In-solid Acoustic Sources Extraction Using a Multi-sensor Architecture

D.T. Pham, M. Yang and M. Al-Kutubi
Manufacturing Engineering Centre
Cardiff University, Cardiff CF24 3AA, UK

Abstract
This paper presents a novel signal feature extraction method based on Location Pattern Matching (LPM) for improving the reliability of pattern matching of received acoustic signals by using multi-sensor architecture. The proposed solution is to solve the reliability problem of signal matching by extracting specific features from the signal that is only associated with the source location not the source information with multi-sensor layout. Experiments showed that this method can increase the tolerance of LPM on various types of impacts, thus the matching accuracy.
Keywords signal feature extraction, pattern matching, multi-sensor architecture

1 Introduction

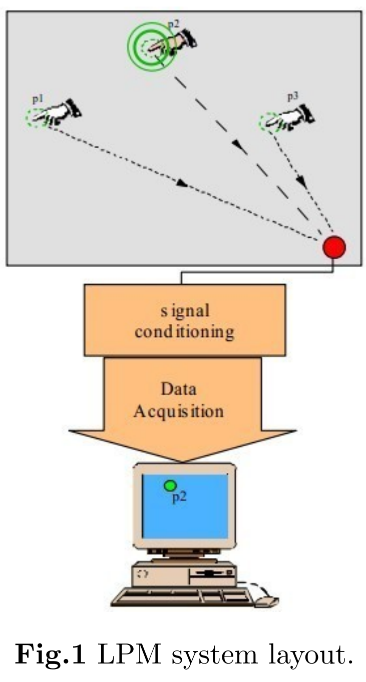


Fig.1 LPM system layout.

According to the well described acoustic time-reversal theory [1], it is possible to reconstruct an acoustic signal at its initial excitation location in a scattering medium by recording the received signal and sending back the time reversed version of the signal through the medium. This implies that the received signal carries its source location signature as a result of scattering in the transmission medium and reflections from its complex boundaries.

In the LPM approach, the uniqueness of features embedded in a received acoustic signal, as a representation of its source location, is employed to determine if it matches one of the predefined locations on a tangible object. This is achieved by creating an array of template of received signals generated from taps on pre-known locations [2-5]. The source location can then be determined by a cross correlation analysis from the best similarity match with a signal in the templates.

Fig.1 shows a block diagram of an LPM system comprising a board, one sensor located at the bottom right of the board, signal conditioning hardware, data acquisition device and a host PC. A tapping on the hardboard is detected by the sensor, conditioned by the signal conditioning hardware, digitised by the data acquisition card and processed by the PC. The matching result can be achieved by a cross correlation analysis which is widely used to measure similarity between two signals or images.

2 Pattern-based Localisation

The theory which location identification from the received signal pattern is based on is that an acoustic wave propagating from its source to a destination carries a specific signature in its pattern associated with the source location as a result of scattering.

When a driving force is applied into a medium, a mechanical wave is generated transporting energy away from the source of disturbance. In a bounded medium the waves propagate until they reach the boundaries and are reflected or absorbed. The reflections cause reverberation on the received signal acquired by a transducer at certain location in the medium. To comprehend the effect of boundary on the magnitude and phase of a received signal, a plane wave is assumed propagating in the x direction, the acoustic wave equation is given by:

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{\nu_p^2} \frac{\partial^2 p}{\partial t^2} \tag{1}$$

where p is the acoustic pressure as a function of time t and distance x.The solution of the differential equation (1) is the propagating wave equation given by [6]:

$$p = Ae^{j(\omega t - kx)} \tag{2}$$

where A is the amplitude constant, ν_p phase speed and k is the wave number. With this wave equation, it can be shown how a transmission medium with one reflection affects the received signal. Assume a simple model of two signal paths from the transmitter to the receiver. One direct path with unity gain and a delay t_d , and the other is reflected with attenuation α and a delay $t_d+\Delta_t$ resulting from the path length difference. The overall transfer function of such a transmission medium $H(\omega)$ can be expressed by:

$$\begin{aligned} H(\omega) &= e^{-j\omega t_d} + \alpha e^{-j\omega(t_d+\Delta_t)} \\ &= \overbrace{\sqrt{1+\alpha^2+2\alpha\cos\omega\Delta_t}}^{|H(\omega)|}\, \overbrace{e^{-j(\omega t_d+\tan^{-1}\frac{\alpha\sin\omega\Delta_t}{1+\alpha\cos\omega\Delta_t})}}^{\theta_h(\omega)} \end{aligned}$$

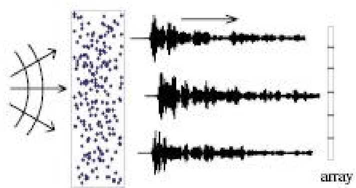
Therefore, multi-path reflections cause the distortion in the magnitude $|H(\omega)|$ and the phase $\theta_h(\omega)$ characteristics of the transmission medium. Since the time delay is the product of the path length difference by the wave-number, it is clear from (1) that the magnitude and phase of the received signal at a certain location will vary as the source location changes. In random medium the received signal is a combination of the direct wave and multiple delayed scattered waves went through reflections and refractions plus the effect of non isotropic material, and therefore the received signal in reality has much more complicated relation to its source location than that given in equation (1). With this phenomenon, signals received from different source locations will have a distinctive feature that can be used to localize an unknown source signal if there is some knowledge about this feature. Practically, location features are obtained from the received signal in the training stage.

3 Focusing in Time Reversal Theory

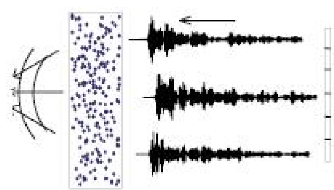
When a source is applied into a medium at certain location and a received signal is recorded by an array of transducers as shown in Fig.1. The received signals are reversed in time and then re-emitted into the medium. The re-emitted acoustic energy wave propagates back through the same medium and goes through all the multiple scattering, reflections and refraction that they underwent in the forward direction and refocuses on the source location. If only an aperture of limited area, called time-reversal mirror, is adopted in a time reversal operation, a small part of the field radiated by the acoustic source is captured and time reversed, thus limiting focusing quality [7].

However, in a bounded medium, multiple reflections along the medium boundaries significantly increase the apparent aperture of the time reversal mirror which means that the transducers can equivalently replaced by reflecting boundaries that redirect part of the incident wave towards the aperture. Thus spatial information is converted into one dimension representation of the time domain and the reversal quality depends crucially on the duration of the time-reversal window, i.e. the length of the recording to be reversed. The heterogeneity of the medium or the boundaries which produces multi-paths contributes to have an aperture that is much larger than its physical size. It has been shown experimentally that in a cavity with specific geometrical property focusing with time reversal can be obtained using one transducer [7].

The time-reversal approach is clearly connected to the inverse source problem.



(a) impulse transmission and reception.



(b) time reversed transmission and impulse localization

Fig.2 Sketch of time reversal focusing in random medium

They both deal with propagation of a time-reversed field, but the propagation is real in the time reversal experiment and simulated in the inverse problem. Moreover, the most important distinction is that the time reversal approach doesnt need knowledge of the propagating medium while the inverse problem method does.

As any linear and time-invariant process, wave propagation through a multiple scattering medium may be described as a linear system with a certain impulse response. If the source sends a Dirac pulse $\delta(t)$ function,the j^{th} transducer of the Time-Reversal-Mirror will receive a signal $h_j(t)$,which is the propagation impulse response from the source to transducer j. Moreover, due to spatial reciprocity,$h_j(t)$ is also the impulse response describing the propagation of a pulse from the $h_j(t)$ transducer to the source. Thus, if the transducer is able to record and time reverse the whole impulse response, the signal generated at the source can be represented by the convolution $h_j(t) * h_j(-t)$.This convolution product, in terms of signal analysis, is a typical matched filter which is a linear filter whose output is optimal in some sense. For whatever the impulse response $h_j(t)$, the temporal result is the convolution between this response and its time

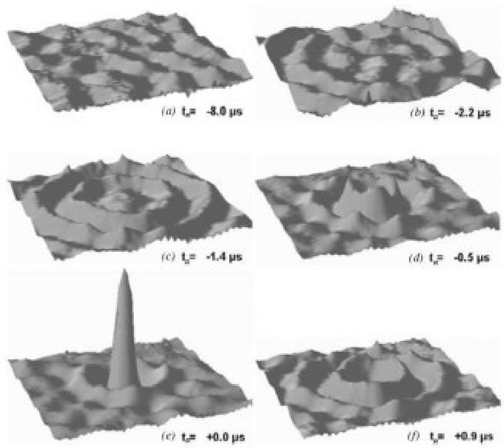


Fig.3 Time-reversed wave field observed at different times around the central point on a square of 15 15mm2 ([FP01]).

reverse version $h_j(t) * h_j(-t)$.which is maximal at time $t = 0$. This maximum is always positive and equals $\int h_j^2 t \, dt$, i.e. the energy conveyed by $h_j(t)$. For an N-element array, the signal recreated on the source can be written as,

$$s(t) = \sum_{j=1}^{j=n} h_j(t) * h_j(-t) \tag{3}$$

Even if $h_j(t)$ are completely random and apparently uncorrelated signals, each term in this sum reaches its maximum at time $t = 0$. So all contributions add constructively around $t = 0$, whereas at earlier or later times uncorrelated contributions tend to interfere destructively one with another. Thus the recreation of a sharp peak after time reversing on N-element array can be viewed as an interference process between N outputs of N matched filters.

4 Realization of Source Localization from Time Reversal Focusing

As described in the previous section, it is possible with time-reversal theory to reconstruct an acoustic signal in its original location in a scattering medium by recording the received signals and sending back the time reversed version of these signals through the medium. This implies that the received signal carries its source location signature as a result of scattering in the transmission medium and reflections from its complex boundaries. With the same assumption of Dirac delta source excitation, the response term $h_j(t)$ of the temporal correlation as given by (1) can be interpreted in LPM as a template obtained in the learning stage. Here $h_j(-t)$ is the applied test signal with negative sign turns the convolution into a cross correlation operation. Therefore, cross-correlation is a focusing process in time reversal but a similarity measure in template matching localization.

Although source reconstruction in time reversal is a transmission process or active, it is comparable to passive source localization with template matching by the operation of cross correlation as illustrated in the previous section. However, time reversal can help to provide explanation of physical limitations for source localization problems.

Ideally, the source reconstruction is achieved with array of sensors surrounding the source origin with element spacing of at least half a wavelength. Practically a limited aperture area is used on the cost of focusing resolution. The smaller the array, the larger the focal spot. As a result of wave diffraction, the waves will refocus to a spot not smaller than the shortest wavelength [5].
Accordingly, the achievable localization resolution using template matching can be increased with more sensors but still limited to the smallest wavelength, and since wavelength is inversely related to frequency, higher accuracy can be obtained with interactions that generate higher frequency signals such as using nail clicks and metallic object than those who generate lower frequencies as finger tap and damping material.

5 Location Feature Extraction for Enhanced Reliability

Ideally to comply with the time reversal theory, an impulsive source is needed for both learning and recognition stages. Practically, it is found that cross correlation is sensitive to the template type which means that similar type of impacts should be used in both stages since the variation in the signal will appear as variation in the location signature. One option to make the system more reliable is to use multiple templates for different type of interactions, but this is impractical as the calibration work will be intensified. Signal filtering improves the resolution but found to be not effective to improve reliability because it filters the frequency components not the location feature. Therefore an attractive novel solution is proposed here to solve the reliability problem by extracting specific features from the signal that is associated with the source location not the source information.

Let an unknown source signal given by $s(t)$ emitted from location i on the surface of a tangible object as shown in Fig.4, and two sensors are receiving the signals $g_1(t)$ and $g_2(t)$. The propagation path from the given source to sensor-1 and sensor-2 can be expressed by a specific transfer function denoted by $h_1(t)$ and $h_2(t)$ respectively. The transfer function is characterized by the complex propagation path and independent on the source signal information. Accordingly, the transfer function for a specific source to receiver path represents the actual source location signature. Treating the transmission medium as a black box of a single input/multiple output time invariant system, the output signal received by the i^{th} sensor can be expressed analytically the convolution integral given by:

$$g_i(t) = \int_{-\infty}^{\infty} h_i(t)s(t-\tau)d\tau \tag{4}$$

For instance considering the output from one sensor only, it is possible to measure the transfer function for a given location by applying an impulse $\delta(t)$ t that location and measure the output. In that case the received signal is the transfer function which in turn can be used as location signature in the template. If the test signal is also an impulse, then the matching process is just the comparison of location signatures and therefore high accuracy is anticipated. Otherwise, if the signal used for test, or for the template, are not an impulse, the resulting received signal will include source signal information plus location information, which accordingly will result in estimation error. This is why LPM works better with impulsive type of impacts. The task now is to develop a technique to extract the only source information from any type of interaction by employing two sensors.

Let the input in the system shown in Fig.5 is a stationary random signal. The

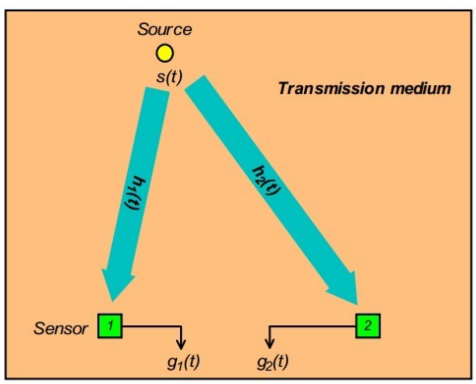


Fig.4 Received signals from two different paths.

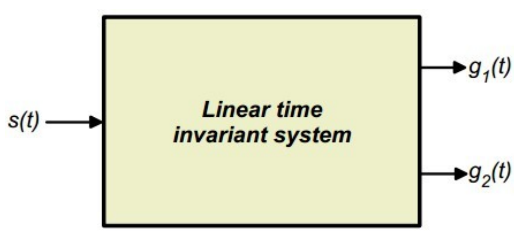


Fig.5 Black box LPM model

input/output relationship in the frequency domain is given by Fourier transform of equation (5) is given by:

$$g_i(f) = s(f)H_i(f) \tag{5}$$

The hypothesis of extracting the location signature involves utilizing a measurable quantity that doesnt require any knowledge about the input excitation or medium transfer function. Thus from the output/output relationship that is given by the cross spectral density function between the two outputs as given by:

$$P_{g_1 g_2} = g_1(f)g_2(f) \tag{6}$$

Then a hypothetical transfer function can be defined as:

$$H_{g_1 g_2}(f) = \frac{P_{g_1 g_2}(f)}{P_{g_1 g_1}(f)} \tag{7}$$

where $P_{g_1 g_2}$ is the autocorrelation function of $G_1(t)$. It can be seen from the above three equations $H_{g_1 g_2}(f)$ is a function of the two transfer functions $h_1(t)$ and $h_2(t)$ which still represents an independent location signature. A related subject in literature is the binaural localization in human which is simulated by the Head Related Transfer Function cue and defined by the ratio of the two output spectrums [7].

By rewriting the complex equation (7) in the form of magnitude and phase as:

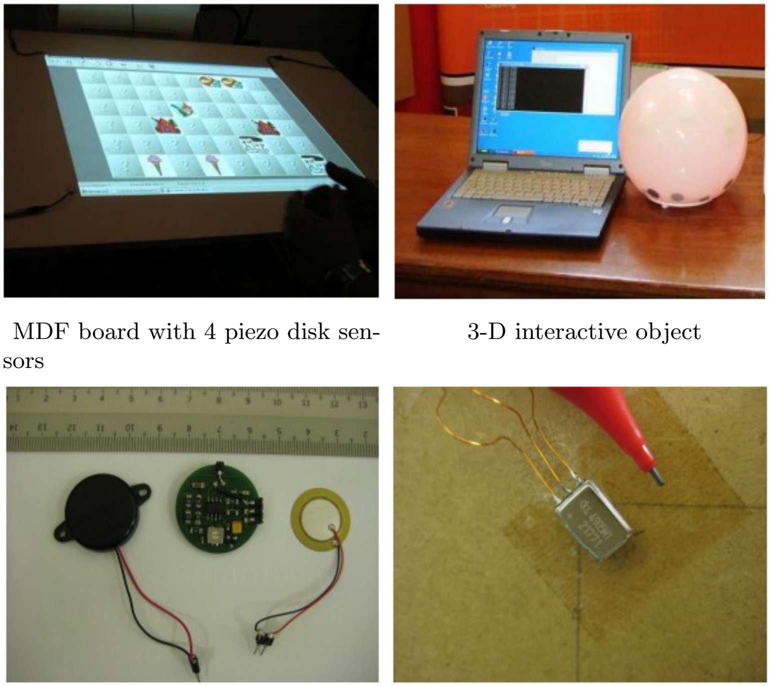
$$H_{g_1 g_2}(f) = A_{g_1 g_2}(f)\angle\varphi_{g_1 g_2} \tag{8}$$

either the magnitude or phase patterns can be used as a location signature information. To use both of amplitude and phase information, the pattern given by:

$$\hat{h}_{g_1 g_2}(t) = \int_{-\infty}^{\infty} \psi(f) P_{g_1 g_2}(f)/P_{g_1 g_1}(f) e^{-j2\pi f t}\, df \tag{9}$$

can be used, where $\varphi(f)$ is a weighting function introduced to compensate the bandwidth variation and depends on the interaction method. With $\varphi(f) = A_{g_1 g_2}(f)$, only phase information are extracted. Obviously, utilizing only the phase or the magnitude pattern is computationally faster than using (9) since it is not necessary to convert them into time domain.

An experimental result was carried out by registering a template from impacts at defined locations generated by pen tip hits on a glass sheet. The test database consists of different interaction types as pen tip hits, nail clicks and finger tapping. Then with the evaluation procedure it is found that highest percentage of correct estimations were obtained using (9), than using the phase information only and lastly when magnitude information were used only.



MDF board with 4 piezo disk sensors

3-D interactive object

Conditioning circuit for Piezo sensors

Piezoelectric microphone on MDF board

Fig.6 Experimental setups and sensors for LPM localization

6 Experimentation

The experimental setup consists of the interactive object, sensors, signal amplifier, data acquisition card and a PC to process the signals. Different object materials and shapes have been tested including metal, glass, plastic, fibre boards and 3-D objects. The suitable sensors were the piezoelectric discs, electret microphones and accelerometers.

For data acquisition, four channel PCI card is used for evaluation and two channel sound card used for demonstrations such as the portable USB sound card and the wireless audio transmitter. Some of them are pictured in Fig.6. The LPM system found to be working well on variety of materials. The Piezo ceramic sounders and electret microphones are the cheapest but can only pick up low frequencies when firmly attached to the surface. The piezoelectric microphone is very sensitive with wide bandwidth response and the most expensive among others. The piezoelectric shock sensor from Murata is the best sensor with sufficient frequency response and a very reasonable price.

7 Conclusion

In this paper, a new method for in-solid acoustic source localisation based on learning and template match is presented. With this Location Pattern Matching technique, it is possible to convert virtually any solid objects into an interactive interface by simply attaching sensors on the objects surface to transfer the acoustic signals resulting from natural interactions through signal conditioning hardware to data acquisition device for sampling before delivering the digital data to a PC where the localization algorithm is running. The reliability problem caused by the variance of impact patterns to the matching process has been improved by the concept of extracting the location signature pattern from received signals using multi-sensor layout.

Acknowledgements

This work was financed by the European FP6 IST Project “Tangible Acoustic Interfaces for Computer-Human Interaction (TAI-CHI)”. The support of the European Commission is gratefully acknowledged.

References

[1] H.Jeong, Y.Jang. (2000), “Wavelet analysis of plate wave propagation in composite laminates”, *Composite Structure*, Vol.49, pp.443-450.

[2] R.Ing, N.Quieffin, S.Catheline, M.Fink. (2004), “Tangible interactive interface usingacoustic time reversal process”, *Applied Physics Letters.*

[3] G.Lai, P Aarabi. (2003), “Active object localization using speaker arrays”, *Proceedings of the IEEE Sixth International Conference on Information Fusion*, Vol.1, pp.70-73.

[4] E Dijk. (2004), “Indoor ultrasonic position estimation using a single base station, PhD thesis”, *Eindhoven University of Technology*, Netherlands.

[5] Bousseljot R. and Kreiseler D. (2000), “Waveform recognition with 10,000 ECGs”, *IEEE Computers in Cardiology Proceedings*, pp.24-27, Cambridge, MA, pp.331-334.

[6] O’Hagan R and Zelinsky A. (1997), “Finger track-A robust and real-time gesture interface”, *Advanced Topics in Artificial Intelligence, Tenth Australian Joint Conference on Artificial Intelligence Proceedings*, pp.475-484.

[7] M.Fink. (1999), “Time reversed acoustics”, *Scientific America*, pp.91-95.

Corresponding author

Author can be contacted at: my@cf.ac.uk.