# Collection and Application of 2D and 3D Panoramic Imagery

**BART BEERS, Waardenburg**

**ABSTRACT**

The use of digital panoramic imagery in various geo-related processes has become common over the last decade. Professional applications are found among situational awareness, inventory of public areas, real estate valuation, assessment of permits and inspecting security risks. More and more the imagery is being used for measuring and map making.

A historical overview is given of the ideas and developments that led to the present system for recording, processing, retrieval and viewing of digital panoramic imagery at CycloMedia in The Netherlands. Attention is paid to the present recording system that delivers geometric correct panoramic images at driving speeds up to 80 km/h and to the application of that imagery. The web application GlobeSpotter is introduced as the means for, among others, retrieval, display and manipulation of panoramic images. Finally the current development of a system for large scale collection of 3D depth panoramas ('Depth Cycloramas') is shown together with an outlook to their application.

## 1. INTRODUCTION

Panoramic photography has come a long way until it could be used for reliable and accurate measurement and interpretation. While at present we see a movement in the direction of increased application of LiDAR point clouds for use in mobile mapping, digital panoramic imagery shows some clear advantages, especially in large scale operations. Interesting results come from the hybrid approach in which georeferenced, geometric correct panoramic images are combined with LiDAR point clouds.

## 2. HISTORY

At the Delft University of Technology, at the then Faculty of Geodesy, the first research on a panoramic measuring system started already in 1980. The goal of the research on this so-called FRANK system was "*The development of a new system for collecting, processing and supplying information for the purpose of for instance systematic road management, green management, facility management, rural and urban structuring and defense purposes. The research is concerned with the analysis of the requirements the FRANK system has to meet and the manufacturing of parts of the functioning model*" (Beers, 1995). The project elaborated on experiences with the Horizon Camera project, an earlier cooperation with De Oude Delft. Other ideas originated from the Movie Map project of the Defense Advanced Research Project Agency (DARPA) and the Massachusetts Institute of Technology (MIT) (Wikipedia, Aspen Movie Map, 2011).

The FRANK system was designed to work with combined intersection and resection using bundle block adjustment with a minimum of control points. To make this work it was necessary to measure numerous triangulation points (see Fig. 1). GNSS and INS were basically non existing at the time. The goal was to reach a mapping precision at centimeter level.

After having developed and tested several prototypes, by 1994 two production systems were built, each consisting of a recording system, a scanning system, a measuring system and an edit system. In the following, some details on these systems will be shown. A discussion on the overall calibrations and the adjustment software will be omitted.
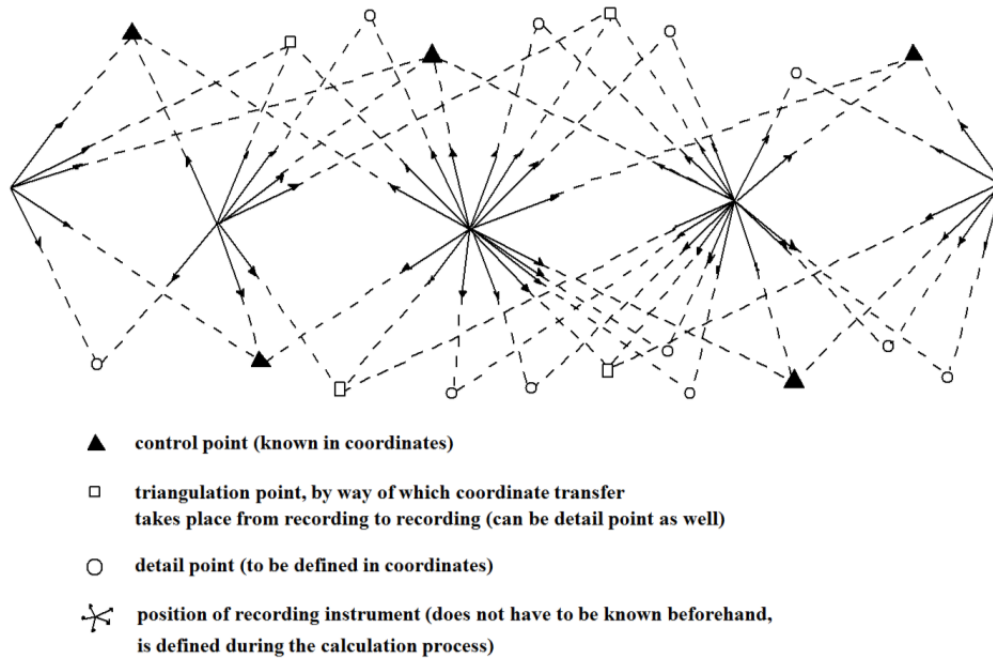
▲     **control point (known in coordinates)**

☐     **triangulation point, by way of which coordinate transfer takes place from recording to recording (can be detail point as well)**

○     **detail point (to be defined in coordinates)**

⚹     **position of recording instrument (does not have to be known beforehand, is defined during the calculation process)**

Figure 1: Combined intersection and resection

## 2.1. Recording System

The images were made using a single, special designed fisheye lens pointing upward, looking around over 360° and looking from the zenith down to 30° below the horizon. A camera suitable for 70 mm wide roll film was connected to this lens. A combination of a reseau platen as a part of the lens system and a vacuum back platen in the camera body ensured sufficient film flatness, avoiding significant metric distortions. Also a provision was made for the projection on the film of fiducial marks and machine readable identification codes for every photographic recording.



Figure 2: Recording system on Mini

The lens-camera combination was mounted in a stabilized platform on a small car (see Fig. 2) as to avoid it to obscure details on the street. The car had to stop on average every 10 meters to make a recording.

## 2.2. Scanning System

To be able to use digital images in the measuring system (former prototypes were optical-mechanical instruments), the films had to be digitized. This was done using special developed scanners with a rotating line array (see Fig. 3). A scanner was able to fulfill an interior orientation by centering the axis of rotation on the principle point of the fisheye image using the fiducial marks. For identification, the machine readable code of each image was read.
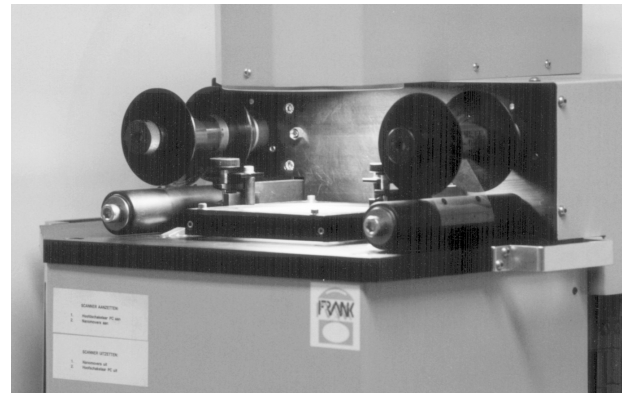
Figure 3: Rotating scanner

A complete scan thus resulted in a black and white panoramic image of 88 MB (22.500 by 4096 pixels), quite a size in a time when the largest available PC hard disk was 1 GB. A scanner also delivered lower resolution 'overview images' of 315 KB (1260 by 256 pixels).

A scanner could work unattended as long as there was roll film loaded. Therefore it was possible to combine several rolls from the camera system.

## 2.3. Measuring System

Employment of the measuring system primarily came down on the measurement of directions to identical points in two or more different images. Besides the metric aspects of this procedure, the correct identification of identical points was equally important. For this, a setup was designed in which, on one computer screen, three panoramic overview images were presented to the operator (see Fig.4). In an overview image, that horizontally showed a selectable part of 180° out of the total image, the operator could identify a point by clicking on it. Then a detail image window was presented in which the point could be identified with the required precision. The extent of a detail image was indicated in the corresponding overview image by a rectangle, that moved along when the operator scrolled the contents of the detail window.
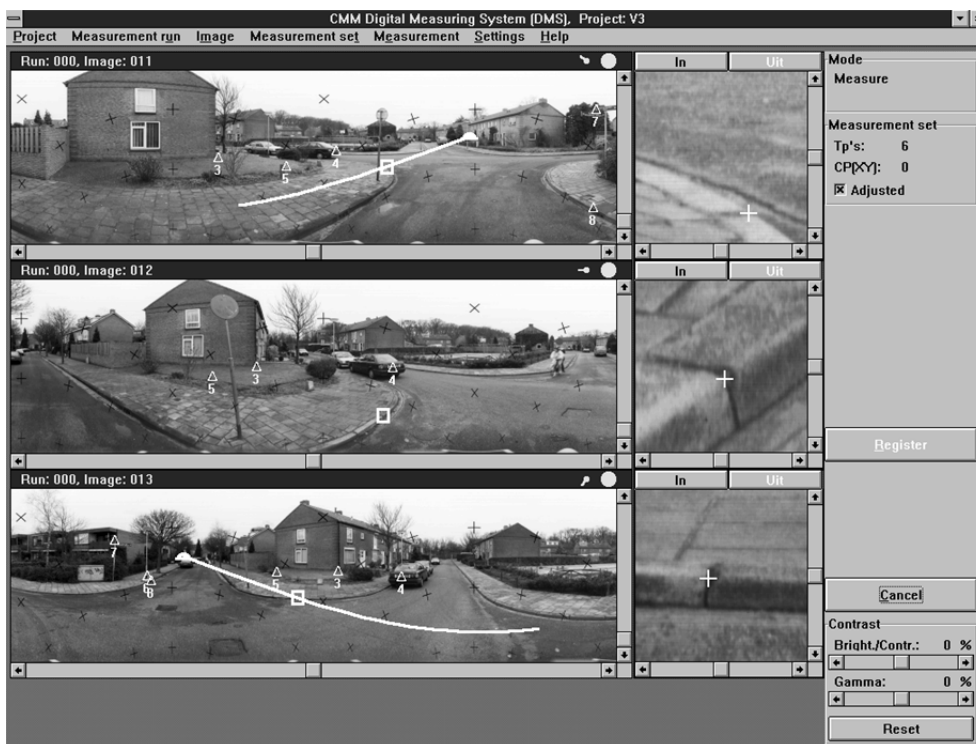
In the measuring system a procedure was used in which the overview images were stored on hard disk and the high resolution images were stored on Digital Audio Tape (DAT) cartridges. The overview images could quickly be displayed in three windows on the left hand side of the user interface. The three corresponding detail images were copied to hard disk to be available



Figure 4: Measuring System screen

for quick random access. In a background operation the following detail image was also loaded. When the operator shifted to a next image triplet (overlapping the former by two images), the subsequent detail image was loaded from DAT to hard disk, overwriting the first detail image. Thus, by the time the operator finished his measurements, the next image triplet was on hand.

## 2.4. Practical experiments and spin-off

Experiments with the FRANK system, conducted in cooperation with the Dutch Cadastre and the municipalities of Rotterdam and Amsterdam concluded that the application of this method in the test projects showed that both technical and economical, the method positively could compete with the then current tacheometric method. Yet, the land surveying practice did not fully accept the system as it could not completely replace the existing methods.

But there already had been developed a spin-off system (see Fig. 5) that employed lower resolution panoramic images (3172 pixels by 610 pixels; 360° horizontal by 110° vertical; non-equal horizontal and vertical resolution; 256 shades of gray) not for measuring purposes, but only for viewing the surroundings.



Figure 5: Image Viewing System screen with three panoramic images and an aerial photo,
looking towards a selected address

This Image Viewing System was used in the first large project for the City of Rotterdam. In 1995 a total of 90.000 fisheye images were recorded with an interval of 20 meters and processed into panoramic images. It was this lower resolution viewing system that, while still maintaining photogrammetric correctness, eventually led to the current CycloMedia system.

## 3. PRESENT RECORDING PRINCIPLE

Over the years several production versions of the recording and processing systems have evolved. Main developments were on the subject of the sensor (film to digital, black and white to color), positioning (from operator assisted positioning to GNNS-INS), recording setup (from single to multi shot, from stop and go recording to distance triggered recording while driving). Of course, the most noteworthy advance has been the enormous increase in productivity of the recording systems as well as the reduction of the recording interval. Remarkably, up to now there have been no major developments in resolution.
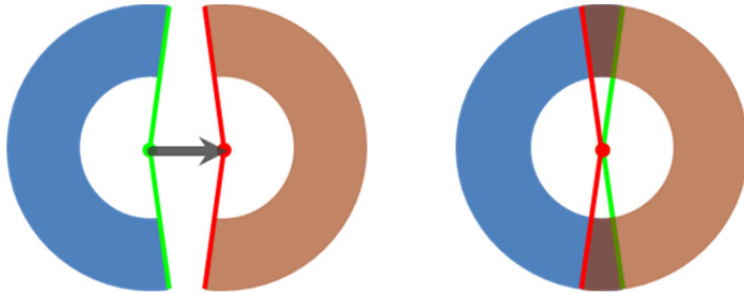
The present system employs a patented method to be able to make multi shot fisheye images while driving and still avoiding parallaxes. This is done using two cameras with fisheye lenses on top of a car, one pointing in the driving direction, the other in the opposite direction (see Fig. 6 and 7). With the car moving forward, the front camera makes an exposure and after that the back camera makes an exposure exactly at the moment the entrance pupil of the back camera coincides with the entrance pupil of the former front exposure.

Figure 6: On the right, the front camera makes the first exposure, both cameras continue to move and…

… the back camera makes an exposure exactly at the moment the entrance pupils coincide.

Fig. 7 shows the principle components of this recording system. At first, the DMI's played an important role in the precise timing of the triggering of both cameras. At present, this role has been overtaken by the IMU, whilst the DMI's keep serving a role providing additional data for position processing.

Fig. 7: Major components of the Recording System

The recording system has been developed to be used in harsh practice and to fulfill photogrammetric requirements with GIS-grade accuracies. As the mathematical relation between pixel coordinates in image space and spatial directions in object space is well known, it is possible to use the panoramic images to precisely measure spatial directions to object points. The car's positioning system delivers 3D position and orientation for every panoramic recording with an accuracy of 10 cm resp. 0.1°, so using the intersection principle with measurements from at least two panoramic images, the coordinates of visible object points can be determined. This workflow has been implemented in the web application GlobeSpotter which among others comprises geodetic adjustment and integrated quality control procedures.

## 4. GLOBESPOTTER

The web application GlobeSpotter offers several functions relating to the use of panoramic imagery:

- Retrieval of panoramic images and aerial photos from a CycloMedia server or from a customer data server using user input such as coordinates, click on map (or aerial photo), address data and postal code data; navigation between panoramic images.
- Display of maps and aerial photos from WMS servers.
- Overlay of address data on the panoramic imagery and on the map.
- Overlay of 2D raster data from WMS servers on the panoramic imagery and on the map.
- Overlay of 2D and 3D vector data from WFS servers on the panoramic imagery and on the map (see Fig. 8).
- 3D point, line and area measurement using the intersection principle with geodetic adjustment and integrated quality control procedures (see Fig. 9). For operator assistance, guiding lines representing spatial directions to points measured in other images, are projected.



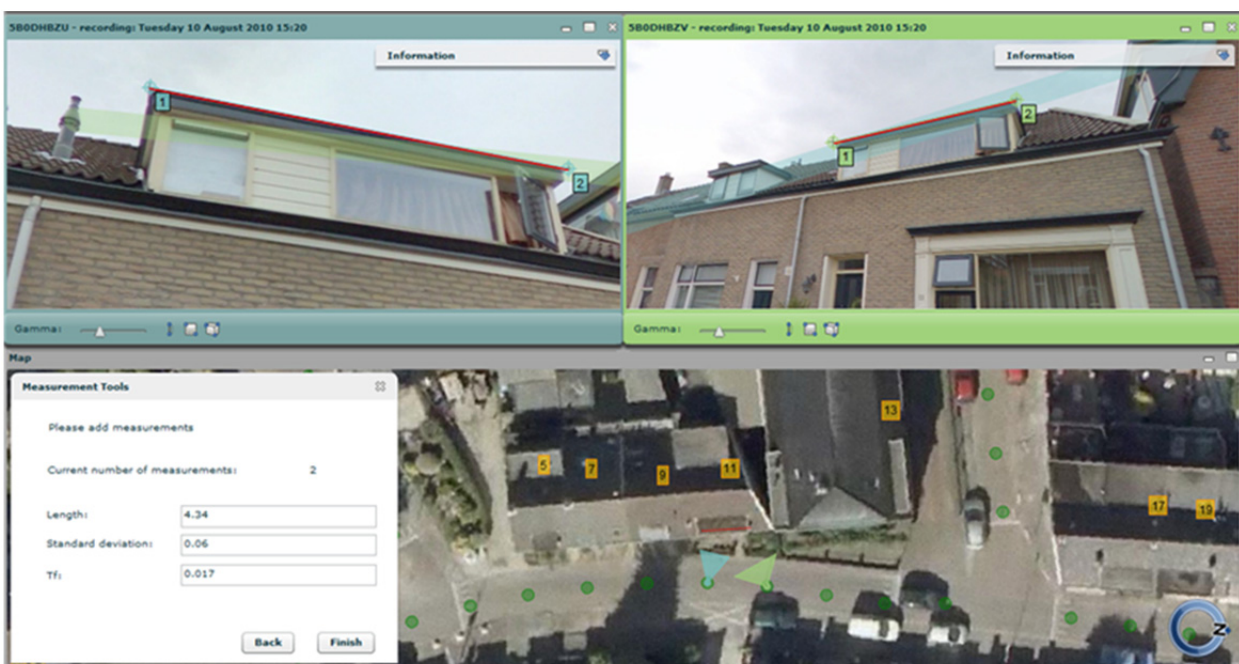Figure 8: Projection of 2D and 3D vector data on panoramic images



Figure 9: Line measurement in GlobeSpotter, showing measurement result, precision and reliability

Figure 10: Projection of a 2D map on a slope

Although GlobeSpotter shows many functions that are available in major geographic information systems, it is not meant to compete with them. Therefore all special functions relating to the panoramic images and the geodetic adjustment are available in an API (Application Programming Interface). Up to now, tens of vendors of GIS have integrated this functionality in their systems, some on a basic level but most up to the incorporation of measuring and projecting.

The combined use of projection of existing 2D and 3D data and measuring (or editing) additional features is a very strong user advantage, for instance for augmenting data, for updating maps and for extended situational awareness. However, in the case of overlaying 2D data we are confronted with a restriction. As the Z-value of the vector points is unknown, the only way to project this data is on the level where the car was driving (which is 2 m below the recording position). And this will only give correct results in the case of flat terrain. Fig. 10 shows the result if this is not the case.

If for every pixel in the panoramic image 3D information were available, we could have draped the 2D data over the terrain, resulting in a much more comprehensible view.

Moreover, the intersection measurement principle requires the user to indicate points in at least two images, which from an efficiency point of view most users would prefer to avoid: there exists a clear demand for single click measuring from panoramic imagery (as opposed to measuring from LiDAR point clouds, which among others requires professional interpretation skills).

This evolved into the idea of producing 3D depth panoramas.


## 5. PRODUCTION OF 3D DEPTH PANORAMAS

At CycloMedia, research on the addition of depth to panoramic images started some years ago by registering adjacent images to form 3D point clouds. Sparse stereo (feature) based approaches lack the point density to generate depth, and dense stereo based approaches suffer from the inability to reliably find good correspondences between images in areas with insufficient texture. However, another source for the addition of depth to panoramic imagery comes from mobile LiDAR systems. Although the point clouds from these systems are usually less dense than those based on stereo from images, they offer reasonably reliable 3D data and do not suffer from loss of information in areas without texture. Because of the advantages of LiDAR data, in further research these point clouds were used for adding depth to panoramic imagery.

In order to obtain depth information for the panoramic images, it is necessary to register them to LiDAR point clouds. When a single mobile mapping system records both panoramas and LiDAR data, registration can be performed by using known calibration parameters (the relative setup between sensors). It is also possible that the panoramas and the LiDAR point cloud were acquired at different times. In this case, registration can be performed using a combination of SIFT (Lowe,

2004) interest point matching, the Iterative Closest Point algorithm (Besl and McKay, 1992) and Bundle Adjustment.

During registration, care has to be taken to reduce influences of parallaxes that occur between data-sets. These parallaxes are an effect of non-coincidence of the image and LiDAR sensors in space and time. They occur in simultaneous and non-simultaneous recording approaches and can be expected to be larger with the latter approach. Parallaxes could be stored in such a way that in a later phase, users of the final data will be warned not to try measuring in these areas.

Before depth information can be projected onto the images as depth layers, there are still several problems to solve. When projecting the LiDAR point cloud onto a depth layer, not every pixel will have a corresponding depth value and occlusion has not been handled.

To give every pixel a depth value and to handle occlusion correctly, after registration we interpolate between points in the LiDAR data by generating a triangle mesh (Carlberg et al., 2008). The depth layer is then generated by sampling depth values from the mesh at each pixel.

There could still be mismatches between image pixels and their corresponding depth values. These mismatches are most obvious near edges, and arise due to the quality of the registration itself, geometric distortions in the point cloud, mesh generation errors, and the limited LiDAR sampling rate. The last depth-to-image mismatch effects can be solved by locally deforming the depth map. This means that the depth map of an image is morphed towards the image itself using, for example, edge information.

Storing depth layers is very similar to storing RGB values, which means there is no need for a new or special storage format like there is for point clouds. This also implies that no specific software like for retrieval is needed. Obvio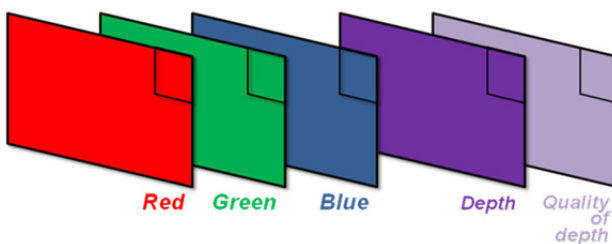usly more storage is needed, the amount depending on the selected compression methodology. In order to use depth layers in combination with the associated panoramic imagery, only extensions to standard viewing tools are needed. It is not necessary to produce completely new software. Therefore users will only need little training to familiarize themselves with the additional options the depth layers provide.

When we also add a layer describing the quality of each pixel (see Fig. 11), it will be possible to calculate the precision of a every point that was measured employing the single click method.



Figure 11: Adding a depth and a quality layer

## 6. ADVANTAGES OF USING 3D DEPTH CYCLORAMAS

Adding depth layers to panoramic images as described above offers several advantages over the use of traditional panoramic images and over the use of LiDAR point clouds where in the process described above discrepancies have been revealed with the geometrically correct panoramic imagery:

- **Navigation**. In comparison with traditional panoramic images, having a depth layer allows easier and more intuitive navigation through the environment. Knowing depth essentially allows navigation in a 3D environment, making the implementation of a user experience using a 3D mouse possible.

- **Single image measurement**. In those cases where traditional panoramic imagery is being used for georeferencing objects, like when performing street inventories, determining the location of an object requires the use of a minimum of two panoramic images and determining the location by means of intersection. When a depth layer is available, such measurements can be performed in a single image with a single click, as now for every pixel X, Y and Z coordinates can be obtained directly (see Fig. 12). An important feature now is the instant availability of the quality of the calculated coordinates. When using colored LiDAR point



Figure 12: A circular measuring mark draped over a 3D depth panorama

clouds (Method 1) for such georeferencing, obviously it is also possible to determine the location of an object with a single click. However, point clouds often lack the resolution for the users to be always able to uniquely identify an object of interest. In case of using an image as a "curtain" in front of the point cloud (Method 2), it is not always clear which object is being measured since the actual measurement is performed "behind the curtain". This is especially the case when measuring narrow objects such as poles, or when attempting to measure edges of objects such as buildings. Offsets between the images and the point cloud might exist, causing the user to unknowingly measure the wrong object.

- **Projection of geographic data**. In metric correct panoramic images it is possible to overlay projections of 2D and 3D geographic data. As we saw before, with 2D data it is not always possible to do this in a proper way. A 3D depth panorama will allow draping the 2D data over its Z-values, thus rendering a more precise view. Moreover, the depth layer allows projection of 2D and 3D maps and designs (not only vector data, but raster/voxel data as well) with removal of occluded features (see Fig. 13). One should be aware of the fact that, in general, maps are simplified compared to reality, while the depth layer is not. This might occasionally lead to removing lines that a user would actually want to see and sometimes showing lines that shouldn't be seen. In that respect it is advised not to remove occluded lines but to shade them.



Figure 13: Removal of occluded lines from an overlayed 2D map

- **Stereo display**. Having depth information available allows the panoramic images to be displayed in stereo over the full 360°. From a 3D depth panorama an image for the left and the right eye need to be generated. Display of the left and right image to the human eyes can be done using numerous techniques (Wikipedia, 2011). When using a single image with its

associated depth layer, the resulting stereo image slightly suffers from imperfections due to occlusion. This is because parts of the scene visible to one of the two eyes are not necessarily visible in the panorama. A more advanced panoramic stereo viewer could use the depth and color information for the occluded parts from neighboring images to overcome this problem. Alternatively it is possible to store multiple depth and color layers per panorama like in layered depth panoramas (Zheng et al. 2007), but in practice it makes storage and use more complicated. Having the possibility of stereo display available can enhance image interpretation. This in turn leads to better and easier object classification. Additionally, improved image interpretation leads to better measurement quality, as it becomes easier to identify objects to be measured. Measurements become more reliable, as identification errors become less likely.

- **Façade textures**. Having the combination of imagery and a depth layer allows for more reliable cut outs of façade textures. In order to avoid the problems of occlusion, it is possible to use data from several neighboring images and depth maps. While this is possible, one should be aware of the fact that the process becomes more complicated, as one needs to merge/blend both the textures and the 3D data from the different panoramas. However because of the good match between imagery and depth layer, the resulting cut-outs will be of higher quality than those based on imagery alone.

- **Backward compatibility**. Besides all mentioned advantages, a 3D depth panorama is backward compatible with a plain panorama. The actual visible information, the RGB values, is not changed. Depth information is only added, no image information is modified or removed. Consequently when applying a sensible file format, existing viewing software for plain panoramic imagery can be used without modification. In that case however the advantages of 3D depth panoramas cannot be exploited.

## 7. CONCLUSIONS

A novel way of combining panoramic imagery with point cloud data is proposed. Compared to two "traditional" methods of using point cloud data, the new approach leads to a more user-friendly use, thus enhancing the possibilities for using the data as well as the quality of interpretation and the quality of measurement results. Compared to traditional panoramic images, the approach leads to enhanced usability of the images, while maintaining the easy visually oriented approach traditional panoramic imagery offers the user.

## 8. ACKNOWLEDGMENTS

## 9. REFERENCES

Beers, B.J. (1995): FRANK – the design of a new landsurveying system using panoramic images. Delft University Press, Delft.

Besl, P.J. and McKay, N.D. (1992): A method for registration of 3-D shapes. IEEE Transactions on pattern analysis and machine intelligence, 14 (2), pp. 239-256.

Carlberg, M., Andrews, J., Gao, P. and Zakhor, A. (2008): Fast surface reconstruction and segmentation with ground-based and airborne LiDAR range data. International Symposium on 3D Data Processing, Visualization and Transmission, pp. 97-104.

Lowe, D.G. (2004): Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), pp. 91-110.

Wikipedia, Aspen Movie Map (2011): http://en.wikipedia.org/wiki/Aspen_Movie_Map (accessed June 23, 2011).

Wikipedia, Stereoscopy (2011): http://en.wikipedia.org/wiki/Stereoscopy (accessed June 23, 2011).

Zheng, K.C., Kang, S.B., Cohen, M.F., and Szeliski, R. (2007): Layered depth panoramas. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), pp. 1-8.