A comprehensive approach to real-time fault diagnosis during automatic grit-blasting operation by autonomous industrial robots

Andrew Wing Keung To, Gavin Paul, Dikai Liu[a]

[a]Centre for Autonomous Systems
University of Technology, Sydney
{andrew.to-1, gavin.paul, dikai.liu}@uts.edu.au

Abstract

This paper presents a comprehensive approach to diagnose for faults that may occur during a robotic grit-blasting operation. The approach proposes the use of information collected from multiple sensors (RGB-D camera, audio and pressure transducers) to detect for 1) the real-time position of the grit-blasting spot and 2) the real-time state within the blasting line (i.e. compressed air only). The outcome of this approach will enable a grit-blasting robot to autonomous diagnose for faults and take corrective actions during the blasting operation. Experiments are conducted in a laboratory and in a grit-blasting chamber during real grit-blasting to demonstrate the proposed approach. Accuracy of 95% and above has been achieved in the experiments.

Keywords: Autonomous industrial robot, Abrasive grit-blasting, Real-time fault diagnosis, Multi-modal sensing, Field Robotics

1. Introduction

Abrasive grit-blasting is commonly performed in manufacturing to prepare steel structures for painting; by removing debris/rust and creating a suitable surface roughness profile for paint adhesion. Inherently grit-blasting creates a dangerous work environment that exposes humans to risks ranging from physical injuries to silicosis caused by long-term exposure to grit dust containing silica [1]. In an effort to improve human safety, autonomous industrial robots are being used for the grit-blasting operations. As such,

grit-blasting robots have been seen used in factories, albeit typically limited to a robotic blasting cell performing scripted blasting routines.

With the recent advancement in robotic research, autonomous robots that can perform in-situ abrasive grit-blasting in a field environment are being developed [2]. Figure 1 shows a robot for autonomous grit-blasting of steel bridge structures. These robots have advanced upon their factory counterparts by possessing the capability of automatically sensing and mapping [3] to build an up-to-date geometric map of the surrounding surfaces. Using the map, the robots autonomously navigate and localise within the environment, and conduct abrasive grit-blasting on the identified surrounding surfaces.
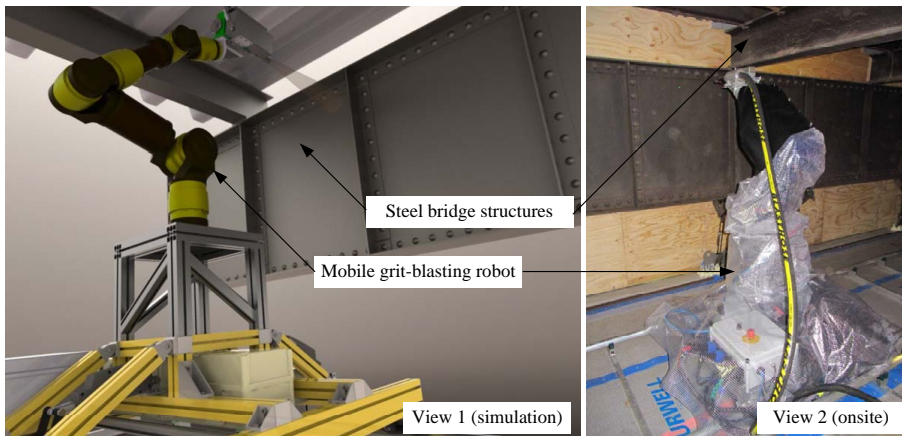

Figure 1: An autonomous robot for grit-blasting steel bridge structures

Due to the harsh operating conditions in a field environment, it is necessary to enable the grit-blasting robot to monitor the operation and diagnose faults that may occur; e.g. loss of ablation from grit running out or from hose blowouts, and missed coverage due to misalignment of the grit-blasting spot on surfaces. Currently, a human operator is still tasked to monitor faults including: 1) no compressed air and no grit exiting hose, 2) blasting without grit, (i.e. compressed air only), 3) grit-blasting occurring unexpectedly and 4) grit-blasting spot misaligned on surface. As such, there has been an increased interest in equipping grit-blasting robots with the capability of fault diagnosis to remove humans from the risks associated with supervising the autonomous grit-blasting operation. However, there currently exists no comprehensive approach available for autonomous industrial robots to perform

real-time fault diagnosis during abrasive grit-blasting.

This paper presents an approach to real-time diagnosis of faults during grit-blasting. Similar to robotic welding which uses vision, depth and acoustic sensors to perform seam tracking [4], weld penetration controlling [5] and on-line monitoring of the welding process [6], it is proposed that the robotic grit-blasting operation can be monitored in real-time using multi-modal sensing. The research conducted in this paper focuses on devising techniques for using multi-modal sensing data to detect the real-time position of the grit-blasting spot and the state-of-blasting, and performing fault diagnosis.

The following is a review of potential techniques that may be used to perform detection within grit-blasting environments. Image processing such as background subtraction techniques [7][8] can be considered for identifying the real-time location of the grit-blasting spot on a surface. Currently, background subtraction has been demonstrated to robustly detect moving objects (i.e. cars and people) in environments with illumination changes [9], when there are camera movements during data capture [10] and under complex weather conditions such as fog and snow [11]. Hence, background subtraction shows good potential for application in a dust laden grit-blasting environment to detect a moving grit-blasting spot on a surface. However, a major drawback of background subtraction is that frame-by-frame contrast changes must exist in order to detect for movements. Given that contrast changes on a grit-blasted steel surface is unobservable, then background subtraction can become ineffective to detect the grit-blasting spot position on such surface. To address this limitation active sensing can be investigated such as a laser vision system [12] that projects and detect for a laser point on a surface, or an object tracking system [13] that uses vision and depth information (RGB-D). By including additional active sensing to aid the background subtraction process, it is possible to detect the location of the grit-blasting spot on a surface at all times and diagnose the positioning of the grit-blasting spot.

In conjuncture with vision-based techniques, sound can also be used to diagnose faults such as an unexpected state-of-blasting (*no-blasting, air-blasting, grit-blasting*) occurring during the grit-blasting operation. A human expert will perceive from the acoustic sounds produced during grit-blasting operation the three different states-of-blasting and can report a fault when there is a discrepancy between the perceived and the planned/expected state-of-blasting. Similarly, a robot can diagnose faults using audio features extracted from the acoustic sounds. Various statistic properties of acoustic

sound [14] in both time-domain [15] and frequency-domain [16] are available as features to distinguish between the different states-of-blasting. In order to efficiently detect the different states-of-blasting, a set of discriminative features must be identified. To achieve this, a feature selection process [17] can be used to select a subset of the audio features extracted in both time and frequency domain. Subsequently, the selected audio features can be used to train a machine learning algorithm (e.g. Support Vector Machine [18], naive Bayes [19], K-nearest neighbours [20]) to classify the different states-of-blasting in real-time.

To summarise, an approach based on vision and sound can be devised to diagnose faults during grit-blasting operation. Experiments will be conducted in both a laboratory and in a grit-blasting chamber during live grit-blasting to test the devised approach. The remainder of this paper is organised as follows: Section 2 presents the overview of the approach, the process for detecting the grit-blasting spot position using RGB-D data, the process for detecting the states-of-blasting using audio data or pressure data, calculating the most probable state-of-blasting based on detection provided by audio and pressure data, and fault diagnosis by comparing detected against the expected. Section 3 presents three experiments to verify the approach, Section 4 presents a discussion and Section 5 presents the conclusion.

2. Proposed Approach

2.1. Approach Overview

An overview of the approach for fault diagnosis is provided in Figure 2. As shown by the figure, fault diagnosis is a cyclic process of on-line detection for the 1) grit-blasting spot position and 2) state-of-blasting, and comparing against the expected position and expected state-of-blasting for that instance of time. If a discrepancy is observed during any instance of comparison, then a fault is detected.
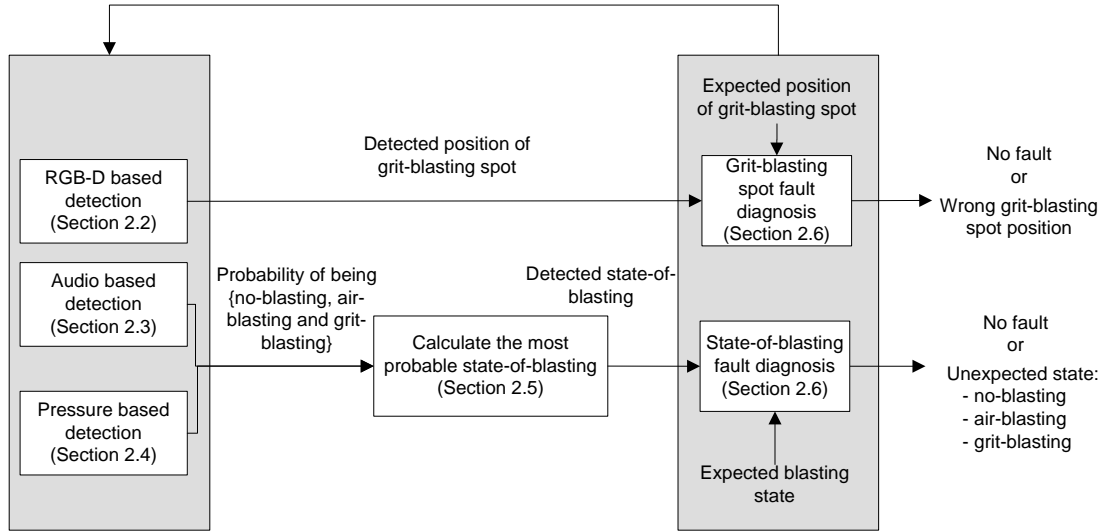
4

Figure 2: Overview of the fault diagnosis approach

Figure 3 illustrates the different sensing modalities used in this approach, where: <u>RGB-D</u> is used for detecting the real-time grit-blasting spot position, <u>audio</u> is used for detecting the state of *no-blasting*, *air-blasting* and *grit-blasting*, and <u>pressure</u> is used for detecting the state of *no-blasting* and *air/grit-blasting*. In essence, RGB-D and audio sensing are selected to imitate the visual and sound perception made by a human operator, and pressure is an already proven modality for diagnosing grit-blasting faults. It is to note that the detection for all the three states-of-blasting can be achieved using audio only, however pressure sensing is included in this approach to provide redundancy for improved robustness. Sections 2.2 - 2.4 will discuss in detail the techniques for performing real-time detection using each modality, Section 2.5 will detail how to combine the audio and pressure detection results, and Section 2.6 will detail the use of the detection results to perform fault diagnosis.
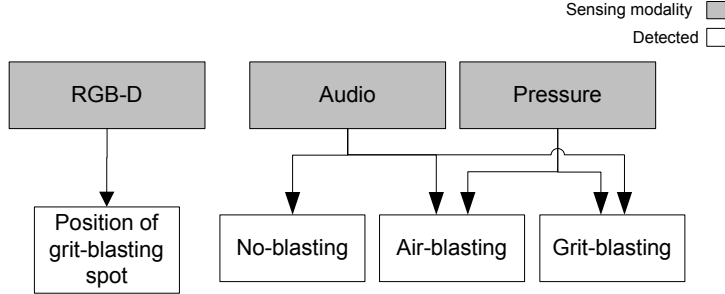
Figure 3: Grit-blasting features detected by each sensing modality

2.2. RGB-D based detection

This section presents a technique to detect the real-time position of a grit-blasting spot by using a RGB-D image stream. Firstly the following conditions has been verified experimentally: 1) a projected laser point can provide a distinct feature within a grit-blasting environment; and 2) a laser pointer can be mounted onto the robot/nozzle such that the position of the projected laser point is always in close proximity to the actual grit-blasting spot on a surface. Based on these conditions, it is possible to detect/track the position of the projected laser point in a RGB-D image stream to approximate the real-time position of the grit-blasting spot on a surface.

The algorithm for detecting the position of a green laser point (can be modified to detect red or blue laser point) is presented in **Algorithm 1**. The following briefly describes the steps in the algorithm:

- Steps 2 - 4 generates an intensity image (*laserPointImage*) that shows high contrast between green and other colours. This is achieved by summing the intensity difference between green/red channel and green/blue channel (step 2), and then multiplying with the grayscale version of the original image (step 4)

- Step 5 generates a binary image (*binaryImage*) by segmenting the intensity image (*laserPointImage*). Where an pixel = 1 if the intensity is greater 75% of max intensity, else 0

- Steps 6 - 8 identifies the blobs in the binary image (*binaryImage*) and calculates the centroid and pixel count of the largest blob

6

- Steps 9 - 15 determines whether to accept or reject the largest blob identified in the binary image as the laser point. If accepted, then the 3D position of the blob's centroid is calculated using the depth information and is returned as an output of the algorithm.

Algorithm 1 Detecting the position of a green laser point

1: **procedure DetectLaserDot**(R, G, B, D)
2:      $green\_RedBlueDiff = (G - R) + (G - B)$
3:      $grayscaleImage = \textbf{ConvertToGray}(R, G, B)$
4:      $laserPointImage = grayscaleImage \times green\_RedBlueDiff$
5:      $binaryImage = (laserPointImage >= 0.75 \times \textbf{Max}(laserPointImage))$
6:      $[blobAreas] = \textbf{IdentifyBlobs}(binaryImage)$
7:      $[maxBlob] = \textbf{IdentifyLargestBlob}(blobAreas)$
8:      $[blobCentroid, pixelNumber] = \textbf{BlobAttributes}(maxBlob)$
9:      **if** $pixelNumber > pixelThreshold$ **then**
10:          $x = blobCentroid.x$
11:          $y = blobCentroid.y$
12:          $laserPosition = calculate3D(x, y, d)$
13:      **else**
14:          $laserPosition = nan$
15:      **return** $laserPosition$

Figure 4 illustrates the laser point detection process using **Algorithm 1**. As shown by the figure, the image region representing a laser point is identified by segmenting for the pixels with the highest intensity in the specified colour (e.g. greenest pixels). Provided that the segmented image region has the adequate number of pixels representing a laser point (user defined), then the centroid of this image region is selected to be the laser point. The corresponding depth value of the centroid is identified from the depth image and is converted into a point in 3D using perspective projection [21]. This point is taken as the detected position of the grit-blasting spot on a surface.
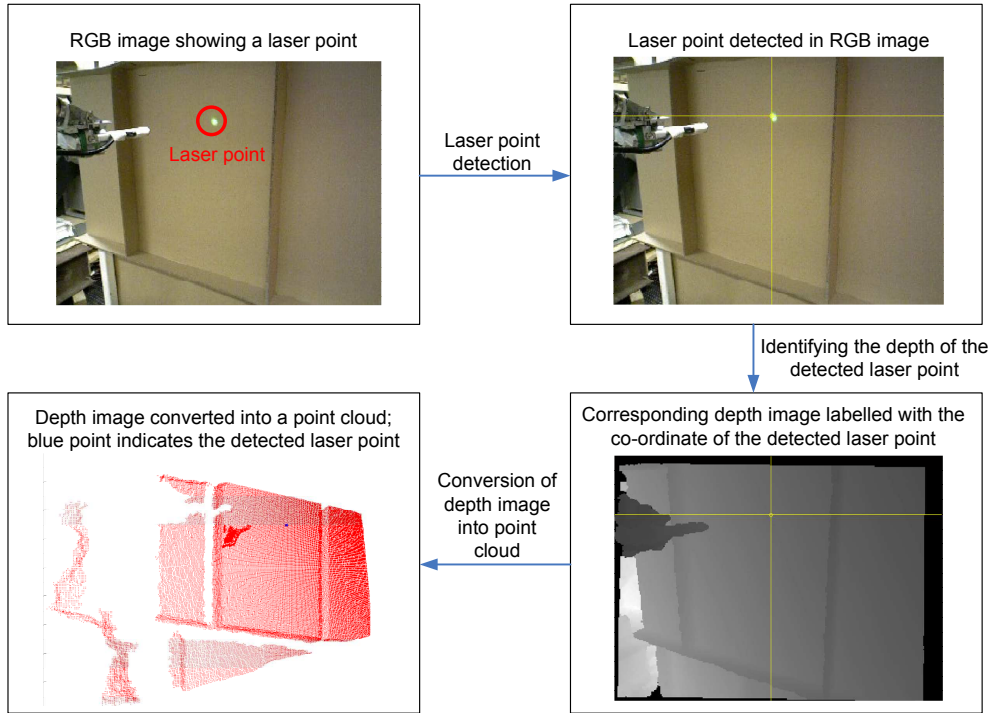
Figure 4: Detecting the position of a laser point in a RGB-D image

Laser point detection can be performed continuously for a live RGB-D image stream to provide the real-time position of the grit-blasting spot. In order for continuous blasting spot detection, tracking of the laser point is needed. A simple tracking method is used that compares the currently detected position against the previous position of the grit-blasting spot. The current position will be accepted if the Euclidean distance between the previous and current positions does not exceed a distance threshold (user defined). Section 2.6.1 details the fault diagnosis performed using the detected positions of the grit-blasting spot and the planned path.

2.3. Audio-based detection

It is proposed that the acoustic sound of particles, or lack of particles passing within the blasting hose can be used detect the real-time state-of-blasting: (*no-blasting*, *air-blasting* and *grit-blasting*). Figure 5 shows the detection process. Firstly, audio features are extracted from a training dataset containing pre-recorded samples of each state-of-blasting to train a machine

learning algorithm (e.g. classifier: nave Bayes, K-nearest neighbour or Support Vector Machine (SVM)). Once trained, the classifier is applied to classify audio features extracted from a live audio feed and outputs the probability of being each of the three states-of-blasting. The process of extracting and classifying live audio features is repeated (i.e. 10Hz) to provide real-time detection of the state-of-blasting. The probabilities produced by each detection cycle are combined with the probabilities from pressure-based detection performed at the same instance of time (refer to Section 2.5). The most probable state-of-blasting deduced from the combined probabilities is used to perform fault diagnosis.
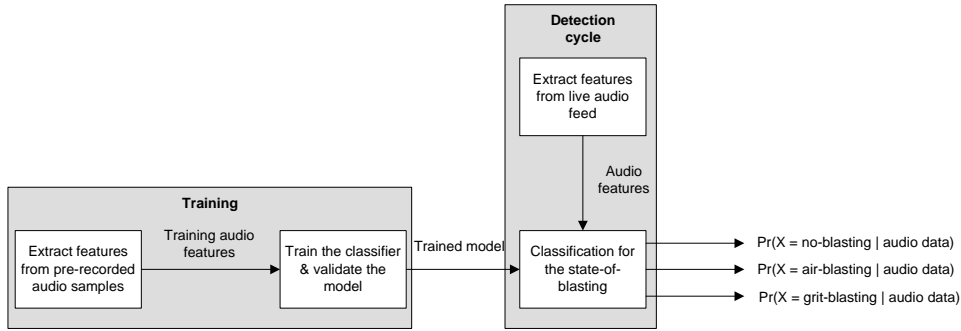

Figure 5: Detecting the three states-of-blasting by using audio data

In order to identify a discriminative set of audio features that can be used to classify the different states-of-blasting, feature selection [17] is performed on 35 audio features commonly used in speech recognition including zero crossing, signal energy, energy entropy, spectral centroid, spectral spread, spectral entropy, spectral flux, spectral roll-off, MFCCs, harmonic ratio, fundamental frequency and chroma vectors [14]. As shown in Figure 6, audio features are extracted from samples taken by time slicing the audio signal into windows, W_t with step size, k_t. The window size and step size can affect detection accuracy, latency and computation times, and should be tuned according to the application requirements.
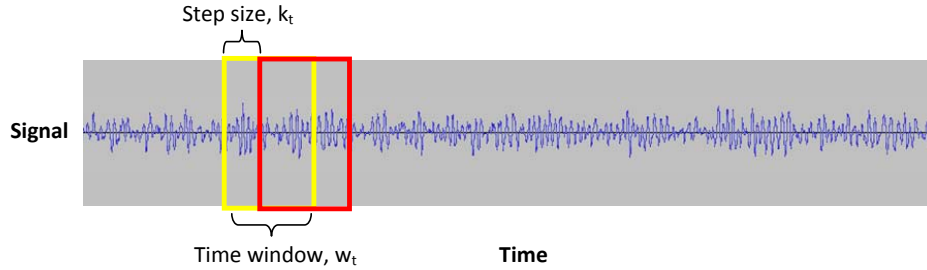
Figure 6: Time slicing an audio signal into samples to extract audio features

Feature selection is performed using the information gain method [22] and listed below are the top 5 of the 35 audio features identified for classifying the three states-of-blasting:

1. Spectral entropy

2. Mel Frequency Cepstrum 2nd Coefficient

3. Mel Frequency Cepstrum 1st Coefficient

4. Spectral roll-off

5. Zero crossing rate

Furthermore, feature selection is also conducted using the wrapper method [23] to identify the best feature subset to train a specific classifier. The wrapper method is applied using the naive Bayes and then the K-nearest neighbour as the base classifier (parametric and non-parametric respectively). As a result, a feature subset of {spectral entropy, spectral roll-off} was identified for the naive Bayes classifier, and a feature subset of only {signal energy} was identified for the K-nearest neighbour classifier.

Based on the feature selection results provided by information gain ranking and wrapper method, it is determined that a naive Bayes classifier trained using audio features {spectral entropy, spectral roll-off} can be used to classify the three states-of-blasting. The following describes a process for calculating the spectral entropy and spectral audio features.

To determine the spectral entropy for the signal s(t) within a time window W_t, the Discrete Fourier Transform (DFT) of the signal S(f) is firstly

10

calculated, and then the Power Spectral Density (PSD) can be calculated using S(f) as:

$$PSD(f) = |S(f)|^2 \tag{1}$$

The PSD is converted into a Probability Density Function by normalising:

$$PSD_n(f) = \frac{PSD(f)}{\sum_{f=-\frac{f_s}{2}}^{\frac{f_s}{2}} PSD(f)} \tag{2}$$

where f_s is the sampling frequency of the signal. Finally, the spectral entropy is calculated using the PSD_n as:

$$\text{Spectral entropy} = \sum_{f=-\frac{f_s}{2}}^{\frac{f_s}{2}} PSD_n(f) log_2[PSD_n(f)] \tag{3}$$

The spectral roll-off point is defined as the Nth percentile of the PSD (Eq.7), where N is usually 95% [24]. The roll-off point is the frequency (f_r) below which the N [1] of the magnitude distribution is concentrated. This can be expressed as:

$$\sum_{f=-\frac{f_s}{2}}^{f_r} PSD(f) = N \times \sum_{f=-\frac{f_s}{2}}^{\frac{f_s}{2}} PSD(f) \tag{4}$$

The spectral roll-off is identified as the value for f_r that satisfies Eq.4 given the selected value for N.

2.4. Pressure-based detection

It has been observed from preliminary investigation that pressure readings captured during *air-blasting* and *grit-blasting* are similar and thus ineffective for distinguishing between the two states. Therefore, pressure-based detection is performed to distinguish between *no-blasting* and the combined state of *air/grit-blasting*. Figure 7 illustrates how detection for the states of *no-blasting* and the combined state of *air/grit-blasting* can be performed by setting a pressure threshold (τ_p).
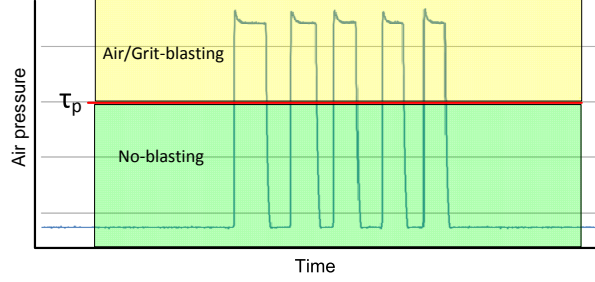
11

Figure 7: *No-blasting* and *air/grit-blasting* by using a pressure threshold τ_p

For each pressure measurement, a discrete probability distribution of the three states-of-blasting can be calculated. Then, the pressure-based results can be combined with the audio-based results (refer to Section 2.5). Given that only the states of *air/grit-blasting* and *no-blasting* can be detected using pressure readings, the probability of being either *air-blasting* or *grit-blasting* is treated as being equal. Thus, the discrete probability distribution of being each state can be expressed as:

$$\Pr(X = \text{no-blasting} \mid \text{pressure data }) = \beta,$$

$$\Pr(X = \text{air-blasting} \mid \text{pressure data }) = \frac{1 - \beta}{2},$$

$$\Pr(X = \text{grit-blasting} \mid \text{pressure data }) = \frac{1 - \beta}{2}, \tag{5}$$

where the value of β is calculated using the pressure reading ρ, the pressure threshold τ_p, and a selected probability value χ. The probability value χ, is selected between [0,1] and is used to ensure that all states in the probability distribution will be greater than zero. The value of β is expressed as:

$$\beta = \begin{cases} \chi, & \rho < \tau_p \\ 1 - \chi, & \rho \geq \tau_p \end{cases} \tag{6}$$

2.5. Calculation of the most probable state-of-blasting

The probability distributions (probability mass functions) produced at each cycle by the audio-based and pressure-based detection is combined to identify the most probable state-of-blasting. In order to take into account the

detection accuracy of each sensing modality, weighting factors are applied. This can be expressed as:

$$\Pr(\text{X=x} \mid \text{audio, pressure data }) = \omega_a \times \Pr(\text{X=x} \mid \text{audio data })+ \\ \omega_p \times \Pr(\text{X=x} \mid \text{pressure data }) \tag{7}$$

where x $= \{$ *no-blasting, air-blasting, grit-blasting* $\}$ and $(\omega_a$ and $\omega_p)$ are the weightings for audio and pressure modalities respectively. The weightings have a value between [0,1] and are determined during a commissioning stage.

- To select a value for ω_a, a cross-validation (10-fold) is performed on the classifier implemented for audio-based detection by using a training dataset. The cross-validation accuracy is normalised and is taken as the value for ω_a.

- To select a value for ω_p, a validation routine is performed using a test pressure dataset with ground truth. The detection accuracy achieved for the test dataset is normalised and is taken as the value for ω_p.

Finally, the most probable state-of-blasting $(\widehat{X})$ in the combined probability distribution is taken as the detected state-of-blasting. Maximum a posteriori (MAP) decision rule is used to select the most probable hypothesis as:

$$\widehat{X} = \arg\max_{X=x} \Pr(\text{X=x} \mid \text{audio, pressure data }) \tag{8}$$

Given the detected (most probable) state-of-blasting calculated using the two sensing modalities, diagnosis for state-of-blasting faults can be performed as discussed in the next section.

2.6. Fault diagnosis

2.6.1. Grit-blasting spot position

Diagnosis for the fault of "wrong grit-blasting spot position", is performed by comparing the Euclidean distance between the detected and expected position of the grit-blasting spot. As shown in Figure 8, the position of the grit-blasting spot is reported as correct when the distance is within a predetermined threshold (e.g. $\tau_e < 50$mm, based on the size of the grit-blasting spot), otherwise a wrong grit-blasting spot position fault is detected.
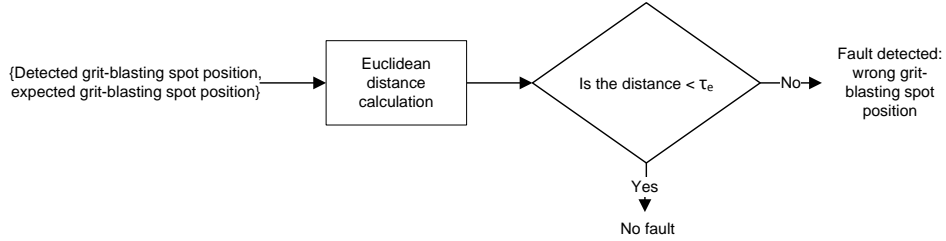
Figure 8: Diagnosing for grit-blasting spot position fault

2.6.2. State-of-blasting

Diagnosis for state-of-blasting related faults is performed by comparing the detected state-of-blasting against the expected. Table 1 shows the comparison chart used to diagnose for the state-of-blasting faults. In this table the expected state-of-blasting (rows) is compared against the detected state-of-blasting (columns) to identify the type of fault that has occurred where: green cells = no fault, yellow cells = a non-dangerous fault and red cell = dangerous fault. Given the identified state-of-blasting faults the robot can respond accordingly.

Expected state	Detected state		
	No-blasting	**Air-blasting**	**Grit-blasting**
No-blasting	No fault	Air is on	Grit is on
Air-blasting	Air is off	No fault	Grit is on
Grit-blasting	Grit is off	Grit is off	No fault

Table 1: State-of-blasting: fault diagnosis chart

3. Experiments

Three experiments are conducted to individually test the proposed detection techniques. For each experiment, the detection technique is firstly verified in a laboratory, and subsequently in a blasting chamber during real grit-blasting.

The experiments are conducted using the following equipments:

- A green laser pointer pen (532nm, 1mW) with a diffuser lens is used for producing the laser points

14

- A 6DoF schunk robot manipulator (shown in Figure 1) is used in the laboratory to position the laser point; the laser point is positioned by a human operator during experiment in the blasting chamber

- A PrimeSense carmine 1.09 is used for capturing RGB-D data

- A computer microphone is used for capturing audio data

- A AST4000 pressure sensor is used for capturing pressure data

3.1. Experiment 1: RGB-D laser point detection and diagnosis of "wrong grit-blasting spot position" fault

Experiment 1 tests the performance of RGB-based detection for a green laser point and fault diagnosis. Detection is firstly conducted in a dust-free laboratory and then subsequently in a dust laden grit-blasting chamber during grit-blasting of a steel plate. Figure 9 shows the setup for detecting a green laser point in the laboratory environment. The green laser pointer is mounted onto the end-effector of a robotic manipulator and a RGB-D camera (PrimeSense carmine 1.09) is positioned to observe the surface with the projected laser point. The laser point is moved through a trajectory over the surface by systematically moving the robot using pre-planned motions. Concurrently, RGB-D data is captured of the surface with the projected laser point.
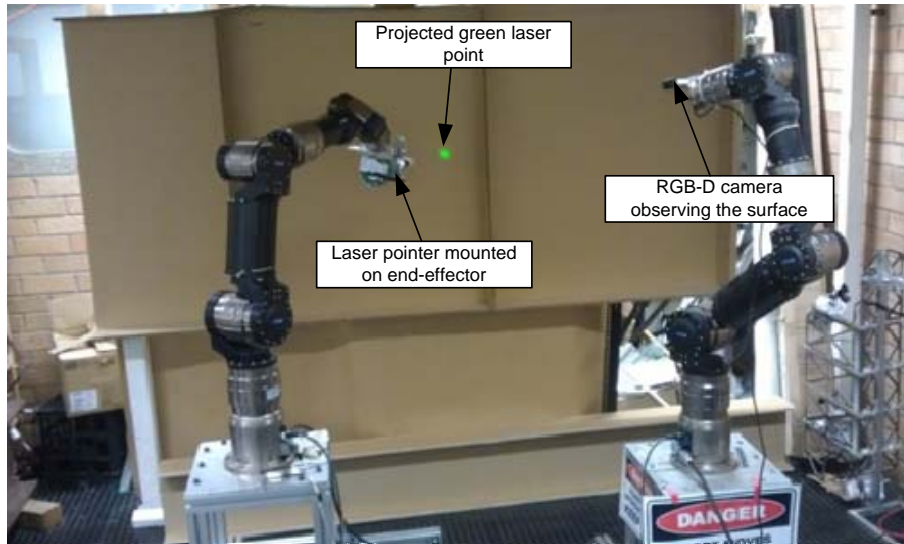
Figure 9: Laboratory setup to collect RGB-D dataset for laser point detection

Figure 10 shows the setup for detecting a green laser point in a blasting chamber during grit-blasting. The green laser pointer is mounted in-line with the blasting nozzle such that the projected laser point and grit-blasting spot are in close proximity. The RGB-D camera (PrimeSense carmine 1.09) is mounted onto a tripod and positioned to observe the surface to be grit-blasted. Grit-blasting of a steel plate is conducted manually by a human operator and RGB-D data is captured concurrently.
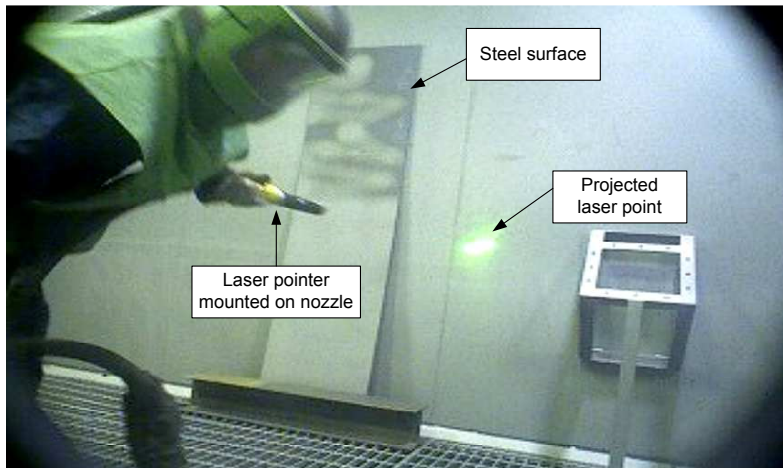
Figure 10: The setup in the blasting chamber to collect RGB-D dataset during grit-blasting

Laser point detection is performed on the RGB-D datasets collected in the

17

laboratory and grit-blasting chamber. Table 2 shows the parameter values set for both datasets. The performance of detecting the laser point is evaluated by identifying the erroneous detections; based on ground truth knowledge of the planned trajectory of the laser point.

Variable	Value
Maximum pixel distance between laser point in current frame and previous frame (distanceThreshold)	50mm
Intensity threshold for segmenting laser point pixels	75% of maximum intensity in image
Area/blob threshold for being a laser point (threshold)	5 pixels

Table 2: Parameter values for RGB-D based laser point detection

Figure 11 shows a sample RGB image extracted from the laboratory dataset and the corresponding point cloud of the observed scene. In this figure, the RGB image is annotated with the detected laser point for this sample, and the point cloud shows the marked location of all the laser points detected in the dataset (blue points). From a total of 422 samples it has been identified that 7 of the detections are erroneous (shown in Figure 11, bottom corner of point cloud). It is to note that the erroneous detections occurred when the laser point was out of view, thus affecting the relative threshold used to segment for the laser point. Overall, a detection accuracy of 98.34% is achieved for the laboratory dataset.
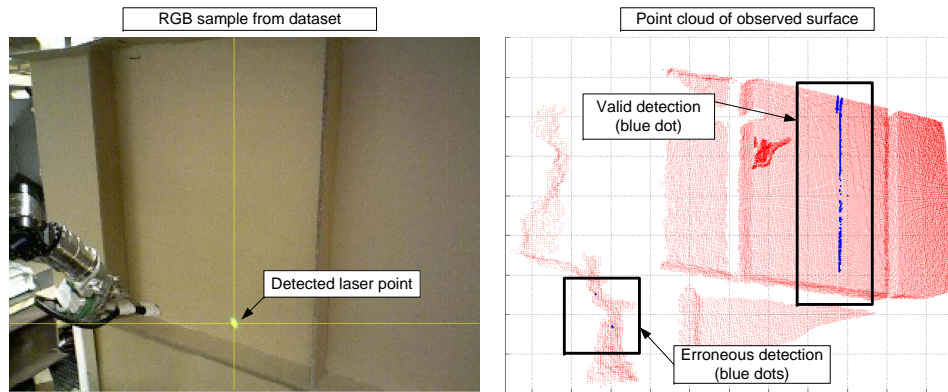
Figure 11: RGB image sample from the laboratory dataset; point cloud (red dots) of the observed surface with detected laser points (blue dots)

Figure 12 shows a sample RGB image extracted from the dataset captured during live grit-blasting in the blasting chamber and the corresponding point cloud of the observed scene. The RGB image is annotated with the detected laser point, and the point cloud shows the marked location of all the detected laser points in the dataset (blue points). From a total of 981 samples, all detections were identified to be valid. Thus, a detection accuracy of 100% has been achieved for the grit-blasting chamber dataset.
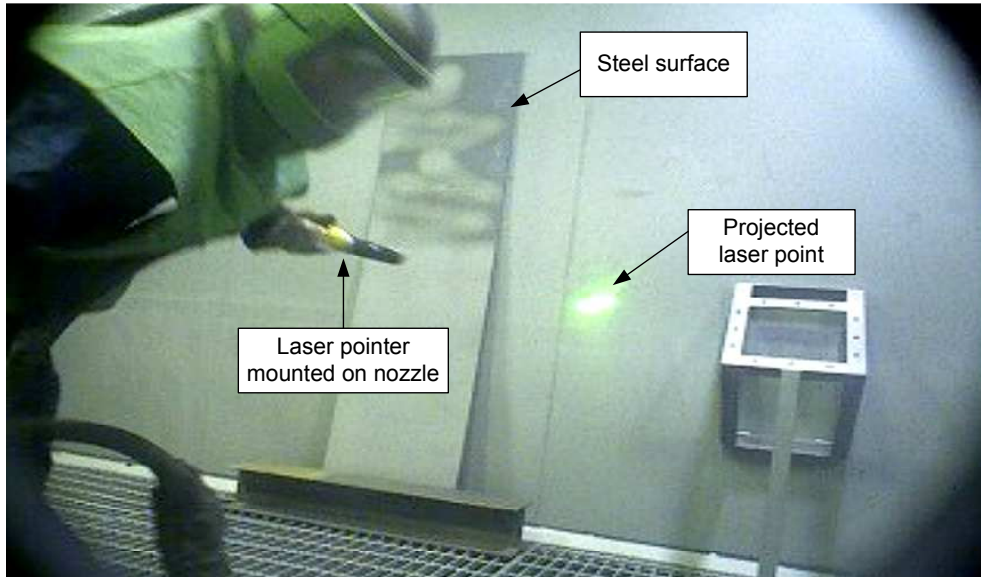
Figure 12: RGB image sample from the grit-blasting chamber dataset; point cloud (red dots) of the observed surface with detected laser points (blue dots)

Fault diagnosis is performed on the laboratory and blasting chamber datasets by using a distance threshold ($\tau_e = 50$mm). Overall, besides the erroneous detections in the datasets all detected positions of the laser point are in close proximity to the expected position (≤ 50mm difference). As a result, fault diagnosis accuracy for the laboratory and dust laden grit-blasting chamber datasets are 98.34% & 100% respectively. To summarise, this experiment has shown that the detection of a green laser point can be performed to a high degree of accuracy and can be used to diagnose for "wrong grit-blasting spot position" fault.

3.2. Experiment 2: Audio-based detection and fault diagnosis of the state-of-blasting

Experiment 2 tests the performance of audio-based detection for the states-of-blasting (*no-blasting*, *air-blasting* and *grit-blasting*) and fault diagnosis. To conduct the experiment, a microphone is attached to the exterior of a blasting hose as shown in Figure 13 . An initial test is performed in the laboratory by using a compressed air line (providing *no-blasting*, *air-blasting*) to verify that undistorted audio data can be captured. Figure 14 shows the data

collected in the laboratory, where the hose pressure was gradually increased from 0 Psi to maximum pressure (90 Psi). As shown by Figure 14, it can be observed that the audio signal amplitude increases relative to the increase in hose pressure; the valve was gradually opened to incrementally increase the pressure. However, at 90 Psi it can be observed that the amplitude of the audio signal exceeds the maximum dynamic range and resulted in audio clipping; due to the microphone being attached too close to the hose exit. Therefore, for subsequent capturing of data in the grit-blasting chamber the microphone is repositioned to the blasting hose outside of the chamber; away from the direct noise created by the hose exit/nozzle.
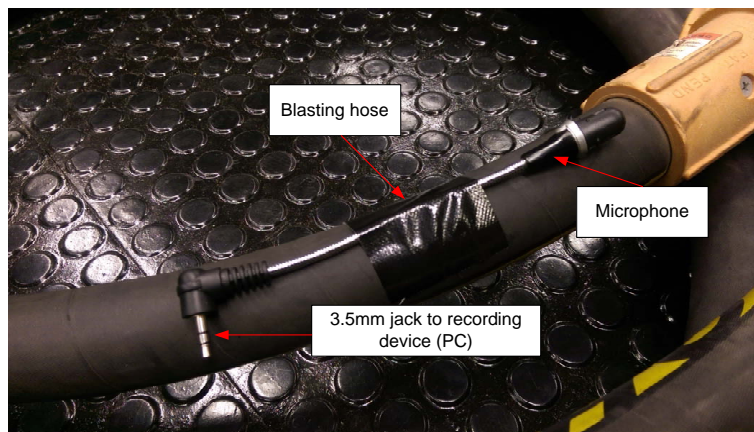

Figure 13: Set up to collect audio data by attaching a microphone to the exterior of the blasting hose
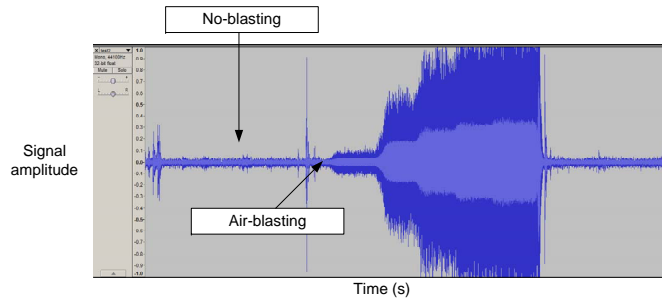
Figure 14: Audio data captured in the laboratory during blasting states: *no-blasting* and *air-blasting*

Figure 15 shows the data collected in the grit-blasting chamber for each states-of-blasting (*no-blasting, air-blasting* and *grit-blasting*); each audio signal is 5 seconds. Features including spectral entropy and spectral roll-off are extracted from each audio signal using a window size $W_t = 0.05$s, step size $K_t = 0.01$s. These sets of data will be used to train a naive Bayes classifier as specified in Table 3.
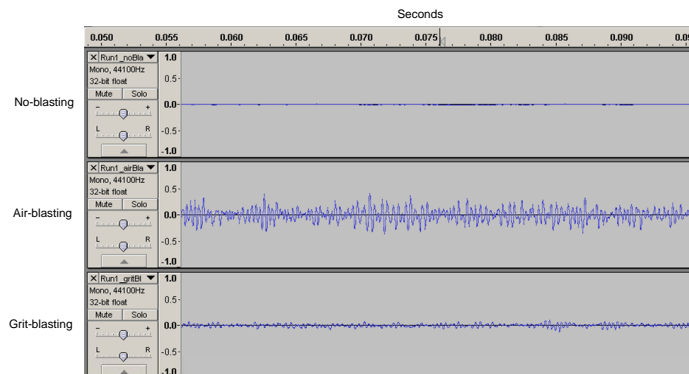

Figure 15: Audio dataset for training of a naive Bayes classifier

The classifier is then verified by applying it to classify a new dataset consisting of 69.63 seconds of audio signal captured during live grit-blasting with transitions between the three states-of-blasting. Figure 16 shows the dataset, where the first row shows the audio signal readings (3070720 samples, sampling rate = 44100Hz) over the whole time duration, the second row

shows the ground truth and detected (most probable) state-of-blasting over the duration of the audio signal (1392 samples; $W_t = 0.05$s, $K_t = 0.05$s, states-of-blasting are represented numerically $1 = no\text{-}blasting$, $2 = air\text{-}blasting$, $3 = grit\text{-}blasting$), and the third row shows the fault diagnosis results.

Overall, from the results shown in the second row of the figure, it can be observed that the detected state-of-blasting (broken blue line) and the ground truth of the state-of-blasting (solid red line) are closely matched and achieve a detection accuracy of 96.3%. For fault diagnosis, a simulated "*no-blasting*" state is assumed for the whole time duration, such that: fault detected = true (1) when *air-blasting* or *grit-blasting* is detected, otherwise fault detected = false (0). The accuracy achieved for this simulated fault diagnosis is 98.8% and is observed higher than detection accuracy because the diagnosis does not distinct between *air-blasting* and *grit-blasting states*. In summary, this indicates that diagnosis for state-of-blasting faults can be satisfactory achieved through audio-based detection.
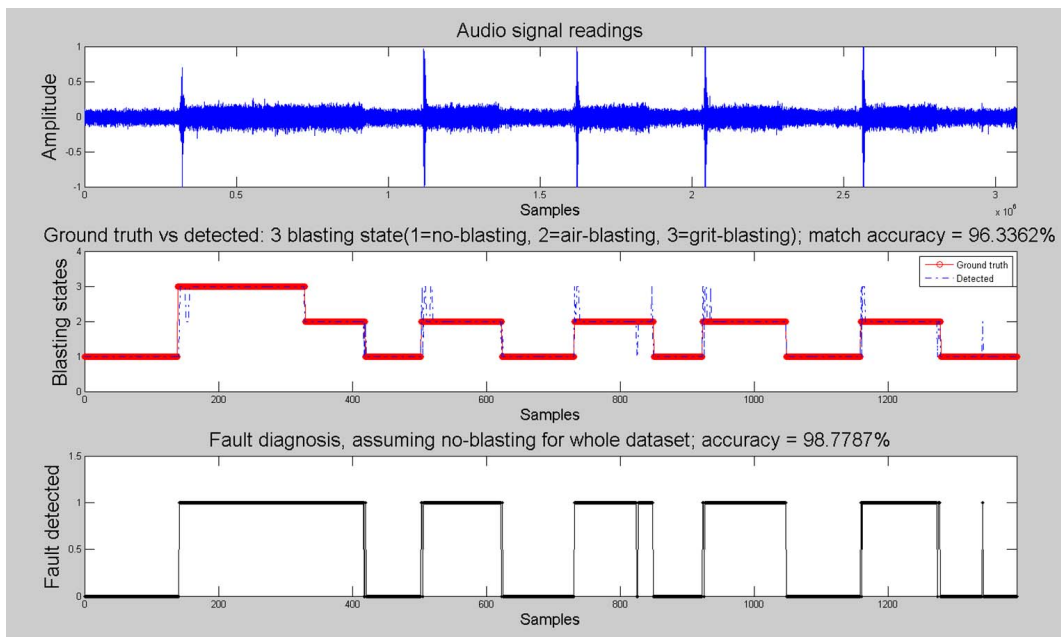

Figure 16: Testing dataset; row1: signal reading samples over the time duration; row2: ground truth and detected state-of-blasting over the time duration; row3: fault diagnosis

Variable	Value
Window size (stWin)	0.05s
Step size (stStep)	0.01s
Classifier	Naive Bayes Classifier
Audio features	Spectral entropy, spectral roll-off

Table 3: Parameters used for audio-based detection

3.3. Experiment 3: Pressure-based detection and fault diagnosis of the state-of-blasting

Experiment 3 tests the performance of pressure-based detection and fault diagnosis for the states-of-blasting (*no-blasting* and *air/grit-blasting*) using data captured by a pressure transducer rated to a maximum of 200 Psi. As shown in Figure 17 the pressure transducer is attached in-line with the blasting hose and signal from the transducer is recorded using an iUSBDAQ (Hytek U120816) at 30Hz . An initial test is conducted in the laboratory using a compressed air line to verify that pressure data can be capture. Figure 18 shows the data collected in the laboratory where the pressure was incrementally increased. From the figure, it can be observed that a step-like increase in signal amplitude was captured; corresponding with the incrementally increase in air pressure. Thus verifying that the transducer is sensitive for capturing pressure readings that can clearly distinguishes between the states of *no-blasting* and *air/grit-blasting*.
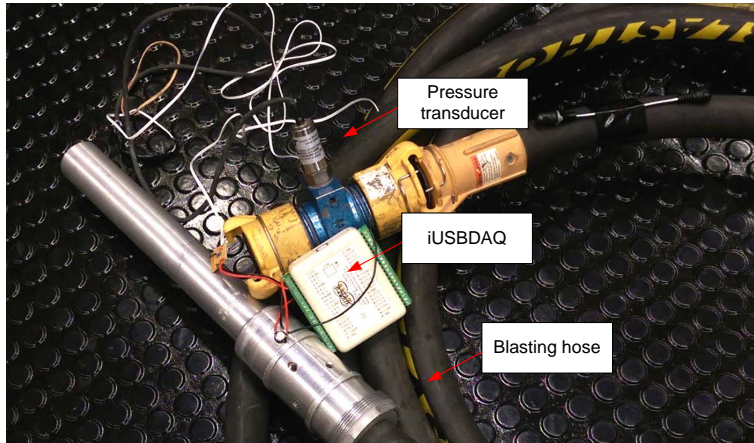
Figure 17: Set up to capture pressure data by attaching a pressure transducer in line with the blasting hose
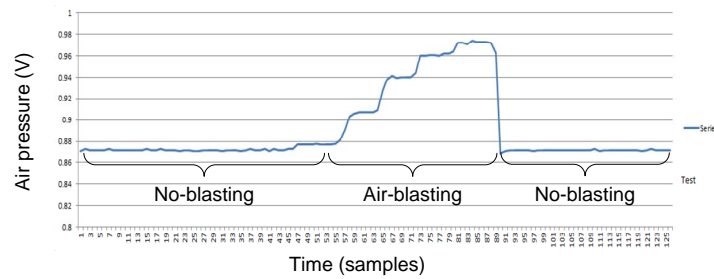

Figure 18: Laboratory dataset of pressure readings

Using the same configuration, a test dataset containing 4175 samples is collected during live grit-blasting where a human operator switches between *no-blasting* and *grit-blasting* at 5 seconds intervals. Figure 19 shows the data collected, where the first row in the figure shows the pressure readings over the time duration, the second row shows the detected and ground truth of the state-of-blasting over the time duration ($0 = $ *no-blasting*, $1 = $ *air/grit-blasting*), and the third row shows the binary fault diagnosis result.

In this experiment, a pressure threshold ($\tau_p = 1$ volt) is set based on the voltage values observed for each state-of-blasting in the dataset. The detection result shown in the second row indicates a close match between

the ground truth (solid red line) and the detected (broken blue line); achieving a detection accuracy of 97.9%. Furthermore, it can be observed that misclassification occurs exclusively during the transitions between blasting states. This can be attributed to pressure change within the hose during blasting state transitions; discharging gradually going from *grit-blasting* to *no-blasting* state. For fault diagnosis, a simulated "*no-blasting*" state is assumed for the whole time duration, such that: fault detected = true (1) when *air-blasting* or *grit-blasting* is detected, otherwise fault detected = false (0). The third row shows the results for the simulated fault diagnosis, achieving an accuracy of 97.9%. In summary, diagnosis for state-of-blasting faults can be satisfactorily achieved through pressure-based detection.
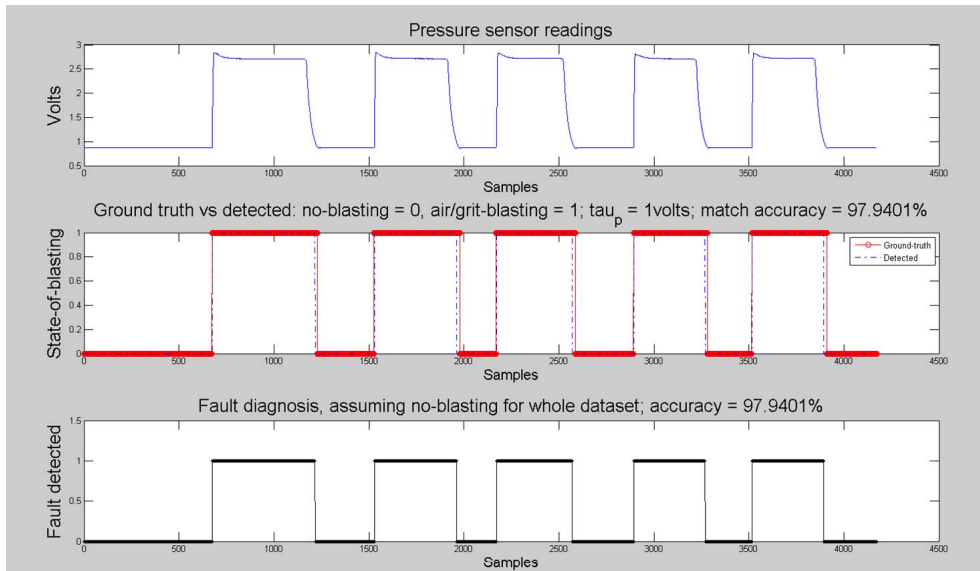

Figure 19: Test dataset; row1: signal over time duration of dataset; row2: ground truth and detected state-of-blasting; row3: fault diagnosis

4. Discussion and future work

4.1. Guideline and limitations for reliable use of the approach

The approach presented in this paper was able to achieve high performance in real-world experiments conducted during a grit-blasting operation. However, a limited range of real conditions were tested in the experiments

and the approach cannot be assumed reliable for all real-world scenarios. The following is a discussion about practical guidelines and limitations that should be considered for reliable use in different real-world applications. The discussion will focus on the RGB-D and audio based detection, given that pressure based detection is already commonly used in grit-blasting operation.

RGB-D based detection during grit-blasting presents the challenge of a dust laden environment. Due to the airborne dust, the projected laser will reflect off the target surface and also off the airborne particles creates an undesired "laser line" linking to the project laser point. Therefore the dust level in the environment should be controlled to reduce the presence of an undesired laser line. Deducing from the real-world experiment, the following steps can be taken to effectively reduce the level of dust: 1) conducting grit-blasting with a dust extraction system to constantly evacuate the dust particles in the air, 2) choosing a suitable grit-blasting garnet (i.e. steel shot) to reduce the level of dust particles released into the air. Nevertheless, it is important to understand the amount of airborne dust that will cause detection to fail and will be studied as part of the future work.

Furthermore, the colour of the laser point (red/green/blue) is another factor to consider for ensuring high accuracy. For example, avoid using a green laser point if grit-blasting is performed in a green environment or on a green painted surface. To address this factor, the following steps can be taken 1) selecting a laser point colour that contrasts with the surrounding environment, 2) ensure that the laser point appears in the captured images with the highest intensity in the intended colour channel. For example in the conducted experiment, the choice of a green laser point proved to be very effective in a grit-blasting chamber which contained mainly grey colours.

Audio based detection during grit-blasting presents the challenge of high ambient noise and/or other grit-blasting operations that may be happening concurrently. The following steps can be taken to reduce capturing noisy audio data 1) attach the microphone on the blast hose away from the grit-blasting end (i.e. the audio data collected during the real-world experiment was captured by attaching the microphone to the blast hose outside of the blasting chamber away from the noise), 2) implement an attachment that secures the microphone to the exterior of the blast hose and also provides insulation from external noise.

4.2. System integration

The experiments in this paper have independently tested the performance of each modality and provide a preliminary assessment of the proposed approach. However, future work is needed to test the complete fault diagnosis approach and involves setting up a system to captures all three sensing modalities during live robotic grit-blasting. It is proposed that this system can be developed using Robot Operating System (ROS) to provide a suitable framework for accessing sensor data and time stamping to a common clock. In this way, sensors publishing at different rates can be integrated and captured into a dataset. Fault diagnosis can be performed using the ROS framework to query the last available reading from each sensor to take the sample in any particular instance of time.

Furthermore, the approach has proposed the use of audio and pressure data to detect for the state-of-blasting. Detection is conducted independently using each modality and then the results are combined in a probabilistic manner to produce the most probably state-of-blasting. Alternatively, the audio and pressure features can be combined into a single step detection. However, independent detection using each modality is preferred for this approach to provide modularity. In this way, either sensing modality can be added and removed readily.

Given the capability to diagnose for the different types of faults during grit-blasting operations, the robot can provide a response accordingly. The responses may include; emergency stopping or pausing the robot, and re-planning the robot's task to automatically rectify the fault. However, this paper does not detail the actual response provided by a robot. It may be desirable for the robot to respond to a fault differently depending on the application, hence the response behaviour will need to be defined for each case when integrating fault diagnosis into the overall robotic system.

5. Conclusion

This paper has presented an approach to diagnosis faults during grit-blasting operation by detecting the state within the blasting line and the position of the grit-blasting spot. Three sensing modalities (RGB-D, audio and pressure) are used as part of the approach and experiments have been conducted to evaluate each sensor's fault diagnosis performance. Experiment 1 demonstrated the use of a RGB-D sensor and a laser pointer to detect for the real-time position of the grit-blasting spot; achieving fault diagnosis

accuracy of above 98.34%. Experiment 2 demonstrated the use of an audio sensor to detect three states of blasting; achieving fault diagnosis accuracy of 98.8%. Experiment 3 demonstrated the use of a pressure transducer to detect between no-blasting and air/grit-blasting; achieving fault diagnosis accuracy of 97.9%. Overall, the three experiments have demonstrated that the proposed fault diagnosis approach is feasible. As part of the suggested future work, a further experiment that integrates all three sensors will need to be conducted in order to evaluate the performance of the proposed fault diagnosis approach as a whole functioning system.

Acknowledgment

References

[1] S. Lahiri, C. Levenstein, D. Nelson, B. Rosenberg, The cost effectiveness of occupational health interventions: prevention of silicosis, American journal of industrial medicine 48(6) (2005) 503 – 517.

[2] D. K. Liu, G. Dissanayake, P. B. Manamperi, G. Fang, N. Kirchner, G. Paul, S. Webb, P. Chotiprayanakul, J. Xie, A Robotic System for Steel Bridge Maintenance: Research Challenges and System Design, In Proceedings of the Australasian Conference on Robotics and Automation (2008, Canberra, Australia).

[3] G. Paul, N. Kirchner, D. K. Liu, G. Dissanayake, An effective exploration approach to simultaneous mapping and surface material-type identification of complex three-dimensional environments, Journal of Field Robotics 26 (2009) 915 – 933.

[4] K. Micallef, G. Fang, M. Dinham, Automatic seam detection and path planning in robotic welding, Robotic Welding, Intelligence and Automation 88 (2011) 23 – 32.

[5] Y. L. Xu, X. F. Kong, S. B. Chen, An optimal design of multifunctional vision sensor system for welding robot, Robotic Welding, Intelligence and Automation 88 (2011) 145 – 152.

[6] N. Lv, S. B. Chen, Investigation on acoustic signals for on-line monitoring of welding, Robotic Welding, Intelligence and Automation 88 (2011) 235 – 243.

[7] A. Pal, G. Schaefer, M. E. Celebi, Robust codebook-based video background subtraction, Int. Conf. on Acoustics Speech and Signal Processing (2010) 1146 – 1149.

[8] S. S. Mohamed, N. M. Tahir, R. Adnan, Background modelling and background subtraction performance for object detection, Signal Processing and Its Applications (2010) 1 – 6.

[9] B. Yang, Y. L. Guo, Y. Y. Ming, A. D. Men, An effective background subtraction under a continuously and rapidly varying illumination, Int. Conf. on Future Networks (2010) 16 – 19.

[10] X. C. Zou, X. B. Zhao, Z. Chi, A robust background subtraction approach with a moving camera, Int. Conf. on Computing and Convergence Technology (2012) 1026 – 1029.

[11] E. Nunes, A. Conci, A. Sanchez, Robust background subtraction on traffic videos, Int. Conf. on Systems, Signals and Image Processing (2011) 1 – 4.

[12] G. C. Karras, K. J. Kyriakopoulos, Visual servo control of an underwater vehicle using a laser vision system, Int. Conf. on Intelligent Robots and Systems (2008) 4116 – 4122.

[13] M. Luber, L. Spinello, K. O. Arras, People tracking in rgb-d data with on-line boosted target models, Int. Conf. on Intelligent Robots and Systems (2011) 3844 – 3849.

[14] T. Giannakopoulos, A. Pikrakis, Introduction to audio analysis: A matlab approach, Academic Press (2014).

[15] F. Ykhlef, L. Bendaouia, Evaluation of time domain features for voiced/non-voiced classification of speech, Int. Conf. on Signals and Electronic Systems (2012) 1 – 4.

[16] F. Zheng, G. Zhang, Z. Song, Comparison of different implementations of mfcc, Journal of Computer Science and Technology 16 (2001) 582 – 589.

[17] E. Frank, M. Hall, L. E. Trigg, G. Holmes, I. H. Witten, Data mining in bioinformatics using Weka, Bioinformatics 20 (2004) 2479 – 2481.

[18] C. J. C. Burges, A tutorial on support vector machines for pattern recognition, Kluwer Academic Publishers, Boston. (2009) 43.

[19] Z. Qin, Naive Bayes Classification Given Probability Estimation Trees, International Conference on Machine Learning and Applications (2006) 34 – 42.

[20] H. Samet, K-nearest neighbor finding using maxnearestdist, Pattern Analysis and Machine Intelligence 30 (2008) 243 – 252.

[21] R. M. Haralick, L. G. Shapiro, Computer and Robot Vision, Addison-Wesley Longman Publishing Co., Inc. (1991).

[22] X. Geng, T. Y. Liu, T. Qin, H. Li, Feature selection for ranking, ACM SIGIR conference on Research and development in information retrieval (2007) 404 – 414.

[23] M. Gutlein, E. Frank, M. Hall, A. Karwath, Large-scale attribute selection using wrappers, Computational Intelligence and Data Mining (2009) 332 – 339.

[24] E. Scheirer, M. Slaney, Construction and evaluation of a robust multifeature speech/music discriminator, Int. Conf. on Acoustics, Speech, and Signal Processing 2 (1997) 1331 – 1334.