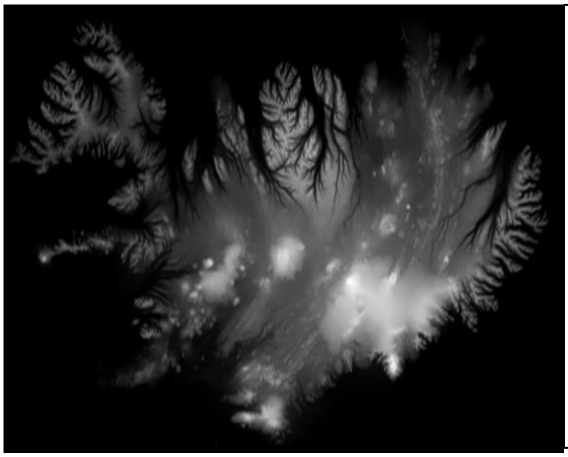
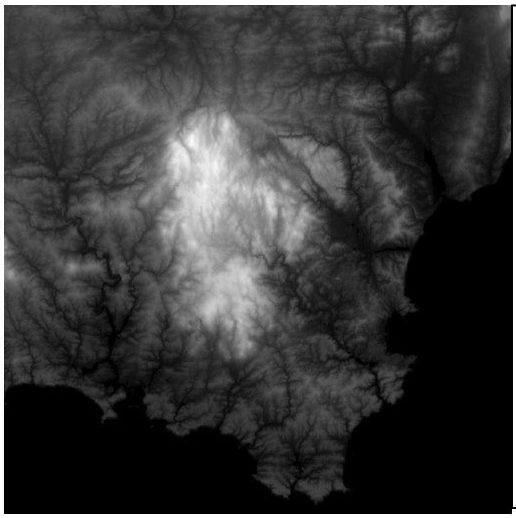
3D Camera Pose Estimation - User Manual & Code Description

Digital Elevation Maps (DEMs)

Below are the grayscale images used to generate 3D terrain for this project. The pixel intensity determines the elevation (height):



My First DEM



My Second DEM

Detailed System Documentation

1. System overview

This app renders a 3D terrain from a grayscale image (height = pixel intensity × scale), and estimates camera pose on that terrain from 2D features. It has two top-level feature modes: Preprocessing (Feature mode ON + Preprocessing ON): capture “key views” and

precompute 2D→3D.
Runtime (Feature mode ON + Preprocessing OFF): match current view to precomputed data and solve camera pose.

2. How the code runs (flow)

At a glance: Load heightmap → build terrain → interact and move camera → optionally record poses → optional picking-based PnP → feature preprocessing (capture views) → runtime matching + PnP → visualize estimated vs. true path.

1) Startup (terrain build)

Loads dem*.png/jpg into a height field; lifts a grid to 3D (y = gray × gHeightScale), builds triangles, computes normals, uploads to GPU buffers.
Creates three windows/contexts (Main, Overview, on-demand Computed View), compiles shaders, sets camera.

2) Free-fly + recording (optional)

Move camera with W/A/S/D/Q/E (+ Shift for speed).
R toggles recording; B appends the current true pose to a recorded path; Ctrl+R enables playback mode (N/V or arrows jump between recorded poses).

3) Picking-based PnP (optional)

With depth picking (Ctrl+B) ON, right-click to pick terrain points.
Build 2D–3D correspondences from MAIN view pixels and picked 3D points; Ctrl+P runs SolvePnP and opens a computed view of the estimated pose.

4) Feature preprocessing (Feature mode ON + Preprocessing ON)

Press B to capture a key view: read the MAIN framebuffer, detect SIFT keypoints/descriptors, cast per-feature rays through the captured pose and intersect the heightmap to assign accurate 3D points; store per-view data (pose, K from FOV/size, image, features, 2D→3D).

5) Runtime feature matching (Feature mode ON + Preprocessing OFF)

Press B to match the live view against stored key views:
Detect SIFT in the live frame; FLANN/BF KNN + Lowe ratio (0.65) + mutual-check; keep only matches whose DB endpoints have mapped 3D.
SolvePnPRansac (AP3P/EPNP) with high confidence (0.995) → refine with Levenberg–Marquardt.
Store gEstCamPos and gEstRwc, draw true-vs-estimated paths in the Overview, and open the computed view window.

6) Visualization

Main: lit terrain, true camera POV.
Overview: top-down orthographic terrain; shows true (black) and estimated (red) oriented triangles, paths (green/red polylines), and per-frame black dots trail while Feature mode is ON.
Computed view: third window showing the estimated camera’s POV when a pose is found.

3. Main code parts

Graphics & scene

Terrain build/upload: heightfield sampling, grid/indices, normals, VBO/EBO/VAO for each context.

Shaders: terrain fragment shader maps height→color (blue→red gradient; widened red band near max), simple line/point shader for paths/dots.

Windows: MAIN (perspective), OVERVIEW (orthographic top-down), COMPUTED (created on demand).

Interaction & state

processKeys(...): keyboard handling (movement, modes, recording, feature controls, picking, SolvePnP, playback).

Modes: gFeatureMode (F), gPreprocessingMode (M), recording/playback, depth picking.

Capture & database (feature preprocessing)

Capture key view: grabs MAIN color buffer, saves pose, FOV/size, detects SIFT (2000), maps each 2D keypoint to 3D via terrain-ray intersection; stores (viewId, keyIndex) and 3D world pos per feature.

Matching & pose (runtime)

5. How to use (keys)

Move: W/A/S/D/Q/E (+ Shift speed).

Recording: R (toggle), B (record a true pose while recording ON), Ctrl+R (playback mode), N/V or arrows (browse recorded poses in playback).

Picking: Ctrl+B (toggle depth picking), right-click in MAIN to pick; Ctrl+S (save pairs), U (undo), C (clear), Ctrl+P (SolvePnP).

Feature mode: F (toggle Feature mode; starts/stops black dot trail in Overview).

Preprocessing: M (toggle Preprocessing). While Feature mode ON:

Preprocessing ON: B captures a key view (builds 2D→3D database).

Preprocessing OFF: B runs matching and pose estimation; shows computed view and true-vs-est paths.

N/M: browse time steps in Overview when paths exist.

K: print current intrinsics for MAIN.

Visuals:

Night-like background/lighting when Feature + Preprocessing ON.

Overview always shows oriented triangles; in runtime after matches, also shows true/est paths and current-step markers; Feature mode draws black dots trail until toggled off.

System Flow Overview

The application consists of several phases: terrain generation, keyframe capture, picking of 2D–3D correspondences, and camera pose estimation.

1. Load grayscale DEM and generate 3D terrain from pixel intensity.

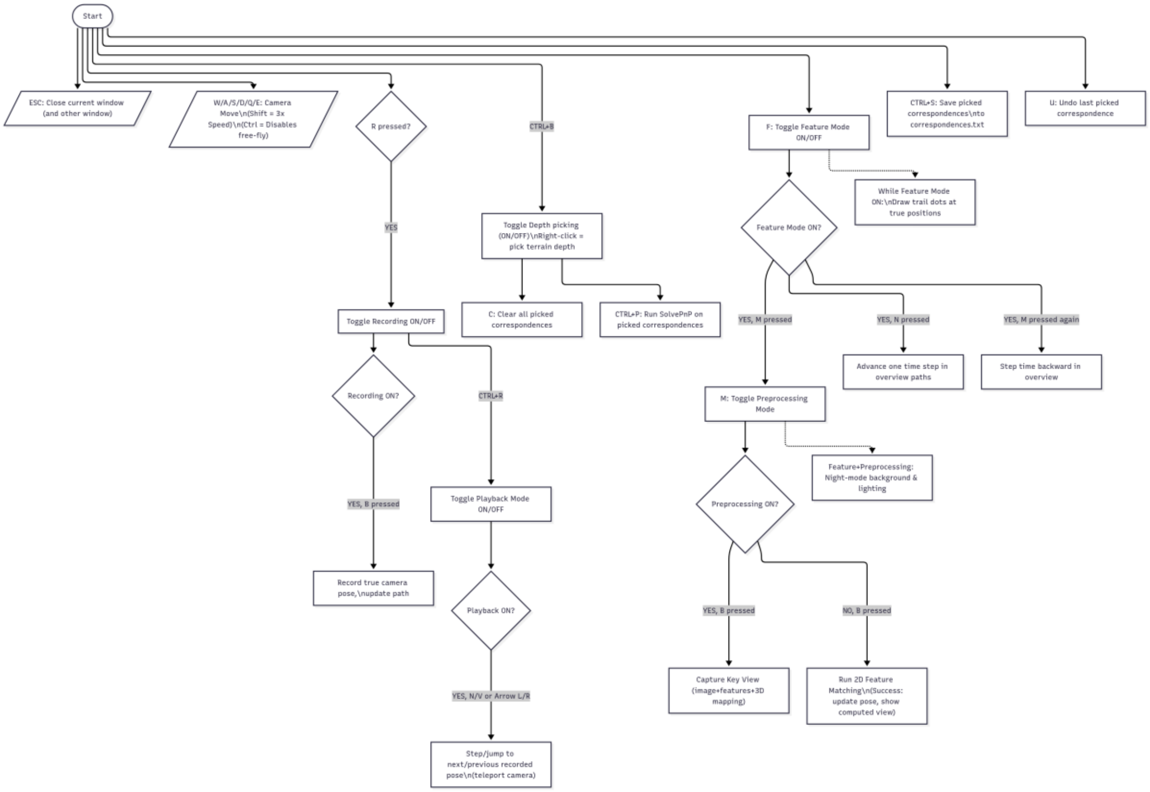
2. Capture keyframes from the scene with visual descriptors (SIFT).

3. Pick corresponding 2D image points and map to 3D terrain points.

4. During runtime, feature matching is applied and solvePnP estimates the camera pose.

Architecture Diagram

See attached diagram for a high-level view of modules and data flow.


Start
ESC: Close current window (and other window)
W/A/S/D/Q/E: Camera Move\n(Shift = 3x Speed)\n(Ctrl = Disables free-fly)
R pressed?
CTRL+B
F: Toggle Feature Mode ON/OFF
CTRL+S: Save picked correspondences\nto correspondences.txt
U: Undo last picked correspondence
While Feature Mode ON:\nDraw trail dots at true positions
YES
Toggle Depth picking (ON/OFF)\nRight-click = pick terrain depth
Feature Mode ON?
Toggle Recording ON/OFF
C: Clear all picked correspondences
CTRL+P: Run SolvePnP on picked correspondences
YES, M pressed
YES, N pressed
YES, M pressed again
Advance one time step in overview paths
Step time backward in overview
Recording ON?
CTRL+R
M: Toggle Preprocessing Mode
Feature+Preprocessing: Night-mode background & lighting
YES, B pressed
Toggle Playback Mode ON/OFF
Preprocessing ON?
Record true camera pose,\nupdate path
Playback ON?
YES, B pressed
NO, B pressed
YES, N/V or Arrow L/R
Capture Key View (image+features+3D mapping)
Run 2D Feature Matching\n(Success: update pose, show computed view)
Step/jump to next/previous recorded pose\n(teleport camera)