What is a Hole? Discovering Access Holes in Disaster Rubble with Functional and Photometric Attributes

Christopher Kong  
Department of Computer Science, Ryerson University, Toronto, Ontario, Canada  
e-mail: c5kong@scs.ryerson.ca

Alex Ferworn  
Department of Computer Science, Ryerson University, Toronto, Ontario, Canada  
e-mail: aferworn@scs.ryerson.ca

Elliott Coleshill  
School of Information and Communications Technology, Seneca College, Toronto, Ontario, Canada  
e-mail: elliott.coleshill@senecacollege.ca

Jimmy Tran  
Department of Computer Science, Ryerson University, Toronto, Ontario, Canada  
e-mail: q2tran@scs.ryerson.ca

Konstantinos G. Derpanis  
Department of Computer Science, Ryerson University, Toronto, Ontario, Canada  
e-mail: kosta@scs.ryerson.ca

Received 14 June 2014; accepted 15 January 2015

The collapse of buildings and other structures in heavily populated areas often results in human victims becoming trapped within the resulting rubble. This rubble is often unstable, difficult to traverse, and dangerous for emergency first responders tasked with finding, stabilizing, and extricating entombed or hidden victims through access holes in the rubble. Recent work in scene mapping and reconstruction using photometric color and metric depth (RGB-D) data collected by unmanned aerial vehicles (UAVs) suggests the possibility of automatically identifying potential access holes into the interior of rubble. This capability would greatly improve search operations by directing the limited human search capacity to areas where access holes might exist. This paper presents a novel approach to automatically identifying access holes in rubble. The investigation begins by defining an access hole in terms that allow for their algorithmic identification as a potential means of accessing the interior of rubble. This definition captures the functional and photometric attributes of holes. From this definition, a set of hole-related features for detection is presented. Experiments were conducted using RGB-D data collected over a real-world disaster training facility using a UAV. Empirical evaluation suggests the efficacy of the proposed approach for successfully identifying potential access holes in disaster rubble.

1. INTRODUCTION

1.1. Motivation

Disasters involving collapsed buildings in urban areas occur for a variety of reasons. Due to the increased population density in these areas, the likelihood of humans becoming trapped in the resultant building rubble is quite high. In response to these events, organized teams or Task Forces of Urban Search and Rescue (USAR) personnel are deployed to locate and extract victims, build support for unstable structures (i.e., shoring), and provide medical care (FEMA, 2009).

When rescue personnel perform triage on a collapsed structure, they first determine areas that are likely to contain trapped victims, and then they formulate a plan to access the structure’s interior. If access holes already exist, these will be evaluated before rubble removal is considered to save time and reduce the chances of creating secondary collapses (FEMA, 2009). Figure 1 shows an access hole that potentially leads into the rubble’s interior.

The terms “hole” or “access hole” are not clearly defined within the USAR nomenclature. The challenge lies in the amorphous nature of holes (e.g., the lack of a prototypical shape, depth, and orientation), thus a definition is left open to the interpretation of each search team. To compound the difficulty of this problem, disaster rubble typically contains many irregularities within the rubble pile. For instance, inconsistency in the size, shape, and types of...
Figure 1. An image of a rubble field from the dataset introduced in this paper. An “access hole” is highlighted by the green bounding box.

Figure 2. Left: A raw RGB image, and right: a per-pixel registered depth image is obtained by a (top) UAV outfitted with an Asus Xtion Pro sensor. The colorized depth map visualizes the depth, where red and green represent smaller and larger depths from the capture sensor, respectively. This figure is best viewed in color.

material constituting the rubble affect what is and is not considered to be a candidate entry hole.

An access hole can be defined through its intended use. In the context of the current work, the goal is to locate holes that are sufficiently large to permit human entry. In this way, an access hole is defined in the context of its functional utility for search and rescue. This is analogous to the functional object recognition paradigm pursued in computer vision (Dickinson, 2009) that models objects, such as chairs, in terms of their function, i.e., the ability to support a human, rather than the particulars of their appearance.

The instability of rubble can prevent USAR teams from safely traversing it to find access holes. This situation has led researchers to investigate ways to minimize risk to human searchers, through the use of unmanned vehicles (Birk, Wiggerich, Bülow, Pfingsthorn, & Schwertfeger, 2011; Ferworn, Tran, Ufkes, & D’Souza, 2011; Finn & Wright, 2012; Murphy, 2004; Onosato et al., 2006). Previous work (Ferworn et al., 2011) demonstrated the ability to equip a UAV with a low-cost, off-the-shelf color camera with per-pixel metric depth information (i.e., an RGB-D sensor) and to capture critical disaster scene information from a safe distance and an alternative perspective. This information can then be used to generate a scene-level model that can quickly provide first responders with important details about the structure of the rubble (Ferworn, Herman, Tran, Ufkes, & McDonald, 2013) and provide input to an automated hole detection system, as pursued in the current paper.

In addition to the challenges posed by dealing with large amounts of visual data, traumatic events can overwhelm an individual, such as a building collapse disaster. This can lead to critical incident stress that can impair the ability of personnel to function and perform tasks involving the detailed observation required for visual search (FEMA, 2009). This paper argues that a system that automates the identification of access holes potentially reduces the cognitive load faced by response personnel.

Current approaches for identifying access holes rely on visual inspection by first responders. If responders are precluded from entering the scene or are not yet present, they must rely on imagery as their main source of information. The ability to collect data far outpaces a human’s ability to deal with those data.

This paper presents a novel vision-based approach for automatically discovering access holes in disaster rubble imagery. The case of holes leading into subsurface voids is examined. An underlying assumption is that an access hole possesses clearly marked boundaries and a salient depth variation from the surrounding area. In addition, the potential access hole possesses a minimum width and aspect ratio to accommodate the entry of an adult human searcher. While the focus in the current paper is on human searchers, other types of search entities are readily accommodated, such as search dogs or robots, by adapting the thresholds used by the approach.

A preliminary version of this work has appeared previously (Kong et al., 2013).

1.2. Related Work

No prior work has directly addressed the concept of automatic hole discovery for the purpose of search and rescue, as pursued here. The related research is relatively limited and can be organized in terms of two domains: (i) the data acquisition platform and (ii) visual object detection.

Search and rescue operations are often time-critical. Response robots benefit USAR operations by providing robust platforms to carry sensors, collect data, and deliver supplies to trapped victims (Murphy, 2000). Sensory data are useful for determining the quality of the environment and potentially assist with locating victims. Prior work has used ground-based robots for autonomous navigation and mapping rubble interior spaces (Mobedi & Nejat, 2012); however, this approach does not attempt to locate access holes for insertion into rubble. Using ground vehicles as platforms for automated, top-down, road inspection has been somewhat successful (Sy, Avila, Begot, & Bardet, 2008). In this work, a sensor collects baseline information about level road surfaces and detects variances that translate to detected surface cracks. Since disaster rubble is often comprised of irregular shapes and materials, as opposed to level terrain, this approach is not directly appropriate for USAR. Work has been carried out using autonomous vehicles for detecting subsurface voids in mining operations (Wilson, Gurung, Paaso, & Wallace, 2009); however, this approach requires heavy equipment, level terrain, and a mobile platform traversing the area of inspection. USAR terrain is inevitably cluttered and chaotic, making effective ground locomotion problematic.

Research in terrain traversability has yielded the concept of “negative obstacles.” Negative obstacles are defined as obstacles below the ground surface that return no sensor data and thus should be treated as holes to be avoided (Heckman, Lalonde, Vandapel, & Hebert, 2007). Early investigations into detecting negative obstacles analyzed ray traces of every pixel, comparing actual range values to expected range ones (determined via the position of the ground plane) to determine the difference (Matthies, Kelly, Litwin, & Tharp, 1995). This method makes the assumption of a homogeneous terrain being traversed, making it unsuitable for USAR. Further work in negative obstacle detection (Sinha & Papadakis, 2013) projects three-dimensional (3D) point cloud data collected directly in front of the sensor to a 2D ground plane to detect gap contours. Detections are then further analyzed for traversability of ground robots in the USAR domain. In contrast to these previous works, which dealt with the avoidance of negative obstacles for terrain traversability, this paper focuses on the suitability of these negative obstacles for insertion of trained search personnel in subsurface voids.

Ground robots are limited in the areas they are able to successfully traverse, since the terrain composition can adversely impact locomotion (Ollero, 2004). This has motivated, in part, the use of UAVs to conduct surveying and reconnaissance tasks (Finn & Wright, 2012). Using a UAV for USAR operations allows rescue personnel to survey areas that would not ordinarily be accessible, and to view the terrain from perspectives unattainable by terrestrial robots (Onosato et al., 2006). This rich information allows responders to carefully plan missions (Birk et al., 2011; Goodrich et al., 2008), and it has proven extremely useful in finding victims in search and rescue missions (RCMP, 2013). Recent work has considered UAVs equipped with an RGB-D sensor to collect data for both terrain mapping and 3D scene reconstruction (Ferwarz et al., 2011). To avoid the limitations of ground robots and to investigate areas inaccessible by a human searcher, the current work employs the use of a UAV to explore remote regions and to collect data.

Rubble characterization is a difficult problem. There have been preliminary investigations attempting to contribute solutions to this problem (Binda, Saisi, & Tiraboschi, 2001; Lombillo et al., 2013; Molino et al., 2007; Onosato, Yamamoto, Kawajiri, & Tanaka, 2012); however, there is no universally accepted categorization method for rubble. Motivated by this prior work, the current paper addresses a specific subproblem of rubble characterization: identifying its absence.
An extensive body of work has accumulated on appearance- and geometry-based object recognition approaches; see the surveys by Grimson, Lozano Perez, & Huttenlocher (1990), Mundy (2006), Dickinson (2009), and Andreopoulos and Tsotsos (2013). Appearance-based approaches map a photometric input pattern to a label of a specific object or class [see, e.g., Dalal & Triggs (2005), Felzenszwalb, Girshick, McAllester, & Ramanan (2010), Lampert, Blaschko, & Hofmann (2008), Krizhevsky, Sutskever, & Hinton (2012), and Girshick, Donahue, Darrell, & Malik (2014)], whereas geometry-based approaches utilize three-dimensional surface descriptions of the input scene to perform object recognition [see, e.g., Koppula, Anand, Joachims, & Saxena (2011), Lai, Bo, Ren, & Fox (2012), and Rusu, Bradski, Thibaux, & Hsu (2010)]. A major advantage of geometry-based approaches over appearance-based ones is their invariance to material properties, viewpoint, and illumination. Further, these approaches simplify the figure-background segmentation problem compared to appearance-based approaches. Three-dimensional recognition has experienced a revived interest in both the robotics and vision communities due to the introduction of commodity priced RGB-D sensors (Newcombe et al., 2011) and the abundant availability of three-dimensional models, e.g., Song & Xiao (2014).

Most closely related to the current work are functional descriptions for object recognition (Grabner, Gall, & Gool, 2011; Stark & Bowyer, 1991; Stark, Lies, Zillich, Wyatt, & Schiele, 2008; Winston, Binford, Katz, & Lowry, 1983), i.e., centering the object model on what one can do with the object rather than its appearance or shape. Many object classes exhibit a large degree of appearance and physical variation; for instance, the number of legs of a chair, while usually four, may vary. Access holes lack a canonical definition for size, shape, or orientation, making detection by appearance or shape a challenging task. For some of these objects, their description could be more easily provided by their function. This idea is adapted in the current work to develop a working definition of an access hole and to use the proposed function of an access hole to detect it. In other words, rather than describing what a hole looks like, it is more productive to define its function.

1.3. Contributions

In light of previous work, this paper makes three contributions. First, a novel definition of an access hole is presented based on a set of features derived from the functional form and photometric characteristics of a collapsed structure. Second, a novel approach is developed to automatically identify access holes in collapsed structures to be used by USAR personnel in accessing the collapsed structure. Analysis is performed on aerial imagery obtained by a UAV outfitted with an RGB-D sensor to identify candidate access holes. Third, a publicly available dataset obtained from a real-world USAR training rubble pile is introduced, where access holes are manually provided as ground truth. A quantitative empirical evaluation of the introduced dataset indicates the potential of the proposed approach for successfully identifying access holes in disaster rubble.

2. TECHNICAL APPROACH

2.1. Access Hole Definition

Before developing an approach to detect access holes, an operational definition is required. In this paper, an “access hole” (or “hole” for short) is defined by its potential for access into a collapsed structure, i.e., its function. In particular, a hole must be deeper in the interior than the surrounding terrain. Furthermore, to be useful for USAR, a hole must be large enough to support entry by a searcher, such as a human, dog, or robot. In the remainder of this paper, a searcher is assumed to be an adult human. This paper has identified three attributes that characterize a hole and allows us to perform access hole detection: (i) depth disparity, (ii) hole size, and (iii) photometric brightness.

2.2. Hole Attributes

The input to the proposed approach is an image pair extracted from an RGB-D sensor consisting of photometric color (RGB) and metric depth. The two images are registered such that they have a one-to-one mapping. To perform detection, candidate regions that potentially contain access holes must be identified from the terrain surrounding it. The proposed approach first oversegments the depth input into regions, i.e., superpixels (Ren & Malik, 2003), with the purpose of isolating regions (i.e., potential holes) exhibiting depth measurement discontinuities along their boundaries. A superpixel is a perceptually meaningful atomic image unit that contains pixels that are similar in some image property, such as depth, color, and texture. It is implicitly assumed that the constituent pixels of a superpixel belong to the same physical entity in the world. An adjacency graph is next created by identifying the neighbors of each superpixel. For each superpixel, a set of geometric and photometric feature scores is assigned, where each score represents the likelihood of a hole. Feature scores for each superpixel are aggregated to realize a final hole detection score. Figure 3 summarizes the data processing flow for the proposed approach to access hole detection.

Depth disparity. Typically, rubble scene imagery is extremely cluttered and unstructured. A hole, the region of interest, must be isolated from the area around it. Due to the heterogeneous nature of rubble, figure-ground separation (i.e., target entity versus background) of holes and rubble from photometric appearance alone is rendered difficult. Fortunately, RGB-D sensors provide an estimate of metric depth information, i.e., the underlying geometry. The depth information is exploited to partition the image into a set of
superpixels along boundaries that exhibit a strong depth gradient. A publicly available superpixel algorithm is used to partition the image. An inappropriate number of partitions results in a contiguous entity (e.g., a hole) either being undersegmented or oversegmented. Undersegmenting an entity produces superpixels that do not respect boundaries of noncontiguous regions, while oversegmentation subdivides a contiguous entity. The assumption is made that every superpixel overlaps with at most one hole, and the set of superpixel boundaries is a superset of real-world access hole boundaries; these are standard assumptions in the use of superpixels in vision applications [see, e.g., Fulkerson, Vedaldi, & Soatto (2009) and Liu, Tuzel, Ramalingam, & Chellappa (2011)].

The absolute depth value of a region does not alone determine if a region is a hole. A hole by definition must be deeper than its surrounding terrain; as such, it is the depth discontinuity between adjacent regions that is important. For each superpixel, an adjacency graph is built to obtain a list of its neighboring regions. A natural way to express the superpixel image is by an undirected graph $G = (V, E)$, where each vertex, $v_i \in V$, corresponds to a superpixel, and the edges, $(v_i, v_j) \in E$, denote the set of neighboring superpixels. Figure 4 shows an example of the superpixel extraction and neighborhood discovery steps.

For each superpixel, $v_i$, its average depth is compared against all other superpixels that share a boundary with it. Superpixels that correspond to a local depth maximum compared to their neighbors serve as access hole candidates for scoring. The higher the mean depth for a candidate region, the more likely it is indeed an access hole. For each superpixel, a relative depth score, $S_{d}$, is calculated. The depth threshold used for scoring is based on data collected from anatomical models (Panero & Zelnik, 1979). This threshold establishes the minimum depth a region must be from its surroundings to be a valid candidate. A linear score is assigned between 0 and 1 for any relative depth between the minimum and maximum thresholds derived from the anatomical model. Any depth greater than the maximum threshold is assigned a score of 1, and any depth less than the minimum threshold is assigned 0.

**Hole size.** An access hole must have an appropriate size for the insertion of rescue personnel or similarly sized entities, e.g., a search dog or a response robot. Two size-based region attributes are computed: (i) width and (ii) aspect ratio.

The width of the region is determined by fitting an ellipsoid around the superpixel from the metric values provided by the depth sensor and projecting the points to a plane. To exclude outliers, points that lie beyond three standard deviations from the mean are filtered and then projected to a plane. An ellipse is fitted to the point cluster and the major and minor axes are computed. This yields a measure of the width and girth of a region in metric units. For a hole to
be considered appropriate for insertion of a searcher, the width of the major axis and girth of the minor axis were adopted based on anatomical data of the average adult human (Panero & Zelnik, 1979). A region width score, $S_w$, between 0 and 1 is assigned, where a higher score indicates a higher likelihood of accommodating a searcher. A score of 1 is assigned to $S_w$ if the measurements of the major and minor axes are both equal to or greater than the anatomical model. If the axis measurements are 50% (or below), a score of 0 is assigned. To minimize missed detections of holes due to partial occlusion or superpixel oversegmentation, a score is applied linearly between 0 and 1 for measurements greater than 50% of the anatomical model measurements.

To limit the candidacy of holes that may be thin and curvilinear, a score for the aspect ratio of the region is introduced. The aspect ratio score is assigned linearly by calculating the ratio of the area of a given superpixel to the area of the bounding box tightly outlining the major and minor axes. The higher the percentage occupied in the bounding box, the better the candidacy of the detected region. A score, $S_r$, is assigned linearly between 0 and 1 based on the percentage of the bounding box occupied by the superpixel.

Photometric appearance. Examining the depth information alone does not provide sufficient discriminatory information about a hole. To account for this uncertainty, photometric brightness derived from the RGB image is incorporated. It is assumed that access holes are poorly illuminated and thus appear darker in the RGB image. To capture this attribute, two feature scores are introduced: i) absolute brightness and ii) relative brightness.

To compute the absolute brightness intensity of a superpixel, the RGB image is converted to the YUV color space (Black, 2009), and the average brightness from the Y-channel (i.e., the luminance) for each superpixel is calculated, where $Y \in [0, 1]$. To determine the threshold for a valid brightness intensity value, a dataset was compiled from images collected via Google Images (Google, 2014). The dataset contains 118 images depicting collapsed buildings and rubble from disaster scenes. Holes were hand-labeled and the mean brightness intensity was collected. A photometric brightness score, $S_b$, is assigned ranging between 0 and 1, where a higher score is assigned to regions lower than a pixel intensity threshold that was empirically determined from the training data.

Since holes are typically darker than the region surrounding them, each region is also scored based on its relative brightness intensity. Using the Y-channel, the difference between the average brightness of a superpixel and the average brightness of all pixels within (directly) neighboring superpixels is calculated. A minimum threshold was determined empirically using the image training set containing the hand-labeled ground truth. A photometric contrast score, $S_c$, is assigned to a given superpixel between 0 and 1.

2.3. Detection

Each superpixel is assigned a final detection score, $S$. Higher scores indicate a stronger likelihood of a superpixel being a hole. The resulting detection score, $S$, is calculated as follows:

$$S = \sum_{S_i \in F} w_i S_i + b,$$

where $F = \{S_d, S_w, S_r, S_b, S_c\}$ is the set of feature scores, $w_i$ denotes the weighting given to the corresponding feature, and $b$ is a bias term. Each detected access hole is represented by a bounding box that tightly outlines the image region. For each image, the final output of the approach consists of the coordinates of the bounding boxes and their corresponding detection score.

3. EXPERIMENTAL RESULTS

3.1. System setup

Evaluation of the proposed approach was performed on a novel rubble scene dataset (Section 3.2). Throughout the evaluation, the various thresholds of the approach were fixed to the same values for all images. The minimum and
maximum depth used for computing the depth score, $S_d$, were based on an anatomical human model (Panero & Zelnik, 1979) and were set to 200 and 1,951 mm, respectively. The same anatomical model was used to set the minimum width and girth thresholds used for computing the size score, $S_s$, and aspect ratio score, $S_r$, set at 655 and 368 mm, respectively. The threshold used to compute the photometric brightness score, $S_b$, was empirically set to the luminance value of 0.274. Similarly, the brightness difference between a superpixel with its neighboring regions used to compute the photometric contrast score, $S_c$, was empirically set to the luminance value of 0.267. To avoid overfitting to the introduced dataset, an equal weighting of $w_i = \frac{1}{8}$ was given to each feature, and the bias term, $b$, was set to zero. This is done to remain unbiased with regard to features that may be stronger in the limited amount of data available for learning the parameters. Consequently, detection scores range between 0 and 1.

### 3.2. Dataset

The proposed access hole detection approach was evaluated on a challenging dataset containing images of a real rubble scene. Data were collected at the Reference Rubble Pile of the Ontario Provincial Police (OPP), located in Bolton, Ontario, Canada (U.C.R.T., 2013). The rubble pile is used for training purposes, and it consists of heterogeneous terrain comprised of concrete, metal, and wood debris fields, purpose-built simulation buildings, shipping containers, and partially crushed and buried vehicles. Commodity RGB-D sensors, such as the Microsoft Kinect and Asus Xtion, are notoriously sensitive to external sources of infrared light (Ferworn et al., 2011). To minimize the corruption of depth estimates for the experiments, data were captured during sunrise or dusk when the influence of the Sun’s infrared emissions was minimal. The dataset is comprised of 254 image pairs consisting of an RGB image and a corresponding registered depth map, with an image resolution of $640 \times 480$. Out of this set, there are 166 RGB-D images that contain 18 unique holes that meet the definition of an access hole. Ground truth was marked by hand-labeling the location of each access hole with a tight bounding box. Figure 7 shows a sample of the data used for evaluation. The image dataset and ground truth are publicly available at http://ncart.scs.ryerson.ca/research/access-hole-detection.

### 3.3. Evaluation

To quantitatively evaluate the detection accuracy of the approach on the introduced dataset, Precision-Recall (P-R), a standard evaluation tool in information retrieval (Rijsbergen, 1979), is used. The curve captures the tradeoff between accuracy and noise as the detection threshold is varied. “Precision” denotes the number of correctly detected holes over the total number of detections, and it is defined as follows:

$$
\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},
$$

where TP denotes the number of true positives (i.e., correctly detected holes) and FP denotes the number of false positives, i.e., the number of detections where no hole is present. “Recall” is the fraction of true positives that are detected rather than missed, and it is defined as follows:

$$
\text{Recall} = \frac{\text{TP}}{\text{nP}},
$$

where nP is the total number of positives present in the dataset. A detection is considered a true positive if there is a spatial overlap greater than 50% with the hand-labeled ground truth. A detection is represented as a (rectilinear) bounding box that spatially outlines the candidate access hole along with the associated detection score, $S$, and the unique identifier of the image pair.

The detection approach was run on the introduced dataset with ground truth. To oversegment the depth image, a publicly available superpixel segmentation algorithm was used. In particular, the entropy rate superpixel (ERS) (Liu et al., 2011) algorithm was used to produce a user-specified number of superpixels with roughly similar sizes and compact shapes. To evaluate the sensitivity of the proposed approach to the number of selected superpixel segments, the detection approach was run using a range of segmentation targets. To summarize the results for each P-R curve, the average precision was computed over the recall interval 0–1. Table I shows the average precision for the approach using 9, 11, 13, 15, 18, and 20 superpixels. The plot shows that the average precision is stable around 0.37, with nine segments achieving the best result at 0.43 average precision. Figure 5 shows a Precision-Recall plot for the proposed approach using nine superpixels. For extended experiments, please refer to Kong (2015).

The motivation behind the proposed approach is to automatically identify and localize access holes for disaster scenarios, thus this paper is focused on high recall for detections with moderate to high precision. The system performs well in this regard as it is able to detect all labeled ground truth holes with ~0.16 precision when recall is 1, i.e., all holes in the ground truth detected. Precision is lowered by the number of false-positive detections. The ultimate goal is to provide detections to response personnel that correctly

<table>
<thead>
<tr>
<th>number of superpixels</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>15</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>0.43</td>
<td>0.37</td>
<td>0.37</td>
<td>0.36</td>
<td>0.38</td>
<td>0.36</td>
</tr>
</tbody>
</table>

---

Kong et al.: What is a Hole? Discovering Access Holes in Disaster Rubble with Functional and Photometric Attributes

identify all access holes with minimal false detections. Since a missed detection can result in the potential loss of life, a high false-positive rate is accepted so that no potential access holes are excluded. Figure 6 shows sample detection outputs. Upon examining the detections, it is found that a number of false-positive detections occur in areas where the geometric features score high, but they are not excluded through the scoring of photometric properties. Nonuniform weighting of the various feature scores via learning may ameliorate some of these issues; however, a lack of sufficient training data currently limits the ability to tune the system without overfitting to the current dataset. Ultimately, these false positives can be rejected by further visual inspection with minimal effort, as compared to evaluating all inputs manually.

Experiments were performed with an unoptimized MATLAB code running on a 64-bit Intel Core I5 2.50 GHz machine with 6 GB of RAM. To detect holes in a single RGB-D image with a resolution of 640 × 480 segmented into nine superpixels, the system requires ∼9 s. Increasing the number of superpixels to 20 yields a runtime of ∼14 s per input image pair. Significant runtime improvements are anticipated via optimizing the code and leveraging parallel computation, e.g., a graphics processing unit (GPU).¹

The intended use for the proposed approach is to create access hole information for first responders in-transit to a disaster scene. The “realistic time-frame” should be considered to be on the order of many hours, e.g., the main body of Canada Task Force 3 (Toronto HUSAR) arrived at the Algo Centre Mall Collapse roughly 14 h after their activation (Belanger, 2014).

Figure 5. Evaluation of the overall approach for detecting holes. The Precision-Recall curve is computed across the entire introduced dataset; the number of superpixels is set to nine.

Figure 6. Sample output of the proposed access hole detection approach. Left: Input RGB image, middle: superpixel segmentation with the ground truth label given in red, and right: detected regions given in green. The first two rows show successful detections, and the last row shows a successful detection with a false detection. This figure is best viewed in color.

¹The intended use for the proposed approach is to create access hole information for first responders in-transit to a disaster scene. The “realistic time-frame” should be considered to be on the order of many hours, e.g., the main body of Canada Task Force 3 (Toronto HUSAR) arrived at the Algo Centre Mall Collapse roughly 14 h after their activation (Belanger, 2014).
Figure 7. A sample of RGB and depth image pairs from the introduced dataset used to evaluate the detection approach. This figure is best viewed in color.
4. DISCUSSION AND SUMMARY

This paper is the first to present an automated system for the detection of access holes in rubble. The approach has shown promising results for detecting access holes using both functional and photometric attributes for the insertion of search personnel into holes within rubble.

A current limitation of the approach is the need for accurate depth maps. Current commodity RGB-D cameras do not work well outdoors in full daylight conditions. This is a well-known challenge within the field robotics community. To date, passive stereo-based algorithms have not achieved the same level of depth accuracy as RGB-D sensors. The consideration of other, more sophisticated sensors is possible in the future. An alternative approach is to investigate ways of improving the depth data estimates, such as integrating the data over time rather than sampling a single frame. This approach could reduce the number of areas with missing or corrupted depth estimates.

The lack of a large real-world dataset is a current limiting factor. While the dataset introduced in this paper takes a first step, it is insufficient for providing examples of the multitude of debris configurations that rubble fields can present. Furthermore, the limited amount of data restricts tuning the weight of features when calculating the detection score. This paper purposely remains neutral regarding these weights to avoid the problem of overfitting performance biased to the current dataset. The availability of other disaster scene datasets would allow for learning the weight parameters and thus improve performance. In addition, a more diverse dataset would also provide a more thorough evaluation of the approach. Overall, as more data become available, improvements in the algorithm may be realized.

The approach presented in this paper can be further developed by improving the identified feature attributes and augmenting the set with additional ones. For instance, holes in rubble tend to have different thermal properties from the terrain surrounding them (Matthies & Rankin, 2003). The use of forward looking infrared (FLIR) sensors to detect secondary thermal effects present around potential holes with humans inside may help reduce errors.

Improving the approach to run in real-time as data are being captured onboard a UAV can provide numerous benefits to search teams. GPS coordinates obtained from the UAV can be transmitted wirelessly to ground crews, allowing USAR teams to mark areas that require further investigation quickly and accurately as they are detected.

There are numerous positive implications of the current contribution. First, the introduced approach may reduce the need for the dangerous task of humans performing initial visual inspection of an urban disaster incident in order to find potential areas of access. Second, the approach may be able to reduce the cognitive load of response workers tasked with identifying access points through visual inspection. Third, a UAV can investigate regions beyond line of sight, i.e., it can search and analyze areas that might not have been accessible before. Finally, significant reductions can be made in the search space of a large collapse to a manageable number of locations, thus saving time. Search and rescue operations are extremely time-critical, as the life expectancy of victims under buried rubble is limited. Identifying and localizing access holes in this way makes better use of limited time.

Since the detection approach is intended for planning purposes, it provides a search team with advanced warning of “potential” access paths that can then be prioritized by human search specialists. The intent is to provide a means of indicating holes that can then be explored or eliminated from further consideration by expert human practitioners (the search team). Furthermore, the intention is to include this information in a physics-aware disaster scene model (Ferworn et al., 2013), with the hole information represented and clearly marked for searchers inside the simulation. It should also be noted that this technique of collecting hole data and rendering a scene model would ideally be used by the advance parties of the task force or the local first responders at the scene. These on-scene teams would then transmit the simulation to inbound task forces whose search teams and structural specialists would use the data as input to form their plans prior to arriving at the scene.

In summary, this paper presented a novel approach for the automated detection of access holes in scenes of rubble. Access holes represent areas of particular interest for first responders. They represent the possibility of accessing subsurface voids where live humans may be hidden. This paper is the first to define the characteristics of an access hole through both functional and photometric attributes inherent to a valid entry point. A novel approach for identifying candidate access holes in RGB-D data was proposed, a real rubble pile dataset was introduced, and an evaluation protocol to validate the approach was provided. Empirical evaluation has shown promising results for detecting access holes.

REFERENCES


mapping and tracking. In IEEE International Symposium on Mixed and Augmented Reality.