How to Integrate Images and 3D Point Clouds: Geospatial Standards and Best Practices

By **Francisco Navarrete Mandly**

1. Capture of Point Clouds and Images

Three-dimensional data acquisition using 3D laser scanners can be performed with static units or mobile systems based on SLAM (Simultaneous Localization and Mapping) technology.

- **Static scanners** perform rotational sweeps from a fixed position, generating high-density, high-accuracy point clouds.
- **Mobile or SLAM-type scanners** acquire data while moving, integrating positioning sensors and real-time mapping algorithms to efficiently generate point clouds.

In addition to three-dimensional coordinates, scanners also record the return signal intensity, providing information on the reflective properties of scanned surfaces.

To complement this purely geometric and radiometric information, many scanners integrate imaging systems that add visual context to the survey.

- **Conventional RGB cameras**: Capture images similar to those of a standard digital camera.
- **Panoramic or spherical cameras**: Use wide-angle or fisheye lenses, or multiple sensors arranged around the unit, to cover a wide field of view.
- **Multispectral cameras**: Capture bands beyond the visible spectrum, such as near-infrared or ultraviolet, useful for vegetation analysis, heritage conservation, or material inspection.

2. Images Generated by Post-Processing Software

In addition to processing LiDAR observations, the post-processing software executes a complete photogrammetric workflow that integrates images captured by the scanner's cameras with data from positioning and orientation sensors (GNSS, IMU, etc.).

In this process the three-dimensional position and orientation of each image are jointly estimated, solving the external orientation parameters and, unless the camera already has a prior calibration, the internal parameters through self-calibration. If the camera has been previously calibrated, these internal parameters can be fixed and the computation focuses on external orientations.

When the system includes multiple cameras, the software stitches and merges the individual images into a common projection, using distortion-corrected images. This ensures continuity of color and geometry so that the images are perfectly aligned with the point clouds for later use in panoramas, orthophotos, or other representations. (Crombez et al., 2015; Sharma et al., 2024)

The most common types of images generated are:

- **Perspective** (pinhole or central projection): Represent the classic geometry of a central-lens camera, analogous to a conventional photograph.
- **Spherical**: Provide a full 360° view of the scene in a continuous spherical projection, ideal for immersive visualization.
- **Cylindrical**: Project the scene onto a cylindrical surface, particularly useful for panoramic documentation of tunnels, corridors, or long façades.

New visualization methods, such as Gaussian Splatting, allow the point cloud to be continuously and realistically represented, generating smooth surfaces with photographic lighting.

3. Image Positioning in Space

The camera position may be expressed in coordinates of a local reference system or in a global system if the scanner integrates GNSS receivers or if ground control points have been used.

Camera orientation is commonly defined by the angles roll, pitch, and yaw:

- **Roll**: rotation about the camera's longitudinal axis (lateral tilt).
- **Pitch**: rotation about the transverse axis (tilt up or down).
- **Yaw**: rotation about the vertical axis (heading left or right).

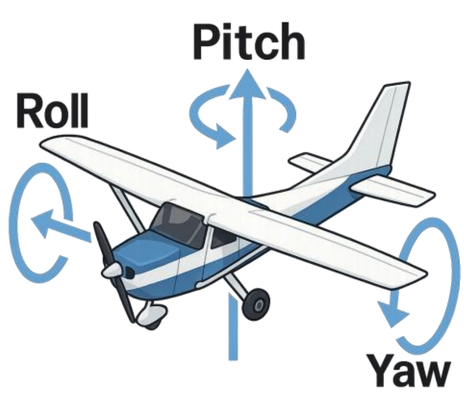

Figure 1. Rotation angles

Although intuitive, these angles have limitations: it is not enough to specify their values; the sign convention and the order of the rotations (Euler rotation matrix) must also be defined. Without this information, different software may interpret the same data differently, causing inconsistencies in image orientation.(Kim & Kim, 2023)

A more robust alternative is the use of **quaternions**, an extension of complex numbers to four components (qx, qy, qz, qw). Quaternions avoid issues associated with Euler angles, such as sign ambiguity or gimbal lock, and allow stable interpolation of orientations. They are widely used in virtual reality, robotics, and 3D graphics.

Beyond accurate position and orientation, it is essential to maintain spatial coherence between point clouds and images so that both coincide exactly within the same reference system. This coherence is especially critical when integrating data in GIS environments or BIM workflows.

4. Standard File Formats for Positioning and Orienting Images

There are several non-proprietary formats capable of storing images together with their position and orientation parameters. However, not all are suitable for images associated with 3D scanners.

EXIF Metadata

The Exchangeable Image File format can store complementary data directly within image files, such as capture date, camera used, exposure settings, and even geolocation information using GPS coordinates and camera orientation.

However, these metadata have important limitations (Acharya R et al., 2023):

- Designed primarily for conventional photography and mobile devices.
- Position is expressed approximately with GPS coordinates, lacking the precision required for surveying applications.
- Orientation is described with basic fields (heading, tilt), insufficient for rigorously representing a camera's 3D pose.
- Cannot adequately represent the position and orientation of oblique images or integrate consistently with point clouds.

World Files

World files (.jgw, .pgw, .tfw, etc.) are text files associated with a raster image that define its location in a coordinate system through translation, resolution, and rotation parameters.

```
5.000000
0.000000
0.000000
-5.000000
454010.000000
4187525.000000
```

Figure 2. World file with 5 m of pixel size, no rotation and upper left corner located at X=454010, Y=4187525

They are widely used in GIS to georeference orthophotos or aerial images in plan view, but they are not suitable for scanner images because:

- Their structure only defines a 2D affine transformation.
- They do not include 3D orientation information.
- They are not designed for oblique or 360° images.

Text Files with Camera Parameters

A common practice in photogrammetry is to use text files listing the position and orientation of each image, expressed as rotation angles or quaternions. These files are flexible but not standardized.

```
image001.jpg 523460.50 4089127.30 134.90 0.237888 -1.141890 92.505228
image002.jpg 523465.00 4089131.60 135.80 0.258252 -1.170447 93.376185
image003.jpg 523469.50 4089135.90 136.70 0.318343 -1.021579 89.016524
image004.jpg 523474.00 4089140.20 137.60 0.328800 -1.005543 89.070848
image005.jpg 523478.50 4089144.50 138.50 0.286861 -1.162192 93.201041
```

Figure 3. File with camera parameters (image, X, Y, Z, omega, phi, kappa)

Interpretation criteria can vary between applications:

- Units may be in meters, centimeters, or even pixels.
- Angles may be in degrees or radians, and the reference system may change.
- The order of rotations and angle sign conventions are not always clearly documented.

Therefore, despite containing all the data needed to align images and point clouds, their lack of standardization complicates interoperability between applications.

E57

The ASTM E57 standard is specifically designed for exchanging 3D laser scanning data and is the most comprehensive option for jointly storing point clouds and images (Huber, 2011). Its structure allows integrating into a single file not only the scan geometry but also the images from each station and all parameters required to position them accurately in space.

E57 supports various image types, from perspective (pinhole or central projection) to spherical or cylindrical panoramas, and explicitly defines the data for their position and orientation.

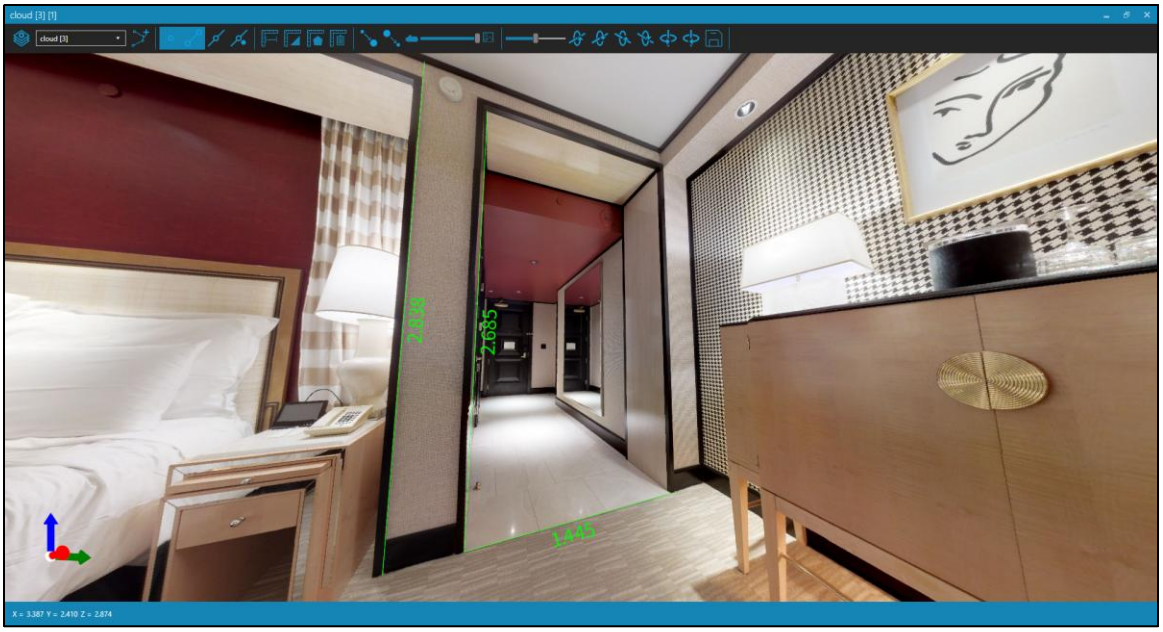

Figure 4. Precise measurement over image and point cloud in the background with Tcp PointCloud Editor

As an open and well-documented format, it facilitates interoperability among different manufacturers and applications, reducing the risk of information loss and ensuring that images can be viewed aligned with point clouds. The open-source library libE57 is a key resource, greatly simplifying the work of developers who need to implement compatibility with this format in their applications, avoiding the need to build E57 reading and writing from scratch.

Despite its advantages, E57 also presents certain limitations:

- Files can become very large when numerous high-resolution images are included, slowing read/write and transfer operations.
- It is not designed for multispectral images, such as those containing multiple spectral bands (infrared, ultraviolet, etc.).
- It lacks detailed support for higher-order lens distortion or very specific camera models.
- It does not natively support time sequences or video, so complementary files or proprietary extensions are required for such data.

These limitations do not detract from the value of E57 as a reference standard, but they should be kept in mind in projects with a large volume of images or with advanced photogrammetric requirements.

5. Conclusions

Due to its capabilities and open nature, **the E57 format stands out as the best practice** for professional exchange of laser scanning data accompanied by georeferenced images.

Manufacturers of 3D scanners are encouraged to natively export both point clouds and images in this format, ensuring that each photograph correctly includes its position and orientation parameters. This approach enhances cross-platform interoperability, improves geometric accuracy, and streamlines workflows in engineering, architecture, and construction projects.

It is also advisable to follow the evolution of the IFC standard, whose version 5 is expected to advance the integration of BIM and GIS environments. This convergence will enable building and infrastructure information models to be directly linked with geospatial data and with point clouds and images captured in the field, reinforcing information continuity throughout the project lifecycle.

Finally, adoption of E57 should not be limited to hardware manufacturers: it is equally important that software developers implement full and rigorous compatibility, ensuring interoperability at the application level and fostering an open ecosystem across the geospatial sector.

Bibliography

Acharya R, S., J, S., S, S., S Aithal, V., & G S, H. (2023). GEO-LOCATING AN IMAGE USING EXIF DATA. *International Journal of Engineering Applied Sciences and Technology*, *8*(1). https://doi.org/10.33564/ijeast.2023.v08i01.007

Crombez, N., Caron, G., & Mouaddib, E. (2015). 3D point cloud model colorization by dense registration of digital images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(5W4). https://doi.org/10.5194/isprsarchives-XL-5-W4-123-2015

Huber, D. (2011). The ASTM E57 file format for 3D imaging data exchange. *Three-Dimensional Imaging, Interaction, and Measurement*, *7864*. https://doi.org/10.1117/12.876555

Kim, S., & Kim, M. (2023). Rotation Representations and Their Conversions. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2023.3237864

Sharma, S. K., Jain, K., & Shukla, A. K. (2024). 3D point cloud reconstruction using panoramic images. *Applied Geomatics*, *16*(3), 575–592. https://doi.org/10.1007/s12518-024-00563-w