GMAW metal transfer mode identification from welding sound

Sipei Zhao (1), Xiaojun Qiu (1), Ian Burnett (1), Malcolm Rigby (2) and Anthony Lele (2)

(1) Centre for Audio, Acoustics and Vibration, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia
(2) Sound Intuition Pty. Ltd., Melbourne, Australia

ABSTRACT
Gas Metal Arc Welding (GMAW) is an arc welding process that forms an electric arc between a consumable electrode and the base metal with a shielding gas to protect the arc. In GMAW, there are various metal transfer modes such as the short circuit mode, the globular mode, the spray mode, and the rotational transfer mode, which show different arc stabilities, weld pool penetrations and spatter production. Identifying the metal transfer mode is critical for process monitoring and quality control of GMAW. In this paper, a method for metal transfer mode identification from the welding sound is presented. A recorder mounted on the welder helmet is used to record the sound signals generated by GMAW under different metal transfer modes, which are analysed in both time and frequency domains. New psychoacoustic parameters based on the auditory perception of an expert welder are extracted to distinguish the metal transfer modes. The Gaussian Mixture Model (GMM) is utilised to identify the metal transfer mode from the welding sound signals and a 10-fold cross validation shows 90% recognition accuracy.

1 INTRODUCTION
Arc welding is used to join two or more materials, through fusion, such as the joint exhibits a sufficient strength and fracture toughness (Tam 2005). Gas Metal Arc Welding (GMAW) is an arc welding process that uses the arc between a consumable electrode and the welding pool with a shielding from externally supplied gas without any application of pressure (Naidu, Ozcelik, and Moore 2003). It has been reported that the expert GMAW welders are able to maintain and direct the welding arc using a combination of their visual and auditory senses, and a series of psychoacoustic experiments with professional welders showed that the welding performance significantly degraded without the acoustic feedback (Tam 2005). Tam and Huissoon (2005) carried out psychoacoustic experiments to study the dependency of welders on acoustic cues in welding and found that application of time delays of welders’ acoustic response to have detrimental effect on their ability to weld and the welder’s response was shown to be completely unstable and erratic when the time delay is 400 ms. The acoustic emissions from a GMAW process contain information on the weld quality and can be utilized for online monitoring, inspection and quality control of welding processes.

Jolly (1969) was among the first to use the acoustic emission as a weld quality monitor for Gas Tungsten Arc Welding (GTAW) and his research shows that defects in the weld can be located approximately, but the types of defect cannot be determined. The acoustic emission from GTAW is spread over a wide frequency range, extending beyond 30 MHz, but the optimum band of frequencies limited by the low frequency machinery noise and the attenuation of higher frequencies in the material is approximately 100 kHz to 2 MHz for ultrasonic sensor (Jolly 1969). Similarly, Hopwood II (1974) also used ultrasonic acoustic sensing to avoid interferences noise in defect detection in the GMAW process.

About 20 years later, Saini and Floyd (1998) explored the welding sound for online quality control of automated welding with various time and frequency domain audio features. The research shows that the time domain parameters enables the detection of deviations from ideal arc while the frequency domain parameters offer some promise for detection of metal transfer mode. Similarly, Wang and Zhao (2001) performed time and frequency domain analysis of the welding sound and found the sound energy below 100 Hz to be promising for detecting the keyhole status in the welding process. Cayo and Alfaro (2009) also compared the time and frequency domain analysis of the welding sound for stability evaluation of the GMAW process and concluded that the time domain analysis presents more clarity than the frequency domain. Ting et al. (2011) used the wavelet package trans-
formation to decompose the welding sound signals into 128 channels within the frequency domain, and to analyse the properties of the sound signals using the wavelet based channel energy.

Grad et al. (2004) evaluated the feasibility of acoustic signals for online monitoring in the short circuit GMAW and showed that the main source of acoustic waves was the arc reignition in short circuit mode. Cudina et al. (2008) predicted the sound pressure from the electrical current using their proposed transfer function model and used the predicted sound signal to monitor the GMAW process to avoid the influence of background noises in the measured sound signals. Yusof et al. (2017) used the Hilbert Huang transform (HHT) based on empirical mode decomposition to filter out the unrelated signal components in the welding sound signals and showed that the application of HHT analysis significantly assisted in identifying hidden information related to the defects. Cayo and Alfaro (2011) calculated the ignition rate and sound level from welding sound signals to assess the quality of the welding. It has been found that when there is no interference in the welding, both the ignition rate and sound level are stationary and the statistical distribution is delimited in an elliptical region, whereas when there is interference, the statistical distribution moves out from the stability region (Cayo and Alfaro 2011).

In addition to the abovementioned studies based on conventional signal processing techniques, machine learning algorithms have also been used for GMAW process monitoring using welding sound signals. Tam (2005) employed artificial neural networks to predict the welding parameters from the measured acoustic signals. Wang and Huissoon (2009) used artificial neural networks with a Bayesian classifier to classify the welding based on the time and frequency domain parameters of the welding sound signal, but no explanation was given on the exact parameters used in the study. Sumesh et al. (2015) employed random forest algorithms to classify 3 categories of welding quality, i.e., good weld, and weld with lack of fusion and burn through, where the inputs for the decision tree algorithm were the statistical features extracted from the welding sound signals; however, the exact features were not presented in their paper. Lv et al. (2017) developed an online welding quality monitoring system via acoustic signals based on back propagation artificial neural networks and found that by training with a large data set, the prediction rate reached 80-90% for the degree of penetration detection. In summary, the significance of acoustic emissions in monitoring the arc welding processes has been investigated with many approaches in existing literature, but most of the abovementioned studies focus on welding quality monitoring for one metal transfer mode, which may fail for other metal transfer modes. In GMAW, there are various metal transfer modes such as the globular, spray, streaming, and rotating transfer modes, which have different arc stabilities, weld pool penetrations, spatter production, porosity population and level of gas entrapment (Kim and Eagar 1993). The mode of metal transfer depends on many operational variables, which include welding current, composition of shielding gas, electrode extension, ambient pressure, polarity and welding material, among which the welding current is the most common variable that the welder adjusts to obtain the desired metal transfer mode (Kim and Eagar 1993). Identifying the metal transfer modes is critical for process monitoring and quality control of GMAW. Although a mode can be selected at initial set up it can inadvertently move from one mode to another due to changes in input parameters during welding. This results in reduced weld quality and increased production cost. Having the welder know or maintain the selected mode throughout the weld maintains its quality, controls production throughput and reduces production costs.

In this paper, the welding sound for different metal transfer modes is recorded by mounting a recorder on a welder helmet. The welding sound signals are examined and compared in both the time and frequency domains first. Two psychoacoustic parameters are then defined and calculated to characterize the difference between each mode based on the auditory perception of an expert welder. In addition, 10 time and 4 frequency domain features are calculated to form a 16-dimensional feature vector together with the 2 psychoacoustic parameters, which are fed into a 3-component Gaussian Mixture Model (GMM) for training and testing. A 10-fold cross validation test is used to demonstrate the recognition rate of the proposed method.

2 WELDING SOUND RECORDING
To record the welding sound, a Tascam Portable Handheld DR-22 Recorder was mounted on the welder helmet, as shown in Figure 1. In the measurements, the gain of the recorder was fixed for all recordings so that the sound levels for different metal transfer modes can be compared. However, the recorder was not calibrated and the absolute value of sound pressure level is not used as an audio feature. Each recording had a duration of ap-
proximately 10 minutes with a sampling rate of 96 kHz. For the welding sound signal processing in the following sections, the recordings were cut into segments of 5 s for audio feature extractions.

The welding sound for different metal transfer modes were recorded for both continuous GMAW and Pulsed-GMAW processes. For the continuous GMAW, the current and voltage are supplied continuously in the welding process. Figure 2 shows the metal transfer modes changing from the short-circuit to globular to spray and finally to rotational, with increasing welding current for continuous GMAW (Tam 2005). The short circuit transfer is a special metal transfer mode where the molten droplet on the wire tip makes direct contact with the workpiece or the surface of the weld pool. It is characterized by repeated, intermittent arc extinguishment and re-ignition. (Wang et al. 2003). The globular transfer, where the droplet diameter is larger than the wire diameter, occurs at low current. Since it is often accompanied by excessive spatter, the globular transfer is only used for welding unimportant parts. The spray transfer, where the droplet diameter is smaller than the wire diameter, occurs at medium and high current. It is a highly stable and efficient process, and is widely used in welding thick steel plates (Tam 2005). Typical manufacturing applications prefer globular and spray transfer modes as they offer better balance between material deposition speed, penetration, and bead aesthetics, whereas the short circuit transfer mode is used in out of position on some heavy structural joining applications and thin materials (Tam 2005).

Typical audio waveforms and the corresponding frequency spectra for the short circuit transfer mode, the globular transfer mode, and the spray transfer mode are illustrated in Figure 3. The frequency spectra were calculated with the Welch method, where the 5 s audios were segmented into 8 sections with a 50% overlap. It can be seen from Figure 3(a) that the audio waveform for the short circuit transfer shows a regular pattern of acoustic pulses, which is also represented by the 48 Hz peak and its harmonics in Figure 3(b). In contrast to the short circuit transfer, the acoustic pulses for the globular transfer in Figure 3(c) are irregular, which is also illustrated by the irregular peaks in the spectrum in Figure 3(d). In contrast, there are rare acoustic pulses for the spray transfer mode in Figure 3(e), so no peaks in the frequency spectrum is observed in Figure 3(f). The difference in the welding sound will be utilized to identify different metal transfer modes in the next section.
Figure 3: Typical audio waveforms and frequency spectra for different metal transfer modes in the continuous GMAW processes, (a) waveform and (b) spectrum for the short circuit transfer, (c) waveform and (d) spectrum for the globular transfer, and (e) waveform and (f) spectrum for the spray transfer mode.

In addition to the continuous GMAW, the Pulsed-GMAW is often used to improve weld quality as well as productivity in thin sheet metal industries, where the welding current or voltage is cycling from a high level to a low level at a desired pulse frequency (Pal, Bhattacharya, and Pal 2010). Two types of metal transfer modes in the
Pulsed-GMAW process, i.e., contact transfer and spray transfer modes, are illustrated in Figure 4(a) and (b), respectively. The typical waveforms and the corresponding frequency spectra for both the contact transfer and spray transfer modes in the Pulsed-GMAW processes are shown in Figure 5.

Figure 4: Illustration of (a) contact transfer mode and (b) spray transfer mode in Pulsed-GMAW processes. (Doodman Tipi, Hosseini Sani, and Pariz 2015)

Figure 5: Typical audio waveforms and frequency spectra for different metal transfer modes in the Pulsed-GMAW processes, (a) waveform and (b) spectrum for the contact transfer mode, and (c) waveform and (d) spectrum for the spray transfer mode.
It can be observed from Figure 5(a) that the acoustic pulses are also regular for the contact transfer mode in the Pulsed-GMAW processes, with a pulse rate of 163 Hz, as shown by the peaks in the frequency spectrum in Figure 5(b). Similarly, the spray transfer mode in Figure 5(c) also shows regular acoustic pulses, with a peak frequency at approximately 156 Hz in the frequency spectrum in Figure 5(d). However, the acoustic pulse amplitude of the spray transfer mode in Figure 5(c) is much lower than that of the contact transfer mode in Figure 5(a). By comparing Figures 4 and 5, it is clear that the acoustic pulse rate in the welding sound produced by the Pulsed-GMAW is much higher than that produced by the continuous GMAW, which can be used to distinguish different modes in the two processes in the next section.

3 METAL TRANSFER MODE IDENTIFICATION

3.1 Method

The proposed method for metal transfer mode identification based on welding sound is depicted in Figure 6. The recorded input audios are divided into training data and testing data, from which various audio features are extracted. In the training process, the audio features are fed into the Gaussian Mixture Model (GMM) to estimate the model parameters that best fit the audio features for different metal transfer modes. In the testing process, the model parameters are fixed and the audio features for testing data are fed into the GMM, which recognises the metal transfer mode automatically in the identification stage.

![Figure 6: Diagram of the metal transfer mode identification method based on the Gaussian Mixture Model.](image)

3.2 Audio Features

Based on the psychoacoustic listening test and the description from an expert welder (Rigby) in the author list, one of the key perception cues to distinguish different metal transfer modes from welding sound is the rhythmic pattern of the acoustic pulses, as shown in Figure 3. To quantify the difference, the pulse rate is defined as the repetition frequency of the acoustic pulses and calculated by estimating the package envelope of the acoustic pulses and counting the number of peaks of the envelope in one second. Because the time interval between each pair of adjacent acoustic pulses is varying with time, both the mean value and the standard deviation of the pulse rate are calculated for each 5 s long audio segment.

Figure 7 shows the mean value and standard deviation of the pulse rate for different metal transfer modes in the continuous GMAW processes based on 10 audio segments for each mode. The mean pulse rate of the short circuit transfer mode is the highest and mean pulse rate of the spray transfer mode is the lowest. In contrast, the standard deviation of the pulse rate is the lowest for the short circuit transfer mode while the highest for the spray transfer mode. These observations are consistent with the audio waveforms illustrated in Figure 3. Similarly, the mean value and the standard deviation of the pulse rate for both the contract transfer mode and the spray transfer mode are also consistent with the audio waveforms.
mode in the Pulsed-GMAW are calculated and shown in Figure 8. It is clear that the mean pulse rate for the Pulsed-GMAW is much higher than that for the continuous GMAW, while the standard deviation of the pulse rate are similar. Both the mean value and the standard deviation of the pulse rates are calculated as the psychoacoustic parameters in the audio features for the metal transfer mode identification.

![Figure 7](image1.png)

**Figure 7:** Comparison of (a) the mean value and (b) standard deviation of the pulse rate for different metal transfer modes of the continuous GMAW.

![Figure 8](image2.png)

**Figure 8:** Comparison of (a) the mean value and (b) standard deviation of the pulse rate for different metal transfer modes of the Pulsed-GMAW.

In addition to the 2 psychoacoustic parameters, 10 time domain parameters and 4 frequency domain parameters are also calculated as the audio features. The definitions of the time domain and frequency domain parameters are presented in Tables 1 and 2, respectively, and these were originally used for bearing fault diagnosis by (Xia et al. 2012). In total, 16 parameters are calculated in the feature extraction step and a 16-dimensional feature vector is formed for each audio signal. These feature vectors are fed into the GMM model for training and testing, as discussed in the next section.
### Table 1: Time domain features extracted from welding sound for metal transfer mode identification

<table>
<thead>
<tr>
<th>Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero Crossing Rate (ZCR)</td>
<td>[ ZCR = \frac{1}{N} \sum_{n=0}^{N-1} \left</td>
</tr>
<tr>
<td>Root Mean Square (RMS)</td>
<td>[ X_{\text{RMS}} = \left( \frac{1}{N} \sum_{i=1}^{N} x_i^2 \right)^{1/2} ]</td>
</tr>
<tr>
<td>Square Root of Amplitude (SRA)</td>
<td>[ X_{\text{SRA}} = \left( \frac{1}{N} \sum_{i=1}^{N} \sqrt{</td>
</tr>
<tr>
<td>Kurtosis Value (KV)</td>
<td>[ X_{\text{KV}} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i - \mu_x}{\sigma} \right)^4 ]</td>
</tr>
<tr>
<td>Skewness Value (SV)</td>
<td>[ X_{\text{SV}} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{x_i - \mu_x}{\sigma} \right)^3 ]</td>
</tr>
<tr>
<td>Peak-Peak Value (PPV)</td>
<td>[ X_{\text{PPV}} = \max(x_i) - \min(x_i) ]</td>
</tr>
<tr>
<td>Crest Factor (CF)</td>
<td>[ X_{\text{CF}} = \frac{1}{N} \left( \sum_{i=1}^{N} x_i^2 \right)^{1/2} ]</td>
</tr>
<tr>
<td>Impulse Factor (IF)</td>
<td>[ X_{\text{IF}} = \max({</td>
</tr>
<tr>
<td>Margin Factor (MF)</td>
<td>[ X_{\text{MF}} = \frac{1}{N} \left( \sum_{i=1}^{N} \sqrt{</td>
</tr>
<tr>
<td>Shape Factor (SF)</td>
<td>[ X_{\text{SF}} = \frac{1}{N} \left( \sum_{i=1}^{N} x_i^2 \right)^{1/2} ]</td>
</tr>
</tbody>
</table>

### Table 2: Frequency domain features extracted from welding sound for metal transfer mode identification

<table>
<thead>
<tr>
<th>Features</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak Frequency (PF)</td>
<td>[ \max \arg (s(f)) ]</td>
</tr>
<tr>
<td>Frequency Centre (FC)</td>
<td>[ FC = \frac{\int_{0}^{\infty} f s(f) df}{\int_{0}^{\infty} s(f) df} ]</td>
</tr>
<tr>
<td>Root Mean Square Frequency (RMSF)</td>
<td>[ \text{RMSF} = \left( \int_{0}^{\infty} f^2 s(f) df \right)^{1/2} ]</td>
</tr>
<tr>
<td>Root Variance Frequency (RVFR)</td>
<td>[ \text{RVFR} = \left( \frac{\int_{0}^{\infty} (f - FC)^2 s(f) df}{\int_{0}^{\infty} s(f) df} \right)^{1/2} ]</td>
</tr>
</tbody>
</table>
In the GMM used in this paper, 3 Gaussian components are used, i.e., $M = 3$. In the training stage, the feature vectors of the training data are employed to calculate the GMM parameters (i.e., $w_i$, $\mu_i$, and $\sigma_i$) that have the highest probability to represent the training data for each mode. In the recognition stage, the GMM parameters are fixed and the probability for each mode is calculated for the testing data, and the mode with the highest probability is chosen as the recognition result. To evaluate the performance, a 10-fold cross validation is used, i.e., the input audios are randomly divided into 10 sets, of which 9 sets are used for training and the remaining 1 set is used for testing. The cross validation is repeated 10 times, with each set of the data used once as the testing data, and the 10 results are averaged to produce a single estimation.

In this paper, 5 metal transfer modes (3 for the continuous GMAW and 2 for the Pulsed-GMAW) are to be identified, and 10 sample audios with a length of 5 s were recorded for each mode. The 10-fold cross validation result showed a recognition rate of 90% (45 correct out of 50 tests). It is noteworthy that the audio signals used in this study are carefully chosen typical welding sounds for each mode and do not display interference from low frequency fluctuations observed in the recording. The exact reason for the fluctuations are unclear at this stage and will be investigated in the future. In addition, both the training and testing data are limited in the current study. We are establishing a measurement system with professional equipment and will build a large scale welding sound database for further comprehensive research, which will be communicated in our future papers. Finally, the physical mechanism of welding sound generation is not studied here, but will be investigated in detail in the future.

4 CONCLUSIONS
This paper presents the results from a preliminary study into identification of the metal transfer mode from welding sound signals. A portable recorder was mounted on the welder’s helmet to record the welding sound for different metal transfer modes, which were examined and compared in both time and frequency domains. According to the auditory perception results from an expert welder, the mean value and standard deviation of the pulse rate are calculated and employed as the psychoacoustic parameters. Furthermore, 10 time domain and 4 frequency domain features were calculated to form a 16-dimensional feature vector as the input to the 3-component Gaussian Mixture Model for training and testing. A 10-fold cross validation shows a 90% recognition rate. Future work will include establishing a measurement system and building a large database for further investigation. The physical generation mechanism of welding sound will also be investigated.

ACKNOWLEDGEMENTS
This research was supported under the Australian Research Council’s Linkage Project funding scheme (No. LP160100616).

REFERENCES


Lv, Na, Yanling Xu, Sichen Li, Xinwen Yu, and Shanben Chen. 2017. “Automated Control of Welding Penetration


https://www.audiolabs-erlangen.de/resources/MIR/TSMtoolbox.


