ANALYSIS OF CROWD-SOURCED FLOODING IMAGES USING COMPUTER VISION TECHNIQUES

Megan A. Witherow, Corresponding Author
Department of Electrical and Computer Engineering
Old Dominion University
2107 Engineering Systems Building, Norfolk, VA 23529
Tel: 757-297-1453; Email: mwith010@odu.edu

Dr. Khan M. Iftekharuddin
Department of Electrical and Computer Engineering
Old Dominion University
231 Kaufman Hall, Norfolk, VA 23529
Tel: 757-683-3741; Email: kiftekha@odu.edu

Dr. Mecit Cetin
Department of Civil and Environmental Engineering
Old Dominion University
135 Kaufman Hall, Norfolk, VA 23529
Tel: 757-683-6700; Email: mcetin@odu.edu

Word count: 3002 words text + 7 tables/figures x 250 words (each) = 4,752 words

July 29, 2016
ABSTRACT
Crowd sourced photographs taken at near ground-level present a new type of data for flooding analysis. In this study, crowd sourced images taken in the city of Norfolk, Virginia are analyzed to extract the inundated area. Photographs of flooded roads are recreated under dry conditions for comparison. Few preprocessing steps are used to normalize the image sets with and without flooding that are then registered considering the image without flooding as reference. Based on the data, two algorithm pipelines are developed to extract the flooded area in the images. The first pipeline obtains the flooded area by subtracting the registered images to determine the difference. From this pipeline, the flooded area with reflections from nearby landmarks on the water is identified. In a second pipeline, the saturation channel of the HSV flooded image is used for threshold-based segmentation of the water areas reflecting the sky. The results from these two pipelines are then combined to capture the inundated area.
INTRODUCTION
With the proliferation of smart camera phones and social media in today’s society, photographs are constantly being taken and shared. This public sharing of photographs presents new opportunities and data sources. When roadways flood, many motorists and members of affected communities take photographs that document the water level change and location. Analysis of these crowdsourced photographs to extract information on the extent of the flooding may provide valuable insights in the environmental and transportation domains.

Previous studies applying image-processing techniques to flooding analysis use remotely-sensed data that provide change information from an overhead perspective (1). Crowd sourced images taken from near ground-level provide a local perspective on roadway inundation. The greater local detail of information available in these near ground-level images of flooded roadways may be useful in the improvement of traffic information, road closure notification, and routing of emergency vehicles during flooding conditions.

Computer vision offers diverse techniques for the processing of the images and extraction of information. In this study, two image processing algorithm pipelines are developed to identify the inundated area of the scene from crowd sourced images taken in the city of Norfolk, Virginia. The images are obtained from (2). The two different pipelines are developed to account for the image content variability of crowd sourced data and the nature of reflections on water. In the primary image processing algorithm pipeline, image subtraction of registered flooded and dry images is used to capture regions of inundation that have reflections of nearby buildings and objects. A separate, secondary pipeline that applies thresholds to the HSV saturation channel is developed for threshold-based segmentation of the flooded area reflecting the sky. The segmented results of these two pipelines are combined to obtain the composite result for the inundated area.

The remainder of the paper will introduce some background on the data, detail the methodology used, share findings, present the conclusion, and discuss avenues for future work.

BACKGROUND
Analysis of floods using image processing is an active research area. However, previous literature is dominated by the use of remotely sensed data acquired from radar or satellites such as MODIS, Landsat, SPOT, IKONOS, QuickBird, GeoEye, RapidEye, EROS A and B (3). These remotely-sensed data have been analyzed for flood monitoring in (4), (5), (6) and (7); for flood risk assessment in (8), (9), (10), and (11); and for flood damage estimation in (12), (13), (14), and (15). Images taken from near ground-level have not yet been explored in the literature.

The data used in this study consists of two sets of images: crowd-sourced photographs of flooded streets and recreated photographs of the same streets under dry conditions. The flooded set consists of color photographs taken in December 2009 from four locations in Norfolk, Virginia near Chrysler Museum of Art, as shown in (Figure 1). The individual photographs are taken from various angles and perspectives in landscape orientation. The dry set of images is acquired in June 2016 from the same four locations, as shown in (Figure 2). The position and angle of the original photographs is recreated as closely as possible. Multiple shots are taken using different cameras and the best photograph is selected for further processing and comparison with the original image. Recreating the appropriate angle and perspective for these images is challenging as many of the original flooded images are taken from the street within a vehicle and the exact height and location of the camera when the photographs are taken is not known.
There are several challenges associated with this data as a result of temporal constraints. Due to the time gap of approximately seven years between the two sets of images, some of the features of the areas in the photographs have changed, as can be seen in the landscaping, growth of trees and other foliage, addition or removal of signs, infrastructure wear and repairs. These changes affect landmarks relevant to the registration or alignment of the flooded and dry condition images. The limited data collection time frame for the second set of images results in a comparison between images taken in the midst of winter and summer. The difference in time of year introduces another challenge to registration in the presence or absence of leaves on deciduous trees and necessitates corrections for seasonal variations. Occlusion from vehicles is also increased in the images taken under dry conditions, since the roads are not flooded.

(a) Location 1: Intersection of South Mowbray Arch & Yarmouth Street
(b) Location 2: Intersections of Yarmouth Street & Grace Street and South Mowbray Arch & Memorial Place
(c) Location 3: Mowbray Arch
(d) Location 4: Mowbray Arch

FIGURE 1 Crowd sourced photographs of flooded streets in Norfolk, Virginia. Images are obtained from (2).
METHODOLOGY
The extraction of the flooded area is accomplished through preprocessing and two image processing algorithm pipelines. The overall pipeline used to capture the inundated area is shown in (Figure 3). The preprocessing pipeline normalizes the images taken under flooded and dry conditions. These normalized images are input into the primary image processing algorithm pipeline, which extracts the flooded area with reflections from nearby landmarks using image subtraction and threshold-based segmentation. A separate, secondary pipeline accepts the original color, flooded image and extracts the areas without reflected landmarks through thresholding the HSV saturation channel. Each image processing algorithm pipeline outputs a binary image of the segmented inundated area. These are combined into a single binary image to produce the final result.
(a) Preprocessing
Before the images can be used to extract information about the flooded area, they must be normalized through few preprocessing steps. The detailed preprocessing pipeline is shown in (Figure 4). Both the images taken under flooded and dry conditions are cropped to remove areas that will neither contribute as landmarks for the registration of images nor provide inundation information. The cropping is done to preserve common features between the two images as much as possible. Following cropping, the images are converted to grayscale.

Contrast correction is performed on the darker of the two images using contrast-limited adaptive histogram equalization (CLAHE). CLAHE transforms each pixel based on the neighboring pixels. Following the comparison of uniform, Rayleigh, and exponential distributions, the exponential distribution is selected for the histogram shape. Through visual inspection, contrast enhancement limits are selected for each image to match the lighter, reference image as closely as possible. In some cases, CLAHE generates noise in homogeneous regions of the image. In these cases, pixelwise adaptive Wiener low pass-filtering is applied to remove the noise so that it does not affect the registration.

Following contrast correction, the larger of the two images is downsampled using the bicubic interpolation method. This ensures that the images have the same dimension for registration.

Due to the constraints of the data, basic techniques are not adequate to register the images. The Scale Invariant Feature Transform (SIFT) flow algorithm (16) from the Image Alignment Toolbox is used for registration. Use of SIFT features allows the algorithm to be robust to the spatial distortions in angle and perspective in the image pairs, warping one of the images as necessary for proper alignment. The image taken under dry conditions is chosen to be warped to avoid distorting the flooded area. Once the images are registered, they are ready for further processing.
(b) Primary Image Processing Pipeline
The primary image processing pipeline captures flooded areas that contain specular reflections of buildings or other objects. The pipeline is shown in (Figure 5). Subtraction of the registered
images is used to identify the areas that are different between the two images. The pixelwise subtraction of intensity values results in a new image that is lighter in areas where the difference is larger and darker in areas where the two images are more similar. This identifies the areas of water with reflections of nearby landmarks because the intensity values are unlike than those for the street. However, areas of water that are reflecting the sky are not well identified by this method because the pixel values of these areas are similar to the pixel values of the street.

A Gaussian low-pass filter is applied to the subtraction result to blur the image, then dilation is performed. These steps reduce the effect of small dark or light patches, such as those caused by tree growth, that are not relevant to analysis of the flooded area. Threshold-based segmentation divides the image into flooded and unflooded portions using an intensity-based threshold that separates these sections of the image.

FIGURE 5  Results of Primary Image Processing Pipeline.
(c) Secondary Pipeline
A secondary image processing pipeline is used to capture the flooded areas that are reflecting the sky. The detailed pipeline is shown in (Figure 6).

FIGURE 6 Results of Secondary Image Processing Pipeline.
The original color, RGB profile image with flooding is converted to HSV or hue, saturation, and value profile. The saturation channel is then taken by itself for further processing. Upper and lower thresholds are defined and evaluated pixelwise to produce a binary image. The binary image is dilated and combined with the result from the primary pipeline.

RESULTS
The preprocessing and processing pipelines described in the Methodology section are applied to four different locations. In all cases, it is observed that the subtraction is capable of only extracting the inundated area where there are reflections from nearby buildings and objects. Segmentation of the remaining area from the HSV saturation channel requires carefully selected thresholds based on trial and error. The saturation channel is not equally effective at capturing the remaining area in all cases so experimentation with dilation is necessary to achieve the best result. Due to the variability of the data, each image must be considered individually to find parameters that are effective for extracting the flooded area. The composite results for four locations in Norfolk, Virginia are shown in (Figure 7).

FIGURE 7 Combined Results From Two Pipelines Showing Flooded street area determined at four locations in Norfolk, Virginia.
CONCLUSION
In this study, techniques for extracting information from crowd-sourced images of flooded streets are explored. Challenges in the recreation of perspective and angle for the dry comparison image suggest that future methods focus on extracting the flooding information without a comparison image or that data be acquired from video taken from a stationary location. Image subtraction alone is found to be inadequate in capturing the entire flooded area. Rather, the information provided from image subtraction is limited to areas where there are reflections from buildings or other objects. This is the result of the intensity values of the water reflecting the sky being too similar to the intensity values of the road itself. A second pipeline is used to capture some of the remaining water area by applying hand-selected thresholds to the saturation channel of the flooded image converted to the HSV profile. Selection of thresholds by hand is time-consuming as trial and error is required to obtain the best result. In future works, some other methods should be pursued for automated identification of these areas. From the combination of these two pipelines, the inundated area is extracted.

FUTURE WORK
In future work, we plan to explore the use of the standard surface-reflectance model and illuminant spectral power distribution estimation to improve segmentation of the flooded area (17). Using the properties of the interface and subsurface reflections (17) may produce better results in the segmentation of water areas, especially those that do not have reflected landmarks. Improving this performance should eliminate the need for hand-selected thresholds and allow for the automation of the detection of flooded areas. It would be interesting if these properties might also contribute to a method of detecting inundated areas without a reference image taken during dry conditions. We plan to investigate the possible use of landmarks in the images for extraction of water depth information. We also plan to explore unsupervised machine learning approaches for robust, automated identification of flooded areas.

ACKNOWLEDGMENT
This project was completed with the support of the Mid-Atlantic Transportation Sustainability University Transportation Center (MATS UTC) as a part of the 2016 MATS UTC Summer Undergraduate Research Program. The authors also acknowledge Skip Stiles and Wetlands Watch Incorporated for providing the flooding images for this project.
REFERENCES


2. Wetlands Watch Incorporated; Address: P.O. Box 9335, Norfolk, Virginia 23505; Phone: 757-623-4835; [http://www.wetlandswatch.org/](http://www.wetlandswatch.org/).


