Innovative sensors for crowdsourced river measurements collection

Krommyda M.[1,*], Tsiakos V.[1], Rigos A.[1], Tsertou A.[1], Amditis A.[1], Georgiou H.[2], Jonoski A.[3], Popescu I.[3], Assumpcao T.[3]

1Institute of Communication and Computer Systems 2 Hellenic Rescue Team of Attica 3 IHE Delft Institute for Water Education

*corresponding author: e-mail: maria.krommyda@iccs.gr

Abstract

Flood risk prediction requires consistent and accurate sensor measurements, usually provided from traditional in-situ environmental monitoring systems. Crowd-sourced data can complement these official data sources, allowing authorities to improve and fill gaps in the hazard assessment process. However, collecting this information from volunteers, with no technical knowledge and while using low-cost equipment such their smartphones and tablets, raises the question of quality and consistency. To alleviate this barrier two tools were developed in the context of H2020 Scent project (grant agreement No. 688930). The Water Level Measurement Tool uses image recognition techniques to extract the water level from images containing a measuring tape. The Water Velocity Calculation Tool uses video processing algorithms to extract the water surface velocity from a video containing a pre-defined floating object moving on the surface of a water body. Each extracted measurement is accompanied by a degree of trust. The tools have been designed so that a high degree of trust can be achieved from images and videos taken from regular smartphones. The crowdsourced river measurements are used to develop improved flood models with a dramatically reduced cost as both the measuring tapes and the floating object are low-cost and re-usable while effectively covering large areas of interest.

Keywords: Flood modeling, innovative sensors, crowdsource data, campaign organization

1. River Measurement Collection

1.1 In-situ environmental monitoring systems

Flood risk prediction has been traditionally based on models that are developed from time-series of data collected over long periods of time from expensive and hard to maintain in-situ sensors available in the case study areas. Such in-situ monitoring systems are available only in limited areas that had been affected by extreme phenomena and require continuous monitoring. The climate change has made the monitoring of the flood events imperative and has raised the question of whether the development of flood models can be disengaged from the in-situ sensors to cover larger areas of interest.

1.2 Crowdsourced data

Due to the climate change and the extreme weather phenomena, more people are affected by environmental issues and become aware of the need to monitor them. As a result, they are willing to offer their time to support the collection of scientific data. The utilization of crowdsourced data comes with the challenge of data quality. The volunteers differ in knowledge and experience in the collection of data, especially when they are required to follow strict guidelines to ensure the accuracy and consistency needed for scientific purposes. The flood models require precise measurements as these will determine the trustworthiness of their results. Aiming to bridge this gap, two tools that extract river measurements from multimedia have been developed.

2. River Measurement Extraction from Multimedia

2.1. Water Level Measurement Tool

Volunteers are asked to capture images of water level measures that have been placed in areas of interest. The images are made available to the Water Level Measurement Tool (WLMT). Then, each image is processed to extract the numbers that it contains. The algorithm examines the location of the identified numbers starting from the lowest part of the image. The first pair of number to be identified is the water level measurement extraction. The identification process is relying on a pre-trained Cascade Classifier, which was trained with a series of images taken at the installed water level indicators at the Kifisos river (Attica, Greece).

2.2 Water Velocity Calculation Tool

Volunters are asked to capture video of a pre-defined floating object (a tennis ball) as it passes infront of them following the river cource. The collected video is first stabilized to ensure that any intentional or unintentional movement of the camera is eliminated as much as possible, removing the associated noise. After the video stabilization the video-frames are examined one-by-one till the tennis ball is located. After that two consecutive frames are examined and the object displacement is calculated. The displacement is calculated by computing the average optical flow of the area of the frame that includes the tennis ball. The average optical flow of the rest of frame is substrated from this calculation to eliminate any camera movement that is still present in the video frames. The process is repeated till a frame without the pre-defined floating object is reached; when such frame is located this is considered the end of the video. The displacement, which has been calculated in pixels is then calibrated using the dimension of the pre-defined object. A tennis ball is 6cm in diameter.

In case that the pre-defined object is not present in the video or it is presented in less frames than the frames per second of the video, giving less than a second of useful material, the video is classified as invalid and no measurement is extracted. In the case of a valid video the water velocity is extracted along with a confidence level. The confidence level is calculated based on the resolution of the video, the duration of the valid part of the video as well as the overall stability of the video.

3. Challenges

3.1. Organizing the Campaign

The Scent project incorporates two pilot-demos for demonstrating the concept of Citizen Observatories (CO) in practice and for demonstrating the use of the applications developed. Participants range in age, professional expertise and familiarity of collecting scientific data. The two pilots have been planned for two distinct use cases and appropriate geographical regions the rural area of the Danube Delta in Romania and the urban area of the Kifisos river basin in Attica, Greece. In both cases, groups of citizens were required to take images and video using the *Scent Explore app* using their mobile phones.

There are several challenges in organizing and running such large-scale open field activities that include non-expert people and span over several months. Such issues include the prompt dissemination of the event, informing the general public and engaging them to get involved and participate, organizing groups per day & site, selecting routes to visit, providing a short first-day training and application installation workshop, minimizing transportation costs and delays as well as guiding them throughout the activity. Regarding safety, Hellenic Rescue Team of Attica (HRTA) was the primary coordinator for ensuring basic guidelines are followed by all during the field activities, as well as first-response emergency medicine provider, if necessary. Some activities were a bit challenging, as the volunteers were asked to go into the river bank and capture pictures and videos for river measurements. In all cases, HRTA teams were escorting and leading the citizen groups, assisting in the proper setup and use of the Scent applications, as well as (primarily) ensuring the safety of everyone in the field.

3.2. Evaluating the data collected

In order to evaluate the quality of the velocity as extracted by the WVCT we had to locate measurements that were collected by people that were at the same location approximately the same time. So, the collected data had to be separated into groups. For this separation, the K-means clustering algorithm [Lloyd, Rigos] was applied and the data were spitted into K = 12 clusters according to their location. The four of these 12 clusters were re-clustered afterwards according to their timestamps, because they contained data collected on different time periods, so finally the collected data were separated into 16 clusters. In figure 1 the spatial distribution of the data along with the clustering centers is presented. Figure 2 presents the box plots that show the range of the collected data. In some cases, there are clear outliers; these were examined further to identify the causes. The main cause was identified as the camera stabilization and the noise in the captured frames.

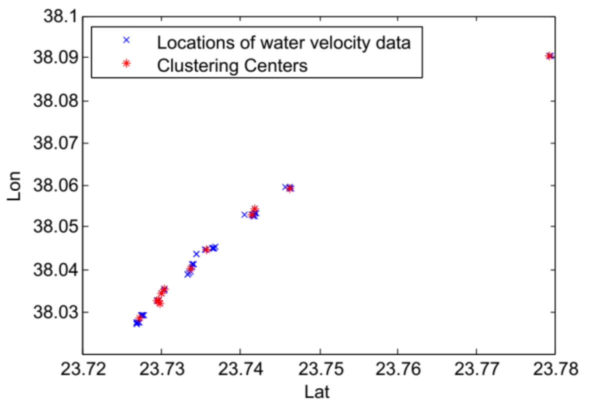

Figure 1. Spatial distribution of collected data

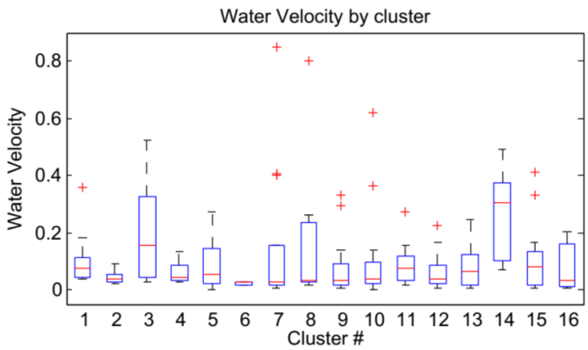

Figure 1. Box plots representing the range of the data

The analysis for the WLMT was been contacted in a similar way, but it is omitted here due to lack of space.

References

S. Lloyd, "Least squares quantization in PCM," IEEE Transactions on Information Theory, vol. 28, no. 2, pp. 129–137, Mar. 1982.

A. Rigos, "Development of polynomial neural networks using orthogonal polynomials," PhD Thesis, University of Aegean, 2017.