Photos and video clips in disaster management

Runde, Detlef; Meiers, Thomas; Berger, Simon

Fraunhofer Heinrich-Hertz-Institute, Einsteinufer 37, 10587 Berlin, Germany
detlef.runde@hhi.fraunhofer.de, thomas.meiers@hhi.fraunhofer.de,
simon.berger@hhi.fraunhofer.de

Abstract: This article deals with the effective support of emergency services and control centres through photos and short video clips in crisis and disaster management. It describes how images are captured, pre-processed, appropriately filtered and grouped into meaningful panoramas, motif groups and event progressions, so that an informative survey of the situation on site is available instead of a confusing flood of images.

Keywords: disaster management, image processing, decision support system,

1. Introduction

A still widespread problem in disaster management is the availability of up-to-date and reliable information about the on-site situation. In many cases, the control centres receive their information from the emergency services on the ground by telephone. In the research project sd-kama (Smart Data Catastrophe Management, funded by the German Federal Ministry of Economics and Energy as part of the technology program "Smart Data - Innovations from Data"), an information platform is now being developed (and implemented prototypically for the Cologne Flood Protection Centre) in which a wide range of data from different sources is made available in real time to emergency staff and disaster managers [1]. Some of these data sources, photos and video clips, will be examined in detail below.

2. Why not just using data from Facebook, Twitter and Co?

The first thought when you think of using photos is certainly to use social media, since many people use them anyway. However, if you talk to the experts from the emergency services and control centres, it quickly becomes apparent that they are sceptical about this. Social media posts, apart from selfies, are often especially spectacular pictures in front of operational areas, but not the photos that would really help the emergency services: where do problems arise, which areas could be endangered in the near future, and what is the state of work, ... Therefore, these data from social media are not regarded as a main source of information. Even in a best-case scenario, they will complement them only partially.

The currently common information transfer concerning the situation on the ground by telephone is often cumbersome, detailed descriptions are time-consuming and it is unclear whether the sender and addressee interpret the information in the same way, e.g. due to previous knowledge. Here a simple, easy-to-use app for informing the control centre, preferably using photos and short video clips, is supposed to provide a remedy; true to the well-known saying, "a picture is worth a thousand words". A user-centred design was used when developing the app, i.e. design and functional scope were defined and evaluated together with the end-user (here: the municipal drainage company of Cologne).

The sd-kama media app has two target groups, the emergency crews (experts) and the general public (everyone, laypeople, "crowd"). Since the expected information from both target groups is likely to be of different quality, the functional range of the app also differs: The categorisation of photos and video clips as well as the sending of text messages is available for registered users, the experts, only.

3. Data acquisition

The app is implemented as an HTML5 web application [2]. It works both via browser and as a stand-alone web app and therefore independently of the operating system on all mobile and stationary devices.

In the event of a disaster, further information is often important in addition to the photos and video clips: This includes the geolocation and time of the recording, an optional categorization (e.g. *information*, *problem*, *emergency*) and the opportunity to add a short note to the media file if it is not self-explanatory. With the help of this additional metadata, it is possible to pre-process the photos and video clips adequately to ensure a helpful visualization within the decision support system at the control centre.

The media app is implemented in compliance with data protection regulations and therefore its use is anonymous by default (privacy-by-default). Before uploading a photo or video clip for the first time, the user is informed in an easy to understand manner, which data will be collected and processed for which purpose. Since the target group are people who want to help, it is assumed that the vast majority of users will agree to the acquisition of their geolocation. Additional options provide the user with incentives to register (also possible by pseudonym). A registration is mandatory for the emergency services.

4. Automated data preparation

After the media files have been uploaded, they need to be processed, on the one hand for legal reasons, and on the other hand to provide the emergency services with the best possible information. Since large amounts of visual data can arise, the main purpose of the analysis described below is to prepare this data in such a way that the task forces can quickly find and view the relevant images and videos. This requires both filtering and generation of genuine added value by a smart combination of different, but contextually related images and video clips.

4.1. Consideration of data protection

Depending on the motifs, the uploaded photo and video footage may contain personal or person related data, e.g. when faces are clearly recognizable. Therefore, the uploaded photos are automatically processed by a face recognition software to pixelate them, even though it has to be accepted that this method is not 100% reliable. All data collected within the scope of disaster management is exclusively intended for the evaluation in regards to hazard prevention and for supporting decisions of the emergency services and the control centre. Photos and video clips will not be published and are only available to a limited group of people within the control centre.

4.2. Quality of data

In order to support the decision-makers in the operations control centre judging the data quality, they are informed about each photo/video clip, whether it comes from an expert or from the "crowd". In addition, the Exif data is evaluated and hints that point to a post-processing of the image are transmitted, together with the image information, to the control centre's decision support system. These include:

- Comparison of recording and digitizing time
- Comparison of GPS data from upload and capturing geolocation
- Comparison of the times of recording and upload
- Checking the Exif item of the image processing software (camera software versus image manipulation software)

4.3. Filtering

In case of an emergency, it may happen that huge amounts of photos and video clips will be uploaded. Then, the challenge is to inform the staff at the operations centres about what is happening on the ground by presenting only the relevant media to support them in making the right decisions, i.e. it is the aim to avoid a flood of photos and video clips.

This is achieved by means of methods that classify images and videos directly after uploading, filter them according to specified criteria and then sort and present them accordingly. In addition to geolocation and time information, quality features and content criteria can also be used as selection criteria.

The quality of a photography is characterized by measuring **contrast** and **sharpness**. If several shots show the same subject, the higher-quality shots will be prioritized.

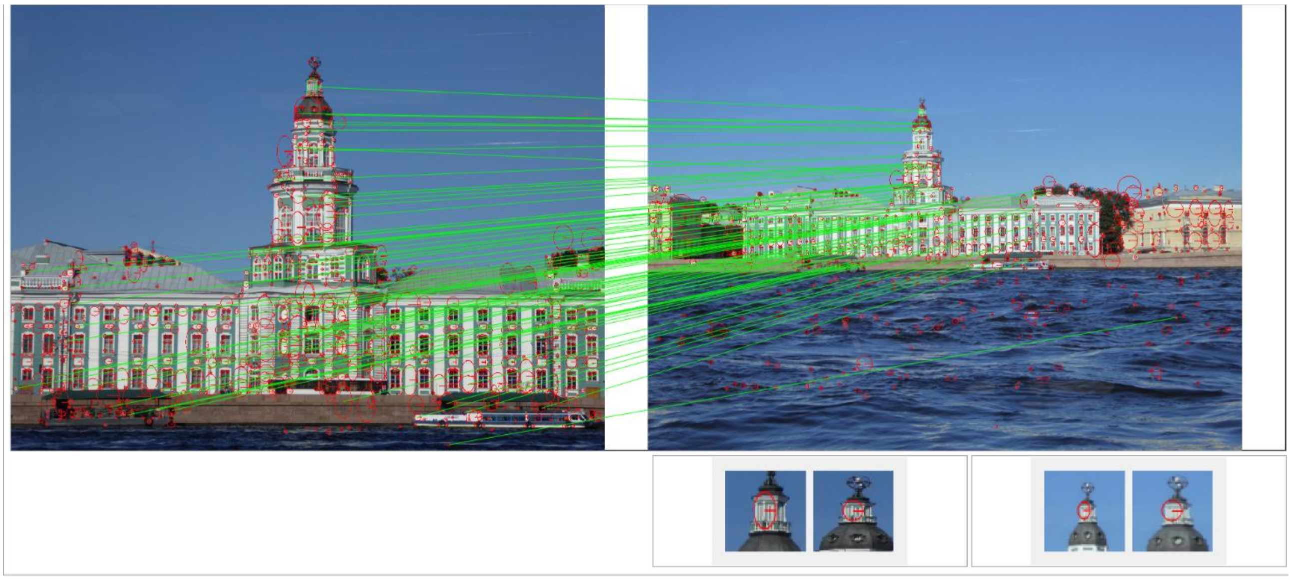

Figure 1: Red ellipses indicate visual features. The four small images at the bottom right show details around the visual features, which are corrected with regard to scaling, rotation and shear in such a way that the clipping in the right and left image becomes comparable. A local descriptor is applied to these corrected clippings, which describes the visual information of the section. Local descriptors can be compared with each other regarding the similarity of the image content. The green lines between the two images indicate correspondence of similar descriptors that are sufficient for a clear projective transformation.

A very important sorting tool is the grouping of images and videos based on their visual content. The extraction of visual characteristics and the determination of suitable descriptors form the basis of this analysis. For the visual features, so-called blobs are used, i.e. image areas in which neighbouring pixels have the same or similar grey values. Their advantage is to be scale-invariant. In a first step, all blobs are determined. Then, according to the scaling, stretching and rotation, a rectangular image area is determined around each blob found, to which a local descriptor is applied (see Figure 1). The local descriptor contains information regarding both the colour and the distribution of the edge- and texture orientation within the image area. If the two image areas are similar, the corresponding local descriptors have a small distance to each other. Comparing two images with each other in terms of visual content means to compare all local descriptors of the first image with those of the second image. If the distance of a descriptor pair is below a threshold value (= high similarity) a correlation is formed. The number of correlated descriptors describes the **degree of similarity** between two images.

This similarity measurement software groups the incoming images on-the-fly according to their motifs. In order to keep the number of image comparisons manageable, comparisons are made only between images of the same (or similar) geolocation. Consequently, almost identical images, which were taken at a relatively short time interval, do not have to be transferred to the decision support system for visualisation.

While these methods described above serve to organize and reduce the flood of information in a reasonable way, the following measures enhance information by creating added value using image processing.

4.4. Generation of value-added information

Even if similar images–as described above–are not presented in the decision support system, they are nevertheless evaluated to determine whether they are suitable for being combined with other images to form a panorama, a motif group or an event progression. Here, it must be ensured that the images remain unaltered in terms of their content: A smooth and visually appealing transition between two combined images may look beautiful, but could result in important image details no longer being visible, an undesirable effect in disaster management. Therefore, this kind of image processing is deliberately avoided. In order to include image information provided as video clips, a temporal video segmentation is carried out (see [3] for a detailed description of the procedure). Representative snapshots of the detected scenes are extracted and used for the above-mentioned value-added information generation as well.

In order to build a panorama, a motif group or an event progression, it is necessary to determine from which camera viewpoint the motifs of two images were taken. In accordance with the principals of projective geometry, points in a plane, taken from different viewpoints, are mapped onto each other by means of a projective transformation. From the matching descriptor pairs of two images (see above), those have to be detected, which can be transformed into each other by a projective transformation . If there are several layer (e.g. house walls) within two images, several projective transformations can be created, from which the predominant one with the most correspondences is selected. A projective transformation can be split into four components: translation, rotation, scaling and shearing.

In principle, you can imagine a wide variety of shooting situations consisting of the scenarios illustrated in Figure 2:

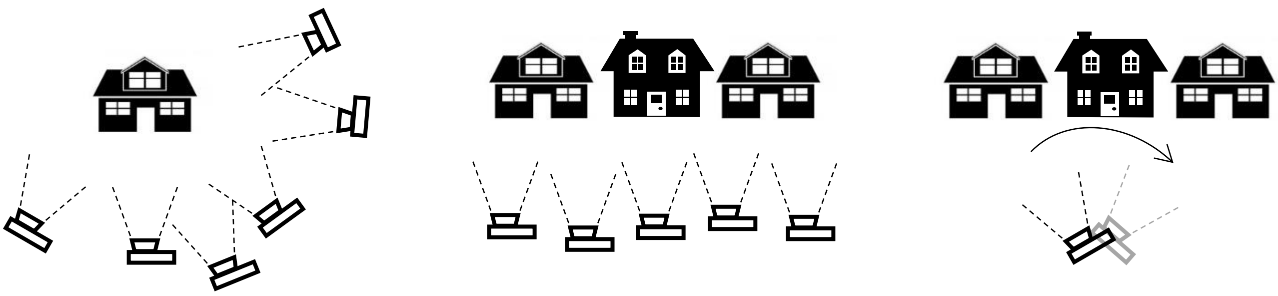

Figure 2: scenarios for constructing panoramas

On the left, an object is taken from different viewpoints. The main part of the projective transformation between the images is a shear. It is possible to arrange the images according to their projective transformations in such a way that you can walk around the object. In the middle and right scenario, several shots are taken along a line in front of the motif (middle) or by panning the camera (left). In both scenarios, we have to deal mainly with a translation. Only little other transformations (rotation, scaling, shearing) are allowed to be present here, as otherwise no continuous rectangular panorama image can be generated.

Images grouped according to the middle and right scenarios shown in Figure 2 can be combined to form a panoramic image. At the operations control centre it is easier to gain an overview with such a **panorama** than with several individual images.

Images from all three scenarios can be grouped into motif groups and sorted in a further step as similarity graphs. The nodes of a graph consist of images and the edges indicate the projective representation between the images. The visualization is carried out by means of a specially developed **navigation viewer**, which allows a spatial navigation within the similarity graph. Figure 3 shows this navigation viewer. Yellow symbols indicate all the images connected by a projective transformation to a source image. By moving the mouse cursor over a symbol, the associated image is presented on top of the original image. With a mouse click, the connected image becomes the new source image. This allows you to navigate through the similarity graph.

Figure 3: Navigation viewer (left: initial image with symbols that draw attention to further images connected by a projective transformation; right: visualization of a connected image when the mouse cursor touches a yellow symbol. A mouse click leads you to the new picture.)

Images assembled into a time-lapse provide a quick overview of the chronological order of damaging events (e.g. the increase in water levels). For this purpose, all images of the same motif from a similar viewpoint, which differ only in terms of low shear, scaling, rotation and translation, are brought into a chronological order and then made available as **time-lapse video**. However, initial experiences showed that this kind of information presentation is too rigid and inflexible. That is why we developed a flexible **event progression viewer** (see Figure 4). The following options are available now:

- Automatic "slideshow" in chronological order (this corresponds roughly to a time-lapse video), with the option to pause/continue it or to navigate forward and backward picture by picture.
- Setting the display duration of the images within the slideshow.
- Select whether to display complete images or only the image area that is present in all images.
- Select specific images using thumbnails, which are visible when positioning the mouse cursor at the time scale.

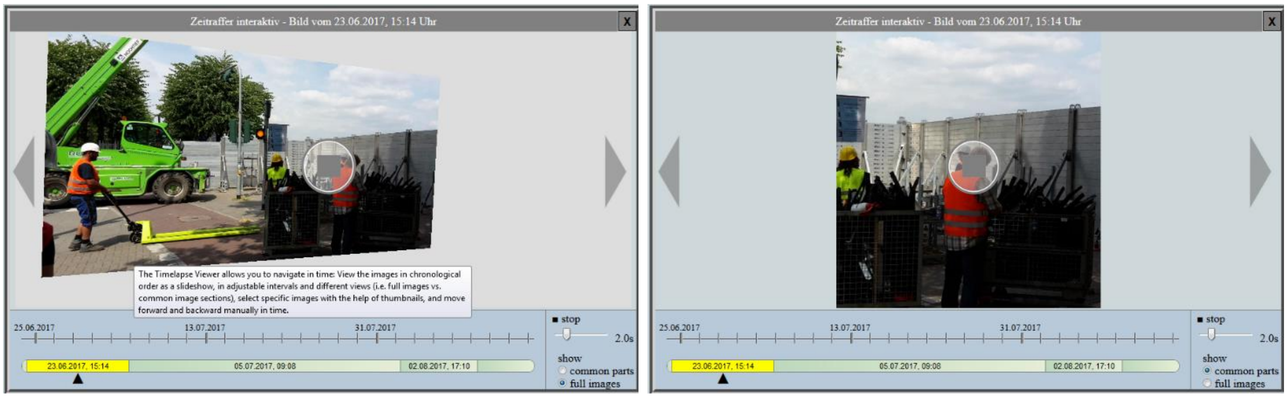

Figure 4: interactive time-lapse (right: presentation of full images, left: presentation of common image parts)

Both described viewers (event progression viewer and navigation viewer) are web-based and integrable into any web interface and web-based decision support system.

All described image group evaluations are highly dynamic, since they can change with each new media file that is added. Image groups can become larger or two separate groups can grow together to one.

5. Transfer to the decision support system at the operations control centre

After the automatic processing of the media (photos and video clips), the decision support system of the operations control centre must be informed about the available information. For this purpose, a JSON data exchange format, standardized within the sd-kama project, is used. The used interface is REST and communication is secured via HTTPS. Only the metadata is transmitted, the media files themselves remain on the original server and are retrieved from there if required. This has the advantage that–especially in the case of the panorama, motif group and event progression formats–the latest presentations are always retrieved.

The system described above, from recording and transmission to automatic image and video processing, is independent of a specific implementation of an information or decision support system. It is only connected via a simple interface protocol. During the project's lifetime, this had the advantage that the decision support system and media evaluation could be developed in parallel, so that results could be quickly presented. After completion of the project, it offers the advantage that the decision support system can also be used without a media module on customer request, but that the media module can also be used in other areas of application without extensive development effort (it only requires a simple interface adaptation). In addition to the media app, it is also conceivable to integrate other media sources, such as drones or permanently or temporarily installed webcams. The media app is currently available in two languages (German/English) and can be upgraded to other languages.

Glossary

Exif	Exchangeable Image File Format – additional image header information
HTTPS	secure Hypertext Transfer Protocol
GPS	Global Positioning System
REST	Representational State Transfer
sd-kama	SmartData – Katastophenmanagement, a research project funded by the German Federal Ministry of Economics and Energy (BMWi) within the framework of the technology program "Smart Data - Innovations from data".

References

[1] sd-kama project-website: https://www.sd-kama.de/ (accessed 28.07.2017)

[2] sd-kama media-app: https://sdkama.hhi.fraunhofer.de/ (accessed 28.07.2017)

[3] Petersohn, C.: Temporal Video Segmentation, Jörg Vogt Verlag, 2010, ISBN 3938860391, 9783938860397, 2010