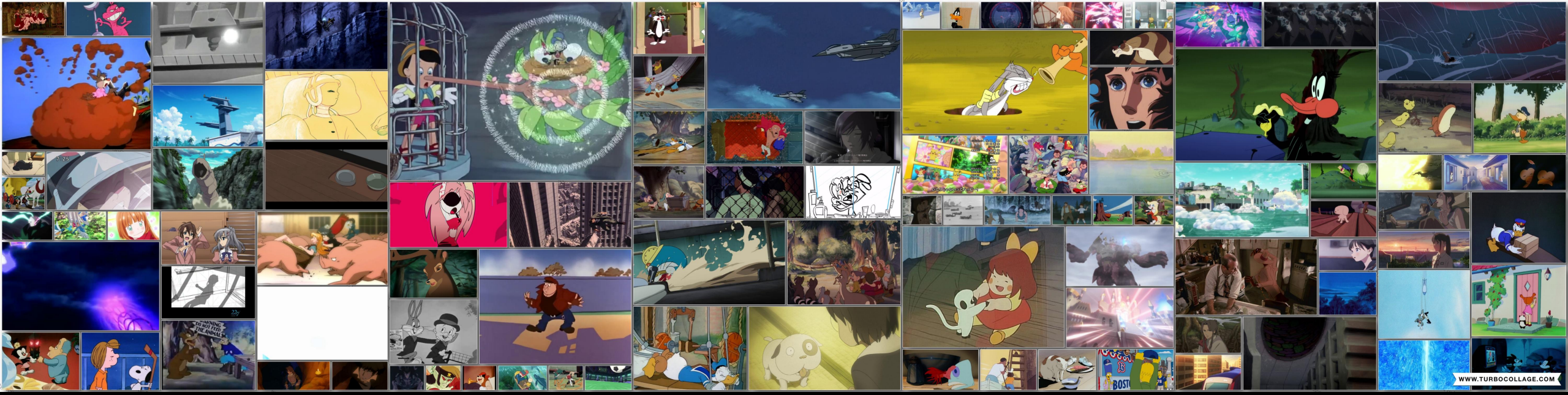
When it comes to

HUGE

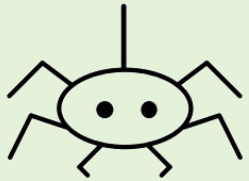
Zhenglin Pan | MEng 22-Fall



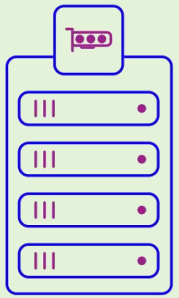
Anita-1M Dataset

- A huge dataset with 1.2M 2D-Cartoon Animation videos.
- Related research: GM, VFI, HCI...
- It is on going

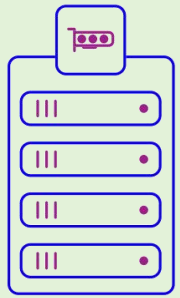
Internet Collection



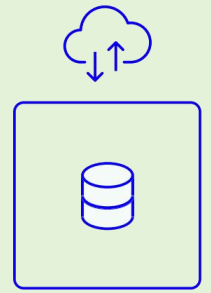
Scene Splitting



Caption Synthesis



Store Data



YouTube

pixiv

63672.mp4

1 video
= multiple clips(scenes)

Splitting

Scene#1

63672_scene3_0.jpg

63672_scene3_1.jpg

Scene#2

63672_scene4_0.jpg

63672_scene5_0.jpg

Scene#3

63672_scene6_0.jpg

63672_scene3_0.jpg

63672_scene3_1.jpg

1 girl, dancing, pom-pom, short shirts

1 girl, rotating, pom-pom, white shirts

LLM

“1 girl is dancing with pom-pom, who wears short shirts and rotating”

Video

63672.mp4

\+

Text

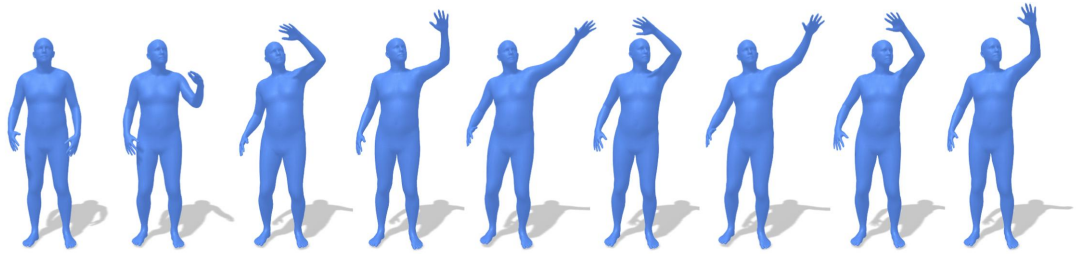

```
{"id": 63672,
"tags": "animated anima_yell! dancing fabri
 "created_at": 1543167118,
 "updated_at": 1543256605,
 "creator_id": 508,
 "approver_id": null,
 "author": "Ashita",
 "change": 541738,
 "source": "#08 (NC) ",
 "score": 142,
 "md5": "9a452fad399523f3698572537337e65c",
 "file_size": 5693587,
 "file_ext": "mp4",
 "file_url": "https://www.sakugabooru.com/d
 "is_shown_in_index": true,
 "preview_url": "https://www.sakugabooru.co
 "preview_width": 150,
 "preview_height": 85,
 "actual_preview_width": 300
}
```


1.2M pairs of Video-Text

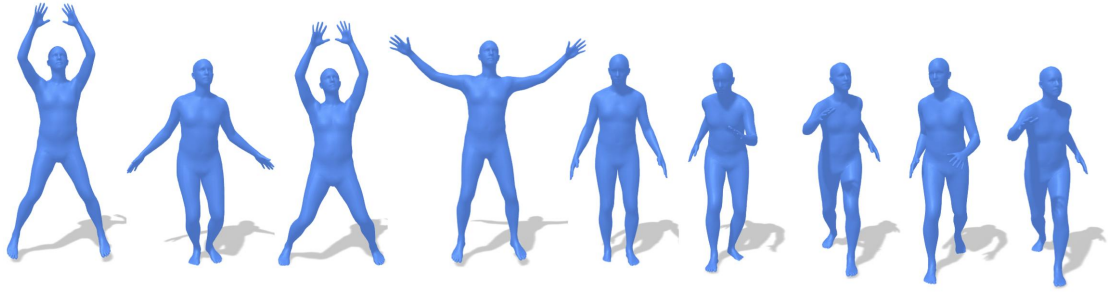
Ultra-max Pro-plus XXXL

HumanML3D[1]

10k samples(10h+ to compile)



1. The person is leaving at someone with his left hand.
2. A person shakes an item with his left hand.
3. A person waves his left hand repeatedly above his head.



1. A person doing jumping jacks and then running on the spot.
2. A person is doing jumping jacks, then starts jogging in place.
3. A person does four jumping jacks then three front lunges.

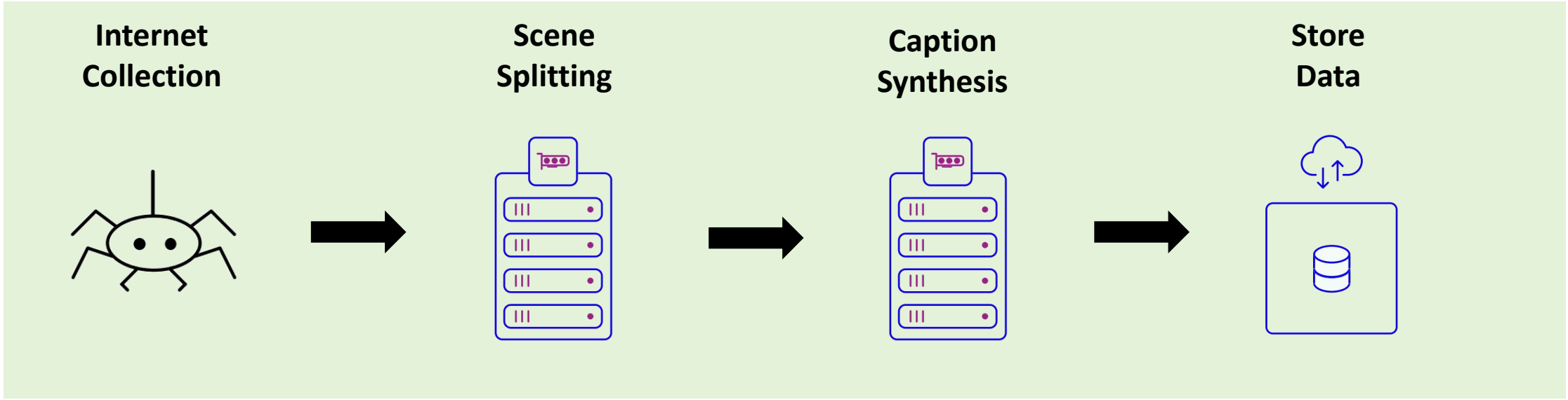
Objaverse-XL[2]

A Universe of 10M+ 3D Objects

arXiv | Google Colab | GitHub | Hugging Face

This is NOT a lesson

I am just sharing...



Download → Store → Process

Download ➡ Store ➡ Process

How to download large data?

Objaverse-XL
A Universe of 10M+ 3D Objects
arXiv Google Colab GitHub Hugging Face

= 100 TB

Distributed Nodes!

Rich Solution: Cloud Service

- AWS[3] Distributed cluster

- Google Cloud Batch data

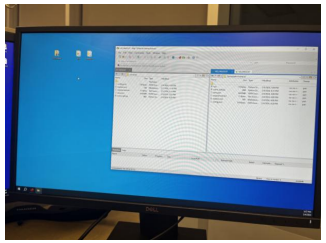
Poor Solution: Mannual Distribution

- Use different PC with different WIFI

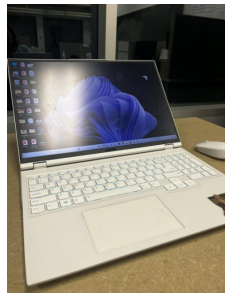
What if we only have 1 computer?

5 Nodes (10T, 5~10d)

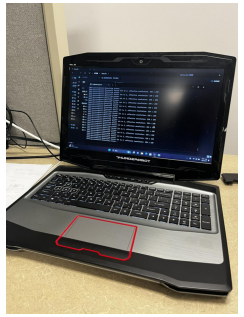
Node#1
1号节点



Node#2
2号节点



Node#3
3号节点



Node#4
4号节点

Node#5
5号节点

Download → Store → Process

How to Store large data?

Distributed Nodes!

Rich Solution 1: HDFS

- Hadoop Hive 

Rich Solution 2: NAS

- NAS + Raid5 

Poor Solution: Cold backup

- Spend some money...

Objaverse-XL

A Universe of 10M+ 3D Objects

arXiv | Google Colab | GitHub | Hugging Face

= 100 TB

16 TB

Sponsored

UnionSine 16TB 3.5" Portable External Hard Drive USB3.0 HDD Storage Compatible for PC, Desktop, Laptop(Black) HD3510

4.5 ★ 3,391

50+ bought in past month

$345.20

prime One-Day

FREE delivery **Tomorrow, Feb 5**

Add to Cart

16 TB

Sponsored

UnionSine 16TB 3.5" Portable External Hard Drive USB3.0 HDD Storage Compatible for PC, Desktop, Laptop(Black) HD3510

4.5 ★ 3,391

50+ bought in past month

$345.20

prime One-Day

FREE delivery **Tomorrow, Feb 5**

Add to Cart

Efficient Data structure

1. Combine your data, reduce IO times

Never use small files!

random reading/writing is **SLOOOOOW**!

1.json

{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}

2.json

{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}

all.json

{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}
{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}
{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}
{"not_aesthetic": 0.5230167508125305,
"aesthetic": 0.4769831895828247}

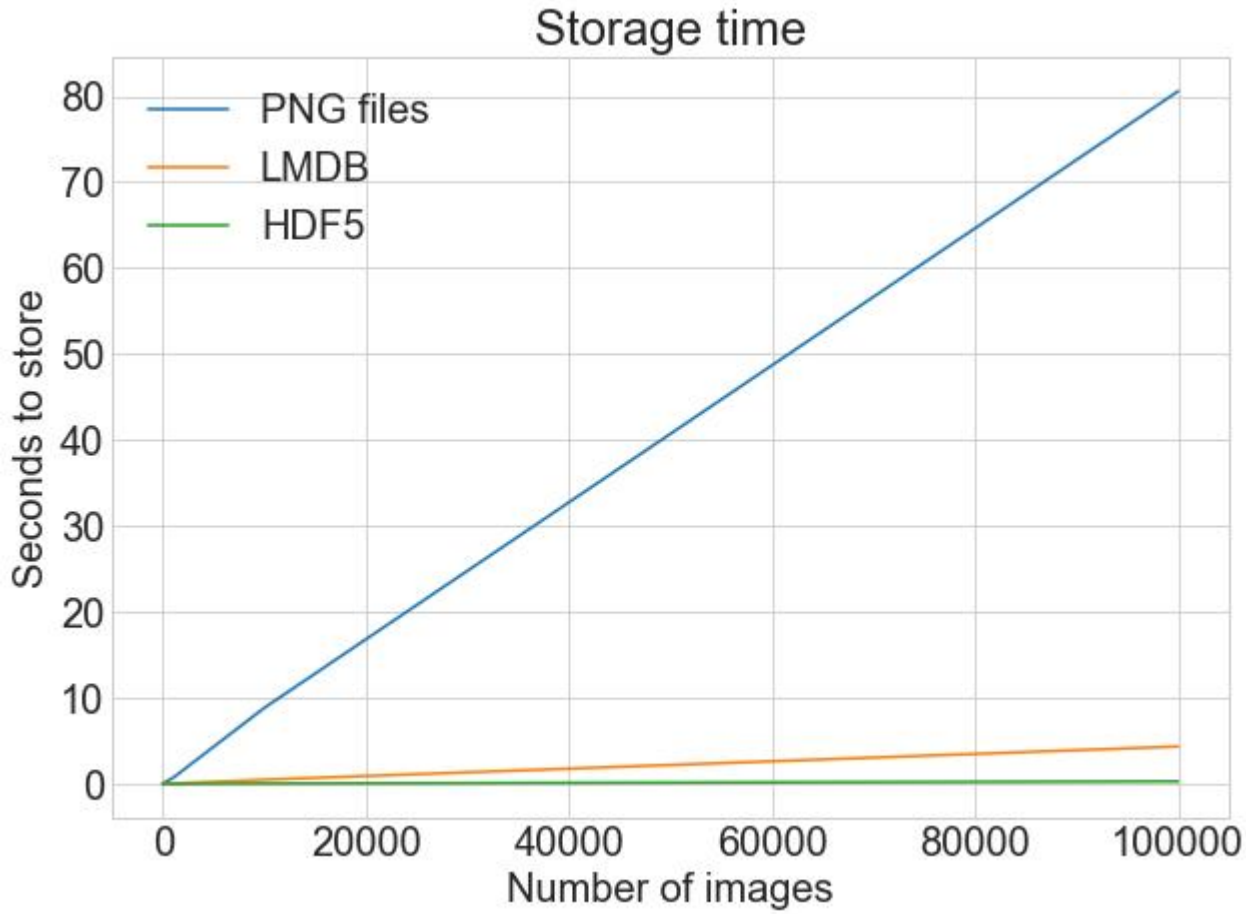
Efficient Data structure

2. Use formats designed for large data

HDF5 = jpeg, npy, txt, ...

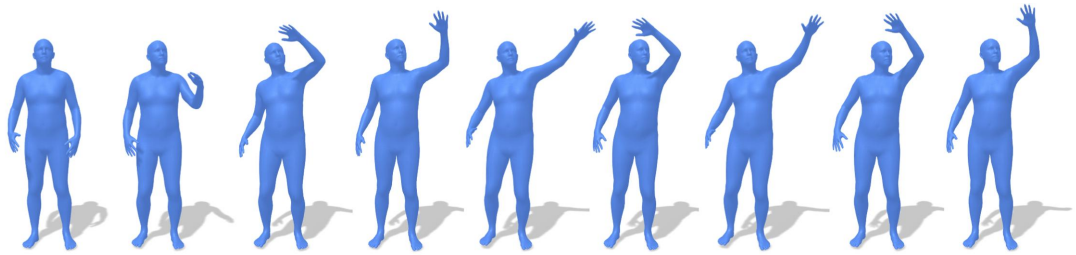
LMDB = dict, jpeg, txt, ...

Parquet = csv, json, txt, ...

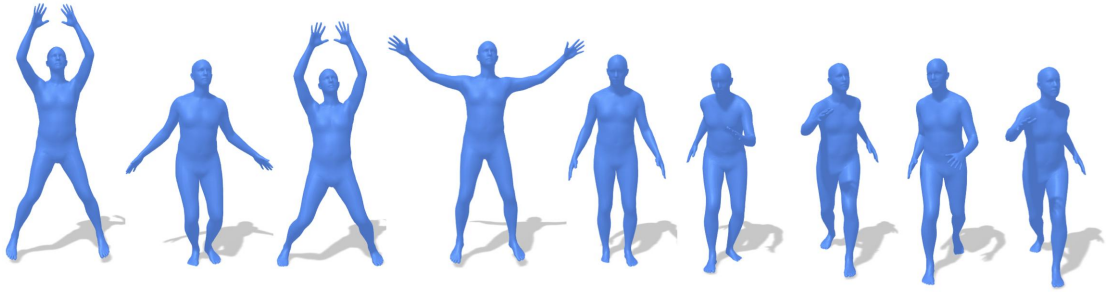


Download → Store → Process

HumanML3D[1]



1. The person is leaving at someone with his left hand.
2. A person shakes an item with his left hand.
3. A person waves his left hand repeatedly above his head.



1. A person doing jumping jacks and then running on the spot.
2. A person is doing jumping jacks, then starts jogging in place.
3. A person does four jumping jacks then three front lunges.

Previous

14,616 npy files
motion

14,616 .txt files
text description

Current

1 hdf5 files
{text:motion}

Download → Store → Process

How to Process large data?

HumanML3D
Ultra-max ProPlus XXXL = 10M files

There’re only rich solutions...

Cluster Level:

- Spark
- SLURM

Unit Level:

- DP / DDP / Hf Accelerate
- torch.multiprocessing
- import process

Refernece

- [1] Guo, Chuan, et al. "Generating diverse and natural 3d human motions from text." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
- [2] Deitke, Matt, et al. "Objaverse-xl: A universe of 10m+ 3d objects." arXiv preprint arXiv:2307.05663 (2023).
- [3] Schuhmann, Christoph, et al. "Laion-5b: An open large-scale dataset for training next generation image-text models." Advances in Neural Information Processing Systems 35 (2022): 25278-25294.