ADVANCED STRUCTURAL DISASTER DAMAGE ASSESSMENT BASED ON AERIAL OBLIQUE VIDEO IMAGERY AND INTEGRATED AUXILIARY DATA SOURCES

N. Kerle[a], R. Stekelenburg[b]

[a] Dept. of Earth Observation Sciences, International Institute of Geo-information Science and Earth Observation, Hengelosestraat 99, P.O.Box 6, 7500 AA Enschede/The Netherlands (kerle@itc.nl)
[b] InnoStack, Ommerbos 51, 7543 GG Enschede/The Netherlands (mail@innostack.com)

Commission VII, TS-PS: WG VII/5

KEY WORDS: Disaster, video, satellite, change detection, integration, texture

ABSTRACT:

Following natural and man-made disasters, comprehensive and reliable information on the nature, extent, and actual consequences of an event is required. Obtaining such information is particularly challenging following unheralded disasters, such as earthquakes or industrial accidents. In those situations, currently operational space-based sensors may not be able to provide timely data. In addition, even high spatial resolution satellites (< 1m) are limited in their capacity to reveal true 3D structural damage at a level of detail necessary for appropriate disaster response.

In this study we investigated the use of oblique aerial video imagery for systematic quantitative damage assessment. Specifically, the following issues were addressed: (i) extraction of individual frames from aerial TV video data, and subsequent enhancement of frame information content using a synthetic aperture approach, and by stacking adjacent frames; (ii) spatial registration of individual frames based on automatically extracted positional information, and (iii) damage analysis based on HIS values and edge elements. We created a working environment that facilitates the video-based damage assessment process and integration with auxiliary data. The limited success of the automatic damage assessment was caused by poor image quality and empirically determined damage threshold values. We provide recommendations for improved image acquisition, and plan further work focusing on generic texture parameters and object geometry.

As an illustration, we used aerial video data acquired in May 2000 over Enschede, the Netherlands, following the explosion of a fireworks factory that severely damaged or destroyed nearly 500 buildings and caused 22 fatalities.

1. INTRODUCTION

1.1 Motivation

Natural and man-made disasters create a need for rapid, comprehensive and reliable information on the nature, extent, and actual consequences of an event. The overall cost of a disaster, both in terms of economic damage and fatalities, depends on how quickly the event is responded to, and how efficiently response activities are managed. Particularly stringent information collection constraints are present following unheralded disasters, such as earthquakes or industrial accidents. The response to virtually every extensive disaster in recent years was delayed by a slow inventory of the event's consequences, in developing countries (e.g. following earthquakes in Koceali [Turkey; 1999], Gujarat [India; 2001], or Bam [Iran; 2003]) as well as developed ones (e.g. following the 1995 Kobe, Japan, earthquake). This is due to a shortage of reliable information coming from the disaster site, difficulties in access, and organizational reasons, such as insufficient preparedness. The utility of geoinformatics in general, and remote sensing in particular, to provide timely and comprehensive information of the post-disaster situation has been repeatedly identified (e.g. Alexander, 1991; Walter, 1994). The reasoning is that a comprehensive damage inventory, prepared at the appropriate (i.e. disaster type-specific) synoptic and detailed scales, provides the prerequisite to direct limited disaster response resources with maximum benefit.

A number of sensors are of potential use to acquire the necessary data, optical spaceborne ones being the first choice due to ease of image interpretation (for example as opposed to radar or laser scanner data), and data acquisition and distribution infrastructure being unaffected by damage on the ground. However, while the potential synoptic coverage of spaceborne systems is certainly an asset, in addition to possible data capture delays due to orbit and pointability restrictions (Kerle and Oppenheimer, 2002), currently operational sensors may not be able to provide a detailed damage assessment in urban areas. Even high spatial resolution sensors (< 1m) are limited in their capacity to reveal true 3D structural damage at a level of detail necessary for appropriate disaster response. This is because the vertical view largely restricts information to building roofs, which may remain intact despite extensive structural damage more readily expressed on building façades. The same is also true for standard aerial photography, which provides the best data in terms of spatial resolution, but which, while typically not being available shortly after a disaster, also tends to suffer from the limitation of vertical viewing.

1.2 The potential of aerial video data

The first data type likely to be available after a disaster in an urban setting is oblique airborne video imagery captured by the news media. The imagery acquired, however, differs substantially from standard air- and spaceborne remote sensing, both in principal purpose and data characteristics. Such video surveys tend to be (i) unplanned in terms of data acquisition scheme, (ii) focus on highly damaged areas at the expense of complete coverage, and (iii) produce uncalibrated data with comparatively low spatial resolution. Commonly used devises used are analogue or digital BetaCams (360 and 720 vertical lines, respectively) or HDTV (High Definition Television) with 1920 columns. All PAL-compatible systems have 576 lines. In addition to the news media, law enforcement agencies are also increasingly using rapid airborne surveys following urban disasters, where also more sophisticated cameras that acquire infrared or thermal imagery are used.

While the acquisition of video data is straightforward, the use of such data poses challenges, in that established image analysis methods can only be used within limits. Reasons for that include the oblique nature of the data, an unstable imaging platform, frequent changes in focal length and, consequently, image scale during data acquisition. In addition they frequently lack camera orientation and location information.

The 1995 Kobe earthquake led to efforts to use non-calibrated data for damage assessment, partially fuelled by delayed overpasses of satellites following the event (Landsat TM: 7 days; JERS: 20 days; Yamazaki, 2001). Improving on previous work by Hasegawa *et al.* (2000b) that only used visual image interpretation, as well as on analysis of multi-temporal satellite data on grounds of impracticality and data unavailability, Hasegawa *et al.* (2000a) explored quantitative analysis methods applied to noncalibrated imagery. Training data of areas showing 3 levels of damage were extracted from individual HDTV frames and used for a classification. Similar methods were applied by Mitomi *et al.* (2000) on earthquakes in Taiwan and Turkey, and by Yamazaki (2001) on the 2001 Gujarat disaster. All studies applied threshold values based on training data to identify damaged areas with varying success, and all were applied to individual video frames only, limiting the practical value of the approach.

1.3 Research aims

A number of issues are addressed in this paper: (i) given the limited resolution and, therefore, detail of standard video data, compounded by further quality loss resulting from data format conversions, we investigated the potential to improve image quality (in terms of higher signal/noise ratio), using a stacking method that incorporates adjacent frames, as well as with a synthetic aperture approach similar to the one described by Gornyi and Latypov (2002); (ii) we applied a damage assessment method to a video stream instead of individual frames based on damage training area characteristics; (iii) we used imagery with encoded GPS data, and relative camera azimuth and inclination information. We extracted this auxiliary information automatically to aid in the mapping of the flightpath as well as spatial registration of the damaged areas. The overall aim of the project reported on here is to provide a video processing environment that combines the above elements, and that allows near-real time processing of video data to detect damage. Results can be displayed related to geocoded pre-event and auxiliary data, in 2 or 3D as applicable.

An important difference to previous work is that our study deals with an industrial accident marked by a radial damage pattern quite different from earthquake damage. Further improvements to our processing environment are planned to make it flexible enough to deal with any type of urban disaster damage. The event addressed in this study is briefly described below.

1.4 The 2000 Enschede fireworks disaster

On 13 May 2000, a series of explosions occurred at a fireworks factory located within a residential area in Enschede, the Netherlands. An area in excess of 40 ha was affected, 22 people killed, and close to 500 buildings severely damaged or destroyed. The force of the explosions led to complete obliteration of buildings close to the site of the factory, while damage severity rapidly declined towards the furthest affected structures, approximately 1 km from ground zero (van Westen and Hofstee, 2001; Figure 1).

Video image acquisition by the police began approximately two hours after the disaster, and was repeated on the following days and augmented by high-resolution vertical aerial photographs 12 days later. These data, however, were not used for the actual damage assessment, which was instead based on ground-based surveys. Disaster response was hindered by outdated map material. Incidentally, a planned aerial survey of Enschede at a scale of 1:18,000 was carried out just 4 hours before the disaster occurred, providing reference data that could have been, but were not, used in the disaster response phase.

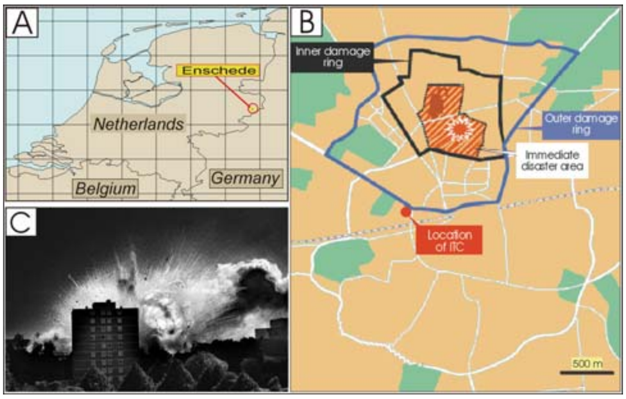

Figure 1. Illustration showing (a) location of Enschede, (b) an overview map of the disaster site, and (c) a photo taken during the explosion of the fireworks factory.

2. DATA AND METHODOLOGY

2.1 Data used

A comprehensive database related to the disaster was compiled, comprising of pre- and post-event, space- and airborne, still and video imagery, in addition to pre- and post-disaster vector data. The principal data sources used for this paper are optical video data acquired by the Dutch National Police Aviation Branch on May 13 and 15, as well as a pre-event Ikonos MS reference imagery acquired on 3 April 2000. Video data were captured with a Sony-HAD camera with 720x576 pixels, as well as an Agema Thermovision 1000 that acquired data between 8-12 µm.

2.2 Methodology

The following principal steps were carried out in this study:

(i) Investigation of frame quality enhancement using AstroStack (www.innostack.com) and a

Synthetic Aperture approach, with quantitative assessment using a the Universal Quality Index (Wang and Bovik, 2002);
(ii) Damage detection in video sequences using empirically derived colour indices, edge characteristics and their variances;
(iii) Automatic extraction of GPS information and camera orientation to map flight path and camera IFOV;
(iv) Mosaicing of videoframes to create a composite view of the disaster area to facilitate orientation.

3. RESULTS

3.1 Frame quality enhancement

The video data acquired, in addition to their inherent comparatively low quality, suffered substantial degradation as a result of a series of conversion steps. The imagery was recorded digitally (720 columns), and transferred to S-VHS tapes (420). These files were later copied to VHS (240) and made available to us. We employed a Sony TRV125 D8 digital video camera for an analogue-digital conversion (back to 720x576 pixels). In total, we obtained 25.5 minutes of coverage recorded on 13 May 2000. Clearly, some of the conversion steps were quite unnecessary, their avoidance likely leading to improved damage assessment results.

AstroStack and a Synthetic Aperture approach

In order to restore some of the lost information and reduce overall noise, an image stacking procedure was performed in AstroStack (www.innostack.com). From a range of adjacent frames a reference frame was chosen, with which the other frames were correlated. Every frames was then shifted in x and y, as well as rotated with respect to the reference frame, to maximise correlation. For this maximisation the Universal Quality Index (UQI, Wang and Bovik, 2002) was calculated for every frame. The UQI also uses a correlation coefficient, in addition to comparing luminance and contrast. The resulting aligned frame series was averaged into a new image with reduced noise.

Gornyi and Latypov (2002) recently described an image enhancement procedure using a synthetic aperture (SA) approach, whereby subpixel-size features were resolved from digital images. The principle involves an image series of an object, which is tracked in sub-pixel increments. Provided that the object itself does not change in-between frames, upon proper alignment of the frames a resolution can be achieved that surpasses that of the recording sensor. Gorny and Latipov's work suggested that, given a number of frames and and subpixel scanning of the subject, a substantial resolution increase can be achieved. To verify their results, and explore the applicability to enhance video data, we first set up a controlled experiment. A picture was produced with lines 1 and 3 pixels wide, with in-between spaces ranging form 1 to 5 pixels (Figure 2). Ten individual images were created with incremental 1-pixel horizontal shifts, played on a computer monitor and recorded with a Sony TRV125 D8 digital video camera. The imagery was processed within AstroStack, and the expected increase in resolution found compared to the individual video frames.

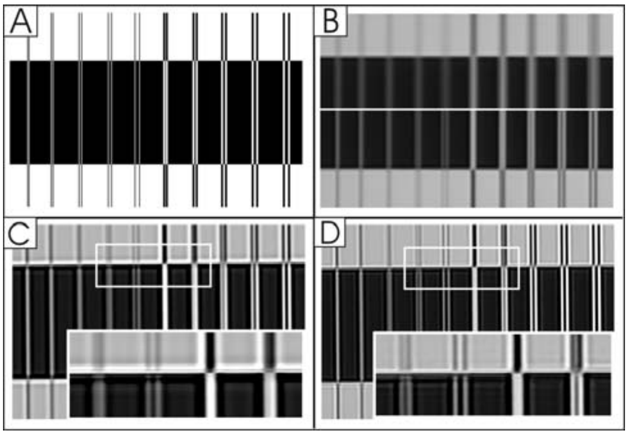

Figure 2. Illustration of synthetic aperture approach to increase the resolution of individual video frames (see text for description and discussion)

A total of 40 video frames (Figure 2b) of the original line array (a) was first simply stacked (top part of [b]), leading to an unfocused image. The lower part of (b) shows the result of stacking after alignment. A point spread function (PSF) derived from the actual lines in the lower part (b) was used for the restoration. The result of the restoration based on (b) is shown in (c). A much clearer restoration is shown in (d), for which the stacked image of (b) was doubled in size before alignment. The results show that details can be extracted that are not resolved in the original video frames.

The SA approach was then applied to the police video data. However, the increase in detail observed in the controlled line experiment was not found. The likely reason for this is the accumulated video quality loss resulting from the aforementioned conversions. Especially the VHS conversion has led to smeared out details and line instability. The conversion to digital also introduced jpg-like artefacts, leading to further noise and reduction in dynamic range. Furthermore, individual elements in the video data contrast much less than the lines in the theoretical experiment. Such a reduction in modulation, however, increases the space between features that can be resolved.

Although with a direct transfer of the original digital data to the computer the need for such resolution enhancements decreases, we expect the SA approach to be useful with higher quality data.

3.2 Automatic damage detection

The actual image processing to detect damage was carried out in a flexible processing environment created by Innostack. The software works with processing blocks that can be connected as required. The graphical user interface (GUI; Figure 3) displays the input images or video, the processed equivalent, as well as command prompt and history list. The GUI is customisable to allow easier execution of pre-defined routines. Our goal is to provide a working environment where post-disaster video data can be processed, spatially registered, and displayed together with co-registered pre-event images or map data as required.

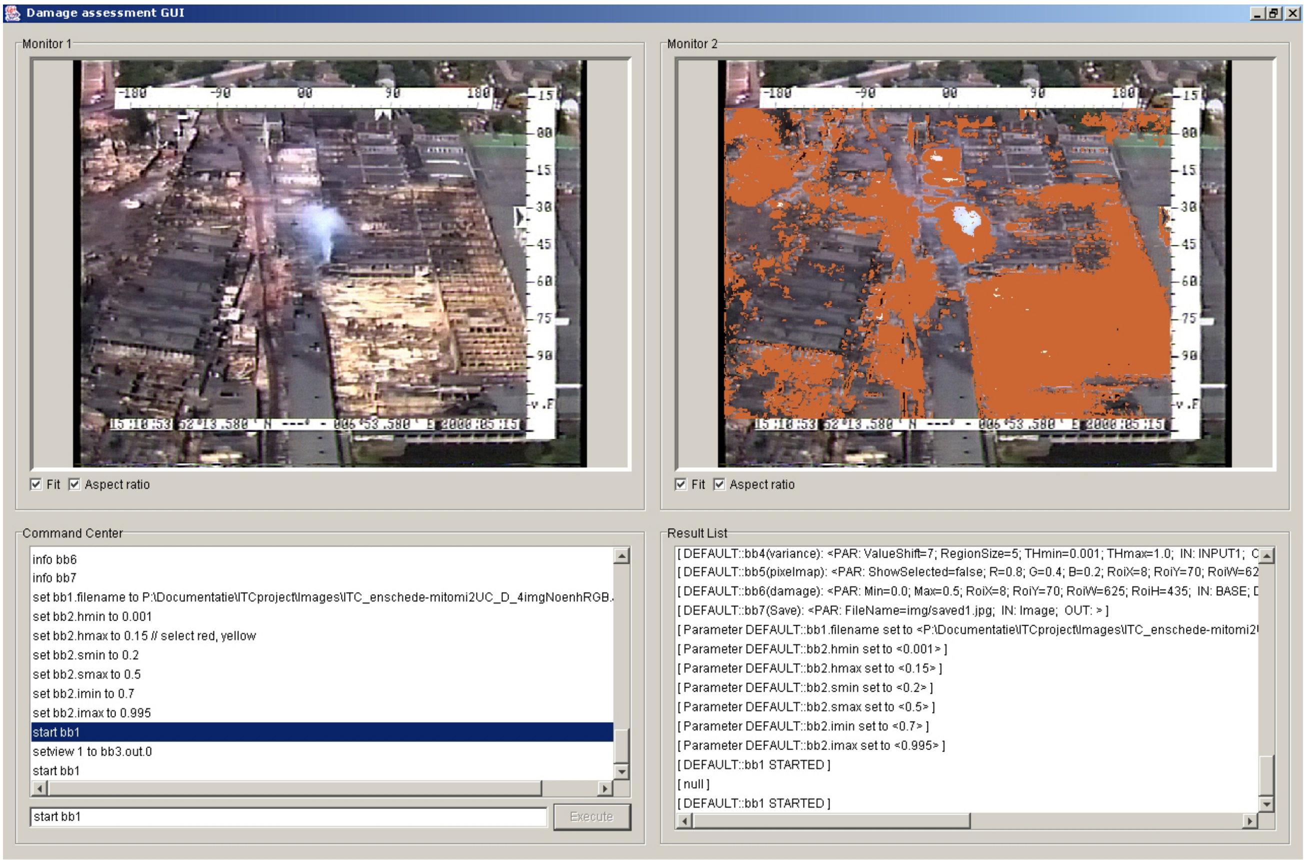

Figure 3. Screen capture of the processing environment for the video processing, with input video on the left, and damaged areas coded in orange on the right

Similar to the approach taken by Mitomi *et al.* (2000), we identified representative damage areas and calculated hue, intensity and saturation, as well as edge and variance values. These values were then applied on a frame-by-frame basis to a video stream within the developed working environment. We limited damage assessment to the area within the horizontal and vertical information bars (625x435 pixels). Table 1 gives an overview of the value ranges used.

Parameter	Threshold range (0.0 – 1.0)
Hue	0.001 – 0.15
Saturation	0.2 – 0.5
Intensity	0.7 – 0.995
Edge (3x3)	0.0 – 0.22
Variance (7x7)	0.016 – 0.11

Table 1. Threshold values used for the damage assessment

From the calculated means and standards deviations of the training areas it became clear that colour bleeding and format-conversions had reduced overall image contrast as well as edge clarity. In particular also the conversion to low-resolution VHS eliminated significant detail. We found that no threshold combination of different parameters was able to detect damage satisfactorily. This approach, however, performed well in a similar study of the 1999 Kocaeli (Turkey) earthquake (Ozisik and Kerle, 2004), which was also implemented in the processing environment presented here. This is likely because the resolution of the input data was higher (720x576), but also because of damage patterns (mostly rubble piles) that were easier to distinguish from non-damaged areas. Applied to the Enschede case data and using low-resolution data, considering also unidentified damage and incorrectly identified damage, only accuracies of 40% were possible using the parameters described above.

3.3 Extraction of spatial information

Spatial data are only useful of they can be spatially referenced. A particular challenge of video data is their erratic nature caused by frequent cuts, and focal length (zooming) and inclination changes. Use of the resulting data, therefore, requires extensive local knowledge. The camera used by the Dutch National Police recorded horizontal GPS information, as well as data, time, and relative camera azimuth and inclination (see Figures 3 and 5 (d)). No data were encoded on flying height and camera zoom factor.

A routine was implemented in the processing environment to extract the available information. Given that the position of the information in the video frame is fixed, three areas of interest were defined, for the date/time/GPS information, and the azimuth and inclination arrows. A total of 12 correlation blocks was then defined (for numbers 0-9, and for the arrows). Given the large number of frames (15 per second), only every 15th frame was processed, to obtain one measurement per second. The azimuth and inclination values range from –180 to 180, and 15 to –120, respectively. The x, y location of the arrows in the processed frames was converted to azimuth and inclination by linear interpolation. The individual correlation blocks were assembled into strings, whereby some elements (e.g. date and parts of coordinates) are unchanged for the whole video and therefore fixed. The following table gives and overview of the initial values and a typical readout (Table 2).

Initial values	Example output string
latitude "0006 ..,...."	Latitude 0006 54.900
longitude "52 ..,..."	longitude 52 13.000
angleH …,..	angleH -14.23
angleV …,..	angleV -33.97
time "1.h..m..s"	time 17h18m23s
date "2000:05:1."	date 2000:05:13

Table 2. Example result of frame information decoding

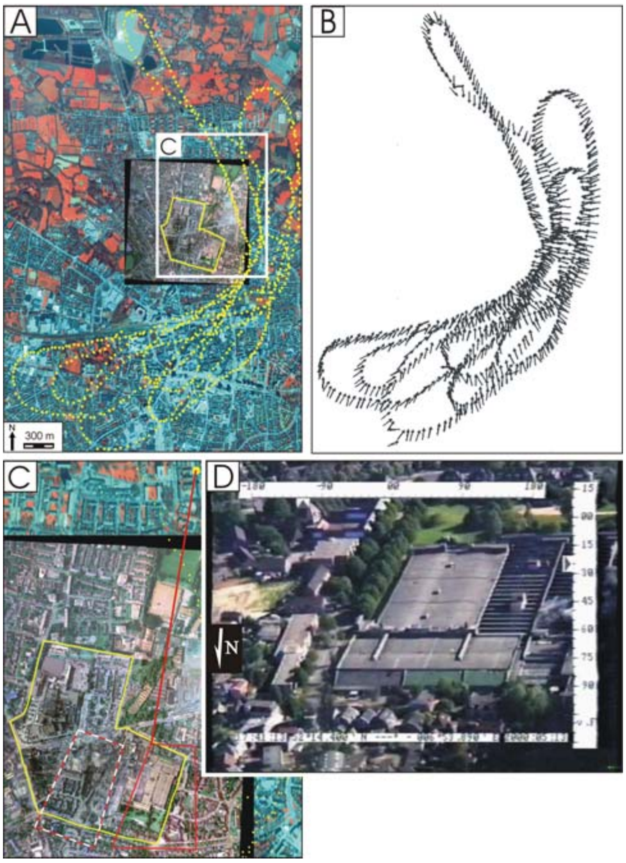

Figure 4. (a) Pre-disaster Ikonos image overlaid with post-event aerial photographs, and with helicopter flightpath of 13 May plotted (yellow dots) based on information extracted from the video data. The immediate disaster area is indicated in yellow. (b) Camera viewing directions calculated from the GPS auxiliary data. These data can also be used to plot the footprints of individual frames, as illustrated in (c). The solid red box shows the approximate location of the frame shown in (d). The box in hatched red is the result of the automatic calculation, based on absolute azimuth, camera inclination, and helicopter location. Focal length and flying height were simulated. The apparent positional difference between the two footprints results from uncertainty in the absolute camera azimuth, as the helicopter orientation is not necessarily identical to the flight direction, further illustrating the need for IMU information.

A problem for the correlation was the low video quality, in particular the horizontal instability between lines, a result of interlacing of two frames carrying half the information each (odd vs even lines). Therefore we first deinterlaced the frames (combining the two half-frames into one), and sharpened the result. The subsequent correlation and string processing then took approximately 4 sec per frame.

The resulting table was then further processed to convert the geographic coordinates into UTM to make them fit our reference imagery, as well as to calculate the absolute flight vector between frames and absolute camera azimuth. We consider the inclination angle to be absolute, although it is dependent on the roll, yaw and pitch of the helicopter. The effect on the IFOV of the camera can be substantial, and should be corrected for with Inertial Measurement Unit (IMU) information.

3.4 Video mosaicing

The erratic nature of video imagery detailed above complicates its use. Unlike with vertical aerial photographs and satellite imagery a simple geocoding is not possible. However, given the value for overview and orientation purposes of such a mosaic, we used RavenView (www.observera.com) to assemble a mosaic of the disaster site based on the police video data (Figure 5). The software also allows a geocoding, although for that a sensor model of the camera used is required.

Figure 5. Mosaic of the Enschede disaster site, comprising 227 video frames (red line approximates outline in Figure 1 b)

4. CONCLUSIONS AND DISCUSSIONS

In this project we investigated the utility of oblique airborne video data for urban post-disaster damage assessment. The main objectives were to enhance the overall quality (lower signal/noise ratio) of the imagery, and detect damage based on hue, intensity and saturation (HIS), as well as edge and variance characteristics in a specially created processing environment. We furthermore explored possibilities to register video data spatially based on encoded GPS information.

The data quality enhancement carried out in AstroStack, based on aligning and stacking, led to a measurable improvement. The SA approach described by Gorny and Latypov (2002) was also verified in a theoretical experiment, but the method did not improve the quality of the video data. This was likely a result of several data conversion steps that led to severe image degradation and colour bleeding, but also of low contrast that is known to lead to reduced separability of small features. The low data quality also led to poor results of the automated damage assessment. No combination of HIS, edge and variance characteristics was able to detect all damage, further hindered by the nature of the damage, which included burnt buildings as well as structures removed by the blast. Nevertheless, we created a working environment suitable for the processing of video streams, and we are planning further work to incorporate more generic damage indicators.

The software was also used to extract GPS and camera orientation information encoded in the imagery, which was used to register the video frames, viewing orientation and the IFOV automatically.

Video data, typically acquired by the news media after a disaster, but increasingly also by law enforcement agencies, are suitable for a rapid post-event inventory of urban damage. The principal advantages over satellite data are timely availability and versatility, i.e. the possibility to obtain optical, infrared or thermal data at required scales and viewing angles. However, our work has shown that, for the data to be of use, the following requirements should be met: (i) as much as possible/practical, image acquisition should cover all affected areas and follow some cohesive flight pattern; (ii) cameras encoding auxiliary information are most useful. In addition to the information available in our imagery, the flying height and zoom factor are also required. If possible, information on helicopter/plane attitude, using an IMU system, should also be collected; (iii) lastly, best results can be achieved if data are recorded digitally at a high resolution (at least S-VHS [720x576], but preferably using HDTV), and transferred directly to the computer.

5. REFERENCES

Alexander, D., 1991. Information technology in real-time for monitoring and managing natural disasters. *Progress in Physical Geography*, 15(3), pp. 238-260.

Gornyi, V.I. and Latypov, I.S., 2002. Possibility for the development of a scanning radiometer with synthesized aperture: Experimental verification. *Doklady Earth Sciences*, 387(8), pp. 962-964.

Hasegawa, H., Aoki, H., Yamazaki, F., Matsuoka, M. and Sekimoto, I., 2000a. Automated detection of damaged buildings using aerial HDTV images, International Geoscience and Remote Sensing Symposium, Honolulu, USA, IEEE, CD-ROM.

Hasegawa, H., Yamazaki, F., Matsuoka, M. and Sekimoto, I., 2000b. Extraction of building damage due to earthquakes using aerial television images, 12th world conference on earthquake engineering, Auckland, New Zealand, No. 1722: CD-ROM.

Kerle, N. and Oppenheimer, C., 2002. Satellite remote sensing as a tool in lahar disaster management. *Disasters*, 26(2), pp. 140-160.

Mitomi, H., Yamzaki, F. and Matsuoka, M., 2000. Automated detection of building damage due to recent earthquakes using aerial television images, Asian Conference on Remote Sensing, Taipei, Taiwan.

Ozisik, D. and Kerle, N., 2004. Post-earthquake damage assessment using satellite and airborne data in the case of the 1999 Kocaeli earthquake, Turkey, XXth ISPRS Congress, Istanbul, Turkey.

van Westen, C.J. and Hofstee, P., 2001. The role of remote sensing and GIS in risk mapping and damange assessment for disaster in urban areas, Second forum catastrophy mitigation : natural disasters, impact, mitigation, tools, Leipzig, Germany, pp. 442-450.

Walter, L.S., 1994. Natural hazard assessment and mitigation from space: the potential of remote sensing to meet operational requirements, Natural hazard assessment and mitigation: the unique role of remote sensing, Royal Society, London, UK, pp. 7-12.

Wang, Z. and Bovik, A.C., 2002. A universal image quality index. *IEEE Signal Processing Letters*, 9(3), pp. 81-84.

Yamazaki, F., 2001. Applications of remote sensing and GIS for damage assessment, 8th International Conference on Structural Safety and Reliability, Newport Beach, USA.

Acknowledgements

Video data of the disaster site were made available by TV Oost and the Enschede Police’s TOL-Team. We particularly thank Hans Kampermann for his support, and providing additional aerial photographs. We further acknowledge the help of Karl Grabmaier in the photogrammetric processing.