Stereo vision for fully automatic volumetric flow measurement in urban drainage structures

Ekaterina Sirazitdinova, Igor Pesic, Patrick Schwehn, Hyuk Song, Matthias Satzger, et al.

Ekaterina Sirazitdinova, Igor Pesic, Patrick Schwehn, Hyuk Song, Matthias Satzger, Dorothea Weingärtner, Marcus Sattler, Thomas M. Deserno, "Stereo vision for fully automatic volumetric flow measurement in urban drainage structures," Proc. SPIE 10332, Videometrics, Range Imaging, and Applications XIV, 103320M (26 June 2017); doi: 10.1117/12.2270233

SPIE.

Event: SPIE Optical Metrology, 2017, Munich, Germany

Stereo Vision for Fully Automatic Volumetric Flow Measurement in Urban Drainage Structures

Ekaterina Sirazitdinova[a,*], Igor Pesic[a], Patrick Schwehn[a], Hyuk Song[a], Matthias Satzger[b], Dorothea Weingärtner[c], Marcus Sattler[b], and Thomas M. Deserno[a]

[a]Department of Medical Informatics, Uniklinik RWTH Aachen, Germany
[b]Seba Hydrometrie, Kaufbeuren, Germany
[c]Research Institute for Water and Waste Management, RWTH Aachen University, Germany

ABSTRACT

Overflows in urban drainage structures, or sewers, must be prevented on time to avoid their undesirable consequences. An effective monitoring system able to measure volumetric flow in sewers is needed. Existing state-of-the-art technologies are not robust against harsh sewer conditions and, therefore, cause high maintenance expenses. Having the goal of fully automatic, robust and non-contact volumetric flow measurement in sewers, we came up with an original and innovative idea of a vision-based system for volumetric flow monitoring. On the contrast to existing video-based monitoring systems, we introduce a second camera to the setup and exploit stereo-vision aiming of automatic calibration to the real world. Depth of the flow is estimated as the difference between distances from the camera to the water surface and from the camera to the canal's bottom. Camera-to-water distance is recovered automatically using large-scale stereo matching, while the distance to the canal's bottom is measured once upon installation. Surface velocity is calculated using cross-correlation template matching. Individual natural particles in the flow are detected and tracked throughout the sequence of images recorded over a fixed time interval. Having the water level and the surface velocity estimated and knowing the geometry of the canal we calculate the discharge. The preliminary evaluation has shown that the average error of depth computation was 3 cm, while the average error of surface velocity resulted in 5 cm/s. Due to the experimental design, these errors are rough estimates: at each acquisition session the reference depth value was measured only once, although the variation in volumetric flow and the gradual transitions between the automatically detected values indicated that the actual depth level has varied. We will address this issue in the next experimental session.

Keywords: 3D reconstruction, stereo vision, particle tracking velocimetry, template matching, water discharge, sewer systems, urban drainage structures

1. INTRODUCTION

Blockages or breakages of sewer lines, inflows of excessive storm water, malfunction of pumping stations or electrical power failures may lead to sewer overflows (SOs). Upon the occurrence of any of the listed above events, untreated waste-water is discharged from a sewer into nearby streams, rivers, or other water bodies prior to reaching sewage treatment facilities. SOs pose a huge risk to the environment, as they contain not only storm water but also untreated human and industrial waste, toxic materials and debris. Reaching water bodies, pollutants threaten public health, endanger aquatic life and impair the use and enjoyment of waterways.[1–3] To avoid this problem, constant monitoring and quantitative and qualitative measurement of water flow in sewers is needed.

Monitoring of sewer flows is challenging. Most of the state-of-the-art systems for volumetric flow measurement used nowadays, such as positive displacement flow-meter,[4] require intrusion into the flow. Taking into account the demolishing nature of waste-water, such systems are usually short-lived in a severe sewer environment and require regular maintenance. Moreover, such equipment is usually expensive, and, thus, its usage is limited.

Further author information: (Send correspondence to E.S.)
E.S.: E-mail: ekaterina.sirazitdinova@rwth-aachen.de, Telephone: +49 241 80-89797

Videometrics, Range Imaging, and Applications XIV, edited by Fabio Remondino, Mark R. Shortis, Proc. of SPIE Vol. 10332, 103320M · © 2017 SPIE
CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2270233

There are also non-invasive methods. One of them is ultrasonic flow-meter.[5] It measures the level in a channel by transmitting a pulse of sound from the the sensor to the flow surface and estimating the time for the echo to return. However, this method is not robust against presence of foam, turbulence, floating debris, oil or grease. Moreover, it is not recommended to use this method on wide channels due to beam spread.

In this work, we proposed a novel solution for automatic measurement of discharge in urban drainage structures. Our system relies on stereo vision and, thus, does not require a contact with waste-water, which makes its maintenance rather simple. Equipped with an infrared lamp, it can provide robust measurements at any time of the day.

2. RELATED WORK

An alternative non-intrusive vision-based solution was already proposed in two complementary works.[6,7] The authors of both papers developed together a system which is able to track volumetric flows in sewers with a help of video cameras. According to the authors, the video camera setup is robust against instrumental loss and does not require frequent maintenance. Volumetric flow rate is measured by combining image-based approaches of water level measurement and surface velocity estimation. Despite relatively high accuracy of water level estimation claimed by the authors (the root mean square (RMS) error varied from 1.33 cm to 4.61 cm compared to the ground truth for different observations), there is a major inconvenience in the proposed technique relying on the detection of water borderline in the images: the described scenario requires calibration relatively to the world. For that calibration, special rulers need to be placed into the scene, and correspondences need to be chosen manually by experts. This complicates the initial setup in the field and originates additional sources of error.

3. MATERIALS AND METHODS

Introducing a second camera to the setup, we exploit stereo vision techniques to recover water level in sewers and to achieve automatic calibration to the real world coordinates. Accompanied with velocity estimation module, our system is able to provide accurate measurements of water discharge in the sewers with known geometry.

3.1 System architecture

Interaction between the components of our system for volumetric flow is organized as follows (Fig. 1): a browser-based client software is used for initial system settings; entered by the user upon installation, location settings (region of interest (ROI), location name and canal geometry) are transferred to the operational web-server, which transfers the data to the storage and to the image processing software running on a single-board computer (SBC) (communication circuit board), which, in turn, communicates with the sensors, or image recording hardware. The recording hardware consists of a lighting element and two original equipment manufacturer (OEM) cameras (Fig. 2a). The recording device is placed into a sewer (Fig. 2b) and image sequences are acquired according to the specified schedule. The produced photographic data is then transferred to the SBC for further processing. The image processing is performed by custom software implemented in C++. The general algorithmic pipeline of discharge computation involves two major steps: depth reconstruction and surface velocity computation (Fig. 3).

The outcomes of the image processing pipeline are interpreted in terms of real world values and compared to the reference values. Based on this step, a decision is made, whether a danger of an overflow event has occurred and, if necessary, notification is sent to the user. All determined values together with recorded images are saved in the storage and are available to the user upon request through the client application.

An additional camera calibration software is used once on the side of the manufacturer before releasing a product item to correct for lens distortion and to determine the relation between the cameras natural units (pixels) and the real world units (millimeters). The results of camera calibration are included in the system configuration file.

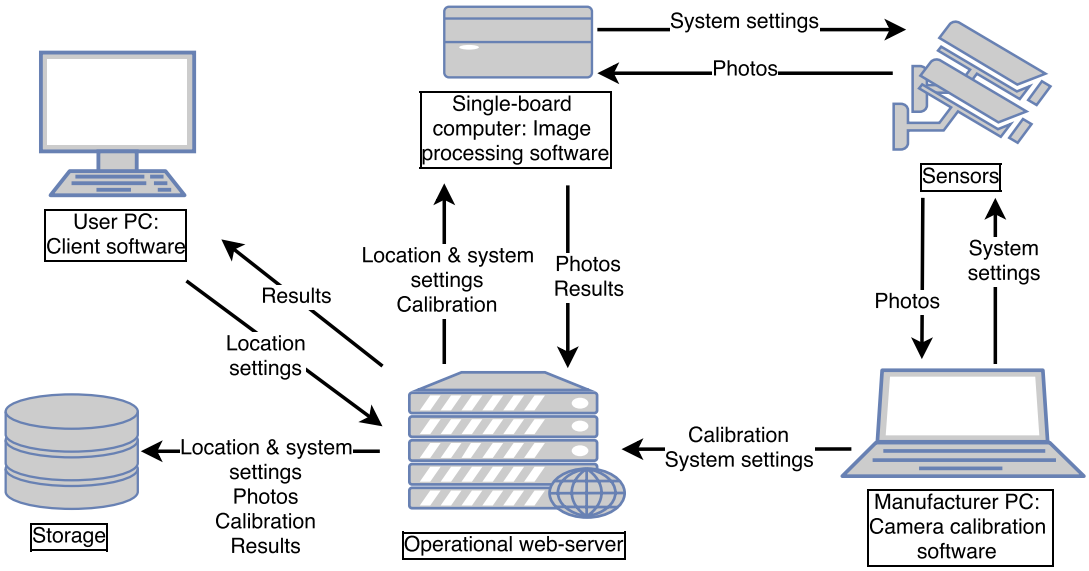


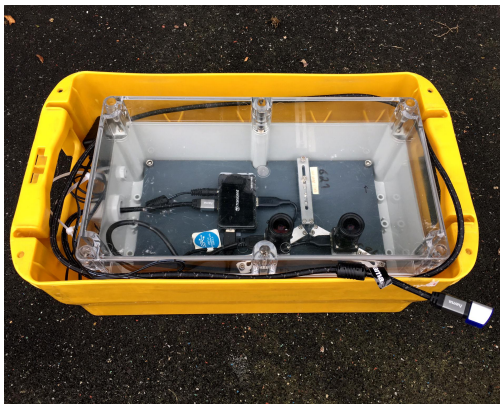
Figure 1: Interaction between the system's components.

3.2 Discharge

In order to measure volumetric flow Q in a sewer system, it is necessary to know the cross-sectional area of the canal A and mean surface velocity $\bar{\nu}$. The value of volumetric water discharge can be then approximated by the product of these two parameters:

$$Q = \bar{\nu}\, A(l). \tag{1}$$

The value of A depends on the current water level l detected automatically and the geometry of the canal. All parameters necessary for the computation of A except l shall be given upon the system installation (due to the fact that the canal profiles are usually standardized, this task is reduced to the selection of the correct profile in the user settings). Thus, the basic milestones of the automatic volumetric flow computation shall be finding the water level l and the velocity $\bar{\nu}$.



(a) Camera prototype.

(b) Placement of a prototype into an open canal.

Figure 2: Recording hardware.

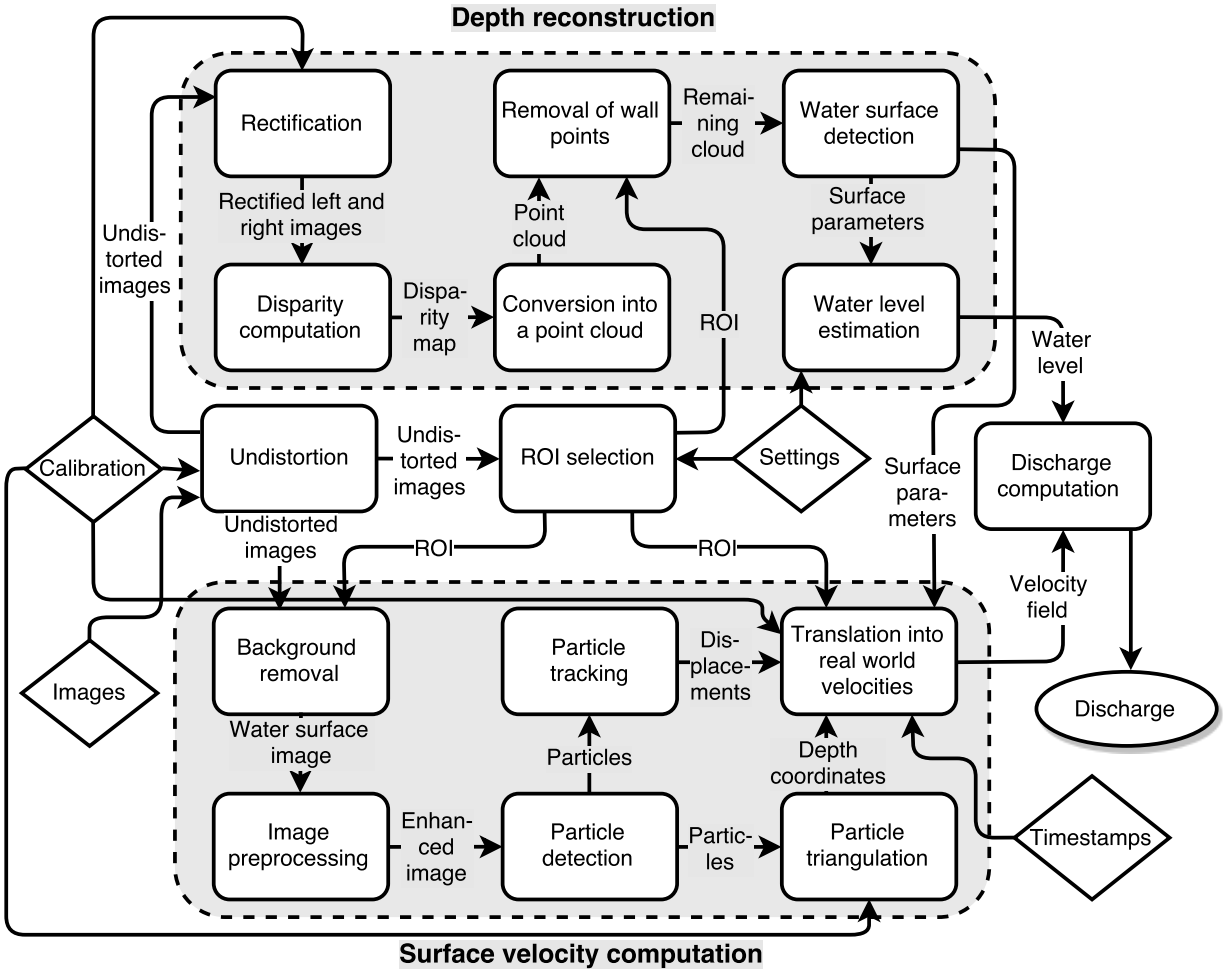


Figure 3: General algorithmic pipeline.

3.3 Camera calibration

Camera calibration is a crucial step needed to be accomplished in order to obtain accurate results of depth estimation and other image processing steps, as well as to produce a reference value for conversion between the image coordinate system and the world coordinate system. With a chessboard camera calibration[8] the intrinsic and extrinsic parameters are estimated for each camera.

The intrinsic parameters are internal parameters of each camera such as focal length f expressed in pixel units, skew coefficient s between the x and the y axis, principal point with coordinates c_x and c_y and image distortion coefficients (radial distortion coefficients k_1, k_2, k_3, k_4, k_5, and k_6 and tangential distortion coefficients p_1 and p_2). Using these parameters, we correct the original 2D images for distortion effects caused by imperfections of optical systems. If not treated, these effects can undermine the accuracy of the follow-up depth estimation and surface velocity measurement. In order to get an undistorted image, each pixel in this image is mapped to its correct position using the Brown–Conrady distortion model, which adds a tangential component to the radial distortion.[9]

The extrinsic parameters define the position of the camera center and the camera's orientation in world coordinates. The position of the camera $\boldsymbol{c}$ is expressed as a negative product of inverted camera rotation matrix $\boldsymbol{R}$ and the position $\boldsymbol{t}$ of the origin of the world coordinate system expressed in coordinates of the camera-centered coordinate system:

$$\boldsymbol{c} = -\boldsymbol{R}^{-1}\boldsymbol{t}. \tag{2}$$

By knowing the positions of two cameras in the real world coordinate system and their corresponding locations in the image coordinates, we can find a factor for conversion of the depth and surface velocity values into metric units. The intrinsic and extrinsic parameters are used to rectify a stereo pair of images to make them appear as though the two images share a common plane which is parallel to the line between optical centers.[10] This step is required for computation of the disparity map which is needed for 3D depth reconstruction.

The intrinsic and extrinsic parameters remain unchanged if the camera set-up does not change, i.e. the focal lengths of both cameras are fixed and the distance between them remains unchanged. Thus, camera calibration needs to be performed only once in the product life cycle for each combination of camera settings. The same parameters will be applied to all images recorded with the same device depending on the chosen settings.

3.4 Depth estimation

The water level l is estimated as a difference between distances from the camera setup to the water surface and to the canal's bottom. The later is measured upon installation and recorded into the system. The distance to the water surface is computed using large-scale stereo matching.[11]

For robust estimation of the distance to the water surface, we apply depth estimation on several pairs of images, where one image pair represents two images recorded simultaneously with the right and the left cameras. We set the number of pairs to 20, but it may vary depending on the allowed computational time: the more pairs are taken, the robuster the result are. The process runs as follows: (i) the images are undistorted and rectified (Sec 3.3); (ii) image contrast is equalized by applying contrast limited adaptive histogram equalization (CLAHE);[12] (iii) a dense disparity map is calculated for the given left-right image pair as it is done in;[11] (iii) the disparity map is reprojected to a point cloud: for each pixel in the disparity map having 2D pixel coordinates (u, v) and a disparity value d, its 3D coordinates (X, Y, Z) in the world coordinate space are computed as

$$\begin{cases} X = Z\,(u - c_x)\frac{1}{f} \\ Y = Z\,(v - c_y)\frac{1}{f} \\ Z = \frac{f\,b}{d} \end{cases} \tag{3}$$

where b is the baseline length (translation from right to the left camera); (iv) the points belonging to canal walls are removed from the model based on the ROI: the remaining points either belong to the water surface, or can be classified as outliers caused by water reflections or resulting from the reprojection error; (v) the parametric water surface plane $D = K\,X + L\,Y + M\,Z$ is approximated using maximum likelihood estimation sample consensus (MLESAC) algorithm.[13] The approximated plane (Fig. 4) should be parallel to the image plane with allowed offset of 5°.

After running all these steps for each image pair, we calculate the trim mean of the water surface plane as an average of values between 20th and 80th quantiles of all values of plane models resulting in a single plane

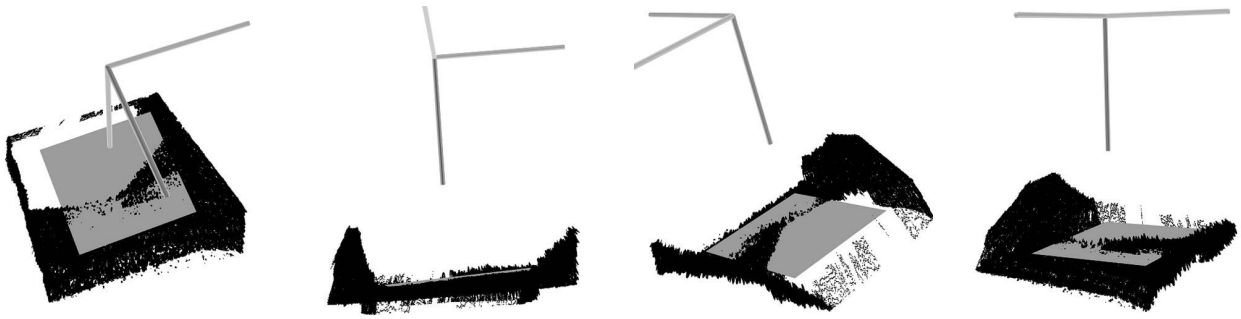


Figure 4: Water surface detected in the reconstructed point cloud shown from different perspectives. The coordinate axis is placed to the point of origin (the camera position).

equation $D' = K'\,X + L'\,Y + M'\,Z$. The distance l from the camera to the water surface is then computed as

$$l = \frac{|K'\,X_0 + L'\,Y_0 + M'\,Z_0 + D'|}{\sqrt{K'^2 + L'^2 + M'^2}} \tag{4}$$

where (X_0, Y_0, Z_0) are the coordinates of the origin.

3.5 Surface velocity computation

The velocity of the flow is calculated by tracking particles of interest in the sequential frames. These can be small pieces of floating waste highly distinguishable from the background water. In this section, we describe, how to make such tracking possible.

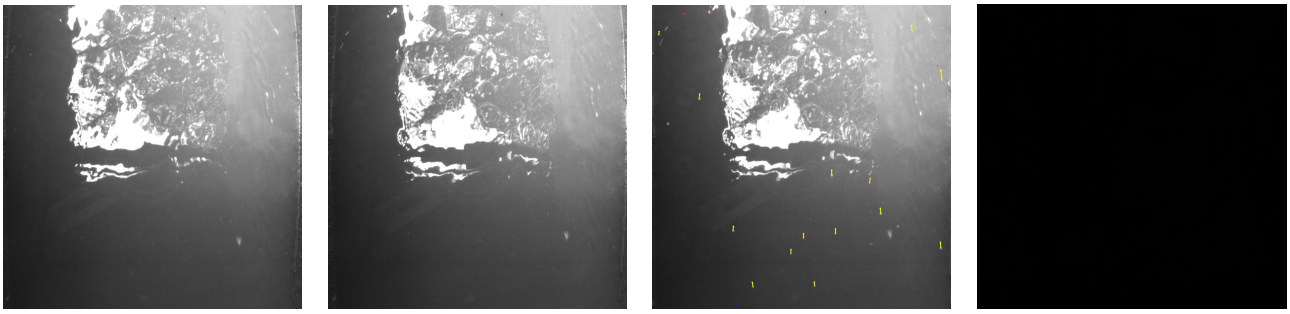
Image preprocessing and particle detection. Images coming directly from the cameras often feature some undesired artifacts. Amongst those are wavy light reflections from natural or artificial light sources, traces of rain or condensate drops, and shadows lowering the intensity of particles. When untreated, they undermine particle tracking by creating false matches, or outliers, or leaving some important particles undetected. With a simple image processing pipeline, we prepare each initial undistorted image cropped to the selected ROI for a robuster particle tracking: (i) CLAHE is applied for image equalization; (ii) gradient edge detection is applied to transform the image into an edge image; (iii) the detected edges are emphasized by grayscale dilation; (iv) the image is binarized with a simple threshold: we set up the threshold value to 4% of maximum intensity in the image. In the case of 8 bit grayscale image, the threshold value is 10. All pixels with intensities above the threshold are turned into maximum intensity, while all other are discarded; (v) all remaining objects in the scene are presented as connected components, which are then classified as reflections and other objects based on their contour length. Reflections, normally having a larger size, are eliminated; (vi) morphological erosion is applied to the remaining objects to set their size back to the original. In the binarized images, particles are detected using simple blob detection based on the connected-component labeling .[14] We label a detected blob as a particle if its area is bigger than 3 pixels.

Particle tracking and estimation of displacement vectors. After the preprocessing steps, displacement vectors can be estimated. Each displacement vector connects coordinates of a detected particle in the first image frame with its corresponding coordinates in the following frame. Such displacements are estimated for all sequential pairs of images. Knowing this displacements in image coordinates and the timestamps indicating when images were taken, we can provide a velocity vector field for each image sequence. In order to find the correspondences between the particles in sequential images, a technique called template matching is used: a function slides through the image and compares the overlapped windows of size $w \times h$ (w – window's width, h – window's height) against the input image using fast normalized cross-correlation.[15] For more effective computations the searched area is narrowed down to a certain size (in our case, we take the area limited by a circle with a radius of 40 pixels with a center in the particle coordinates). Probabilities of different possible matches are estimated, the window with higher matching probability is localized and the particle correspondences are established. The matches with a probability lower than 30% are discarded. In the Fig. 5c, it is seen that, with a chosen set of parameters, only few correspondences between particles are detected in the pair of sequential images. However, after processing the whole image sequence, all detected correspondences together form a vector field, which is sufficient for a robust velocity computation (Fig. 5d). The velocity $\bar{\nu}$ is calculated as average value in the velocity vector field after statistical outlier removal.

4. EVALUATION

4.1 Dataset

Six sessions were recorded on six different dates. Each session consists of ten image series. Each series was recorded over the 8 s interval with a frame rate of 25 images per second. Five sessions were recorded at location A and one – at location B. In the beginning of each session, a reference depth value l_{ref} was measured with a ruler. Additionally the value of volumetric flow Q_{ref} was recorded: in the first two sessions only once and in the rest four – before each series. Both locations have rectangular profiles.



(a) (b) (c) (d)

Figure 5: Particle tracking: (a) and (b) are originals of two sequential images before undistortion; (c) found correspondences between the particles in the images (a) and (b); (d) the vector field over the whole sequence of images.

4.2 Pipeline

The following values were computed automatically using the developed image processing software:

- Average velocity $\bar{\nu}$,
- Distance from the sensor to the water surface l.

In the rectangular profile, knowing the width of the channel w_{channel} and the total distance from the sensor to the channel bottom l_{total}, and having automatically estimated the velocity $\bar{\nu}$ and the distance from the sensor to the water surface l, the volumetric flow Q was estimated as

$$Q = \bar{\nu}\, w_{\text{channel}}\, (l_{\text{total}} - l) \tag{5}$$

There were no reference values for surface velocity available. However, it was possible to compute theoretical surface velocity $\bar{\nu}_{\text{theor}}$ from the reference volumetric flow Q_{ref} and depth l_{ref} as follows:

$$\bar{\nu}_{\text{theor}} = \frac{Q_{\text{ref}}}{w_{\text{channel}}\, l_{\text{ref}}} \tag{6}$$

The absolute velocity error E_{vel} was computed as a difference between $\bar{\nu}_{\text{theor}}$ and $\bar{\nu}$. The absolute depth error E_{dep} was computed as a difference between the estimated depth $(l_{\text{total}} - l)$ and reference value l_{ref}. The relative depth error values were computed as ratios of absolute error values and the values of the measuring ranges (which is in our case the distances from the camera to the channel's bottom).

4.3 Results

With the respect to the measuring range, the minimal depth error E_{dep} was estimated as 0.11% (or 0.21 cm), the maximum depth error was 4.12% (or 8 cm), while the average error was 1.69% of the measuring range (3 cm). The RMS error of the total set was 3.4 cm. The error of surface velocity E_{vel} resulted in the interval between 1 cm/s and 13 cm/s with an average error of 5 cm/s and the RMS error of 9 cm/s. We excluded from this range two first sessions which are lacking the information about individual values of volumetric flow for each series. However, due to the experimental design, these errors are only rough estimates: at each acquisition session the reference depth value was measured only once, although the variation in volumetric flow and the gradual transitions between the automatically detected values indicated that the actual depth level has varied. Other possible sources of error: inaccuracies in calibration (with a very low probability) and presence of time gaps between two synchronous images due to a possible problem of the sensors or recording software. These issues are the topic of our ongoing work.

5. CONCLUSION

Considering the results of preliminary evaluation, we already can see that our system performs similarly to the one of Nguyen et al.[6] and Jeanbourquin et al.[7] which to our knowledge was the only development in the field so far. Considering the fact that we managed to achieve fully automatic depth and surface flow estimation, we can recommend our system as an industrial solution for volumetric flow measurement.

REFERENCES

[1] Burkhardt M, Kupper T, Hean S, Haag R, Schmid P, Kohler M, et al. Biocides used in building materials and their leaching behavior to sewer systems. Wat Scien & Techn. 2007;56(12):63–67.

[2] Government, US. Effects of wastewater and combined sewer overflows on water quality in the Blue River Basin, Kansas City, Missouri and Kansas. General Books; 2011.

[3] Rodríguez RA, Gundy PM, Rijal GK, Gerba CP. The impact of combined sewage overflows on the viral contamination of receiving waters. Food and Envir Virol. 2012;4(1):34–40.

[4] Spitzer DW. Industrial flow measurement. Resources for measurement and control series. ISA-The Instrumentation, Systems, and Automation Society; 2005.

[5] Lynnworth LC. Ultrasonic measurements for process control: Theory, techniques, applications. Orlando, FL, USA: Academic Press, Inc.; 1989.

[6] Nguyen LS, Schaeli B, Sage D, Kayal S, Jeanbourquin D, Barry DA, et al. Vision-based system for the control and measurement of wastewater flow rate in sewer systems. Wat Scien & Techn. 2009;60(9):2281–2289.

[7] Jeanbourquin D, Sage D, Nguyen L, Schaeli B, Kayal S, Barry DA, et al. Flow measurements in sewers based on image analysis: Automatic flow velocity algorithm. Wat Scien & Techn. 2011;64(5):1108–1114.

[8] Zhang Z. A flexible new technique for camera calibration. In: Proc IEEE Pattern Anal Mach Intell. Washington, DC, USA: IEEE Computer Society; 2000. p. 1330–1334.

[9] Brown DC. Decentering distortion of lenses. Photometr Eng. 1966;32(3):444–462.

[10] Fusiello A, Trucco E, Verri A. A compact algorithm for rectification of stereo pairs. Mach Vis Appl. 2000 Jul;12(1):16–22.

[11] Geiger A, Roser M, Urtasun R. Efficient large-scale stereo matching. In: Kimmel R, Klette R, Sugimoto A, editors. Proc ACCV 2010. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 25–38.

[12] Zuiderveld K. In: Heckbert PS, editor. Graphics Gems IV. San Diego, CA, USA: Academic Press Professional, Inc.; 1994. p. 474–485.

[13] Torr PHS, Zisserman A. MLESAC: A new robust estimator with application to estimating image geometry. Comp Vis Image Underst. 2000 Apr;78(1):138–156.

[14] Suzuki S, Abe K. Topological structural analysis of digitized binary images by border following. Computer Vis Graph and Im Proc. 1985;30(1):32–46.

[15] Lewis J. Fast normalized cross-correlation. Vis Interf. 1995;10(1):120–123.