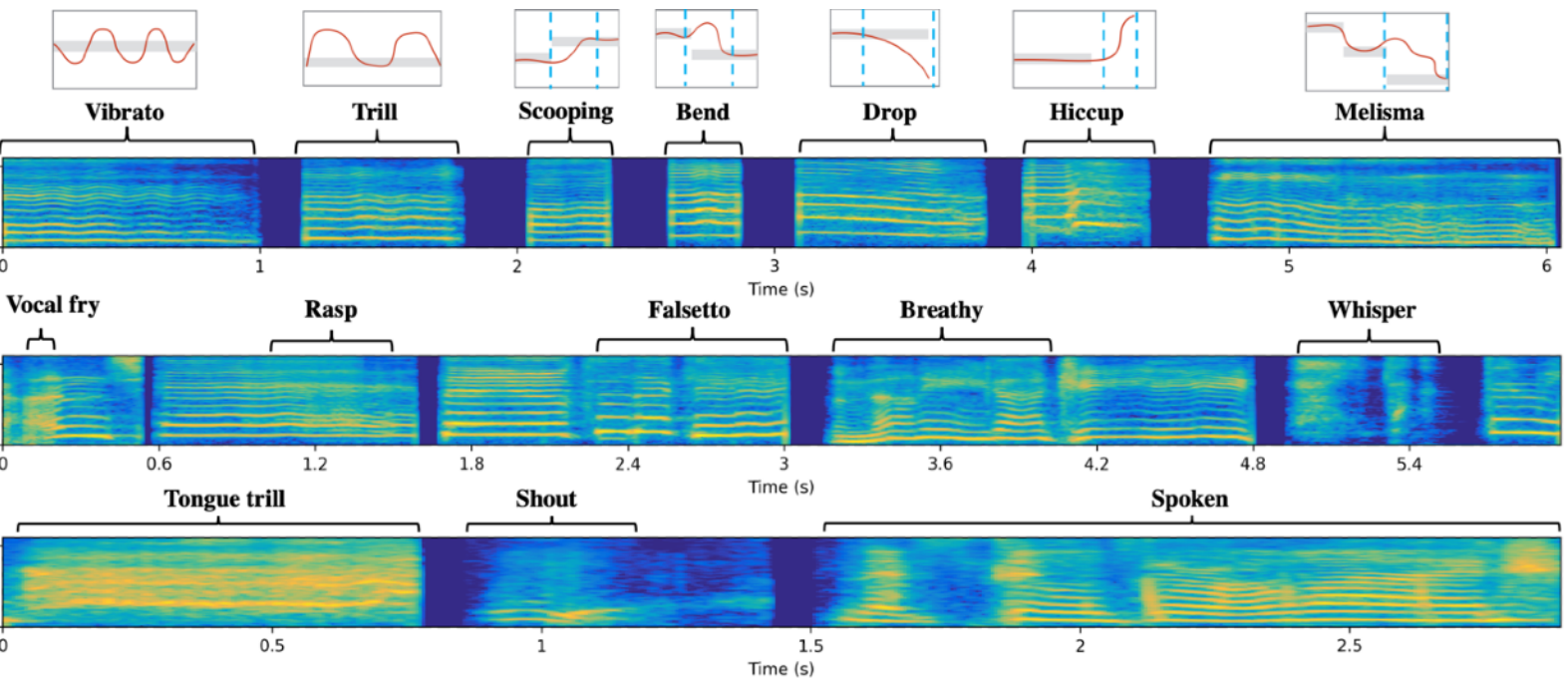
We made annotations on J-POP solo singers' 168 repertoires to analyze singing techniques.

Singing technique is an important component in vocal performance, which is realized by fluctuation of pitch, timbre etc. to render performance.



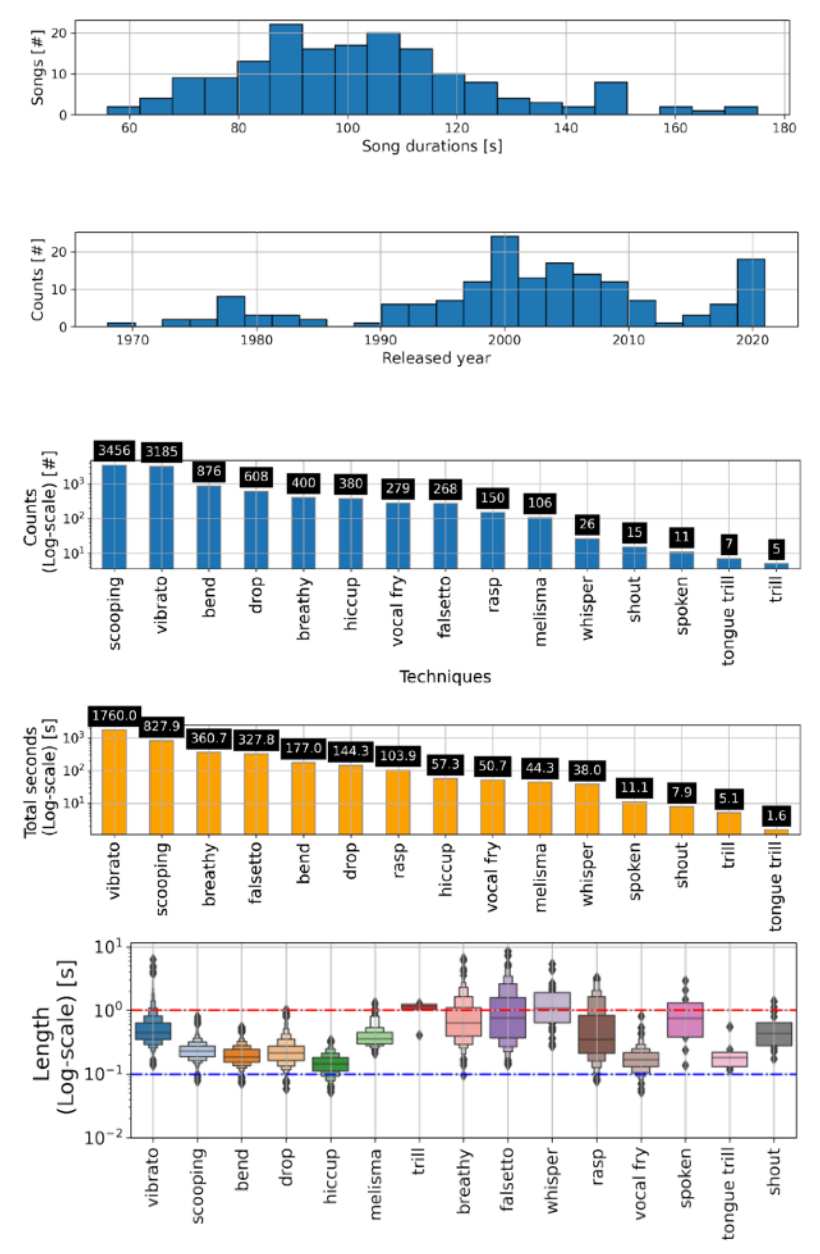
It contains following annotation:

- **Singing technique annotation with timestamp**
- **Pitch of melody**
- URLs of Various music streaming, etc...



We annotated **15 singing techniques** with various types of fluctuation. We determined the targets based on survey and fact-finding listening.

Dataset Overview



Singer ID	vibrato	scooping	bend	drop	hiccup	melisma	trill	breathy	falsetto	whisper	vocal fry	rasp	spoken	tongue trill	shout
ado	66	54	7	5	44	0	2	8	24	0	11	8	0	2	0
al_oostuka	10	7	14	15	1	0	2	1	20	0	1	20	0	0	0
aiko	64	82	37	31	6	4	0	1	8	0	4	18	0	0	0
aimyon	4	96	45	30	0	5	0	4	4	0	0	2	0	0	0
akira_fuse	132	43	10	6	5	5	0	10	2	0	2	1	0	0	0
aya_matsuura	58	73	48	22	47	0	0	18	8	0	0	3	3	0	0
ayaka	69	104	32	4	1	3	0	49	12	0	1	17	0	0	0
ayaka_hirahara	108	84	16	1	0	8	0	16	10	4	0	6	0	0	0
ayumi_hamasaki	80	74	16	6	25	2	0	0	1	0	1	18	0	0	0
chara	10	35	4	8	8	4	0	10	10	16	2	8	1	0	1
chihiro_onitsuka	81	103	18	1	3	8	1	27	3	0	0	11	0	0	0
creepya	0	90	11	67	0	0	0	14	0	0	2	14	0	0	0
flumpool	69	107	3	8	2	0	0	10	3	0	0	1	0	0	0
gackt	56	72	15	23	5	0	0	0	2	0	2	3	0	0	0
hikaru_utada	75	97	23	8	0	2	0	25	16	0	5	18	4	0	0
ikimono_gakari	53	111	18	2	13	0	0	9	7	0	0	1	0	0	0
judy_and_mary	15	35	3	40	0	0	0	10	0	0	0	1	0	0	0
kazumasa_oda	19	49	3	0	1	0	0	4	0	4	0	0	0	0	0
ken_hirai	141	182	63	1	1	10	0	14	24	0	1	18	0	0	0
kenshi_yonezu	69	75	29	5	5	1	0	1	3	0	1	1	0	0	0
koji_tamaki	74	39	14	7	1	0	0	16	6	0	2	6	0	0	0
kumi_koda	55	77	22	7	1	0	0	10	0	0	0	0	0	0	0
laric_en_ciel	96	124	12	44	68	0	0	8	16	0	3	17	0	0	0
lisa	150	110	25	2	27	1	0	22	10	0	3	10	0	0	0
masaharu_fukuyama	60	55	16	9	0	1	0	2	0	0	0	2	0	0	0
masayoshi_yamazaki	88	77	18	12	0	15	0	4	1	0	0	0	0	0	0
mayuki_nakajima	123	23	19	0	0	0	0	40	0	0	15	0	0	0	0
momoe_yamaguchi	56	60	10	29	1	2	0	4	0	0	0	3	0	0	0
mr_children	62	68	9	44	9	1	0	1	6	0	25	2	0	0	0
naotaro_moriyama	141	90	18	2	2	2	0	18	25	0	0	4	0	0	0
noriyuki_makihara	55	83	25	0	2	0	0	2	2	0	0	0	0	0	0
official_hige_dandism	78	86	0	0	22	0	0	40	0	0	1	10	0	0	0
southern_all_stars	75	97	33	13	33	7	0	1	1	0	26	3	1	3	5
spitz	53	117	14	31	0	0	0	1	2	0	0	0	0	0	0
sukima_switch	64	104	22	0	2	2	0	3	10	0	0	0	0	0	0
t_m_revolution	74	43	13	4	13	0	0	1	0	0	0	2	0	0	0
tsuyoshi_nagabuchi	112	60	24	64	0	1	0	0	0	0	20	12	2	2	0
yo_hitoto	72	83	51	4	9	7	0	3	12	0	0	0	0	0	0
yoasobi	11	69	41	0	1	3	0	11	15	1	0	2	0	0	0
yuki_koyanagi	154	112	29	13	15	7	0	10	7	0	1	15	0	0	0
yumi_matsutouya	59	25	11	0	0	0	0	1	0	0	0	0	0	0	0
yutaka_ozaki	81	121	35	25	4	0	0	15	4	0	21	0	0	0	0

Singer-wise Statistics:

We gathered **42 singers** and each singer has **4 songs** in the dataset. we analyzed the singer-wise counts.

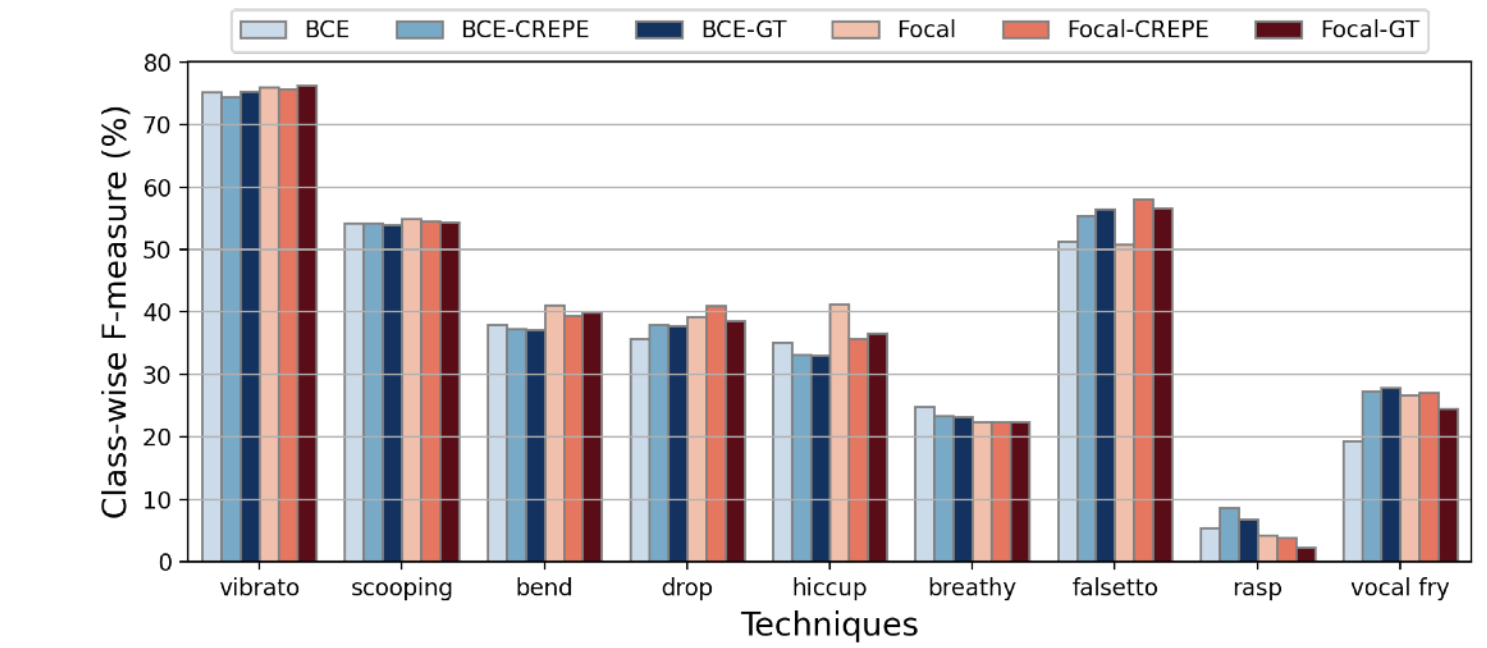
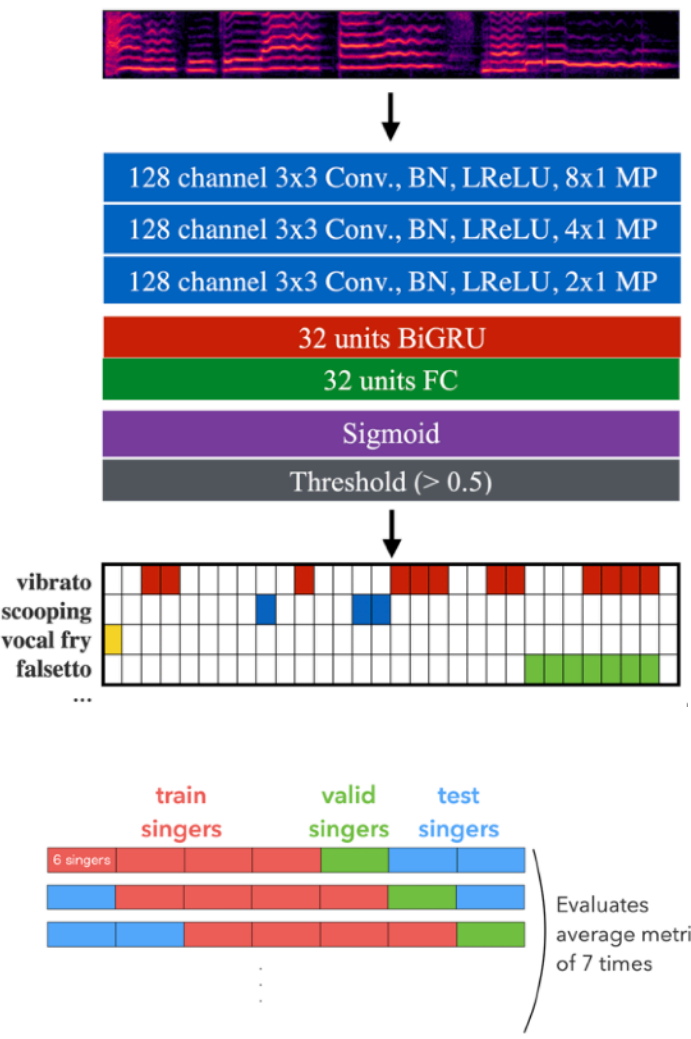
Findings:

- Most common techniques are scooping, vibrato, bend and drop. especially, every singer used scooping.
- Several singer used hiccup, rasp, vocal fry. They might have a power of characterization of singer.

• There are singers who don't use vibrato so frequently

Singing Technique Detection

We also tackled automatic detection towards complete analysis of vocal rendering.



Results: Best model: adding ground truth pitch + Focal loss

- Pitch improves performance on falsetto, Focal loss improves short techniques (bend, drop etc.)
- Breathly, rasp and vocal fry are relatively low -> due to the difference of timbre by singer?
- Common errors: too short detected regions, confusing pattern from fine fluctuation etc.

* [Kim+ ICASSP18]: CREPE: A Convolutional Representation for Pitch Estimation, [Lin+ ICCV 17] Focal Loss for Dense Object Detection