MultiEgo Dataset Instruction

The dataset contains 5 scenes: talking, statement, concert, sword, and presentation. Each scene provide video, camera intrinsic, camera poses, timestamp, and a sparse point cloud of the first frame scene.

The file construction is as follows:

```
scene
|-cam1
| |-<scene>-cam1.mp4
| |-intrinsic.txt
| |-camera_poses.txt
| |-sampletime.txt
|-cam2
|-cam3
|-cam4
|-cam5
|-sparse
  |-camera.bin
  |-images.bin
  |-points3D.bin
  |-points3D.ply
```

where `<scene>-camx.mp4` is the egocentric video of the performer x in the scene. If frame extraction is performed on all videos, it is recommended to reserve 25 GB of storage space.

`intrinsic.txt` is the intrinsic matrix of the camera x, in the format as:

$$\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

`camera_poses.txt` is the camera poses matrix of the frames in the `<scene>-camx.mp4`. The camera poses are represented as camera-to-world transformations in the world coordinate system. The pose in the format as:

$$\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

`sampletime.txt` is the capture time of the acquisition system. The data in `sampletime.txt` is in the unit of nano-second.

The `sparse` directory contains COLMAP binary files for all images, including intrinsic camera parameters (`camera.bin`) and world-to-camera extrinsic transformations (`images.bin`). The `images.bin` file names follow the naming convention `camx_frame_00000.png`. Additionally, we provide sparse 3D point clouds reconstructed from the first frame's images and extensive images, stored in `points3D.bin` and `points3D.ply`.

Data Loading Example

a data loading pipeline example: Modified from `dataset_readers.py` in 4DGaussian

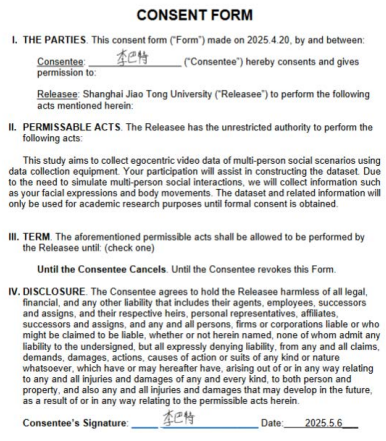
```
#
# Copyright (C) 2023, Inria
# GRAPHDECO research group, https://team.inria.fr/graphdeco
# All rights reserved.
#
# This software is free for non-commercial, research and evaluation use
# under the terms of the LICENSE.md file.
#
# For inquiries contact  george.drettakis@inria.fr
#

import os
...

class CameraInfo(NamedTuple):
...

class SceneInfo(NamedTuple):
...

def getNerfppNorm(cam_info):
...


def getTimeScale(scene):
    timescale=[]
    for num in range(1,6):
        with open(f'path/to/{scene}/cam{num}/sampletime.txt','r') as f:
            txt=f.readlines()
        time=[]
        for i in range(1,len(txt)): # the first line is an annotation
            time.append(float(txt[i]))
        timescale.append(time)
    timescale=np.array(timescale)
    timescale/=np.max(timescale)
    return timescale

def getImageFolder(scene,cid,iid):
    return f'/path/to/video/frames/{scene}/cam{cid}/frame_{iid:05d}.png'
    # In this example, the frames is in {scene}/cam{x}/frame_00000.png

def getCandIid(name): # get cam and image id from image_name in colmap .bin file
    split=name.split('_')
    cid=int(split[0][-1])
    iid=int(split[2])
    return cid,iid

def readColmapCameras(scene,cam_extrinsics, cam_intrinsics, images_folder):
```

```
    # Read the entire timeline before the loop
    timescale_all=getTimeScale(scene)

    cam_infos = []
    for idx, key in enumerate(cam_extrinsics):

        sys.stdout.write('\r')
        sys.stdout.write("Reading camera {}/{}".format(idx+1, len(cam_extrinsics)))
        sys.stdout.flush()

        # scale the camera and image
        scale=0.5

        extr = cam_extrinsics[key]
        intr = cam_intrinsics[extr.camera_id]
        height = int(intr.height*scale)
        width = int(intr.width*scale)


        uid = intr.id
        R = np.transpose(qvec2rotmat(extr.qvec))
        T = np.array(extr.tvec)

        if intr.model in ["SIMPLE_PINHOLE", "SIMPLE_RADIAL"]:
            focal_length_x = intr.params[0]*scale
            FovY = focal2fov(focal_length_x, height)
            FovX = focal2fov(focal_length_x, width)
        elif intr.model=="PINHOLE":
            focal_length_x = intr.params[0]*scale
            focal_length_y = intr.params[1]*scale
            FovY = focal2fov(focal_length_y, height)
            FovX = focal2fov(focal_length_x, width)
        elif intr.model == "OPENCV":
            focal_length_x = intr.params[0]*scale
            focal_length_y = intr.params[1]*scale
            FovY = focal2fov(focal_length_y, height)
            FovX = focal2fov(focal_length_x, width)
        else:
            assert False, "Colmap camera model not handled: only undistorted datasets
(PINHOLE or SIMPLE_PINHOLE cameras) supported!"

        # get cam num and frame id from image_name
        cam_num,img_id=getCandIid(os.path.basename(extr.name).split(".")[0])
        # get image path
        image_path = getImageFolder(scene,cam_num,img_id) # os.path.join(images_folder,
os.path.basename(extr.name))
        image_name = os.path.basename(image_path).split(".")[0]
        image = Image.open(image_path).resize((width,height))
        image = PILtoTorch(image,None)
        # get timestamp (or automatic allocation)
        time= timescale_all[cam_num-1,img_id-1] # float(img_id/len(timescale_all[0]))
```

```
        cam_info = CameraInfo(uid=uid, R=R, T=T, FovY=FovY, FovX=FovX,
image=image,camera_id=cam_num,image_path=image_path, image_name=image_name, width=width,
height=height,time = time, mask=None)
        cam_infos.append(cam_info)
    sys.stdout.write('\n')
    return cam_infos

def fetchPly(path):
    plydata = PlyData.read(path)
    vertices = plydata['vertex']
    positions = np.vstack([vertices['x'], vertices['y'], vertices['z']]).T
    colors = np.vstack([vertices['red'], vertices['green'], vertices['blue']]).T / 255.0
    # no such normals
    normals = np.vstack([0, 0, 0]).T
    return BasicPointCloud(points=positions, colors=colors, normals=normals)

def storePly(path, xyz, rgb):
...

# the boundaries of different scene
bound={'talking':[[-15,-5,-20],[25, 7, 14]],
       'statement':[[-15,-8,-25],[12, 6, 11]],
       'concert':[[-12,-15,-17],[15,7,12]],
       'sword':[[-10,-16,-5],[16, 5, 20]],
       'presentation':[[-10,-6,-3],[8, 5, 12]]}

# generate random point cloud
def randomPCD(scene):
    num=1e5
    xyz_scale=bound[scene]
    x=np.random.uniform(xyz_scale[0][0],xyz_scale[1][0],num)
    y=np.random.uniform(xyz_scale[0][1],xyz_scale[1][1],num)
    z=np.random.uniform(xyz_scale[0][2],xyz_scale[1][2],num)
    colors = np.random.randint(0, 256, size=(num, 3))
    normals = np.zeros((num, 3))
    xyz=np.array([x,y,z]).T

    return BasicPointCloud(points=xyz, colors=colors, normals=normals)


def readColmapSceneInfo(path, images, eval, llffhold=8):
    # get scene
    scene=path.split('/')[-1]

    try:
        cameras_extrinsic_file = os.path.join(path, "sparse/0", "images.bin")
        cameras_intrinsic_file = os.path.join(path, "sparse/0", "cameras.bin")
        cam_extrinsics = read_extrinsics_binary(cameras_extrinsic_file)
        cam_intrinsics = read_intrinsics_binary(cameras_intrinsic_file)
    except:
        cameras_extrinsic_file = os.path.join(path, "sparse/0", "images.txt")
        cameras_intrinsic_file = os.path.join(path, "sparse/0", "cameras.txt")
        cam_extrinsics = read_extrinsics_text(cameras_extrinsic_file)
```

```
        cam_intrinsics = read_intrinsics_text(cameras_intrinsic_file)

    reading_dir = "images" if images == None else images
    cam_infos_unsorted = readColmapCameras(scene,cam_extrinsics=cam_extrinsics,
cam_intrinsics=cam_intrinsics, images_folder=os.path.join(path, reading_dir))
    cam_infos = sorted(cam_infos_unsorted.copy(), key = lambda x : x.image_name)

    if eval:
        train_cam_infos = [c for idx, c in enumerate(cam_infos) if idx % llffhold != 0]
        test_cam_infos = [c for idx, c in enumerate(cam_infos) if idx % llffhold == 0]

    else:
        train_cam_infos = cam_infos
        test_cam_infos = []

    nerf_normalization = getNerfppNorm(train_cam_infos)

    ply_path = f"/path/to/random/pointcloud/{scene}/randomply.ply"

    bin_path = os.path.join(path, "sparse/0/points3D.bin")
    txt_path = os.path.join(path, "sparse/0/points3D.txt")
    if not os.path.exists(ply_path):
        print("Converting point3d.bin to .ply, will happen only the first time you open the
scene.")
        try:
            xyz, rgb, _ = read_points3D_binary(bin_path)
        except:
            xyz, rgb, _ = read_points3D_text(txt_path)
        storePly(ply_path, xyz, rgb)

    ## choose one
    # pcd=randomPCD()
    pcd = fetchPly(ply_path)


    scene_info = SceneInfo(point_cloud=pcd,
                           train_cameras=train_cam_infos,
                           test_cameras=test_cam_infos,
                           video_cameras=train_cam_infos,
                           maxtime=0,
                           nerf_normalization=nerf_normalization,
                           ply_path=ply_path)
    return scene_info
def generateCamerasFromTransforms(path, template_transformsfile, extension, maxtime):
    ...
### no changes followed
```

consent forms



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

Consentee: 李巴特 ("Consentee") hereby consents and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

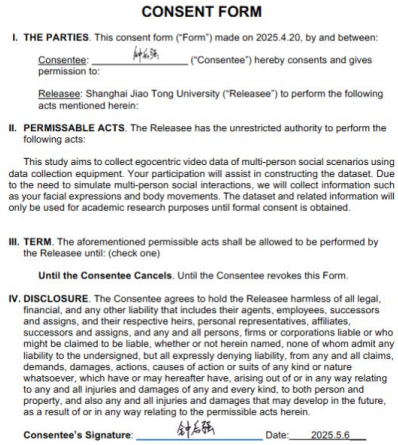
This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 李巴特 Date: 2025.5.6



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

Consentee: 钟石琪 ("Consentee") hereby consents and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

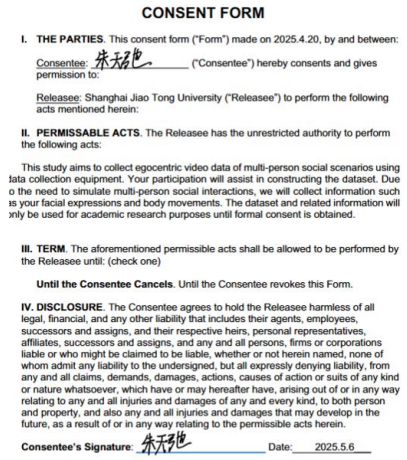
This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 钟石琪 Date: 2025.5.6



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

Consentee: 朱玥艳 ("Consentee") hereby consents and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

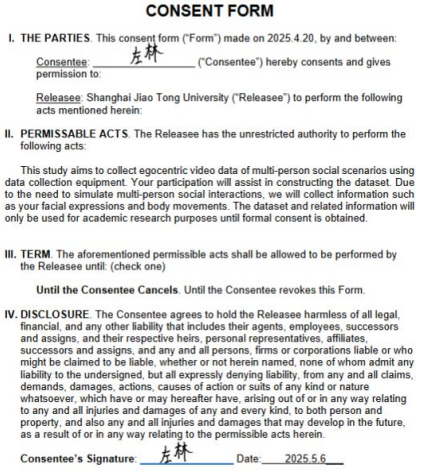
This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 朱玥艳 Date: 2025.5.6



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

Consentee: 左林 ("Consentee") hereby consents and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

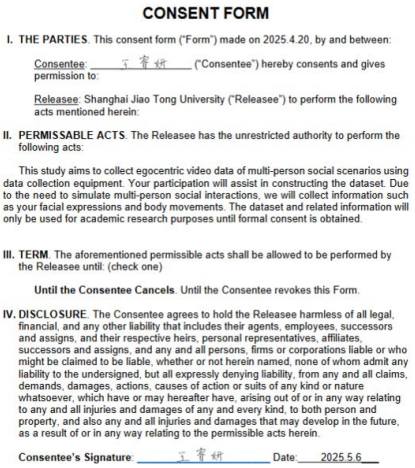
This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 左林 Date: 2025.5.6



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

Consentee: 王宥妍 ("Consentee") hereby consents and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

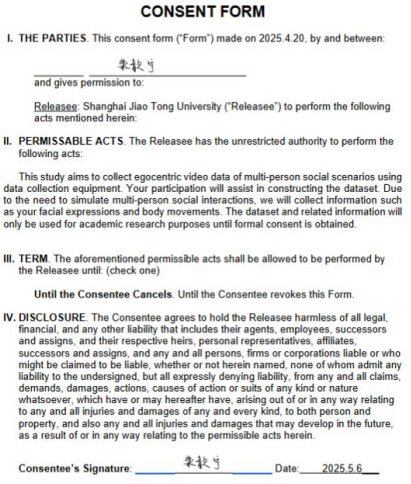
This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 王宥妍 Date: 2025.5.6



CONSENT FORM

I. THE PARTIES. This consent form ("Form") made on 2025.4.20, by and between:

朱软兮 and gives permission to:

Releasee: Shanghai Jiao Tong University ("Releasee") to perform the following acts mentioned herein:

II. PERMISSABLE ACTS. The Releasee has the unrestricted authority to perform the following acts:

This study aims to collect egocentric video data of multi-person social scenarios using data collection equipment. Your participation will assist in constructing the dataset. Due to the need to simulate multi-person social interactions, we will collect information such as your facial expressions and body movements. The dataset and related information will only be used for academic research purposes until formal consent is obtained.

III. TERM. The aforementioned permissible acts shall be allowed to be performed by the Releasee until: (check one)

Until the Consentee Cancels. Until the Consentee revokes this Form.

IV. DISCLOSURE. The Consentee agrees to hold the Releasee harmless of all legal, financial, and any other liability that includes their agents, employees, successors and assigns, and their respective heirs, personal representatives, affiliates, successors and assigns, and any and all persons, firms or corporations liable or who might be claimed to be liable, whether or not herein named, none of whom admit any liability to the undersigned, but all expressly denying liability, from any and all claims, demands, damages, actions, causes of action or suits of any kind or nature whatsoever, which have or may hereafter have, arising out of or in any way relating to any and all injuries and damages of any and every kind, to both person and property, and also any and all injuries and damages that may develop in the future, as a result of or in any way relating to the permissible acts herein.

Consentee's Signature: 朱软兮 Date: 2025.5.6