https://doi.org/10.1016/j.tust.2020.103355

GPR pattern recognition of shallow subsurface air voids

Tess X.H. Luo, Wallace W.L. Lai*

Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Hong Kong
* Corresponding author: ZS615, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. email: wllai@polyu.edu.hk

Abstract: Countless subsurface voids in urban areas of cities threaten people's lives and property. A workflow for automatically identifying subsurface voids from ground penetrating radar (GPR) data was developed in this study. The workflow consists of 3 stages: locating voids automatically from C-scans, then verifying voids from corresponding B-scans, and finally making judgements based upon the previous 2 sets of results. This study adopted 2 (Lai et al., 2016) approaches: approach 1 quantified the GPR response of air voids using forward modelling, while approach 2 used workflow prototyping and validation with inverse modelling. Forward simulations indicated that different ratios of void size to GPR signal footprint could result in a variety of patterns in B-scans: they can be hyperbolas, cross patterns, bowl shaped patterns and reverberations. With a database of void patterns of both C-scans and B-scans established, in approach 2 the workflow uses a pyramid pattern recognition method – with pixel value or gradient being used for feature identification – to search automatically for air-filled void responses in GPR data. The workflow was tested using 2 laboratory and field experiments and the results were promising. The constraint values proposed by the 2 experiments were validated with another site experiment. Given the huge workload involved in city-scale subsurface health inspections, a standardized workflow can help improve efficiency and effectiveness of subsurface void identification.

Keywords: Ground Penetrating Radar; subsurface air void; pyramid method; pattern recognition

1. Introduction

Metropolises such as Hong Kong have complex utility networks buried underground. These utilities ensure that cities function effectively have a crucial influence on the daily lives of citizens (Lai et al., 2017; Lai et al., 2016). Among these utilities, water supply network degradation has become a growing issue, and urban hazards like the sudden bursting of water mains and seepage in busy roads are frequently reported (Farley et al., 2001). Subsurface pipe leakage causes underground wash-out, leaving an air void when water is drained and finally resulting in road collapse when the damaged area can no longer support the heavy load of the pavement structure and its traffic. Over time, countless numbers of air voids develop beneath a city's road network, which threaten the safety of the citizens' lives and property. Conducting regular health checks on road surfaces in order to diagnose subsurface voids can help predict potential urban hazards and contribute to effective and efficient utility management.

The management of these utility networks is difficult and complex. Non-destructive testing (NDT) and near surface geophysics (NSG) enables operators to see the unseen subsurface world without excavations. Ground penetrating radar (GPR) is an NSG technique that is based upon the propagation, reflection and measurement of electromagnetic (EM) waves (Annan, 2004; Jol, 2009). GPR has proven to be a time and cost-efficient method in infrastructure inspection applications as it can provide high-resolution imaging of the subsurface world(Annan, 2002). There is a substantial body of past research that has demonstrated GPR's capabilities in detecting subsurface air voids.

Although thanks to the development of antennas and control units, the survey time required by GPR has been significantly reduced, it is still difficult to carry out city-wide GPR surveys. The complicated subsurface environment distorts GPR responses, and the analysis of GPR patterns is therefore still mainly reliant upon human visual interpretation. If large-scale GPR surveys are conducted, there might be dozens or even hundreds of GPR profiles requiring analysis. Recent research has focused upon automatically interpreting GPR responses using pattern recognition techniques (Al-Nuaimy et al., 2000; Ayala-Cabrera et al., 2011; Gamba & Lossani, 2000; Ghasemi & Abrishamian, 2007; Pasolli et al., 2009; Xie et al., 2013). Applications of automatic recognition techniques mainly focus on the more typical kinds of GPR responses: namely, hyperbolas. In contrast to hyperbolic reflections generated by point reflectors, reflections from voids have no fixed morphology, as the subsurface voids themselves come in various sizes and shapes. The methods used in past research, such as Hough transform (Simi et al., 2008), support-vector-machine (SVM) (Xie et al., 2013), and neural network (Gamba & Lossani, 2000), all demand sufficient training in order to obtain a valid template. In contrast, due to its developmental maturity and computational efficiency, a much simpler pyramid-based pattern matching approach is explored in this study.

The flexible template of pyramid-based matching is also more suitable for the detection of subsurface voids of various shape. This study was therefore aimed at developing an automatic void identification workflow using pyramid pattern recognition.

The essential criteria for defining a void concern its extent and nature–i.e., whether it is filled with air or water. Before conducting pattern recognition, the GPR responses of various kinds of subsurface voids were firstly investigated and quantified using forward modelling to predict how air-filled subsurface voids would look in 2D (B-scan) and 3D (C-scan) representation of GPR data. Next, templates were constructed to facilitate further pattern recognition. Then an integrated pattern recognition workflow was developed and prototyped in order to conduct inverse modelling for void identification. Controlled experiments were then carried out to validate this workflow.

2. GPR response of air-filled voids

2.1. GPR principle

The electromagnetic waves produced by GPR instrumentation are emitted from a transmitting antenna. The signal then spreads out into the ground in a downward conical form and penetrates through any subsurface layers present. When the EM wave encounters any electrical parameter contrast in the ground, it is backscattered and received by receiving antenna. The main electrical parameters of any medium, i.e. the permittivity, the permeability and the conductivity, are generally a function of frequency. Contrasts in permittivity are the main cause of reflections of the radiated EM waves. The change in conductivity affects absorption of the radar signal by the medium, whilst the variation in permittivity determines variations in the impedance characteristic of the medium (Annan, 2004; Benedetto & Pajewski, 2015).

A, B and C-scans are three forms of GPR data presentation from 1 to 3 dimensions, respectively. An A-scan denotes a single waveform recorded by GPR at a particular position. When moving antennas along a traverse, a set of A-scans form a vertical section through the ground, and this section is called a B-scan. C-scans map 3D data volumes comprising multiple data 'slices', each of which is a horizontal plan for a plane at a certain depth. A C-scan is formed by stacking multiple B-scans collected in this horizontal plane, and then plotting the amplitudes of the recorded data at a given time (Goodman & Piro, 2013). C-scans show reflection intensities from an overall perspective, while B-scans offer full waveform information in vertical sections. The generation of C-scans follows a standardized workflow developed by Luo et al. (2019).

2.2. Air-void pattern

The inner boundary of an air-filled void causes reverberation or 'ringing' of the electromagnetic waves (Kofman et al., 2006). Although many kinds of buried infrastructure may cause reverberation patterns in B-scans, for example manholes and metal plates, the reverberation signals from these objects start from time-zero, whereas those from air voids continue attenuating in a time/depth window that is not close to the surface (Lai et al., 2017).

The significant permittivity contrast between soil and air also produces a strong reflection intensity. In a C-scan, reflection intensities (pixel value) of air voids are significantly higher than that of the background. Subsurface voids are normally irregular in shape and present as a local but discontinuous object in a C-scan image at a certain depth, whereas a continuous reflection in a C-scan is most likely generated by an underground utility instead (Lai et al., 2017).

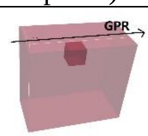
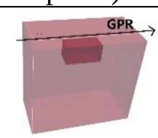
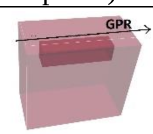
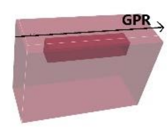
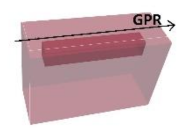
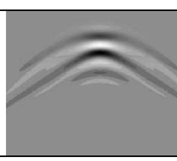
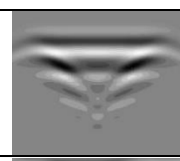
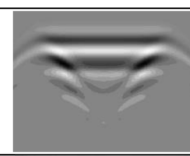
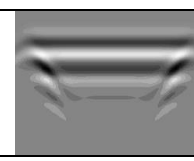
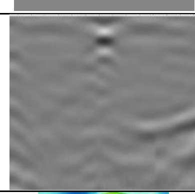
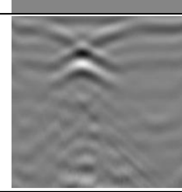
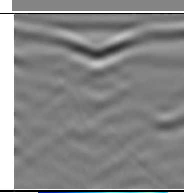
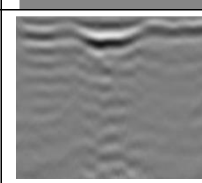
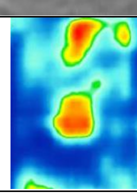
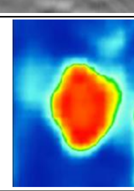
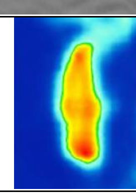
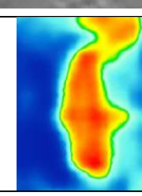
Much research has been conducted aimed at studying the GPR response of voids. Clemeña et al. (1986) and Plati and Dérobert (2015) investigated the feasibility of detecting voids in concrete with GPR. Xu et al. (2010) used GPR to detect several common types of subsurface voids and found that cracks yield hyperbolic responses. Lai et al. (2017) validated the reverberation pattern of voids through three case studies. Casas et al. (1996) indicated that when the void is very small compared to GPR wavelength, diffracting hyperbolas happen in GPR radargrams, while, in contrast, big voids cause irregular signals with chaotic reflections and a reduction in the received frequency. Kofman et al. (2006) simulated GPR responses using GPRMax and pointed out that reverberation occurs only when the void size is sufficiently large when compared with the GPR wavelength. It is believed that the strong reflections with reduced frequency with air voids are caused by the reverberation of the EM wave (Kofman, 1994). All the above research has demonstrated that the general appearance of air voids in GPR data, that is attenuating reverberation. And the GPR response of different voids is subject to a number of factors, they might be geometry, size and depth, etc., but clear definition is still lacking. Besides, the use of automatic recognition is highly dependent upon a prior knowledge of void presentation; therefore, testing and definition of void presentation are an essential foundation upon which pattern recognition can be based. In order to ascertain the presentation of voids for use in pattern recognition, the GPR response of voids with diverse morphologies were further investigated using the GPRMax simulation.

2.3. Forward modelling: GPR response simulation

Drawing on the above research, a void's extent within a survey traverse has a significant impact on the GPR response pattern, so simulations must focus on quantifying the relationship between void size and GPR response. Since subsurface voids would be collapsed before they are excavated, ground truths of voids morphology are hard to be obtained. Simulations were conducted with GPRMax, a Finite-Difference Time-Domain method (FDTD) simulation package that allows users to simulate the GPR response to the subsurface world (Giannopoulos, 2005; Warren et al., 2016).

Subsurface voids with varying horizontal spread were created in the Underground Utility laboratory of the Hong Kong Polytechnic University (PolyU) (Wu, 2015). The tank was filled with garden soil. GPR profiles were collected with an IDS 600MHz system using a 10cm profile spacing within an orthogonal grid, and C-scans of each void were generated. Meanwhile, GPRMax simulations were conducted, imitating the laboratory environment; such that voids or varying horizontal and vertical size were located at shallow depth (10cm) within a soil environment. Simulated signals were transmitted and received by a 600MHz common offset antenna unit with 15cm antenna offset, according to the specification of IDS 600MHz antenna. The laboratory experiments and simulations produced similar GPR responses. Four typical void patterns were identified in the B-scan data – hyperbola, cross, bowl shape, and reverberation - and these patterns appeared in succession as the void's spread grew, as can be seen in the 3^{rd}-4^{th} rows in Table 1. In this way, the pattern templates were created in a relative homogenous environment, free from any interference from scatterers.

Table 1 Forward modelling of voids with different horizontal spread

Void spread (cm)	10 (1* void spread)	20 (2*void spread)	40 (4*void spread)	50 (5*void spread)	60 (6*void spread)	80 (8*void spread)
Void Photo						
FFZ (600MHz)	10cm					
Simulated Pattern						
Lab Experiment (B-scan)	No data	No data				
Lab Experiment (C-scan)	No data	No data				
Pattern	hyperbola	cross	cross	bowl	bowl	reverberation

It is noticeable in Table 1, that in parallel with the increase of void horizontal spread, the hyperbolic reflections from the void's edges also become more widely separated. When two hyperbolas overlap, cross patterns occur. Whether two signals can be distinguished depends on GPR spatial resolution, while GPR horizontal resolution is determined by the footprint of a GPR beam. The radar beam width determines the smallest footprint, therefore a narrower beam width leads to a smaller footprint, and provides better spatial resolution. Another important criteria that affects the spatial resolution is the spatial sampling in imaging. The spatial sampling must be smaller than the radar footprint, so as to make the best of the radar beam width. In this study, the radar footprint is within dozens of centimeters, which is significantly larger than the spatial sampling (200 scans/m), thus only radar footprint is calculated to describe the spatial resolution of GPR C-scans. There is a preliminary piece of research that established a simplified model of radar horizontal resolution, based on the relationship between feature size and radar resolution (Annan & Cosway, 1992). The footprint is usually estimated as the First Fresnel Zone (FFZ), and there are various equations to estimate the FFZ in the far field (Leckebusch & Peikert, 2001; Leucci & Negri, 2006; Leucci et al., 2003). Among these equations, Equation [1] was chosen for this study for its stable performance (Pérez-Gracia et al., 2008). The estimation of spatial resolution is always rough as it must take into account a number of factors, such as antenna design, frequency, beam angle and host medium properties.

$$\mathrm{r}(v,z,f) = \sqrt{\left(\frac{v^2}{16f^2} + \frac{vz}{2f}\right)} \qquad [1]$$

Where r is the radius of FFZ, z denotes depth, f is the standard for dominant frequency, and v is GPR wave velocity.

Because the footprint is dominated by wavelength and depth, and the void pattern is wavelength dependent, the footprint is calculated as 10cm (Table 1) in order to take the influence of void depth into consideration. It is observed that when the void spread is smaller than r, the GPR response is presented as a point reflection – a hyperbola. When the void spread increases to 2-4 times r, a cross pattern occurs in the radargram. The bowl-shaped pattern is expected when the void spread is larger than 4 times r. These rules can be applied in an inversion study, in which the void size is estimated from the GPR response. In terms of C-scans, voids with different spread present as localised strong reflections of different size, as shown in the bottom row of Table 1.

It is also noted that the vertical thickness of the void does not have an obvious impact on GPR response, unless it is significantly smaller than wavelength (Casas et al., 1996). Along with the increase of void vertical spread, the pattern in the radargram did not change noticeably, except that the reverberation range along the time window may increase (Kofman et al., 2006). A database consisting of both simulated and experimental patterns was established to provide templates for further automatic pattern recognition.

3. The void identification workflow prototype

The developed workflow includes 3 stages: 1) roughly locate the void and estimate void size from a C-scan; 2) inspect the corresponding B-scans across the suspected void location and estimate the void size based on the matched pattern; and 3) cross-validate the results of the two previous stages, and select the most convincing void size and void material, while giving priority to the C-scan image.

The detailed workflow is shown below in Figure 1:

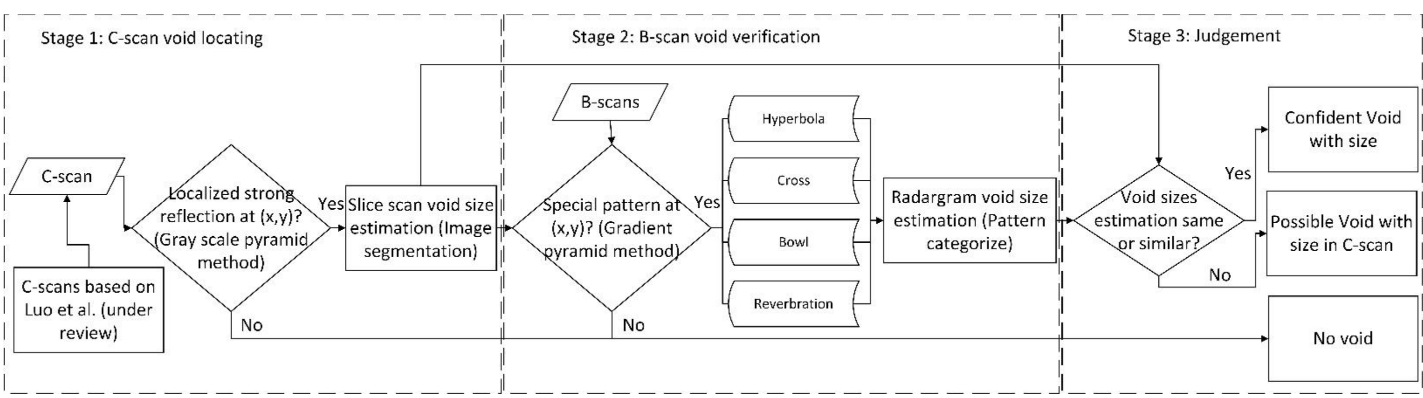


Figure 1 A 3-stage pattern recognition for void identification

Stage 1: C-scans void locating.

C-scans, as a kind of 3D representation, provide a general but intuitive view of the subsurface. Given that the dielectric properties of air and the garden soil host medium are different, the GPR reflection intensity of voids, which the C-scan is mapping, is often strong enough to be visible. It is therefore relatively straightforward to establish the approximate position of air voids within a survey area. And as shown in Table 1, voids are presented as local region with high reflections.

Furthermore, any suspected voids in C-scans can be extracted using the image segmentation method. The size of the void is estimated by multiplying the scale of the pixel size to the ground distance and by counting the number of pixels.

Stage 2: B-scans void verification.

C-scans only provide information based on the normalized reflection intensity. Many kinds of local reflectors may generate similar patterns in C-scans if they have similar dielectric properties. A further verification using B-scans is therefore essential. Given that B-scans display the full GPR waveform, and subsurface voids generate reverberation-like responses, confidence can be enhanced by examining whether a localised response occurs in the suspected void's position. Since different void characteristics, such as size and material, may yield various reflection patterns, a void's description can be estimated.

Stage 3: Judgement.

In this stage, 2 sets of results (from C-scan and B-scan) are presented and compared. A void is most likely to exist when it is identifiable in both C-scans and B-scans. In addition, further evidence is provided if the void size estimations in the two previous steps are similar. In this case, C-scans and B-scans both support the same judgment and hence the highest identification confidence is given to the void. On the contrary, if no void pattern is found in B-scans that corresponds with the location of a suspected void in C-scans, then the anomaly may result from another kind of reflector. If a void is found in both C-scans and B-scans, but the sizes estimated by the two steps are significantly different, then the size estimated from the C-scan is adopted. This is because C-scans offer an overall perspective view of the survey area, while B-scan estimation relies entirely upon the void's spread along the profile section.

According to the designed workflow, it is critically important to define whether a specific void pattern is identifiable in the data. The pyramid method is applied in this stage in order to conduct automatic pattern recognition. In align with the workflow, a decision-support program, which integrates automatic void recognition from both C-scan and B-scan data, was developed with LabVIEW. LabVIEW is "a system design platform and development environment for a visual programming language from National Instruments" ("LabVIEW – See it. Solve it.s," 2018).

3.1. Automatic pattern recognition using pyramid method

In stage 1 and stage 2 of the workflow, pattern recognition is conducted in both B-scans and C-scans using the pyramid-based pattern recognition method. The pyramid-based method is integrated within the LabVIEW system. It was observed that the greyscale pyramid method performs best for void identification in C-scans, while the gradient pyramid is more suitable for B-scan pattern recognition.

Pattern recognition includes 2 phases. Firstly, the algorithm learns the characteristics of a specific pattern using a template. A description of the pattern, including the region of interest, is then constructed. Next, in the matching phase, the algorithm searches for specific patterns within a data image.

In the learning phase, the pixels of both data images and template images are resampled to construct the pyramid. The resolutions of both data images and templates are reduced to 4 lower-resolution levels using Gaussian pyramids. In a Gaussian pyramid, the original image is continually convolved and subsampled in one-octave step with a Gaussian kernel, then the resulting image of this process is a low-pass filtered copy of the original image (Adelson et al., 1984; MacLean & Tsotsos, 2008). This process is repeated 4 times to obtain a sequence of smoothed images, which constitute the representation model built in the learning phase. Since the pyramid representations are built for both template images and data images during the matching stage, the algorithm conducts pattern correlation computation with a coarse-to-fine approach, whereby the search starts from the highest pyramid level which has lowest resolution. The degree of correlation between the image information of the data pyramid and template pyramid at same level is then estimated. The correlation score is repeatedly computed for the different pyramid levels until the desired score is achieved. Two kinds of image information can be used to compute the correlation: grayscale values representing GPR reflection intensities, and gradients that describe the pattern geometry and edge information. The process is demonstrated in Figure 2.

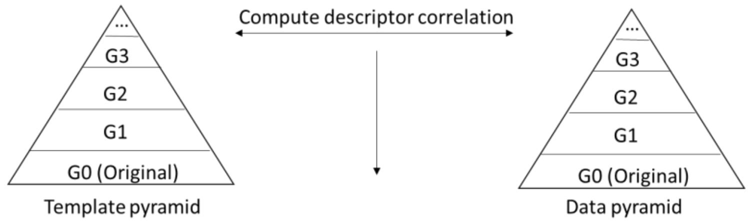


Figure 2 Coarse to fine approach for matching template pyramid with data pyramid

In the search phase, when the greyscale image's pixel value is considered to be the image information, the normalized cross correlation (NCC) is calculated to establish the correlation score between the template and data images, as shown in Equation [2]:

$$C = \frac{1}{n}\sum_{x,y}\frac{1}{\sigma_f \sigma_t}\left(f(x,y)-\overline{f}\right)\left(t(x,y)-\overline{t}\right) \qquad [2]$$

where n is the number of pixels in the template image $t(x,y)$, $\overline{f}$ is the average pixel value of the data image $f(x,y)$ and $\sigma$is the standard deviation of f. It should be noted that the correlation is computed based on a template image that should be smaller than the data image, and the template image is moved across the data image in order to find the optimal matching score.

Therefore, the greyscale pyramid method places emphasis upon the distribution of normalized pixel grey values. It is helpful when the pattern presents as a specific greyscale shade but has no particular shape or sharp edges. As a consequence, the greyscale pyramid is sensitive to brightness change. As discussed in the previous section, air voids present as regions with locally high reflection intensity, which occurs due to the comparatively significant reflection contrast between an air void and the host material, and, thus the greyscale pyramid method is more suitable for C-scan detection.

If the gradient is used as the image information, vector correction is then applied to calculate the correlation between the template image and data image. The vector component of images describes the gradient direction of each pixel. The vector correlation calculation is shown in Equation [3]:

$$C = \frac{Cov(t,f)}{\sigma_t \sigma_f} \qquad [3]$$

where $Cov(t,f)$ is the covariance between template image $t(x,y)$ and data image $f(x,y)$, and $\sigma$is the standard deviation of f. The same moving kernel is applied in order to find the optimal position with the highest matching score.

As the gradient pyramid focuses on filter edge vectors, it is more suitable for templates that have a clear structure and obvious edge. Even though it is insensitive to an intensity change, the gradient pyramid demands higher image resolution, because the strength and reliability the of edges reduces at very low resolutions. As discussed above, for pattern recognition in B-scans, changes in void size lead to different GPR response patterns. These patterns mainly differ in terms of their shapes and structures, and in this case the gradient pyramid possesses better recognition performance.

When searching for rotated matches, a coarser angle is preliminarily adopted and then the rotation is refined with smaller angle step sizes ("NI Vision - Pattern Matching Techniques," 2018; Pavlidis, 2013). Since templates are analysed across all pyramid levels, the approach is scale and angle invariant.

3.2. Image segmentation with Otsu's method

As shown in Figure 1, void size is estimated from C-scans using the image segmentation technique. Unlike traditional remote sensing images that are composed of multiple bands, GPR C-scans present only single band reflection intensities. In addition, due to the diffraction of GPR signals at the edges of an object, its boundary is not necessarily sharp in GPR C-scans. Histogram thresholding, therefore, is the most straightforward approach for object extraction from C-scan images. There has been much research conducted on image segmentation, including the following: Laplacian and gradient counts of greyscale values focused upon the identification of maximum degrees of difference (Gou et al., 2013); Gaussian determination of edges based on image frequency (Permuter et al., 2006); and K-means and Otsu's clustering-based threshold definition (Lee et al., 1990). Among these methods, Otsu's method is adopted in this study for its computational efficiency. Besides, Otsu's method evaluates global image pixel value distributions, which requires no preliminary knowledge of the object reflectance, while, in contrast, some other methods take adjacent pixel values into consideration.

Otsu's method is clustering-based and is widely used to conduct image segmentation in an automatic and unsupervised manner (Sezgin & Sankur, 2004). The objective of image segmentation is to define the threshold at which pixels can be classified into two groups: foreground pixels and background pixels (Lee et al., 1990). The algorithm firstly computes an image's histogram and probabilities for each intensity level. When two initial classes are established, the algorithm iteratively computes the class mean and class probability. Then an optimum threshold can be obtained through minimizing the weighed sum of within-class variances in order to calculate the maximal inter-class variance (Otsu, 1979; Sezgin & Sankur, 2004). More details concerning the algorithm can be found in Otsu (1979).

In addition to the tail of the GPR response generated by a buried object, a coarser profile spacing and interpolation in the C-scan processing would also both lead to an imaged object appearing to be larger than its real size. Therefore, image erosion using a 6*6 structure element is also introduced during image segmentation. The size of a structure element is defined based upon the resolution of GPR C-scans: given that the spatial resolution of GPR C-scans is normally within centimetres, the structure element should not be larger than the size of the smallest detectable objects, otherwise they might be removed from C-scans.

3.3. Inverse modelling: Cross validation and decision-support system

The interface of the designed workflow prototype is shown in Figure 3. In the designed LabVIEW interface, three stages of void identification workflow are separated into three sub-interfaces. Based upon the simulations and experimental results in section 2, two databases of void patterns for C-scans and B-scans were constructed, as shown in Table 1.

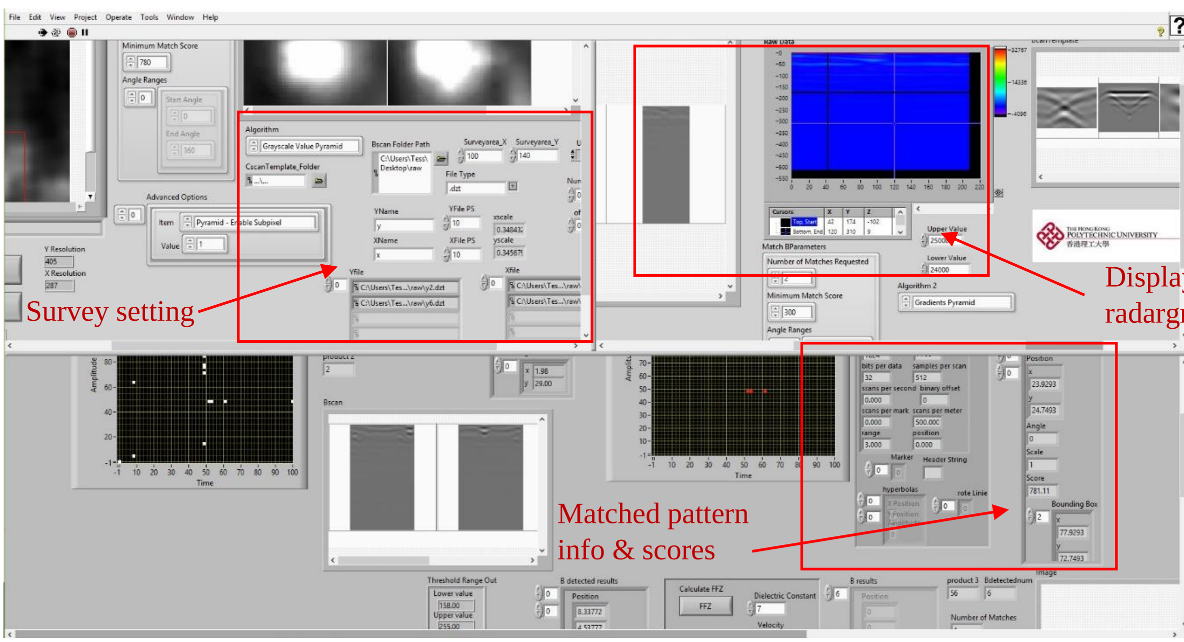
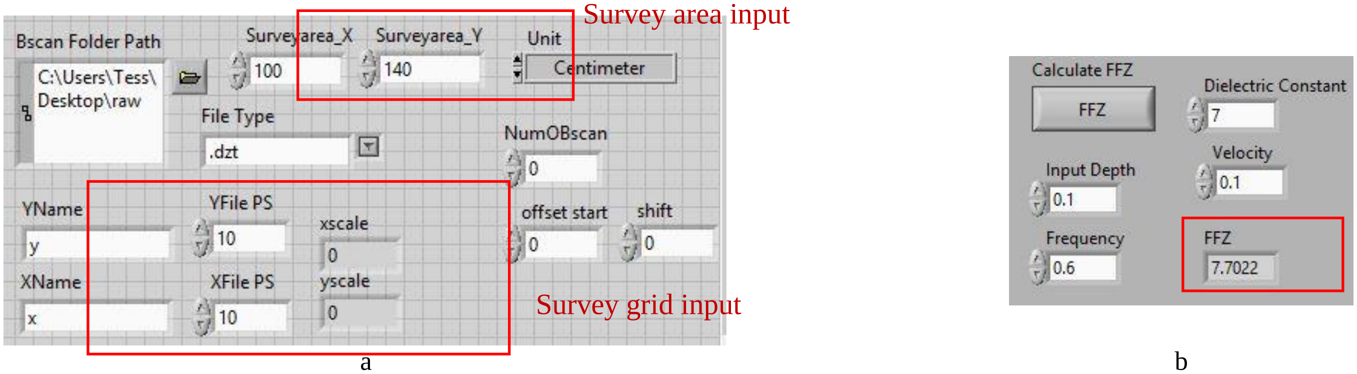


Figure 3 Illustration of interface of designed workflow prototype

For stage 1, the survey C-scan corresponding to the suspected void's depth is input, and the greyscale pyramid method is selected as the desired algorithm. The program will retrieve void patterns from the C-scan database and search for the optimal match. As discussed in section 3.2, the matching score is computed in order to determine the correlation between the template and the inputted data image. A constraint can be defined to filter out results with lower matching scores. Through numerous experiments, it was possible to demonstrate that abnormal reflections in C-scans were easily detected with higher matching scores, so that a higher threshold – above 700 out of 1000 – on the matching score is preferred if an optimal matching result is desired. This is because in C-scans reflection scatters may frequently occur, and they are very unlikely to present an identifiable structure.

Having input the survey grid information, as shown in Figure 4 a, the C-scan image is georeferenced and the coordinates of pixels are transformed into a real world coordinate system. The position of a void in terms of a real world coordinate system is thus defined and can be displayed in a plan, and the estimated void sizes in the real world should match. With the help of Otsu's method (inter-class variance), the program computes the void's edges based on the histogram distribution. A bounding box depicts the whole area of the detected void and the centre of the bounding box refers to the centre of the void. The radargram profiles that are closest to the centre point are indexed. For each void, two radargram profiles are selected for a gridded survey using x and y direction GPR survey profiles. These selected radargrams are then read and transferred to pixel-image format for further B-scan pattern recognition.



a b

Figure 4. a presents the survey grid information setting interface while b shows the interface of the radar footprint calculation

In the B-scan verification step (stage 2 in Figure 1), the approach is mainly the same as that adopted for identifying void location in C-scans, only in this case the gradient pyramid method is adopted when matching patterns. For each selected B-scan, the program visits the B-scan patterns in the database one by one in order to find a match. The GPR reflection is affected by various factors, such as adjacent reflectors, profile orientation, and material properties. It is therefore difficult to have templates that match with GPR responses. Therefore, a lower constraint on matching score – 200 out of 1000 for instance – is suggested to avoid missing patterns. For all detected patterns, only those falling within the bounding box defined by the C-scan pattern are considered valid. The detailed flow of this loop is shown in Figure 5. According to the regularity of the GPR response on air void, as discussed in section 2, void spread along the direction of radargram profile is estimated.

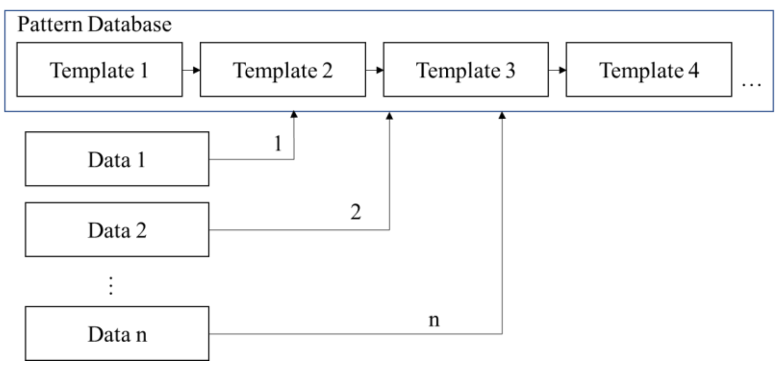


Figure 5 Illustration of the matching scheme of both C-scan and B-scan pattern recognition

By inputting GPR parameters as in Figure 4 b, the FFZ is calculated according to Equation [1], and then the void size can be estimated based on the relationship between void response and void spread. Armed with the estimates of void position and void spread from both C-scans and B-scans, the operator can then make preliminary judgements regarding the void. The whole program imitates the human judgement process, but leaves the heavy cross-checking work to the program once the GPR survey setting information has been input.

4. Case experiments

Two experiments were conducted to test the workflow: a laboratory experiment within a controlled environment and an outdoor site experiment. Then another validation experiment was conducted to test the constraint values. There were 2 criteria used for evaluating the precision and accuracy of the void pattern recognition:

1) the pattern detected was generated by the air void;
2) the type of the pattern matched with the void size.

For each criterion, the results can be categorized into 4 classes (Powers, 2011):

True Positive (TP): void existed, and workflow claimed it existed	False Positive (FP): void did not exist, but workflow claimed it existed
False Negative (FN): void existed, but workflow claimed it did not exist	True Negative (TN): void did not exist, and workflow claimed it did not exist

The recognition sensitivity was evaluated using a true positive rate (TPR) and false positive rate (FPR), as described by Equations [4] and [5] respectively, and a higher TPR means a higher sensitivity (Fawcett, 2006).

$$\text{TPR} = \frac{TP}{(TP + FN)} \quad [4]$$

$$\text{FPR} = \frac{FP}{(FP + TN)} \quad [5]$$

The selection of a matching score constraint is critical in recognition performance, especially for B-scan patterns. Multiple adjustments may be necessary before an optimal constraint figure is established. The use of a receiver operating characteristic curve (ROC) was used to find the best constraint. The ROC plots the TPR against FPR, and the data point – constraint – that is closest to the top left-hand corner (TPR =1, FPR = 0) denotes the perfect result; however, all results are correct, and none are left out. Then the distance of each value point from the perfect result point (0,1) was calculated using Equation [6] in order to find the smallest variance and thus the optimum constraint (Fawcett, 2006).

$$\text{Optimum} = \min_{i} \sqrt{(FPR_i{}^2 + (1 - TPR_i)^2} \quad [6]$$

where i denotes each data point in the ROC plot.

4.1. A laboratory experiment

Similar to the experiments in the forward modelling (section 2.3), another void was created in the soil tank of the Underground Utility Survey Laboratory in PolyU. The void spread measured 15*25*15cm in length, width and height respectively, and was surrounded by dry garden soil whose relative permittivity was defined as 7. A GPR survey was conducted using a 900MHz GSSI unit within an orthogonal grid. The void spread is therefore significantly larger than the GPR footprint, so that reverberations and cross like response was expected. The survey site measured 100*140cm and the profile spacing was 10cm. Having undergone basic signal processing, such as de-wow, gain, bandpass, time zero correction and velocity analysis, the GPR profiles were stacked together to construct C-scans based on the standard workflow for C-scan generation. A surface void was imaged in C-scans images, as shown in Figure 6. The reflection intensity of the void is significantly higher than that of the background, and the edge of the void is blurred.

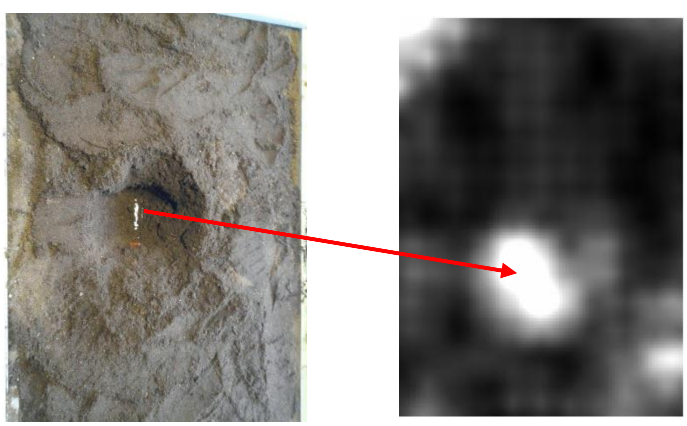
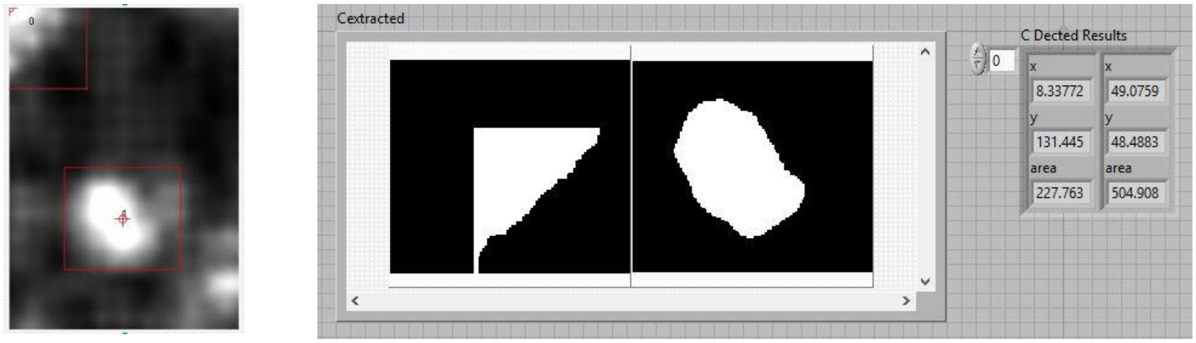


Figure 4 The surface void dug in the Laboratory and its C-scan

In the C-scan pattern recognition, 2 air voids were identified: one is the target void in the centre of the survey site (47, 48), and another is a manhole in the top left-hand corner, with the coordinate as (10, 130) of the survey site. They both present as strong local reflections in C-scans. The position of void centre and void size estimation are displayed in Figure 7.

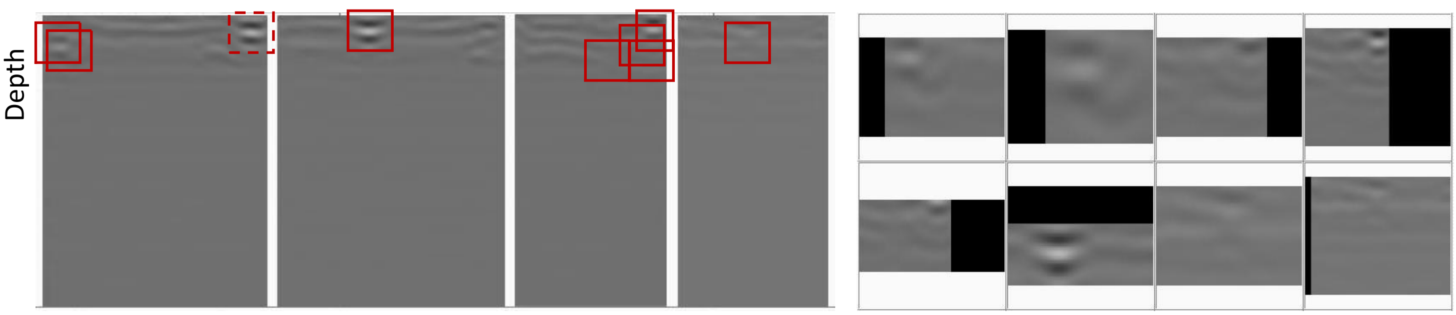


a b

Figure 5 a shows 2 identified voids in the laboratory experiment as in red rectangle and b presents the extracted voids with their size and position information. Remarks: column x and y present the coordinates of the void centre, while ‘area’ shows the void size estimation (cm^2) of each detected void

Even though the void shape was not a regular circle like that of the template, it was successfully identified. The constraint matching score was 780 in C-scan pattern recognition, only the 2 largest anomalies were extracted, and smaller scatters were excluded. When compared with the void’s true size, the void size estimations were slightly bigger. The inhomogeneous reflections surrounding the void contributed to incorrect boundary definition during void extraction, and some adjacent responses were also in error included, as indicated by red rectangles in Figure 7 a.

For the 2 detected “voids”, 4 GPR radargrams were indexed. A total of 15 constraint values were tested, ranging between 150-350. The performance results with the ROC are displayed in Figure 9. According to Equations [4-6], when the constraint was set to 200 (shown as a red dot in Figure 6), the workflow performed the best, which resulted in 8 events being recognized as void responses, among which 5 were true positive, as displayed in Figure 8. With the matching constraint set at 200, the TPR and FPR were 0.8 and 0.59 respectively.



a b

Figure 6 a are 4 indexed B-scans and b shows all recognition result in laboratory experiment. Remarks: solid red rectangles are TPs, dashed red rectangles are FNs, and dotted rectangles are FPs.

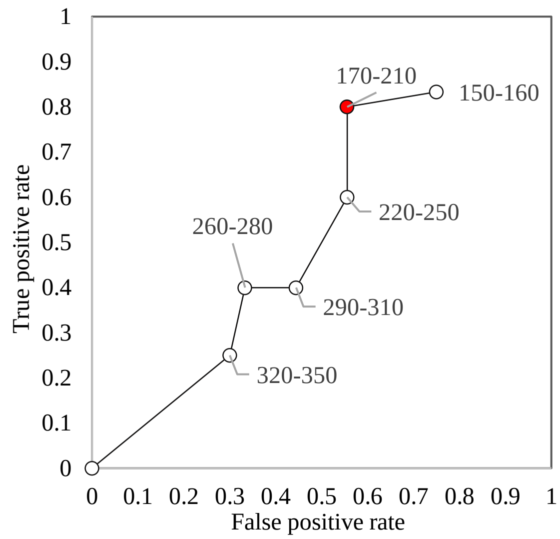


Figure 7 ROC of void recognition from B-scan in laboratory experiment, the optimal point is highlighted with red.

Figure 10 displays the positions of detected results from both C-scans and B-scans; they are closely distributed, which means the void positioning results were promising. Table 2 summarizes the size estimation results from both C-scans and B-scans. It is obvious that the size estimation from B-scans was not stable enough: although they indeed highlighted the void, the identification was perhaps based on the incorrect types of pattern templates and led to an incorrect size estimation.

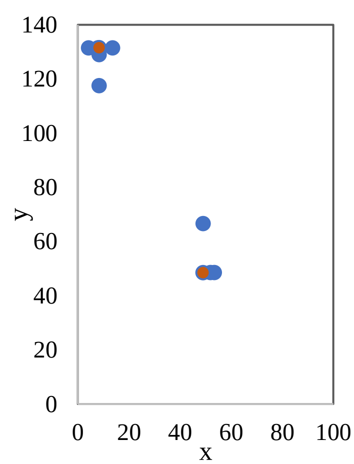


Figure 8 Position of identified responses from both C-scans (red) and B-scans (blue) in laboratory experiment

Table 2 Summary of void size estimation in laboratory experiment (cm)

Void	C-area	B-pattern	B-spread	B-area
		1	7.7022	46.5929
		3	23.1066	419.336
V0	227.762	1	7.7022	46.5929
		2	15.4044	186.372
		4	46.2132	1677.34
		1	7.7022	46.5929
V1	504.908	1	7.7022	46.5929
		2	15.4044	186.372

Remarks: Column “B-pattern” refers to the kind of pattern with which the detected response was identified and matched. Number 1-4 in “B-pattern” represent hyperbola, cross, bowl and reverberation like patterns.

4.2. A site experiment

The tank in the laboratory was filled with homogenous soil, such that beyond the reflection from the void, there was hardly any signal disturbance. But in the real world, the subsurface environment is very complex and multiple types of objects in various shapes and materials are buried there. Real world site experience is therefore necessary to validate the workflow, and the survey of a seawall platform in Tai O, Hong Kong provided just such an opportunity.

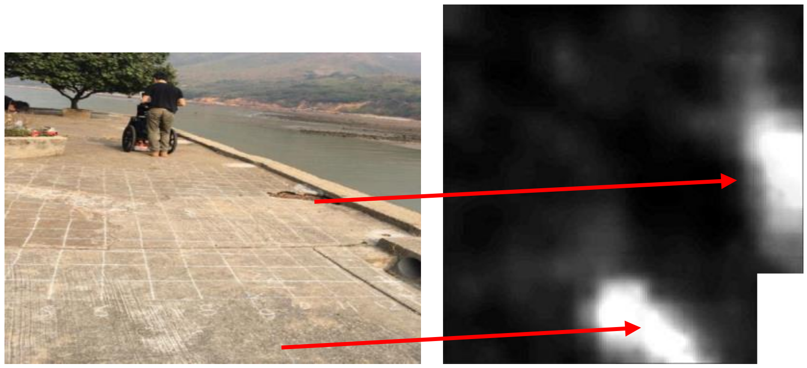
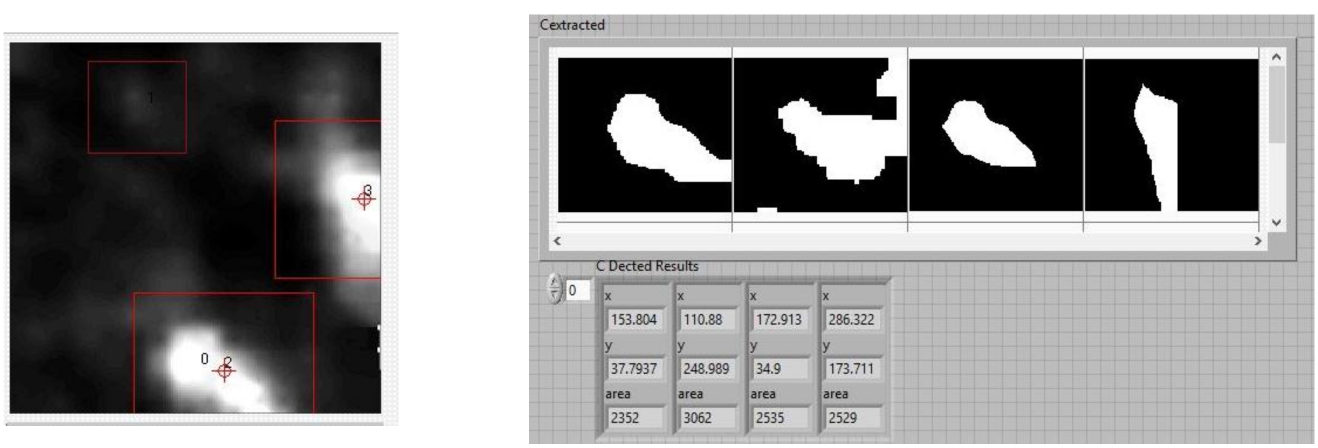


Figure 9 Tai O site and its C-scan at 40cm deep

The site is near the seashore and subject to the threat of seawater infiltration. Voids are likely to have occurred and there were indeed voids found there, as shown in Figure 11. The site area measures 280*320cm in size. GPR data were collected by traversing the grid in both the x and y directions using a profile spacing of 20cm. A 400MHz GSSI antenna and RADAN SIR-4000 control unit were utilized in this survey.

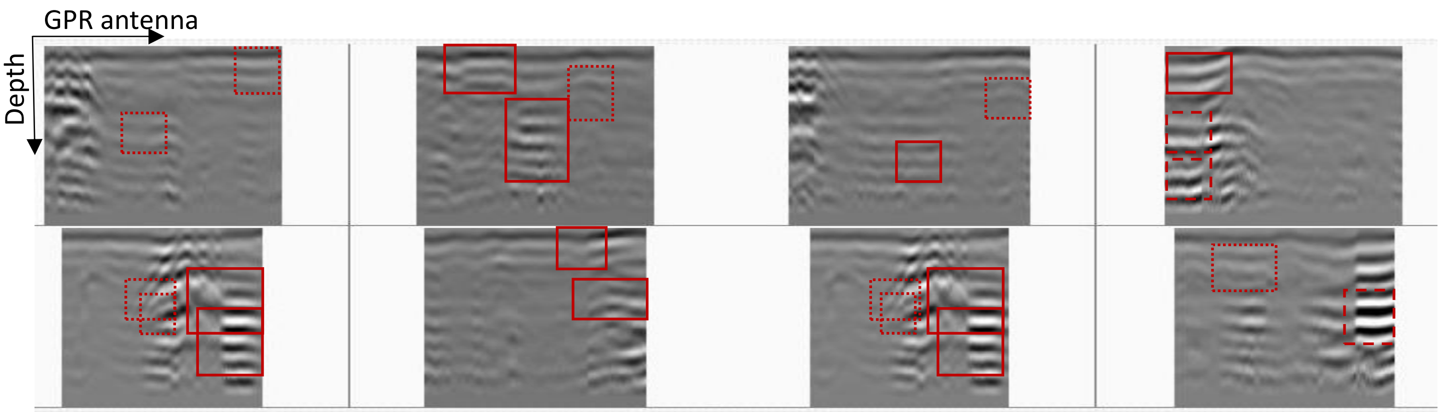
Standard 2D and 3D processing were conducted on the GPR profiles. C-scans were generated as per Figure 7, and 2 voids were clearly imaged: one was already known, but the other was previously unknown. The shapes of the two voids were even more irregular than the void created in the laboratory. In this experiment, the matching score constraint for C-scans was kept the same at 780. The recognition result illustrates that the object shape does have certain effects on the result of greyscale pyramid pattern recognition. Since the void template in C-scans is almost circular in shape, an area of high reflection in the Tai O site C-scan was incorrectly identified as two voids (void-0 and void-2). The scattering around the hypothetical voids was also recognized to exist within the void. According to previous research conducted at this site, 2 voids were confirmed: one was a visible and known surface void that was identified as void-3 in this study, while the other was an invisible but known shallow void that was detected as void-2 in this study (Lai et al., 2017). According to a site record drawing, there was a group of vertical utilities located in the position where void 0 was identified, and they also generated a strong reflection in the C-scan. In contrast, void-1 was not found in any previous records. The C-scan recognition results and size estimations are displayed in Figure 12 below.



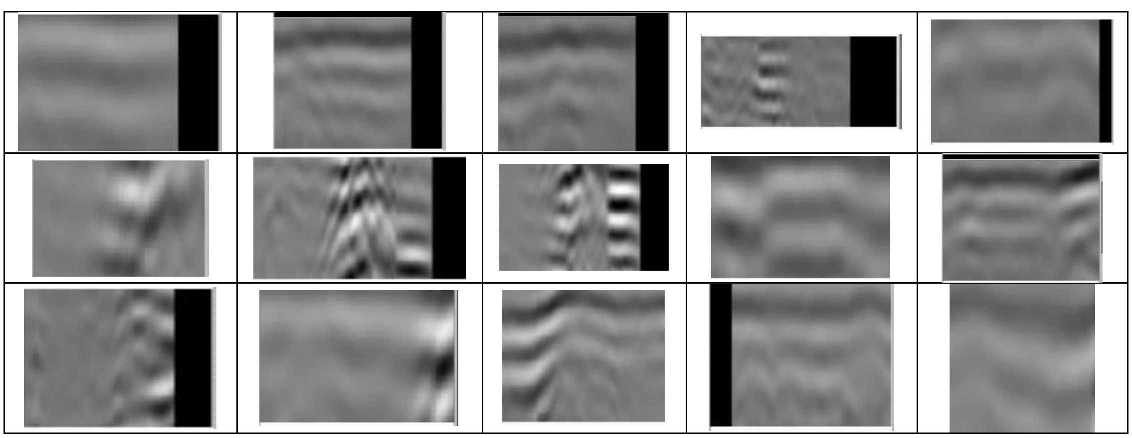
a b

Figure 12 a shows identified voids at Tai O site in red rectangles and b presents the extracted voids with their size and position information. Remarks: column x and y present the coordinate of the void centre, and column area shows the void size estimation of each detected void

There were 4 voids extracted from the C-scans, which led to 8 GPR profiles being indexed as shown in Figure 13. In terms of B-scan recognition, the matching score constraints were explored from 150 to 350, by an increasing step of 10. The ROCs demonstrating the performance of various constraints are displayed in Figure 14. It is obvious that the value of 180 yet again provided the optimal performance. As shown in Figure 13, 15 events were recognized as void patterns, and 10 of them were correctly identified. The TPR is 0.75, which is comparatively lower than that of the laboratory experiment.



a



b

Figure 11 a are 8 indexed B-scans and b shows all recognition results in Tai O experiment. Remarks: solid red rectangles are true positives, dash red rectangles are false negatives, and dotted rectangle are false positives.

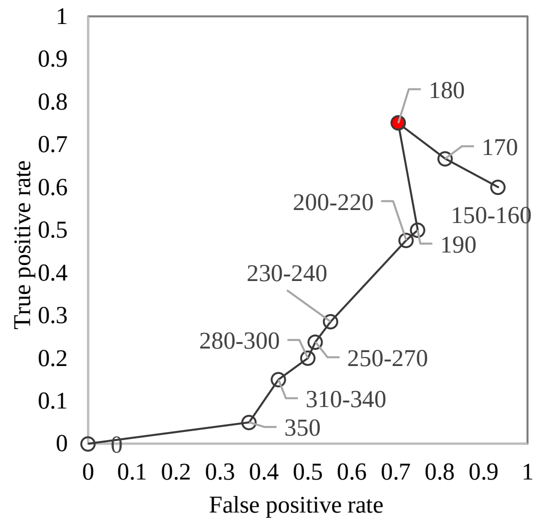


Figure 12 ROC of void recognition from B-scans in the Tai O experiment. Remarks: optimal point is highlighted in red.

Figure 15 displays the positions of detected results from both C-scans and B-scans: they are rather sparsely distributed. Multiple patterns were identified along a traverse. The void size estimations from both C-scans and B-scans are displayed in Table 3. Since excavation was not permitted within the historical site, it was not possible to ground truth the voids' existence, not to mention confirm their sizes. With visual inspection, the voids extracted from C-scans by image segmentation were closer to the reality. The sizes estimated from B-scans were variable; however, the response of a single void was successfully matched with multiple templates, which was similar to the results of the laboratory experiments. In conclusion, size estimation from B-scans is unreliable.

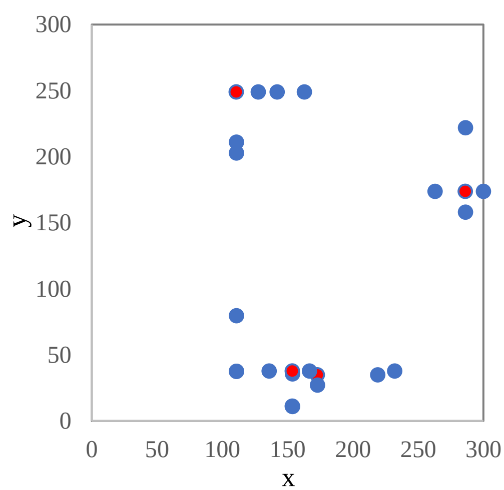


Figure 15 Position of identified pattern from both C-scan (red) and B-scans (blue) in the Tai O experiment

Table 3 Summary of void size estimation in Tai O experiment (cm)

Void	C-area	B-pattern	B-spread	B-area
V0	2352	1	19.1421	287.787
		2	38.2843	1151.15
V1	3062	1	19.1421	287.787
		4	114.853	10360.3
		1	19.1421	287.787
		2	38.2843	1151.15

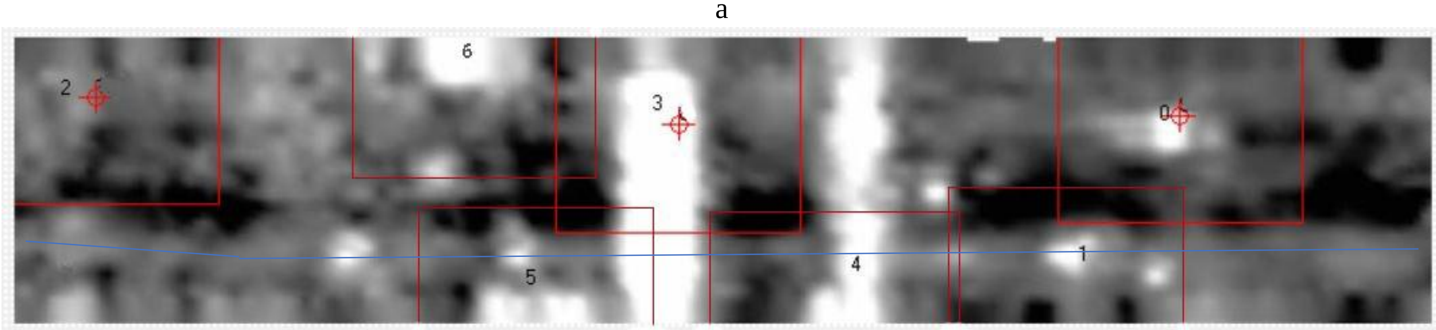
		3	57.4264	2590.08
V2	2535	1	19.1421	287.787
		2	38.2843	1151.15
		2	38.2843	1151.15
		3	57.4264	2590.08
		1	19.1421	287.787
		1	19.1421	287.787
		2	38.2843	1151.15
		3	57.4264	2590.08
		1	19.1421	287.787
V3	2529	1	19.1421	287.787
		3	57.4264	2590.08
		1	19.1421	287.787
		4	114.853	10360.3

Remarks: Column “B-pattern” refers to the kind of pattern with which the detected response was identified and matched. Number 1-4 in “B-pattern” represent hyperbola, cross, bowl and reverberation like patterns.

4.3. The site validation on constraint values

The constraint values suggested in the laboratory and site tests, around 700 in the stage 1: void locating from the C-scan and 200 in the stage 2: void verification from the B-scan, were further validated with a real site case. The validation site, which was 7*30 meter large, was located at Lamma Island, Hong Kong, as shown in Figure 16.a. The GPR C-scan of 1 meter deep is shown as Figure 16.b.

a



b

Figure 13. a. is the photo of the validation site at Lamma Island; b shows the result of the void locating from the C-scan.

With the 300 constraint, in total 7 areas were identified as suspected voids – they presented as local-high reflection in the C-scan. Among these 7 areas, voids labelled as 0, 1 and 5 were metallic utilities pits, voids labled as 3 and 4 were concrete utility pits, and they could be seen from the ground. Void 2 was an unknown reflector. An excavation was dug in void 6, and a piece of concrete plinth was found: the high reflections in void 6 was generated by the concrete plinth. Small scatterers and the horizontal utility (marked as blue line in Figure 16.b) were excluded successfully.

To further investigate the GPR patterns of the areas with high reflections, a section of B-scan that across both void 3 and void 6 was indexed (Figure 17). It can be seen that the GPR reflections at two void areas (blue rectangles) were very similar to the pattern of large voids – the plain reverberation. However, these two "voids" were proved to be either the concrete plinth or the utility pit.

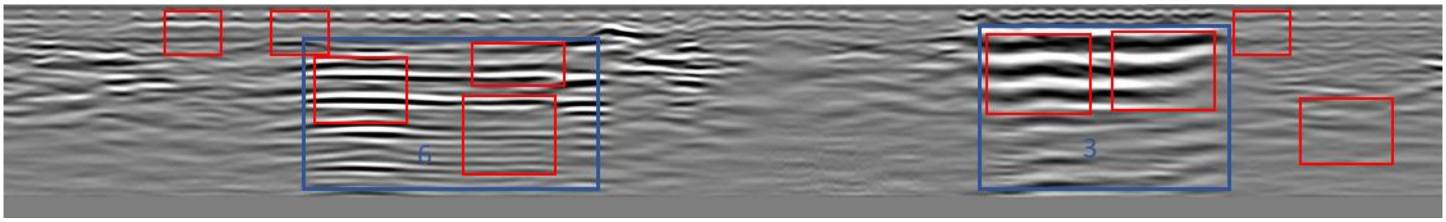


Figure 14 the result of the void pattern verification from B-scan, Remarks: blue rectangles indicating the areas of suspected void 3 and void 6 in Figure 16, and red rectangles circle the identified GPR responses.

When the constraint was 250, the result of the void verification from B-scan was shown as red rectangles in Figure 17. Some small scattering that fell outside the suspected void areas were identified, thus they were given lower identification confidence. Other matched reflections were located within the void areas, therefore the void 3 and the void 6 met the two criteria of the void identification workflow: present as high reflection in C-scan and the reflection in B-scan match the specific pattern. The dielectric constant of the concrete plinth is around 6 ("ASTM D6432," 2011), which is much larger than that of the air. The validation test in Lamma shows the limitation of the proposed system: when the GPR reflection of the object is similar to that of voids, the system may not be able to distinguish them.

5. Discussion

5.1. Reasonable recognition workflow and promising void positioning

The three experiments demonstrate that the subsurface void positioning of the workflow is promising, in that subsurface voids in 2 tests (the laboratory and Tai O) were successfully and precisely identified, and in the Lamma validation experiment, the constraint values were proved proper that only void-like patterns were identified. Similar to seeing a doctor – comprehensive medical examinations are of vital importance before a diagnosis can be concluded. C-scans can deliver a comprehensive view on the subsurface world, and the position and shape of subsurface objects can be roughly delineated. For civil engineering application, it is essential to conduct a 3D GPR survey in order to ensure important information concerning the underground situation is not omitted. Specifically, in the Tai O experiment, an underground structure was mistakenly identified as void-0 in the C-scan as it also presented as local strong reflection. But in the B-scan recognition stage, no proper pattern was matched with this void, and it was therefore identified with reduced confidence. The workflow, which roughly locates voids using C-scans and further verifies them with B-scans, is logically designed in that many local objects yield strong reflections, just as voids do, but they also generate various waveforms and present as different patterns in B-scans.

- Tolerances of void positioning and void size estimation

However, full coverage C-scans are generated using interpolation to fill in blank areas within survey profiles, which means that some pixel values in C-scans are a reflection of processing rather than true measurements. Added to that, the indexed B-scans are not necessarily positioned to cut across the centre of the void, since this position may be located in gaps among the GPR survey profiles. The maximum deviation of B-scan positions from the void centre is half a profile spacing. If the shape of a subsurface void is very irregular, then the patterns identified in B-scans may not describe the void's geometry. Besides, during the B-scan recognition stage, the void size is determined by predefined multiples of the GPR footprint, and void size estimation is therefore not a continuous measurement and may only result in one of several predefined numbers. This further illustrates the point that void size estimation from B-scans is unreliable.

- Conservative but efficient approach

The whole recognition process, from the initial stage of C-scan scanning, to the final void diagnosis stage, takes less than 10 minutes, which is significantly more convenient than the currently used visual interpretation approach. In view of the time-consuming nature of GPR data interpretation, the application of automatic recognition in GPR survey deserves further investigation in order to improve its reliability. The pyramid method applied in this study possesses both strengths and weaknesses. In B-scan recognition, it efficiently identified most GPR responses from air-filled voids in both experiments when using a lower constraint, although the experimental data was visually significantly different from the void template. Using lower constraints may lead to not only higher TPRs but also higher FPRs. Considering the hazardous nature of subsurface voids, it is better to be conservative in order to avoid missing voids.

- Using greyscale or gradient as image information

In this study, greyscale value and gradient were used for feature recognition in C-scans and B-scans, respectively. This approach was proven to be suitable through experimental testing. If the gradient approach was used in the C-scan void recognition, some round shaped local anomalies with weak reflection intensities were incorrectly identified as voids, because the template used in C-scans is round shaped and the gradient method focuses on pattern structure. Similarly, if the greyscale approach was used in the B-scan void recognition, more reflections with strong intensities were in-correctly identified as voids, while actual void responses with weaker reflection intensities were overlooked. Such cases resulted in lower TPRs and higher FPRs.

5.2. Variety and confusion: GPR responses from voids

Both the designed workflow and case experiments illustrate that in a GPR survey, an air void can only be defined by satisfying two criteria. The first criterion is the existence of non-continuous strong reflections in C-scans, while the second is the presence of special patterns with decaying amplitude and later time windows in B-scans. These two criteria reflect the findings of previous research (Lai et al., 2017).

In both the C-scan and B-scan recognition stages, the Tai O site experiment has higher FPRs, which means more irrelevant responses were identified as voids. The laboratory experiment was conducted using a relatively homogenous background medium and, beyond the two identified voids, there were hardly any scatterers. In contrast, it can be observed from Figure 13a that diverse and complicated response events were present in the Tai O site radargrams. Some of these response events were generated by multiple unknown objects rather than voids. In the Lamma experiment, a known utility pit and a concrete plinth were identified as voids, because they have similar reflection patterns.

In addition, the void's extent in the Tai O site was large enough to yield reverberation patterns, but very often some cross or hyperbola responses with weaker reflection intensities were found near to the targeted reverberation response. These cross or hyperbola like GPR responses were interference generated by the void's top and bottom edges. As these responses occur at the edges of voids instead of within them, they might be confused with void patterns and contribute to incorrect void positioning. The interference response also resulted in a single void being identified by multiple templates, which in turn produced various void size estimations.

With B-scans, it is difficult to describe a void's appearance using a series of templates. Obvious differences can be observed by comparing templates and experimental data. Many factors, including inhomogeneous background material, equipment specifications, survey settings such as traverse interval and time window, and interference from adjacent objects, contribute to a variety of GPR responses to voids. Moreover, the validation experiment in Lamma shows that some other kinds of objects may yield similar patterns in GPR survey data, which also leads to confusion. For instance, both the subsurface void and manhole in this study's Lamma experiment were filled with air and had an interface with the garden soil, and they both, perhaps not surprisingly, presented quite similar patterns in both the C-scans and B-scans. Beyond the flexible matching method used in this study, the use of intelligent searching and learning schemes deserves further exploration. In particular, the machine learning technique can be applied to continuously train the classifier using both simulation or real GPR measurements, so as to improve the precision and accuracy of void identification.

6. Conclusions

This study has developed a workflow for automatically identifying subsurface voids in GPR data. The workflow imitates the human judgment process, and integrates a pyramid pattern recognition technique in order to search for GPR responses generated by air-filled voids. The GPR responses of air-filled voids were investigated in advance. A void appears as a local anomaly with high reflection intensity in C-scans. Different ratios of void size and GPR signal wavelength result in different patterns in B-scans: they can be hyperbolas, cross patterns, bowl shaped patterns and reverberations. A database of void patterns for both C-scans and B-scans was established, and the pyramid pattern recognition method – with pixel value or gradient used for feature identification – was used to search for such GPR responses automatically. In this way, a preliminary knowledge of void location and void size can be obtained without human intervention. Three case experiments were conducted and produced promising results. Known voids were successfully identified, although some errors existed. In view of the demanding and heavy workload involved in subsurface health inspections, the workflow has proven to be efficient and effective. The study result raises the possibility of conducting city-scale full coverage subsurface health inspections.

Acknowledgement

The authors wish to thank Wu Ho Yi for his contributions in void experiments in lab. Great appreciation also goes to Mick Atha for his efforts in English editing. Authors would also like express gratitude to reviewers for their constructive comments.

Reference

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., & Ogden, J. M. (1984). Pyramid methods in image processing. *RCA engineer, 29*(6), 33-41.

Al-Nuaimy, W., Huang, Y., Nakhkash, M., Eriksen, A., Fang, M. T. C., & Nguyen, V. T. (2000). Automatic detection of buried utilities and solid objects with GPR using neural networks and pattern recognition. *Journal of Applied Geophysics, 43*, 157-165. doi:S0926-9851 99 00055-5

Annan, A., & Cosway, S. (1992). Simplified GPR beam model for survey design. In *SEG Technical Program Expanded Abstracts 1992* (pp. 356-359): Society of Exploration Geophysicists.

Annan, A. P. (2002). GPR—History, Trends, and Future Developments. *Subsurface Sensing Technologies and Applications, 3*(4).

Annan, A. P. (2004). *Ground Penetrating Radar Applications, Principles, Procedures*. Mississauga, Canada: Sensors and Software.

ASTM D6432. (2011). In *Standard Guide for Using the Surface Ground Penetrating Radar Method for Subsurface Investigation*.

Ayala-Cabrera, D., Herrera, M., Izquierdo, J., & Pérez-García, R. (2011). Location of buried plastic pipes using multi-agent support based on GPR images. *Journal of Applied Geophysics, 75*(4), 679-686. doi:10.1016/j.jappgeo.2011.09.024

Benedetto, A., & Pajewski, L. (2015). *Civil Engineering Applications of Ground Penetrating Radar*. London: Springer.

Casas, A., Lazaro, R., Vilas, M., & Busquet, E. (1996). *Detecting karstic cavities with ground penetrating radar at different geological environments in Spain.* Paper presented at the Proceedings of the 6th International Conference of Ground Penetrating Radar.

Clemeña, G. G., Sprinkel, M. M., & Long, R. (1986). *Use of Ground-Penetrating Radar for Detecting Voids Underneath a Jointed Concrete Pavement*. Retrieved from

Farley, M., Water, S., Supply, W., Council, S. C., & Organization, W. H. (2001). Leakage management and control: a best practice training manual.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters, 27*(8), 861-874.

Gamba, P., & Lossani, S. (2000). Neural detection of pipe signatures in ground penetrating radar images. *IEEE Transactions on Geoscience and Remote Sensing, 38*(2), 790-797.

Ghasemi, F. S. A., & Abrishamian, M. (2007). A novel method for FDTD numerical GPR imaging of arbitrary shapes based on Fourier transform. *NDT & E International, 40*(2), 140-146.

Giannopoulos, A. (2005). Modelling ground penetrating radar by GprMax. *Construction and Building Materials, 19*(10), 755-762.

Goodman, D., & Piro, S. (2013). *GPR remote sensing in Archaeology* (Vol. 9). New York: Springer.

Gou, S., Zhuang, X., Zhu, H., & Yu, T. (2013). Parallel sparse spectral clustering for SAR image segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*(4), 1949-1963.

Jol, H. M. (2009). *Ground Penetrating Radar Theory and Applications*: Oxford: Elsevier.

Kofman, L. (1994). *Use of ground penetrating radar in geotechnical applications.* Paper presented at the Proceedings of the 12th Conference of the Israel Association for the Advancement of Mineral Engineering, Eilat.

Kofman, L., Ronen, A., & Frydman, S. (2006). Detection of model voids by identifying reverberation phenomena in GPR records. *Journal of Applied Geophysics, 59*(4), 284-299.

LabVIEW – See it. Solve it.s. (2018). Retrieved from http://www.ni.com/en-us.html

Lai, W. W., Chang, R. K., & Sham, J. F. (2017). *Detection and imaging of city's underground void by GPR.* Paper presented at the Advanced Ground Penetrating Radar (IWAGPR), 2017 9th International Workshop on.

Lai, W. W. L., Chang, R. K. W., Sham, J. F. C., & Pang, K. (2016). Perturbation mapping of water leak in buried water pipes via laboratory validation experiments with high-frequency ground penetrating radar (GPR). *Tunnelling and Underground Space Technology, 52*, 157-167. doi:10.1016/j.tust.2015.10.017

Leckebusch, J., & Peikert, R. (2001). Investigating the true resolution and three-dimensional capabilities of ground-penetrating radar data in archaeological surveys: measurements in a sand box. *Archaeological Prospection, 8*(1), 29-40.

Lee, S. U., Chung, S. Y., & Park, R. H. (1990). A comparative performance study of several global thresholding techniques for segmentation. *Computer Vision, Graphics, and Image Processing, 52*(2), 171-190.

Leucci, G., & Negri, S. (2006). Use of ground penetrating radar to map subsurface archaeological features in an urban area. *Journal of Archaeological Science, 33*(4), 502-512. doi:10.1016/j.jas.2005.09.006

Leucci, G., Negri, S., & Carrozzo, M. T. (2003). Ground Penetrating Radar (GPR): an application for evaluating the state of maintenance of the building coating. *Annals of Geophysics, 46*(3).

Luo, T. X., Lai, W. W., Chang, R. K., & Goodman, D. (2019). GPR imaging criteria. *Journal of Applied Geophysics, 165*, 37-48.

MacLean, W. J., & Tsotsos, J. K. (2008). Fast pattern recognition using normalized grey-scale correlation in a pyramid image representation. *Machine Vision and Applications, 19*(3), 163-179.

NI Vision - Pattern Matching Techniques. (2018). Retrieved from http://zone.ni.com/reference/en-XX/help/372916P-01/nivisionconcepts/pattern_matching_techniques/

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics, 9*(1), 62-66.

Pasolli, E., Melgani, F., & Donelli, M. (2009). Automatic analysis of GPR images: A pattern-recognition approach. *IEEE Transactions on Geoscience and Remote Sensing, 47*(7), 2206-2217.

Pavlidis, T. (2013). *Structural pattern recognition* (Vol. 1): Springer.

Pérez-Gracia, V., González-Drigo, R., & Di Capua, D. (2008). Horizontal resolution in a non-destructive shallow GPR survey: An experimental evaluation. *NDT & E International, 41*(8), 611-620.

Permuter, H., Francos, J., & Jermyn, I. (2006). A study of Gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recognition, 39*(4), 695-706.

Plati, C., & Dérobert, X. (2015). Inspection Procedures for Effective GPR Sensing and Mapping of Underground Utilities and Voids, with a Focus to Urban Areas. In *Civil Engineering Applications of Ground Penetrating Radar* (pp. 125-145): Springer.

Powers, D. M. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging, 13*(1), 146-166.

Simi, A., Bracciali, S., & Manacorda, G. (2008). *Hough transform based automatic pipe detection for array GPR: Algorithm development and on-site tests.* Paper presented at the Radar Conference, 2008. RADAR'08. IEEE.

Warren, C., Giannopoulos, A., & Giannakis, I. (2016). gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. *Computer Physics Communications, 209*, 163-170.

Wu, H.-Y. (2015). *Characterization of Shallow Subsurface Air-filled Voids by Ground Penetrating Radar.* (Bachelor of Science), The Hong Kong Polytechnic University, Hong Kong.

Xie, X., Qin, H., Yu, C., & Liu, L. (2013). An automatic recognition algorithm for GPR images of RC structure voids. *Journal of Applied Geophysics, 99*, 125-134. doi:10.1016/j.jappgeo.2013.02.016

Xu, X., Zeng, Q., Li, D., Wu, J., Wu, X., & Shen, J. (2010). GPR detection of several common subsurface voids inside dikes and dams. *Engineering Geology, 111*(1-4), 31-42. doi:10.1016/j.enggeo.2009.12.001