

Historical Images for Surface Topography Reconstruction (HISTORY)

Experiment dataset

Last updated: 4 July 2025

Table of content

Test sites	1
Datasets	1
Iceland	1
Casa Grande, USA	4
Data preparation	7
Default image preprocessing	7
Aerial images	7
KH-9 MC images	8
KH-9 PC images	8
Auxiliary data	8
Reference DEMs	8
Stable mask	8
GCPs	9

Test sites

Two test sites have been identified:

- Iceland

Approximate location: South Iceland

Bounding box (LL, UR): -19.7565, 63.4240, -17.6719, 63.8650

Type of terrain: rocks, sand, low vegetation, glaciers, lava, coastal areas.

- Casa Grande, USA

Approximate location: South of Casa Grande, Arizona, USA. Site of a [photogrammetric test range](#).

Bounding box (LL, UR): -111.9204, 32.6514, -111.5971, 32.9870

Type of terrain: semi arid with sand and low vegetation, small urban areas

Datasets

Note: The horizontal CRS of the high resolution reference DEM (local metric system) are used for georeferencing, to avoid resampling of the data. The heights of all products have been converted (when needed) to ellipsoidal heights.

Iceland

General geographic context	
Horizontal CRS (for all georeferencing, unless stated otherwise)	ISN 2016 (EPSG: 8088)
Bbox of “zoom” area, i.e. aerial image extent (xmin, xmax, ymin, ymax)	2662800, 124980, 2765310, 173700
Bbox of “large” area, i.e. satellite image extent (xmin, xmax, ymin, ymax)	2484780, 108990, 2808510, 421800
Image datasets	
Aerial images	<ul style="list-style-type: none"> • date: 1980/08/22 • image count: 125 • image IDs: F-9*.tif • camera type: Wild RC 10 (B/W) • GSD: 0.9 m • source: Natural Science Institute of Iceland, data available for download here: https://www.lmi.is/is/vefsjar/korta-og-loftmyndasofn/loftmyndasafn
KH-9 MC images	<ul style="list-style-type: none"> • date: 1980/08/22 • image count (pre-processed): 4 • image IDs: <ul style="list-style-type: none"> ◦ DZB1216-500280L003001 ◦ DZB1216-500280L004001 ◦ DZB1216-500280L005001 ◦ DZB1216-500280L006001 • camera type: KH-9 MC frame camera (B/W) • GSD: 6-9 m • source: USGS EarthExplorer - Declass 2 dataset
KH-9 PC images	<ul style="list-style-type: none"> • date: 1980/08/22 • image count (pre-processed): 6 • image IDs: <ul style="list-style-type: none"> ◦ D3C1216-200533A021 ◦ D3C1216-200533A022 ◦ D3C1216-200533A023 ◦ D3C1216-200533F021 ◦ D3C1216-200533F022 ◦ D3C1216-200533F023 • camera type: KH-9 PC (panoramic)

	<ul style="list-style-type: none"> • GSD: 0.6-1.2 m • source: USGS EarthExplorer - Declass 3 dataset
Auxiliary data	
<p>High resolution reference DEM</p> <ul style="list-style-type: none"> - for zoom extent only - file "reference_dem_zoom.tif" 	<ul style="list-style-type: none"> • vertical CRS: WGS84 ellipsoid • type: DSM (vegetation, buildings) • posting: 2 m • accuracy: <ul style="list-style-type: none"> ○ Horizontal: +/-2 m ○ Vertical: +/- 0.5 m • source: https://gatt.natt.is/geonetwork/srv/eng/catalog.search#/metadata/e6712430-a63c-4ae5-9158-c89d16da6361
<p>Low resolution reference DEM</p> <ul style="list-style-type: none"> - for large extent - file "reference_dem_large.tif" 	<ul style="list-style-type: none"> • vertical CRS: WGS84 ellipsoid • type: DSM (vegetation, buildings) • posting: 30 m • accuracy <ul style="list-style-type: none"> ○ Horizontal: 6 m ○ Vertical: 4 m • source: https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.032021.4326.1
<p>Stable mask</p> <ul style="list-style-type: none"> - used for large/zoom DEMs coregistration - files "reference_dem_zoom_mask.tif" & "reference_dem_large_mask.tif" 	<ul style="list-style-type: none"> • Includes ESA worldcover categories 20, 30, 60, 100 (shrubland, grassland, bare / sparse vegetation, moss & lichen) • Excludes Randolph glacier inventory V7 outlines
GCPs	<ul style="list-style-type: none"> • count: 18 • horizontal CRS: WGS84 (lon, lat) • vertical CRS: same as reference DEM
Figures*	

Figure large extent

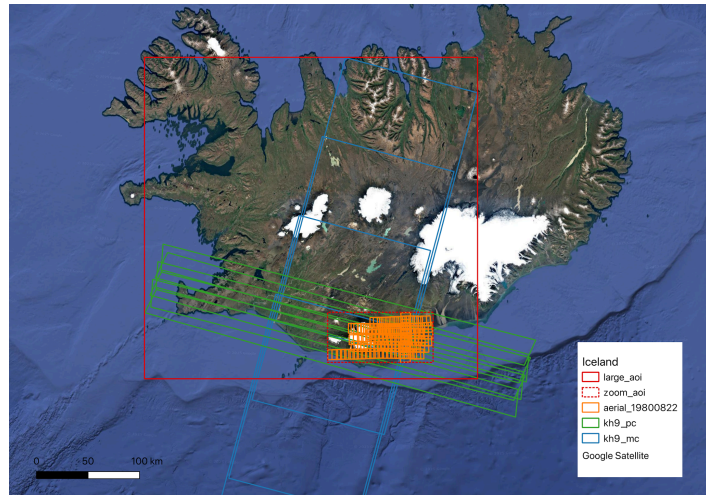


Figure zoom extent

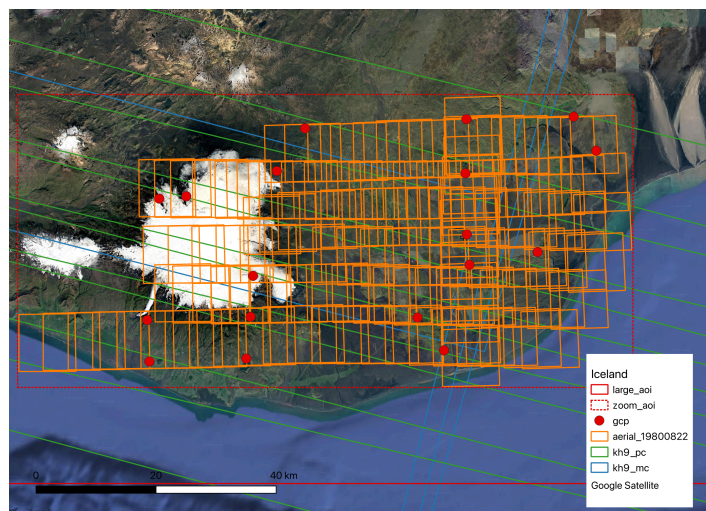
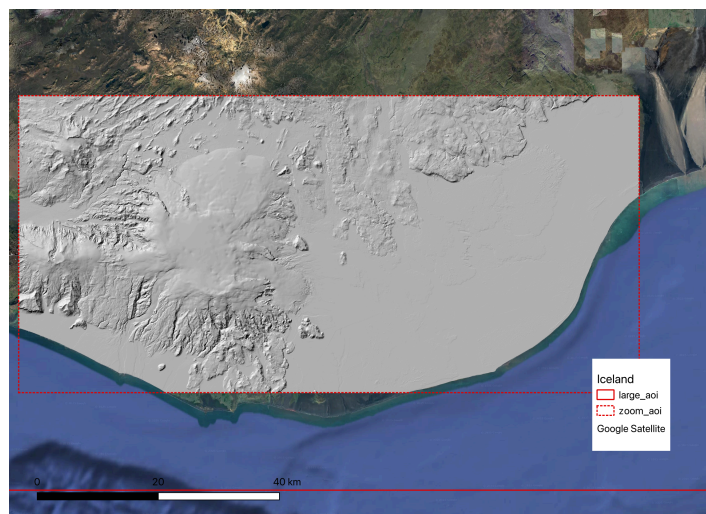


Figure zoom reference DEM



*Link to full res figures:

<https://drive.google.com/drive/folders/1omuOonnKoDrpstbpu9fjMy5sp-2kXbCC>

Casa Grande, USA

General geographic context	
Horizontal CRS (for all georeferencing, unless stated otherwise)	NAD83 / UTM zone 12N (EPSG: 26912)
Bbox of “zoom” area, i.e. aerial image extent (xmin, xmax, ymin, ymax)	414000, 3613020, 444000, 3650010
Bbox of “large” area, i.e. satellite image extent (xmin, xmax, ymin, ymax)	261990, 3405030, 612990, 3880020
Image datasets	
Aerial images	<ul style="list-style-type: none"> • date: 1978/09/06 • image count: 37 • image IDs: ARBCSRD000100*.tif • camera type: Wild RC 10 (raw: RGB, preprocessed: B/W) • GSD: 0.7 m • source: USGS EarthExplorer - aerial photo single frame dataset
KH-9 MC images	<ul style="list-style-type: none"> • date: 1978/03/25 • image count (pre-processed): 4 • image IDs: <ul style="list-style-type: none"> ◦ DZB1214-500055L004001 ◦ DZB1214-500055L005001 ◦ DZB1214-500055L006001 ◦ DZB1214-500055L007001 • camera type: KH-9 MC frame (B/W) • GSD: 6-9 m • source: USGS EarthExplorer - Declass 2 dataset
KH-9 PC images	<ul style="list-style-type: none"> • date: 1978/03/25 • image count (pre-processed): 6 • image IDs: <ul style="list-style-type: none"> ◦ D3C1214-100097A013 ◦ D3C1214-100097A014 ◦ D3C1214-100097A015 ◦ D3C1214-100097F012 ◦ D3C1214-100097F013 ◦ D3C1214-100097F014 • camera type: KH-9 PC panoramic (B/W) • GSD: 0.6-1.2 m

	<ul style="list-style-type: none"> source: USGS EarthExplorer - Declass 3 dataset
Auxiliary data	
<p>High resolution reference DEM</p> <ul style="list-style-type: none"> for zoom extent only file "reference_dem_zoom.tif" 	<ul style="list-style-type: none"> vertical CRS: WGS84 ellipsoid type: DTM (no vegetation, no buildings) posting: 1 m accuracy: <ul style="list-style-type: none"> Horizontal: not specified, but well < 1 m Vertical: +/- 0.1 m source: USGS - https://rockyweb.usgs.gov/vdelivery/Datasets/Staged/Elevation/1m/Projects/AZ_MaricopaPinal_2020_B20/
<p>Low resolution reference DEM</p> <ul style="list-style-type: none"> for large extent file "reference_dem_large.tif" 	<ul style="list-style-type: none"> vertical CRS: WGS84 ellipsoid type: DSM (vegetation, buildings) posting: 30 m accuracy <ul style="list-style-type: none"> Horizontal: 6 m Vertical: 4 m source: https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.032021.4326.1
<p>Stable mask</p> <ul style="list-style-type: none"> used for large/zoom DEMs coregistration files "reference_dem_zoom_mask.tif" & "reference_dem_large_mask.tif" 	<ul style="list-style-type: none"> Includes ESA worldcover categories 20, 30, 60, 100 (shrubland, grassland, bare/sparse vegetation, moss & lichen) Excludes groundwater/aquifer subsidence mask (downloaded here)
GCPs	<ul style="list-style-type: none"> count: 34 horizontal CRS: WGS84 (lon, lat) vertical CRS: same as reference DEM
Figures*	

Figure large extent

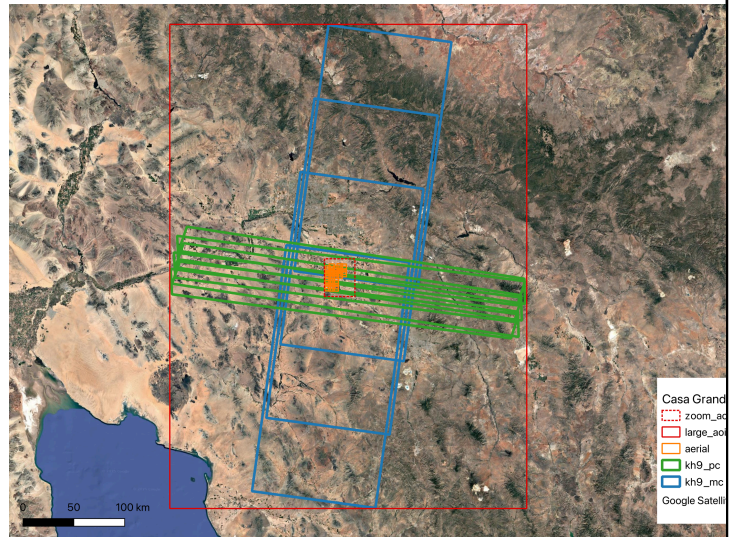


Figure zoom extent

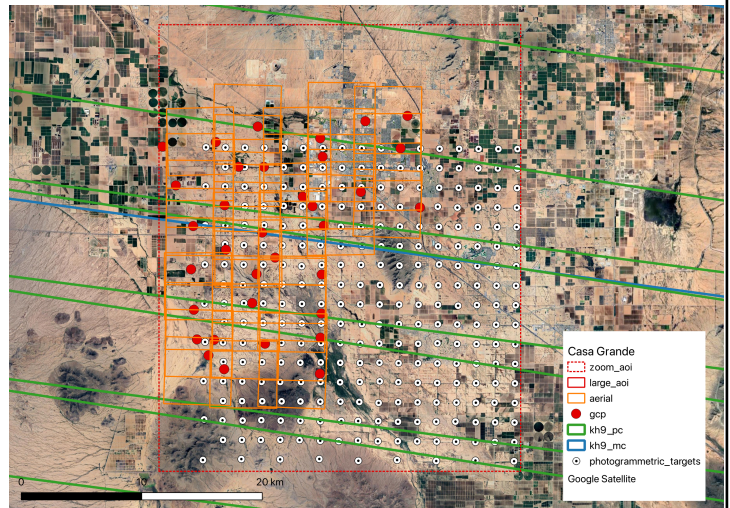
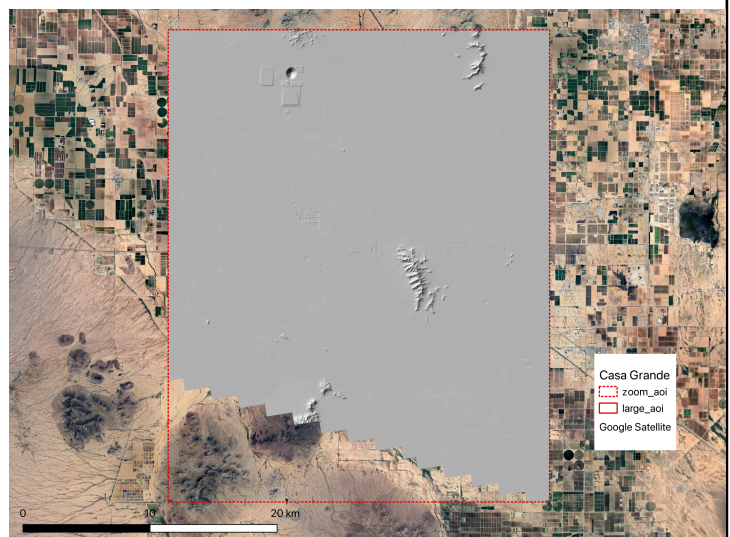


Figure zoom reference DEM



*Link to full res figures:

<https://drive.google.com/drive/folders/1omuOonnKoDrpstbpu9fjMy5sp-2kXbCC>

Data preparation

Default image preprocessing

Aerial images

The preprocessing method mainly follows [Knuth et al. \(2023\)](#). In brief:

- For Casa Grande, RGB images are converted to grayscale with OpenCV's "imread_grayscale"
- Create fiducial templates by manually selecting them on one image.
- Use fiducial templates to perform template matching on all images to get the coordinate of all fiducial markers. Filtering mismatches by using the matching score. Compute center of valid detected fiducial markers.
- Compute an affine transformation to align the detected markers with the calibration report positions.
- Crop images at fixed dimensions around the previously computed center.
- Apply CLAHE contrast enhancement

Data preprocessing notebooks are available here:

https://github.com/shipp/history/blob/main/notebooks/casa_grande_aerial.ipynb

https://github.com/shipp/history/blob/main/notebooks/iceland_aerial.ipynb

Note: For the Icelandic images, the fiducial markers were invisible on 2 images (F-9426.tif and F-9458.tif). We manually identified the mid-side pseudo-fiducial markers on one reference image which was successfully aligned, and on the 2 problematic images, to calculate the corresponding affine transformation between each image and the reference.

KH-9 MC images

The preprocessing method follows the method of [Dehecq et al. \(2020\)](#). In brief:

- the position of the reseau markers (dark crosses) is detected at subpixel resolution using a convolution with a cross-like kernel.
- outliers are rejected by removing markers that are not regularly spaced (1 cm)
- a degree 3 polynomial transformation is first applied to remove rotation and scaling, then a Thin Plate Spline interpolation is used to calculate the distortion at each pixel.
- the two pieces are stitched together by matching the areas of overlap using NCC.
- the image is cropped to a fix distance from the outermost markers to keep a constant image dimension among the images
- the reseau markers are filled with inpainting.

KH-9 PC images

The preprocessing method is similar to that of [Ghuffar et al. \(2023\)](#). In brief:

- The image pieces are joined into a single composite image using ASP's [image_mosaic](#) command. It uses IP matching and affine transformation to align the pieces together.
- The film corners are selected manually.
- The image is rotated and cropped so that the film edge is horizontal.

Data preprocessing notebooks are available here:

https://github.com/shipp/history/blob/main/notebooks/casa_grande_kh9pc.ipynb

https://github.com/shipp/hypp/hypp/history/blob/main/notebooks/iceland_kh9pc.ipynb

Auxiliary data

Image footprints and camera extrinsics

The image footprints and metadata were downloaded from USGS EarthExplorer for the satellite images and Casa Grande aerial images. For the Iceland aerial images, the footprints were manually generated by Joaquin Belart, using a pre-existing, approximate location of the camera centers. In QGIS (version 3.4), using the digitizing tools, a generic, approximate footprint of one of the images, was drawn, which was then copied successively over all the camera centers.

The camera extrinsics were pulled from the footprint files and converted into a CSV file.

Camera intrinsics

The camera intrinsics CSV files were filled manually based on the calibration report or information from the data provider.

Reference DEMs

The high resolution DEMs were downloaded as raw GTiff tiles from the provider, mosaicked without resampling and the vertical datum updated (for Casa Grande only, from NAVD88 geoid to ellipsoid).

The low resolution Global Copernicus 30m DEM was downloaded as raw GTiff tiles from OpenTopography, reprojected on the same horizontal CRS as the high-res DEM and converted from original EGM2008 heights to ellipsoid heights. It was then coregistered to the high-res DEM by applying a horizontal and vertical shift, calculated using xDEM's implementation of the Nuth & Kääb (2011) algorithm. Pixels outside the "stable mask" (see below) are excluded during coregistration.

Stable mask

The ESA worldcover raw GTiff tiles were downloaded and mosaicked, and reprojected on the same CRS and grid as the two reference DEMs. Anything but shrubland, grassland, bare/sparse vegetation, moss/lichen (respective values of 20, 30, 60 and 100) are masked as unstable.

For Iceland, the Randolph glacier inventory (RGI) v7 outlines were rasterized to the DEM grids and masked.

For Casa Grande, the subsidence mask was rasterized to the DEM grids and masked.

The DEM and stable mask processing scripts are located at

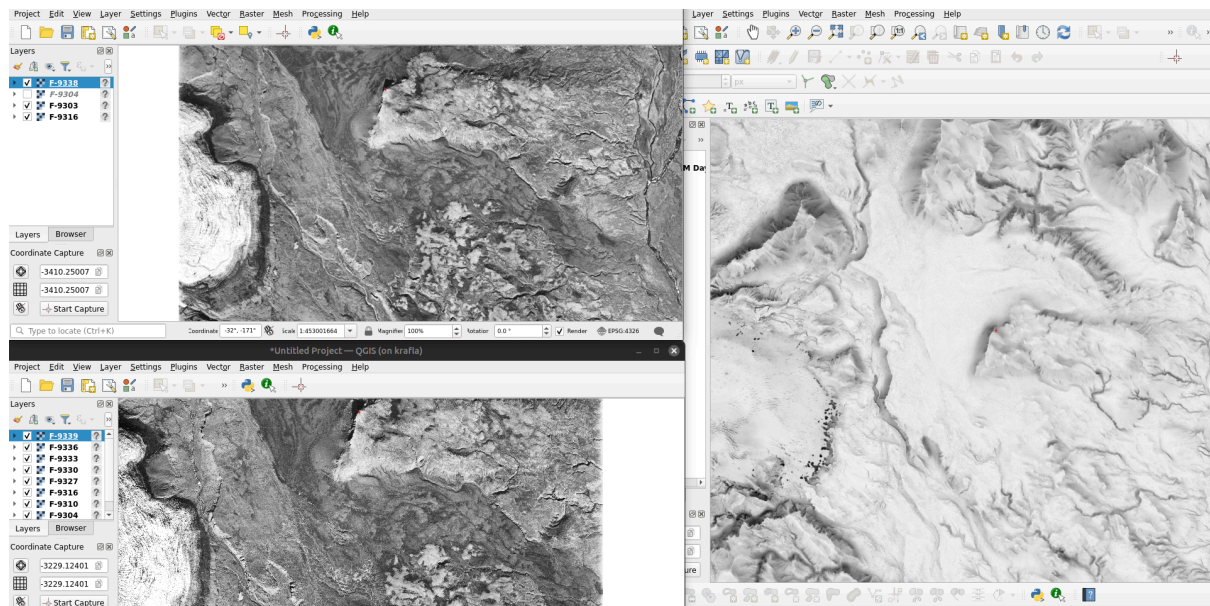
https://github.com/shipp/hypp/hypp/history/tree/main/src/history/aux_data (yet to be added to the repo as of 3 July 2025).

Ground Control Points

GCPs are provided only for the aerial dataset. They have been manually picked by Joaquin Belart using the data provided in this experiment, as well as the orthomosaic from Bing Maps in QGIS (©Microsoft, data accessible in QGIS via XYZ tiles using <http://ecn.t3.tiles.virtualearth.net/tiles/a{q}.jpeg?g=1>).

For picking of GCPs, the QGIS plugin "Coordinate capturer" was used. This allows extracting pixel coordinates of the aerial photographs (rows and columns), as well as geographical coordinates of the reference data (longitude and latitude). The measurements were gathered

into an ASCII list. Then, the elevation values were extracted from the reference DEM using the command *geodiff* in the Ames StereoPipeline software (version 3.6, Beyer et al., 2018).



Example of settings in QGIS (version 3.4) for manual picking of GCPs. Left: two preprocessed images of the Iceland dataset (F-9338.tif and F-9339.tif). Right: Multi-directional hillshade obtained from reference DEM (reference_dem_zoom.tif).

For the aerial photographs of Iceland, the reference DEM was visualized as a hillshade in QGIS. The Bing maps were also used as a guide when the terrain was lacking enough relief, but the final coordinates were always extracted from the hillshade in order to maintain the positional accuracy from the reference DEM.

For the aerial photographs of Casa Grande, the primary source for picking up GCPs was the Bing mosaic. The area has limited relief, so a hillshade does not allow a straight-forward recognition of features from the old photographs. In Casa Grande, the reference DEM used for retrieving elevations was the Copernicus GLO30 DEM.

In both cases, we ran a preliminary bundle adjustment of the aerial photographs with the hand-picked GCPs for a first check of the results. This was done in Agisoft Metashape (version 2.1). By doing this, we could identify a few (typically 3 to 5) GCPs that presented high residuals (> 5 meters in XYZ) in the bundle adjustment. These GCPs were re-located in another place, while ensuring a homogeneous distribution throughout the block of aerial photographs.