INSTRUMENTAL REQUIREMENTS FOR VIRTUAL REALITY

Prof. O. Kölbl, M. von Moos
Ecole Polytechnique Fédérale de Lausanne, Switzerland

Commission II, Working Group II/6

KEY WORDS : Virtual reality, automatic aerial triangulation, video photogrammetry, study of impact.

SUMMARY

Photogrammetric measuring methods represent a precious base for photo montage and the production of video scenes in which fictive objects have to be integrated. This technique is of special interest for planning and studies of impact of new constructions. The article gives an overview of the possible techniques of production of video scenes with integrated artificial objects conceived with CAD tools. This integration requires the orientation of the photographs, preferable by automatic aerial triangulation and tools for the combination of video images with artificially produced film sequences. Some examples are presented concerning the study of impact of a transmission tower, the integration of a new railway line into the landscape and an analysis of traffic safety on a dangerous road section.

1. SCOPE OF THE STUDY

With the progress of digital image processing, orthophotos became a common medium for the presentation of GIS data. However othophotos are static and give a specific view not always in accordance with the view to which a human observer is accustomed. Therefore town models based on the creation of spatial objects became a habit, especially for buildings. These types of presentations make it possible to generate even moving scenes with the help of corresponding softwares like Microstation, Autocad or Masterpiece. However, the graphic presentation is in most cases quite limited, as our environment is much more complex and simple cubes hardly satisfy a critical observer. On the other hand, a more sophisticated presentation including the modelling of windows and details of the roofs is time consuming and requires extensive field surveys. Furthermore, our landscape is not only formed by buildings but also includes the vegetation and is animated by human beings. If one wants to avoid overly detailed modelling, it is more efficient to use real video scenes and to model only those elements which do not exist. Let us recall in this context that most animations are needed to study the impact of a new building or a new construction, or other technical precautions.

2. REQUIREMENTS FOR THE COMBINATION OF VIDEO IMAGES

2.1 Combination by "Paste and Copy"

Filming a landscape or actions is quite common and has also been developed as a well-known art. However, combining such scenes with fictive objects requires certain precautions, leading to photogrammetric techniques, whenever high geometric precision is required such as for fitting of objects into the perspective. It is true that the film industry has been using techniques to combine film scenes for a long time. Very often, rather simple geometric manipulations are used for such combinations. Pasting and copying can already produce very surprising effects. If one wants to project a moving person in another film, one has to eliminate the background of the person and adapt the scale of the person which can even vary, giving the impression of moving back or forth. Various software products allow this type of manipulation. The elimination of the background is rather easy if one films the person against a homogeneous screen in a colour, preferably dark blue, which is not found in the clothes. By spectral manipulation, one can then simply transform the background into transparency, so that it will no longer interfere with the underlayed scene. With the help of special effects, one can even add a shadow to the inserted person.

As long as only raster pictures are used for the modelling of the objects, then the possibilities for geometric transformations remain limited to plane operations, like scale changes, stretching or rotations within the image plane, but it is impossible to turn the object round and see faces which are not present in one single image. This would already require the combination of more then one picture. Therefore simple image combination is not satisfying for extended objects to be observed from different view angles, but this simple approach has undeniable advantages for many tasks.

2.2 Combining Constructed Object with Film Sequences

Another common technique in the film industry is the combination of constructed objects with natural scenes. In principle, the same techniques as described above can be applied in this context. However, one can also go a step further and apply a proper photogrammetric approach; that means that one determines precisely the camera position used for filming the video scenes and then projects the artificial objects into the video images.

The orientation of photographs is more or less routine in photogrammetry, only we developed techniques appropriate for one single image are a few, but the treatment of a video sequence with some 1000-10,000 images requires a considerable degree of automation. Let us recall in this context that a video sequence of only 3 minutes generates some 4000 images : 3 minutes * 60 seconds * 25 images/second, which might have to be treated. Furthermore the pictures have a rather narrow view angle and might show angular variations of up to 0.1 - 1 degree between subsequent images when taken from a moving car without special precautions.

On the other hand, we can assume that the projection centres undergo only slight variations from one image to the next. If we assume that the video sequence is filmed from a moving car, then the projection centre might change from one image to the next between 0.15 to 1 m depending on the speed (10-60 km/h) of the camera. Additionally, most video cameras are equipped with zoom lenses, which might have been manipulated meaning a change of the focal length or rather of the principal distance between successive images. Finally possible vibrations of the camera should be corrected as they disturb the impression.

3. THE PHOTOGRAMMETRIC ASPECTS OF VIDEO IMAGES

3.1 Image Acquisition

Video cameras have been widely circulated in recent years, also for amateur applications. Most of the models use CCD elements for image recording. Then the raw signals have to be composed to form an image at the exit of the CCD sensors. The images can be stored in two different ways, either analog or digital. The analog technique is still the most common one, based on the TV-standards developed in the 50s for image transmission. They depend heavily on precise synchronisation and suffer from the ageing of the storage medium. Only recently (about 1996) has it become possible with these cameras to also store the images in digital form with special DV cassettes (Digital Video). Prior to storage, the images are compressed, in general according to the JPEG procedure (Discrete Cosine Transformation) to about 1/5 of the original data volume.

The video cameras used for these studies belong to the new generation: The Sony DXC-D30 (cf. fig. 1) is a camera for professional applications, which supplies high resolution images with low distortion. The JVC GR-DVM1 camera (cf. fig. 1) belongs to the family of the "mini DV" for amateurs, equipped with DV cassettes (cf. figure 1). The JVC is equipped with a zoom lens 1:10 (f = 4.5 mm until 45 mm, pixel size ≈ 5.5 μm, image format about 3.5 x 2.7 mm).

3.2 Image Preprocessing

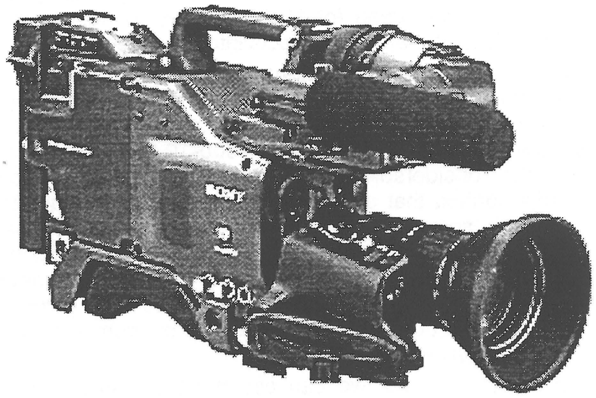
The images recorded on tape cannot be transferred directly into an aerial triangulation program; rather, it is necessary to transform them into a standard format and store them on the mass storage device of a computer. The Institute of Photogrammetry used the standard routines of the image library of Silicon Graphics for this task. Unfortunately, it was not possible to transfer the images directly in their numeric form; the digital data had to be transformed into analog signals over a Frame Grabber and were then transferred to the mass storage of the computer. Silicon Graphics offers a large variety of facilities for image transformation, including routines for the generation of images in Tiff-format. There are also special procedures for the compression of video sequences (MPEG- Moving Picture Expert Group); however it was not possible to take advantage from these routines.

3.3 Positioning of the Projection Centre of the Camera

Coming from photogrammetry, one quickly has the idea that the question of how to determine a camera position has already been largely resolved. In fact, in survey planes, GPS and INS are very often used for the determination of the camera position. Furthermore, we can think of car navigation systems and realise that a similar problem is again solved by GPS, a compass and the counting of the wheel rotations. Practical tests with 3 GPS mounted on a car also showed that its position as well as its orientation can be determined quite easily in this way. When installing the video camera rigidly on the roof of the car, one can effectively record the camera position and the camera angles every second; it is then possible to compute the effective camera position for every image (25 images per second) by simple interpolation. However, the use of GPS requires open terrain, and an uninterrupted signal for several minutes at least. Furthermore, one can only determine the position of the antenna properly when operating with a hand-held camera, in this case the angular information is difficult to determine.

3.4 Aerial Triangulation with Video Images

A very rigorous approach for the determination of the orientation elements is definitely the technique of aerial triangulation. In theory, video images can be used as other images for aerial triangulation. Furthermore, the technique of automatic aerial triangulation might allow one to run this operation completely automatically. However, the great overlap of the images and the narrow opening angle of the camera fundamentally change the possible precision of the orientation elements. Furthermore, the images are often displaced in the direction of the taking axis, contrary to standard aerial triangulation.



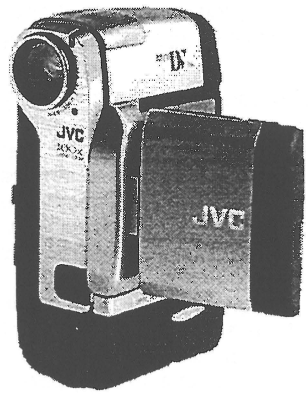


Fig. 1
Digital video cameras Sony DCX-D30, a professional camera, and JVC GR-DVM1, a digital video camera for numerical recording.

Precision requirements

Let us recall in this context that it is practically impossible to obtain the 5 orientation elements with reasonable accuracy when doing only relative orientation with photographs overlapping by more than 90%. However, the main objective of our exercise is to obtain an obviously proper coincidence between the image content and the fictive objects to be superimposed. That means that it would be sufficient to place a number of control points around the position of the fictive object and to determine the orientation elements in function of these points as well as possible. For a more or less linear object, one would need only 2 points, and a simple scale adaptation would already be satisfying. The other orientation elements could be chosen freely. Even when in full control, one quickly realises that ϕ and ω are heavily correlated with the shift of the projection centre in X- and Y-direction. A simple calculation shows that for a taking distance of 100 m, an image rotation of 7° can be compensated by an image shift of 14 m by tolerating an image error of only 1 pixel in the extreme corner (assumptions: focal distance 50 mm, image size 3.5 x 3.5 mm or 500 pixels à 7 μm). In this case we used the well-known parallax formula $\Delta x = c(1+x^2/c^2)\Delta\phi$ and computed the $\Delta\phi$ for a $\Delta x = 7\mu m$, omitting the linear term. That means that for narrow angle cameras the projection centre must be known to a precision of only 10% of the taking distance, the rest can be compensated by the angular orientation elements. It is understood that the focal distance enters in a linear way into this formula and for a focal length of 25 mm one would obtain a $\Delta\phi$ of 3° and a displacement of 5% of the taking distance. Under these circumstances, it is possible to determine the position of the projection centre by approximate methods and then limit aerial triangulation itself to the orientation angles. Whenever the photos are taken from a stable platform, one can even omit the image rotation κ from the adjustment algorithm. That means that only one point would have to be transferred from one image to the next.

Procedures of Automatic Aerial Triangulation

Currently the most common packages for automatic aerial triangulation are Match AT (Info, Ackermann), Phodis AT (Zeiss) and HATS (Helava). For practical reasons, it was not possible to use Match AT for the triangulation of video images, but tests have been made with the Helava program HATS and the triangulation program of Zeiss Phodis AT. The problem with Phodis AT is the requirement of large images of at least 1500 x 1500 Pixels in order to obtain the necessary size of the image batches. HATS features a special parameter set developed specially for video images and the process of point transfer operated to our satisfaction. Of course it was still necessary to eliminate a number of mismatches after the process of point transfer. The possibility of limiting the number of orientation elements in the adjustment program, so that effectively only the angles were determined, also proved favourable.

Inner Orientation

It is understood that video cameras are not metric cameras, with a well-defined principal point and a calibrated principal distance. Here again, it is important to state that the inner orientation must be just precise enough not to disturb noticeably the adjustment; that means that deviations of 1 pixel or 0.2 % of the format can be tolerated. According to our experience (Kölbl, 1976), the principal point can be assumed precisely enough in the centre of the image. The principal distance is variable, whenever the zoom is used; however, its value is only important for objects with a significant depth extension. Whenever the object is more or less close to a plane one can correct an error of the principal distance with the help of the outer orientation elements. Furthermore, various software packages for aerial triangulation like HATS allow one to introduce the elements of inner orientation as unknowns. Problems have so far resulted only from distortion, such as with Softimage, which does not allow one to introduce a distortion, consequently residual errors of 1-2 pixels remained in the corners of the images, which could not be corrected.

4. IMAGE PROCESSING AND IMAGE COMBINATIONS

Various software packages are available for the creation of video sequences. In the beginning we already pointed out that our main objective is the combination of video sequences with projected objects. A main tool for this task is the software package "Softimage" with "Eddie"; Softimage is a drawing tool, whereas Eddie allows the proper combination of video sequences. Products similar to Softimage are 3D Studio Max and Explorer; they all find their principal application in the film industry, whereas in engineering and planning one mainly uses CAD tools like Microstation (Benthley and Intergraph). A very interesting solution is also offered by "Perspective Scene" integrated into the Helava software Socet.Set for digital photogrammetry.

4.1 CAD Tools

Microstation and other CAD tools offer the possibility of constructing 3D objects and presenting rather realistic visualisations by object rendering; the selection of the colours gradated according to the intensity of a fictive illumination and the possible texturing of objects allows one to create very efficient views of constructed objects. Many CAD-tools then make it possible to define a path of a camera and the calculation of a sequence of images, which can then be stored and presented as film. This approach has the advantage, that one can work with a drawing tool extensively used for planning; there is no need for exporting and restructuring the data. The tool is efficient as long as one is limited to the constructive elements; the incorporation of raster images as texture elements is possible but limited to static applications. However, it is not easily possible to integrate moving objects like cars or people.

4.2 Image Draping

In order to create more realistic views, one frequently drapes raster images over a reference object. The simplest approach is to create a DTM and to drape an orthophoto over the meshes; this scene can then be observed from different oblique view angles, which can be combined to a film sequence. Intergraph also supplies standard products for image draping, such as various software tools like "ISI".

A more refined tool for image draping is "Perspective Scene" by Helava; this package is integrated into the software for photogrammetric plotting. It allows the draping of various images onto a wire frame model, which can include constructions like houses. The software automatically chooses the raster images which are optimal for the texturing of an object surface; even oblique images can be used for texturing. This tool was used for the analysis of road security, as will be explained later.

4.3 Softimage

Softimage can also be considered as a CAD software package, but has many facilities for the animation of the scene content. Softimage was founded in 1986 by Daniel Langlois, Montreal, a movie and animation designer. The firm which started with a capital of 50'000 $ and counts today 300 employees. The software is used for special effects and animation in numerous movies like Jurassic Park, the Lion King and The Mask. In June 1994 Softimage was purchased by Microsoft. The original version runs on Silicon Graphics, but it is currently ported to Windows NT. The operator has 4 views to design his objects in 3D (cf. fig. 2). As usual he can use points, lines and surfaces; the surfaces are subdivided into finite elements; the objects can be rendered or can obtain a texture and can be deformed temporarily according to a definable time function. The flexibility goes so far that even human persons can be constructed and seem to move around quite naturally. One can even order a library of different figures and animate a scene in this way.

In a similar way a camera can be defined and can be moved through space according to a well-defined path; during the movements, the angular orientation and the focal length can vary.

The software allows the import and export of data under others in DXF-Format, which facilitates communication with other software packages

4.4 Combination of Image Sequences

Softimage can be combined with the software tool "Eddie" in order to combine different film sequences in raster format; that means that elements constructed with Softimage can be projected into a film sequence. If one properly defines the position or the path of the camera, then the constructed elements are properly combined with the existing images; special tools allow operators to define fore- and background and masking for proper superposition.

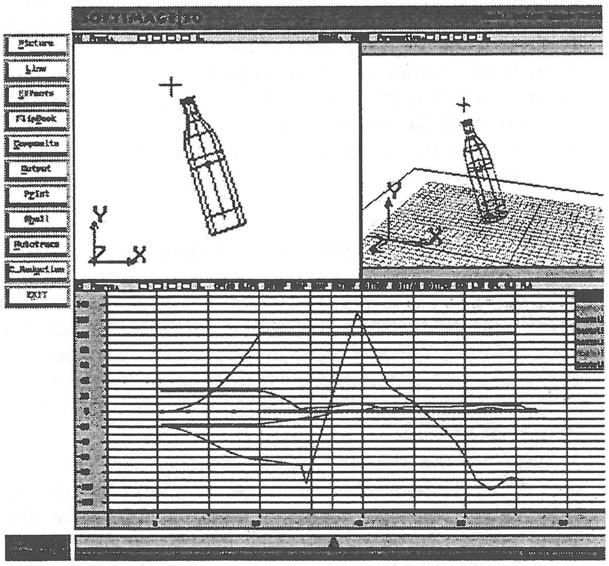


Fig. 2
Working field for Softimage, the 2 graphical windows are reserved for the construction of objects, the lower window allows the definition of object movements in function of time for the video sequences. Each curve models a specific parameter of the object or of the camera.

5. EXAMPLE OF APPLICATIONS

According to the general aim, the applications concentrated on projects for which the visual impression is of importance. Up to now we have treated a study of the impact of a high voltage transmission tower into the landscape near Martigny, and the impact of the new railway location line of "Alptransit" near Biasca; furthermore we studied the traffic security on a road section, along which several accidents occur.

5.1 Study of Impact of a Transmission Tower

The photographs for the example of the transmission line have been taken from a moving car with the JVC camera. The tower was modelled with Microstation and then imported as raster image directly into Eddie. In this case, we gave up on a correct orientation of the images and localised the tower with the help of marked image details; in order to keep the effort for the image localisation small it was placed properly in every 10^{th} image only. In principle Eddie has a function enabling it to stick an object to an image detail; however, this detail must be quite distinct in order to obtain a good recording.

The resulting images show the essential features (cf. fig. 3); however, the images are not stable enough and the tower seems to tremble, which is quite disturbing. This rather humoristic effect of trembling might considerably disturb an unbiased observer and there is the danger that such an involuntary aspect obstructs the proper aim of the video sequence.

5.2 Alptransit (von Moos, 1998)

In order to facilitate the image orientation, we mounted the camera for the following investigation on a tripod and limited the image movement to a slow lateral rotation. The firm Passera & Pedretti graciously made available to us the projected railway location in all details, including bridges, cuts and damsites in 3D Autocad format. The data have been imported into Microstation for some editing and then exported to Softimage; all these transfers of data were realised over the DXF-format.

Again, each 10^{th} image has been orientated, using control points. The Helava routine for the automatic transfer of tie points was also tested and proved efficient. The result was a stable image showing the new project well. The railway location was only slightly edited and is presented in rendered colours; furthermore a few trees had to be corrected to improve their visibility.

The result was very satisfactory; the colourful presentation of the railway location facilitated the recognition of the project and its amplitude, but did not disturb the general impression. Of course by choosing a better texture and copying a real train into the scene a complete natural scene would have resulted; however it would have been difficult to explain to the observer where to see the projected railway (for technical reasons, the images can only be shown as film and are not reproduced here).

5.3 Analysis of Traffic Security with the Help of Virtual Reality (Gigon, 1997)

A somewhat different task had to be solved for the road Cossonay - La Sarraz, a section on which a considerable number of traffic accidents occur. Again there was the intention to use the working tools of virtual reality to find an explanation for the problematic traffic conditions. We therefore tried to simulate the view of a car driver, in using Perspective Scene of Helava. The aerial photographs

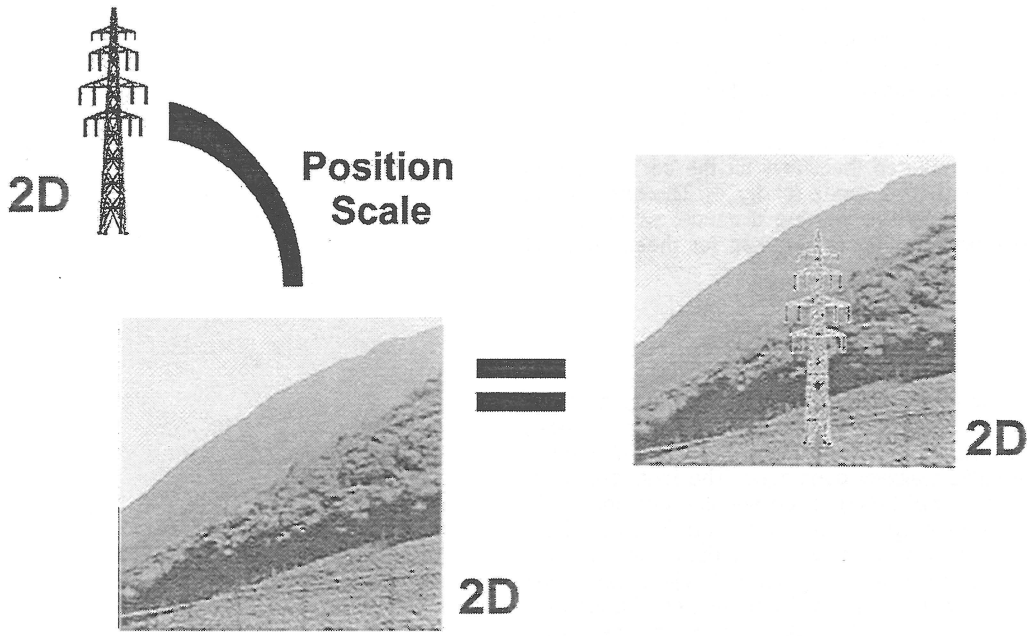


Fig. 3
Principle of the combination of a constructed object (transmission tower) with an image video scene.

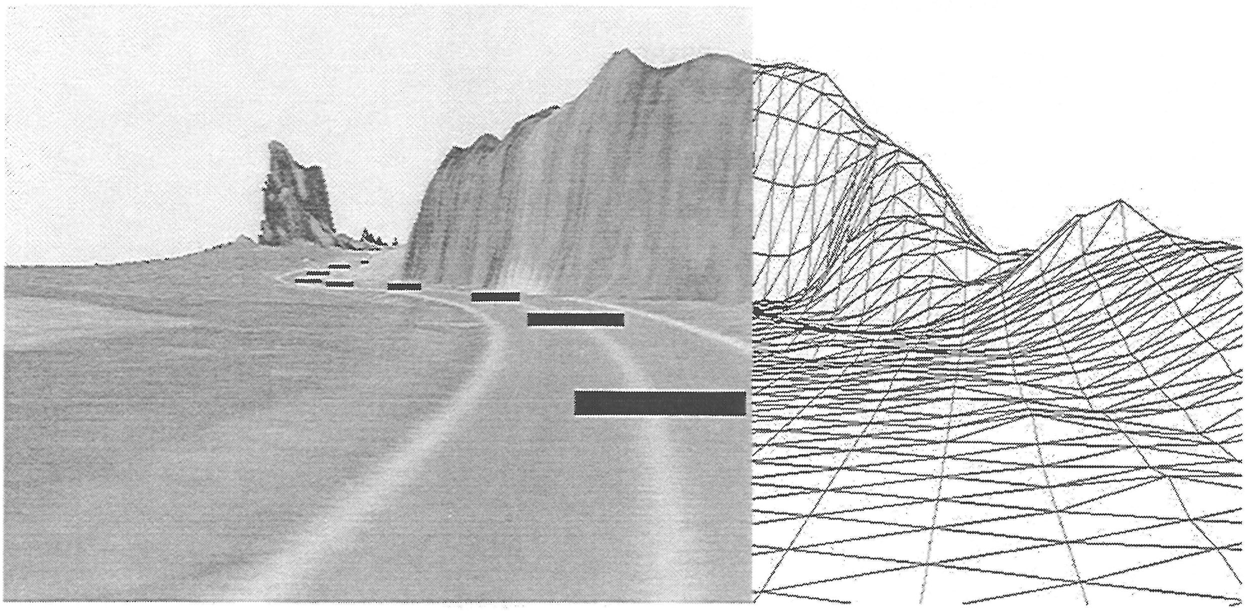
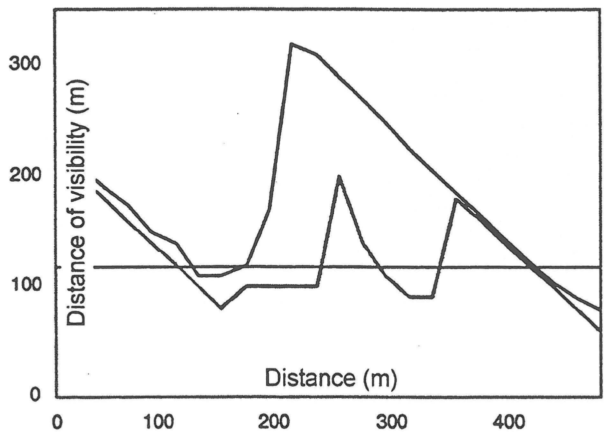


Fig. 4 ⇧
Oblique view of a road section created with Perspective Scene (Helava) with the help of a digital terrain model and the draping of an orthophoto. Black bars introduced every 20 m allow to evaluate the viewing distance.

Fig. 5 ⇨
Distance of visibility on the above road determined with oblique views.
Lower curve : effective visibility determined with Perspective Scene.
Upper curve : visibility derived from the construction elements.
Horizontal curve : tolerance of visibility



were taken with the help of the Linhof hand-held aerial camera; they were used to produce a precise digital terrain model and a digital orthophoto, covering some 10 different images. The orthophoto was then draped over the digital terrain model, which included trees and other obstacles of sight. Then we introduced bars every 20 m on the road and simulated the views for the car driver with the help of Perspective Scene (cf. fig. 4). The bars served to estimate the effective viewing distance, which is partially quite below the 80 m required for these types of roads (fig. 5).

6. CONCLUSIONS

Photogrammetric working methods offer a number of efficient tools in order to produce video sequences showing the impact of new projects in civil engineering or to study specific visibility conditions. The high degree of automation of aerial triangulation should open this technique up to routine applications of filming. The possibility of precisely locating the camera standpoint or its paths enormously facilitates the integration of objects according to their construction elements. Data transfer with the help of DXF and the use of Softimage with Eddie greatly facilitated the generation of combined video sequences.

REFERENCES

Kölbl O., 1976. Metric or Non-Metric Cameras. PE&RS, 42(1), pp. 103-113 and Archives of the XIII Congress ISP, Helsinki, Finland, Comm. V, WG V/2.

Gigon B., 1997. Développement d'une méthodologie basée sur des moyens photogrammétriques pour la prise en compte de la visibilité dans l'analyse de tracés routiers, travail de diplôme EPF-Lausanne.

Von Moos M., 1998. Intégration d'un projet d'ouvrage d'art dans une séquence vidéo, travail de diplôme EPF-Lausanne.

Authors' address :

Prof. Dr Otto Koelbl
EPFL - Photogrammétrie
GR-Ecublens
CH - 1015 Lausanne / Switzerland
Tel. ++41-21-693.27.75
Fax ++41-21-693.57.20
Otto.Koelbl@epfl.ch

Matthias von Moos
EPFL - Photogrammétrie
GR-Ecublens
CH - 1015 Lausanne / Switzerland
Tel. ++41-21-693.37.77
Fax ++41-21-693.57.20
Matthias.VonMoos@epfl.ch

http://dgrwww.epfl.ch/PHOT/index.fr.html