

AnimaSync V3 데이터셋 설명서

0. 한 줄 요약

각 행에 적힌 **대사 + 감정 + 강도(VAD)** 정보를 보고, 함께 제공되는 **mp3 음성**을 들으면서 그에 어울리는 **바디 모션**을 녹화하시면 됩니다.

1. Excel 시트 구조

`dataset_export.xlsx` 파일을 열면 4개 컬럼이 보입니다.

컬럼 이름	설명
No	데이터셋 번호 (1 ~ 2,013)
scenario	시나리오 ID + 상황 + 말투
내용	turn 단위 대화. 화자 · 대사 · 감정 · VAD
mp3 path	turn별 음성 파일 경로 (repo 루트 기준 상대경로)

행 읽는 법

내용 컬럼과 mp3 path 컬럼은 줄(turn) 단위로 매칭됩니다. 내용의 첫 번째 줄 = mp3 path의 첫 번째 줄에 해당하는 mp3 파일.

예시 (1행, daily_003)

내용 컬럼:

A: 점심 뭐 먹죠 [struggle V=-0.18 A=+0.01 D=-0.10]

B: 음... 모르겠어요 매일 같은 거 먹어서 [struggle V=-0.27 A=+0.04 D=-0.20]

A: 새로 생긴 태국 음식점 어때요? [neutral V=+0.07 A=+0.01 D=+0.12]

mp3 path 컬럼:

data/audio_preview/daily_003_t0_struggle.mp3 ← A의 첫 대사 음성

data/audio_preview/daily_003_t1_struggle.mp3 ← B의 첫 대사 음성

data/audio_preview/daily_003_t2_neutral.mp3 ← A의 두 번째 대사 음성

3. 시나리오 종류 3가지

종류	개수	화자 수	turn 수	설명
solo_	253개	1명	1 turn	짧은 한 문장. 감정 하나만 표현
daily_	1,397개	2명 (A-B)	2~6 turn	일상 대화. 화자가 교대로 발화
long_	363개	1명	3~4 turn	모놀로그. 한 화자가 연속 발화하지만, 문장 단위로 turn이 나뉘어 있고 turn마다 감정이 바뀔 수 있음

시트 맨 위 10행은 솔로(solo) 예시 모음

처음 10개 행은 일부러 **다양한 감정의 1문장 솔로 시나리오**를 모아둔 워밍업/레퍼런스 영역입니다.

이 10개를 먼저 작업해 보시면 감정 → 모션 매핑 감이 잡힙니다. 그 다음 11번부터 본격 데이터(daily/long)로 들어가시면 됩니다.

#	시나리오	감정	의도
1	solo_007	neutral	가장 평범한 baseline
2	solo_027	joy	차분한 긍정
3	solo_011	sadness	가라앉는 부정
4	solo_037	anger	강한 부정 + 외향적

#	시나리오	감정	의도
5	solo_073	surprise	갑작스러운 각성
6	solo_045	excitement	강한 긍정 + 외향적
7	solo_116	crying	슬픔의 극단
8	solo_052	laughter	기쁨의 극단
9	solo_020	shy	위축된 긍정
10	solo_069	fluster	당황·긴장

4. 같은 시나리오의 변형판 (_p0, _p1)

시트를 보시면 daily_ 시나리오에서 같은 상황의 대화가 약간씩 다른 버전으로 여러 번 나옵니다.

...

daily_003 → 원본

daily_003_p0 → 같은 상황을 다른 단어로 표현 (paraphrase 0)

daily_003_p1 → 같은 상황을 또 다른 단어로 표현 (paraphrase 1)

...

작업하실 때 알아두실 점

- daily_003, daily_003_p0, daily_003_p1은 같은 상황의 변형이므로 **거의 비슷한 바디 모션**으로 녹화하셔도 됩니다.

5. 감정 16종 설명

V3는 5개 베이스 감정 + 11개 세부 감정 = 총 16개를 다룹니다.

5-1. 베이스 감정 (Base, 5종)

emotion	한국어	핵심 느낌	모션 힌트
neutral	평온	평소 상태	자연스러운 기본 자세, 작고 일상적인 손짓
joy	기쁨	따뜻하고 안정된 긍정	어깨 살짝 올라감, 입꼬리, 부드러운 손짓
sadness	슬픔	가라앉음	어깨 처짐, 시선 아래로, 동작 느려짐
anger	분노	강한 부정 + 공격성	어깨 단단함, 손짓 강하고 빠름, 턱 긴장
surprise	놀람	예상 밖 자극	순간 정지, 눈썹 위로, 몸 살짝 뒤로

5-2. 세부 감정 (Sub, 11종)

emotion	한국어	핵심 느낌
excitement	흥분/들뜸	joy의 강한 버전
laughter	웃음	폭발적 즐거움
gratitude	감사	긍정 + 상대 향함
agreement	동의·수긍	긍정 + 호응
shy	수줍음	긍정인데 위축됨
fluster	당황	갑작스러운 각성 + 통제 잃음
crying	우는 중	sadness의 강한 버전
struggle	고민·답답함	정체된 부정
sulk	뻘뻘	약한 anger + 회피
refusal	거절	부정 + 단호함
apology	사과	위축 + 미안함

6. VAD가 뭔가요?

각 turn에는 [감정 V= \pm 0.00 A= \pm 0.00 D= \pm 0.00] 형태로 3개의 숫자가 붙어 있습니다. 이걸 VAD라고 합니다. 같은 감정이어도 강도와 색깔이 다른데, 그 차이를 -1 ~ +1 범위의 3개 축으로 표현한 게 VAD입니다.

6-1. 3개 축의 의미

축	풀네임	의미
V	Valence	긍정/부정
A	Arousal	활성도/에너지
D	Dominance	주도성/통제력

6-2. 각 축의 값 범위

값	V (Valence)	A (Arousal)	D (Dominance)
-1 (낮음)	매우 부정적	매우 차분, 졸림	위축, 통제 받는 입장
0 (중간)	중립	평상시	중립
+1 (높음)	매우 긍정적	매우 격렬, 흥분	자신감, 통제하는 입장

→ VAD 값의 절댓값이 클수록 더 강하게 표현해 주세요.

7. MP3 파일 안내

7-1. 파일 위치

...

/dataset/AnimaSync-mic-fix/data/audio_preview/

...

시트의 D 컬럼에 적힌 경로(예: data/audio_preview/daily_003_t0_struggle.mp3)가 그대로 파일명입니다.

7-2. 파일명 규칙

...

{시나리오ID}_t{turn번호}_{감정}.mp3

...

예시

- daily_003_t0_struggle.mp3 → daily_003 시나리오, 0번째 turn, struggle 감정
- solo_073_t0_surprise.mp3 → solo_073 시나리오, 0번째 turn, surprise 감정
- long_001_t2_anger.mp3 → long_001 시나리오, 2번째 turn, anger 감정

7-3. 음성 특징

- TTS(Text-to-Speech)로 생성된 합성 음성입니다 (ElevenLabs).
- 대본의 감정에 맞춰 톤·속도·강세가 조절되어 있습니다.
- ***mp3를 먼저 들어보시고*** 모션을 결정하시는 것을 권장합니다.

8. 데이터셋 통계 요약

항목	수치
총 시나리오 수	2,013개
solo (1문장)	253개
daily (다중 화자 대화)	1,397개

항목	수치
long (모놀로그)	363개
총 turn 수 (mp3 개수)	약 6,800개
감정 종류	16개
VAD 차원	3축 (V/A/D), 각 -1 ~ +1
