3D RESTITUTION AND RENDERING THROUGH HIGH RESOLUTION IMAGERY:
STATE OF THE ART AND NEW CHALLENGES

C. Valorge

CNES, DSO/OT/QTIS/HR, 18, av. Belin – 31401 Toulouse Cedex 4 – France
Christophe.Valorge@cnes.fr

KEY WORDS: DEM/DTM, 3D, altimetric restitution, rendering, virtual reality, SPOT5, Pléiades

ABSTRACT:

Estimating terrain relief through stereoscopic pairs of images is now a well-known technique, leading to commercial services, as far as low frequency variations of altitude are concerned (edges smoothing, loss of building shapes…). In this case, the relief can be represented by a grid of altitudes, called DEM: this is a « 2.5D » representation of the landscape.

Concurrently, recent advances in real time 3D rendering, especially led by computer graphics, virtual reality and video games techniques, make it very easy, even on a standard laptop, to virtually fly over the landscape made by the image draped over the DEM.

These two points will be illustrated with SPOT5 results.

However, dealing with high resolution imagery (i.e. better than 1 meter) proves to be difficult in these fields. At this scale, « real 3D » appears throughout the entire image, with specific features: some walls can be seen, parts of the landscape are hidden by other parts (vegetation, buildings…), shadows change the radiometry of small but numerous parts of the landscape… To summarize, the appearance of objects in 3D makes it impossible to use our « traditional » 2D image processing tools and we have to reconsider all our methods. In terms of relief estimation for instance, next years' challenge will be to extract not only grids of altitudes but objects shapes and sizes, real 3D models of the landscape.

In terms of representation of the landscape, one can understand that a 2D projection can not be sufficient and that one has to represent it in real 3D : a 3D model textured with the image can thus be visualized with the same « virtual flight » techniques as in “2.5D”. This textured 3D model (with or without shadow correction) will also be useful (and necessary) to generate other views of the landscape, orthoimages for instance.

In this part, illustrations will be given from the future Pléiades program.

To summarize, being able to build this kind of 3D modelling of a landscape from high resolution images will be a pre-requisite of most uses of high resolution images. This is why, in the framework of the Pléiades program preparation, CNES is building an « accompanying » methodological program aiming at defining the appropriate modelling of the landscape and the corresponding information extraction techniques.

1. INTRODUCTION

Estimating terrain relief from spatial stereoscopic pairs of images using image matching algorithms is an old and validated technique. It has been used for 15 years on SPOT1 to SPOT4 to operationally and commercially produce Digital Elevation Models, i.e. altitudes given on a regular ground grid.

Recent availability of imaging satellites able to acquire along-track stereo-pairs (thanks to their agility like IKONOS or QuickBird or with a specific and dedicated instrument like HRS on-board SPOT5), now allows truly operational DEM generation.

As the acquisitions are quasi-simultaneous, both images are available at the same time (no need to wait for days for a second pass after the first to obtain the second image) and as there are no changes in the landscape and similar illumination conditions, matching techniques prove to be very efficient (over 95 % good correlated points).

Even for high resolution sensors however, these techniques are still limited to the generation of regular altitude grids, unable to model vertical walls, arches of bridges… and the complexity of urban geometry in general. This is what one can call a “2.5 D representation” of the landscape.

Even if not optimal, this kind of representation can fully benefit from state-of-the-art techniques in real time rendering, thus allowing a new way of analyzing images. The image warped over the DEM contains the same information as an anaglyph representation but can be presented in a much more “natural” and intuitive way: the flight over this virtual landscape.

SPOT5 THX images (2.5 m resolution in colour) draped over HRS or PXS DEM represent the ultimate case of such 2.5D representation: at this resolution, buildings and vegetation are seen in “read 3D” in the image but not represented as such in the DEM. One can understand that, especially for high-resolution sensors, real 3D geometric modelling is necessary to fully benefit from the information of the image.

This is one of the next challenges concerning the use of high resolution images.

2. STATE OF THE ART IN 3D RESTITUTION AND RENDERING

2.1 Relief estimation from images

We present in the next sections the state of the art in relief estimation with 10m or 2.5m images (using SPOT5 as example) or with 1m images (using simulated Pléiades images).

2.1.1 Relief estimation with SPOT5

Launched on the fourth of May 2002, SPOT5 is designed to ensure the continuity of SPOT services and to offer customers and users improved performances in terms of capacity of
acquisition, image resolution and related services (Fratter, 2001; Meygret, 2002, Gleyzes, 2003). It carries four payloads: two “High Resolution Geometry” instruments (HRG, evolution of the former HRVIR on-board SPOT4), the “High Resolution Stereoscop y” instrument (HRS) and the low-resolution mission VEGETATION2 (identical to VEGETATION1 on-board SPOT4) for daily coverage.

2.1.1.1 Along-track stereo with HRS
The HRS instrument is made of two telescopes allowing a 20° forward view and a 20° aft view over a 120-km swath, respectively (Figure 1).

![Figure 1: HRS viewing capacity](image)

This concept of along-track stereo view enables both panchromatic images of the pair to be obtained within an interval of 90 seconds in a single pass. The resulting B/H ratio is slightly higher than 0.8. The instantaneous field of view of each detector is 10 m but the sampling rate is 5 m along the track allowing higher altimetric precision of the DEM. As forward and aft image acquisitions are exclusive, the length of a continuous stereo strip cannot exceed 600 km. However several segments can be linked together for mass production. The altimetric accuracy of the DEM obtained with HRS has been evaluated during in-flight commissioning phase (Rudowski, 2003; Nonin, 2003) and the main conclusions are as follows:

In non-mountainous areas, the altimetric accuracy of the DEM is better than 5m RMS with more than 95% correctly matched points. In mountainous areas however, this performance can be lowered to 10m RMS (because of the relatively high B/H ratio, steep slopes can remain unseen on one of the images) with a simultaneous decrease of the matching success to 85%, which never the less remains much better than with non-simultaneous stereoscopy.

These excellent results allow SPOTIMAGE and IGN to commonly and operationally produce huge surfaces of DEM, as in the Reference3D project (Bernard; 2003a and 2003b). Further investigation and quality assessment of HRS DEM will be conducted under the HRS-SAP initiative (Baudoin; 2003).

2.1.1.2 Across-track stereo with HRG
The HRG instruments benefit from noticeable improvements of the SPOT4’ HRVIR concept: the colour (XS) images have now a 10m resolution and the panchromatic images can either have a 5m (HM) or 2.5m (THR) resolution. As with SPOT1 to SPOT4 images, the low B/H ratio (1.8 %) between color and panchromatic images allows high performance image fusion, thus giving access to 5m (PXS) or even 2.5m (THX) color imagery (Latry, 2000 and 2003).

![Figure 2: HRG viewing capacity](image)

Like HRS, the altimetric accuracy of the DEM obtained with HRG across-track stereoscopy has been evaluated during in-flight commissioning phase (Rudowski, 2003; Nonin, 2003) and the main conclusions are as follows:

The major factor contributing to performance remains (as for SPOT1 to 4) the time lag between the two acquisitions, the shorter, the better, especially in terms of percentage of correctly matched points. At a lesser degree, the B/H ratio also influences the performance, with the optimum between 0.2 and 0.9, which is quite easy to respect.

In these best conditions, the altimetric accuracy is estimated to be better than 7m RMS with the 10m colour bands (same result as with the former 10m panchromatic bands of SPOT1 to SPOT4), better than 4m RMS with the 5m panchromatic bands and shows a further 10% improvement with the 2.5 m panchromatic band (see Nonin; 2003).

2.1.1.3 Along-track stereo with HRG
Although weak, the 1.8 % B/H ratio between the panchromatic band and the colour bands allows relief estimation thanks to a highly precise matching algorithm (Vadon, 2003).

The advantage of this low value of B/H ratio is that both images look really similar (similar view angle and only 2s time lag), especially when correlating the 5m panchromatic band with a combination of the B1 and B2 bands. Operationally of course, it also benefits from the “classical” advantages of along-track stereoscopy: simultaneous and immediate availability of all data.

The main drawback is that any error in either instrument modelling (focal plane distortions) or satellite attitude estimation will have a large impact on the DEM quality: a 1 µrad error due to attitude jitter, for instance, will be interpreted as a 45m elevation oscillation in the DEM… To avoid this, special algorithms filtering along-track oscillations (due to attitude jitter) and across-track systematic features (due to viewing direction mis-modeling) are used.

Primary results show an altimetric accuracy in the range of 10m RMS, with more than 99% correctly matched points.
2.1.2 Relief estimation with Pléiades-HR

Pléiades-HR is the high-resolution optical component of the Pléiades programme, which also includes an Italian high-resolution X-band component called COSMO-SkyMed. Due to launch in mid-2008, Pléiades-HR will have a 70cm resolution in panchromatic and 2.8m in colours (B, G, R and NIR) over a 20 km swath at nadir. Its high agility (25s for 60° in either roll or pitch) allows along-track stereoscopy with the same instrument and even tri-stereoscopy from the same track (see figure 3).

Its geo-location accuracy is specified to be better than 10m CE at 90%.

![Figure 3: Pléiades-HR viewing capacity](image)

During Pléiades conception phase, a large study, focused on 3D extraction capability with this future imagery, was conducted by 5 different bodies (IGN, ISTAR, Alcatel, S&DE, Thalès IS) under CNES coordination. The main conclusions are as follows:

Over urban areas (main interest of high-resolution stereoscopy), it was shown that a quasi-nadir view is mandatory to have access to the “bottom” of the streets between buildings. Concerning the “ideal” B/H ratio, it is a trade-off between the percentage of well-matched points (higher with low values of B/H) and the altimetric precision (better with higher values of B/H), with an optimum between 0.2 and 0.4. But the best results are obtained with 3 images, one nadir and the others giving B/H ratios around 0.1 and 0.4.

In these best cases, the altimetric accuracy can be estimated around 3m RMS with more than 90% correctly matched points.

2.1.3 Discussions on these results

At first sight, it is surprising to realize that altimetric accuracy is not proportional to resolution: whereas one can obtain sub-pixel accuracy on 10m or 5m images, it’s no more the case with 2.5m resolution and even worse (4 pixels RMS) with 70cm resolution!

In fact, this apparent paradox is due to the complexity of the objects (especially buildings) that we try to measure with these photogrammetric matching techniques, which leads us to the following question:

2.1.3.1 What do these DEM really represent?

It is now commonly admitted that the quality of a DEM not only relies on its planimetric and altimetric accuracy: the spacing between the points at which altitudes are given must also be taken into account (not to mention the confidence attached to each individual measurement).

For instance, it’s easy to understand that the GLOBE DEM is not appropriate in mountainous areas because of its 1 km posting. The same idea applies for a 30 m spaced DEM on an urban area.

When the resolution of the images used to generate the DEM improves, a new question arises: which altitude do we measure through the matching algorithm? Very high resolution images (aerial 25 cm images for instance) tend to give access to the altitude of the highest point of the landscape: the ground for the roads and parking lots, the roof of the buildings, the top of the trees… The answer is less obvious with 10m, 2.5m or even 1m images: the estimated surface is an intermediate between the ground and the top of the objects of the scene, as the measure is in fact not a pinpoint altitude (the correlation coefficient gives an average value on each image chip).

An illustration is displayed on figure 4, where we compare altitude profiles of the same urban landscape estimated with different images (green : SPOT5, 5m – red : SPOT5, 2.5m – blue : aerial, 25cm).

![Figure 4: altitude profiles from different image resolutions](image)

Note that as our photogrammetric methods only give access to the surface over both ground and above ground features, our DEM quality assessment used in all cases reference data corresponding to this surface: it was either DEM obtained through photogrammetric methods using higher resolution images or the combination of a traditional DTM (altitude of the ground only) with a 3D model of the buildings.

2.1.3.2 Evaluation of “real 3D” extraction capacity

Having noticed that altimetric accuracy is unable to account for real improvements in terms of sharpness of the estimated relief, a first solution can be simply to display the DEM obtained over urban areas to compare them visually (figure 5, next page).

Unsurprisingly, the conclusion is that there are, in fact, strong differences in terms of DEM sharpness between 2.5m, 5m and 10m imagery but that 2.5m resolution is unable to allow real 3D estimation, especially over European, dense urban landscape. For this purpose, 1m imagery at least is mandatory.

If we want to quantify this “sharpness” of the DEM, the only solution for high resolution imagery is to change our modelling of the landscape and go one step further: it is mandatory to consider the objects of the scene and to process them as such.

To do so, in terms of quality assessment, we have to be able to identify the different objects of the scene (buildings, bridges, trees…), obtained from the reference data and to proceed to dedicated estimations, depending on the type of these objects. For instance, the altitude of the buildings, as the difference between the mean altitude of the roof (excluding values near the edge) and the mean altitude of the ground (excluding values near the base of the buildings) can be compared with the reference value. Similarly, we can use the Hausdorff distance to qualify the estimated planimetric position of the edges of the buildings.
Even if better than “raw” altimetric accuracy estimation, this interesting approach really quickly appears limited because of limitations of the DEM itself: how can we compare the altitude of a two-slope roof? How can we estimate the quality of the shapes of the buildings? To do so, we have to model the objects as “seen” in the DEM; in fact, this object-level comparison needs an object-level representation of the scene. And as it would be a pity to develop vectorization algorithms of the DEM only to assess their quality (even if the problem is easier to solve thanks to the reference data), the real challenge is now to go from this 2.5D representation of the scene to a real 3D representation. This will be discussed in the “new challenges” section.

2.2 3D rendering of these DEM draped with images

3D navigation over a virtually rendered landscape is becoming a popular application, especially for defence, airport and territorial management and communication towards general public and managers. Products of this application can either be video files displaying a chosen flight over the scene or simply 3D views of the landscape (figure 6).

2.2.1 Specific needs of this kind of application

Logically, this new and specific application generates its specific needs in terms of DEM and image absolute and relative qualities.

Concerning the image, as the navigation aims at being realistic, it must be in natural colours, i.e. blue, green and red. Concerning SPOT5, for instance, it implies both panchromatic-colour fusion (to obtain 5m or 2.5m resolution colour images) and conversion from “false” colours (green, red and NIR) to natural colours (blue, green and red).

Besides, as the DEM is unable to account for 3D objects (like buildings and trees), the image is preferably acquired near nadir to avoid walls visibility disturbing natural vision (false 3D). Concerning the DEM, it can be noted that low spatial frequency altimetric errors are undetectable during 3D navigation so that altimetric biases and drifts can be much larger than with other applications. On the other hand, this application is very sensitive to correct high spatial frequency representativity of the DEM: hills, ridges and valleys (particularly if there are visible roads or rivers) must be correctly extracted.

And because this application combines imagery and DEM, relative quality must also be ensured: specific DEM artefacts become really disturbing, especially those affecting planar surfaces as sea, lakes but also airports, stadiums, parking lots, roads (see figure 7)…

The relative positioning of the image with respect to the DEM must also be of pixel accuracy, throughout the scene (to avoid rivers not flowing in the bottom of the valleys, for instance). Finally, the respective resolutions of the image and the DEM must be compatible: there is no use for a 30m DEM when viewing a VEGETATION 1km image and, reciprocally, 3D navigation over high resolution images (from 2.5m to sub-metric resolution) in which the human eye naturally detects 3D objects, is hindered by low frequency DEM, unable to represent 3D objects (see figure 8).
With these “specifications”, one of the best DEM to use to drape SPOT5 images is in fact the along-track HRG DEM: it is a by-product of the panchromatic/color fusion process; even if not absolutely highly accurate, its high frequency restitution is good; the good similarity between the images gives a high percentage of correctly matched points and, last but not least, both the merged image and the DEM are natively in the same geometry: such a 3D navigation is thus potentially available on every HM-XS or THR-XS acquisition!

Of course, 3D navigation over HRS or across-track HRG DEM is possible but is far less straightforward: it needs more imagery and one has to superimpose these images in a common geometry.

Furthermore, it’s also possible to drape SPOT5 images over other DEM (IGN reference DEM of France, for example), provided that relative positioning image/DEM has been achieved.

2.2.2 State of the art in 3D navigation

Even if numerous image processing and visualising tools are able to drape images over DEMs, the size of the images delivered by our satellites (1.6 Go for one SPOT5 THX scene) largely exceeds standard memory capacities and a real time navigation over this kind of scene needs specific handling of the data in order to only keep in memory what is displayed on screen.

As an outcome of a R&D study with the CS-SI company, CNES participated to the development and improvement of a software called VirtualGeo. Its main characteristics are:

- Available on a PC PIII or Athlon 1000 MHz (or higher) with 256 Mo RAM, 10 GB hard-drive (SCSI preferred) and fully compliant with OpenGL 1.3 3D graphic card (Nvidia / GeForce for instance),
- Real-time 3D visualization software of massive textured terrain databases using raster images
- Intuitive, 6 degrees of freedom navigation using keyboard, mouse or 3D control peripheral (joystick, Spacemouse®...)
- Visualization of lighting with cast shadow at a given time
- Sophisticated algorithms used to pre-compute DTM, that allows VirtualGeo to maintain terrain topology as well as 30 FPS (Frame Per Second) during visualization,
- A fine multi-resolution texture management,
- A “dynamic loading” and a “pre-fetching” algorithm that allow VirtualGeo to load only useful data in the main memory. It avoids lag during data loading, because VirtualGeo is able to determine which area will be displayed in the next 5 frames and checks that data are in memory
- AVI export,
- Stereoscopic visualization capacity.

Before performing the navigating, one must pre-process the images and DEM to convert them to VirtualGeo internal format, with a CNES software called TerrainDesigner: the files are split into super-imposable tiles, the image is compressed, the DEM converted into an irregular network of triangles (TIN)...

The only requirement for input data is that the DEM and the image are in the same geometry (rows and columns must be parallel but they have neither to be at the same resolution nor to have the same location for their first pixels). The user only has to give the image and DTM sizes and resolutions and to indicate one tie point between them.

Thanks to these two pieces of software, 3D navigation is easy to perform over any image associated to a “2.5D” DEM.

In this case, the image is vertically projected onto the DEM so that quasi-vertical triangles of the DEM are draped with quasi-uniform textures, leading to poor rendering of facades (figure 9): in this field also, one can feel that real 3D rendering will need further improvement...

3. NEW CHALLENGES IN 3D RESTITUTION AND RENDERING

3.1 Real 3D restitution

As a preparation for the Pléiades-HR program, CNES is undertaking different methodological studies aiming at making image use easier for non-specialists. One of the items of these studies deals with “real 3D” restitution from Pléiades images. The reviews concerning this study are still on-going but some first elements can be considered:

Real 3D estimation is already a major problem in aerial photogrammetry: conclusions drawn after the first Pléiades study converge with those obtained in aerial altimetric restitution (need of a nadir view, need to multiply the number of images to automate the process and improve the coverage...). The Pléiades studies will thus be conducted in cooperation with these experts. The underlying modelling of the scene must go further than a simple grid of altitudes: at first order, it can be considered as...
In fact, the need for real 3D modelling not only comes from 3D virtual flights but is a general need for all applications. The first level of modelling of the scene can be an irregular network of triangles (TIN), each of them associated with the corresponding texture in the image. At this first level, much work has to be conducted to be able to extract this kind of information from the images and to be able to use it (change of geometry, virtual flight...).

The second level of modelling of the scene is a real object-level. At this level, the first work is to establish a complete catalogue of the different objects of any scene, then to identify/adapt existing modelling of these objects and create the missing ones. After this, the techniques of extraction of these objects from the image will have to be studied.

5. REFERENCES


