Proceedings of the 6th International Conference on Natural Language Processing and Knowledge Engineering
IEEE NLP-KE 2010 Aug. 21-23, 2010, Beijing, China

Recognition of Abnormal Vibrational Responses of Signposts using the Two-dimensional Geometric Distance and Wilcoxon Test

Michihiro JINNAI
Kagawa National College
Takamatsu, Kagawa, Japan
jinnai@t.kagawa-nct.ac.jp

Yukio AKASHI
Nexco Engineering Ltd.
Takamatsu, Kagawa, Japan
yukio.akashi@w-e-shikoku.co.jp

Shinsuke NAKAYA
Nexco Engineering Ltd.
Takamatsu, Kagawa, Japan
shinsuke.nakaya@w-e-shikoku.co.jp

Fuji REN
University of Tokushima
Tokushima, Tokushima, Japan
ren@is.tokushima-u.ac.jp

Minoru FUKUMI
University of Tokushima
Tokushima, Tokushima, Japan
fukumi@is.tokushima-u.ac.jp

Abstract:

In expressway companies, workers have been impacting signposts using wooden hammers and estimating the degree of the corrosion by listening to the sound. In order to automate this, we have been developing software that recognizes an abnormal impact vibrational response due to corrosion. This software extracts sonograms from impact vibrational waves using the LPC spectrum analysis, and matches images of the sonogram between a standard and an input impact vibrations using the Two-dimensional Geometric Distance. Then, the software distinguishes the abnormality of the input impact vibration using Wilcoxon rank-sum test. We have measured the impact vibrations of five normal signposts and five abnormal signposts, and carried out the automatic recognition experiments. As a result, the software has recognized correctly in all cases. We have verified the effectiveness of the proposed method.

Keywords:

Similarity measures; distance functions; pattern matching;

1. Introduction

On expressways, signposts have broken owing to corrosion and have fallen from elevated roads. Accidents have happened. In expressway companies, in order to prevent these accidents, workers have been impacting signposts using wooden hammers and estimating the degree of the corrosion by listening to the sound. The skilled workers must be used to interpret the impact sound of the signpost. As an improvement, we have been developing software for detecting abnormal impact vibrations of signposts so that non-skilled workers can inspect the signposts. This software extracts sonograms from the impact vibrational waves using the LPC spectrum analysis, and matches images of the sonogram between a standard and an input impact vibrations using a similarity scale. Then, the software distinguishes the abnormality of the input impact vibration using Wilcoxon rank-sum test of a statistical analysis.

The similarity scale works as follows, if a human would recognize two patterns as similar to each other, the computer software outputs a small value, and if a human would recognize the two patterns as dissimilar, then the computer software outputs a large value. In conventional pattern matching, the similarity scales known as the Euclidean distance and cosine similarity have been widely used to measure likeness [1],[2]. Conventional similarity scales compare the patterns using one-to-one mapping. The result of the one-to-one mapping is that the distance metric is highly sensitive to noise, and the distance metric changes in a staircase pattern when a difference occurs between peaks of the standard and input patterns. To improve these shortcomings, various similarity scales have been proposed for comparing two patterns in speech recognition [2-10], pattern classification [11] and image retrieval [12-15].

In our previous paper [16], a new similarity scale called the Geometric Distance was proposed. At that point, we initially developed a mathematical model for the similarity scale. In order to realize the mathematical model, an algorithm based on one-to-many point mapping was proposed. In this research, the algorithm is expanded to the "Two-dimensional Geometric Distance (2-d GD)". Using the 2-d GD instead of the conventional similarity scales, we have been developing the software for detecting abnormal impact vibrational responses of signposts. In this paper, we introduce the principles of the 2-d GD, and describe the automatic inspection software for the signpost using the LPC spectrum, the 2-d GD and Wilcoxon test.

978-1-4244-6899-7/10/$26.00 ©2010 IEEE

2. The sonogram of impact vibrational response of signpost

Figure 1 shows the sonograms (time-frequency-power). These sonograms are extracted from the impact vibrational responses measured by a sensor. Where, we set the vibration sensor on an upper position, an intermediate position and a bottom of the signpost, and impact on the opposite side of the sensor using wooden hammer, respectively. Since the signposts are eroded at the bottom in many cases owing to puddles of rainwater, we assume that the signpost has the abnormality at the bottom in this experiment as shown in Figure 1. These sonograms have been calculated using the method of Linear Predictive Coefficient (LPC). We have set the analysis conditions of the vibrational wave with the 51.2kHz sampling frequency, 16bit quantization, 3.93msec frame width, 0.098msec frame period, 151 total frames, 10 order of LPC, 1Hz to 12000Hz frequency range, 46.9Hz frequency resolution, and –40dB to 0dB logarithmic power spectrum. In this research, in order to avoid the initial impulse response of the waveform, the sonogram is extracted from only the natural vibrational wave of the signpost. Therefore, if we assume 0msec at the time of impact, the sonogram is extracted from the vibrational wave in the range of 0.79msec to 18.67msec.

From the upper and middle sonograms shown in Figure 1, it is evident that the peak frequencies are 2579Hz and 2532Hz respectively and the two patterns are similar to each other. On the other hand, from the lower sonogram shown in Figure 1, it is evident that the peak frequency is 6516Hz and the pattern is different from the other two. Moreover, as a result of comparing these three impact sounds aurally, we have confirmed that the sounds of the upper and middle sonograms are similar to each other and the sound of the lower sonogram is different from the other two.

3. The similarity scale and its robustness

In sound recognition, a known sonogram stored in a PC memory is called here the "standard pattern", and a comparison sonogram is called "input pattern". The degree of likeness between the standard pattern and the input pattern is evaluated using a similarity scale. If the similarity of the standard and input patterns is close, then those two patterns are considered to be in the same category and the input pattern is recognized and classified. The similarity is often measured as a "distance" between the two patterns. Conventionally, the similarity scales known as the Euclidean distance and cosine similarity have been widely used. Section 3.1 describes the shortcomings that are found in the conventional similarity scales. Furthermore, Sections 3.2 and 3.3 describe new similarity scales called the "One-dimensional Geometric Distance (1-d GD)" and the "Two-dimensional Geometric Distance (2-d GD)" that have been developed by us for improving the shortcomings.

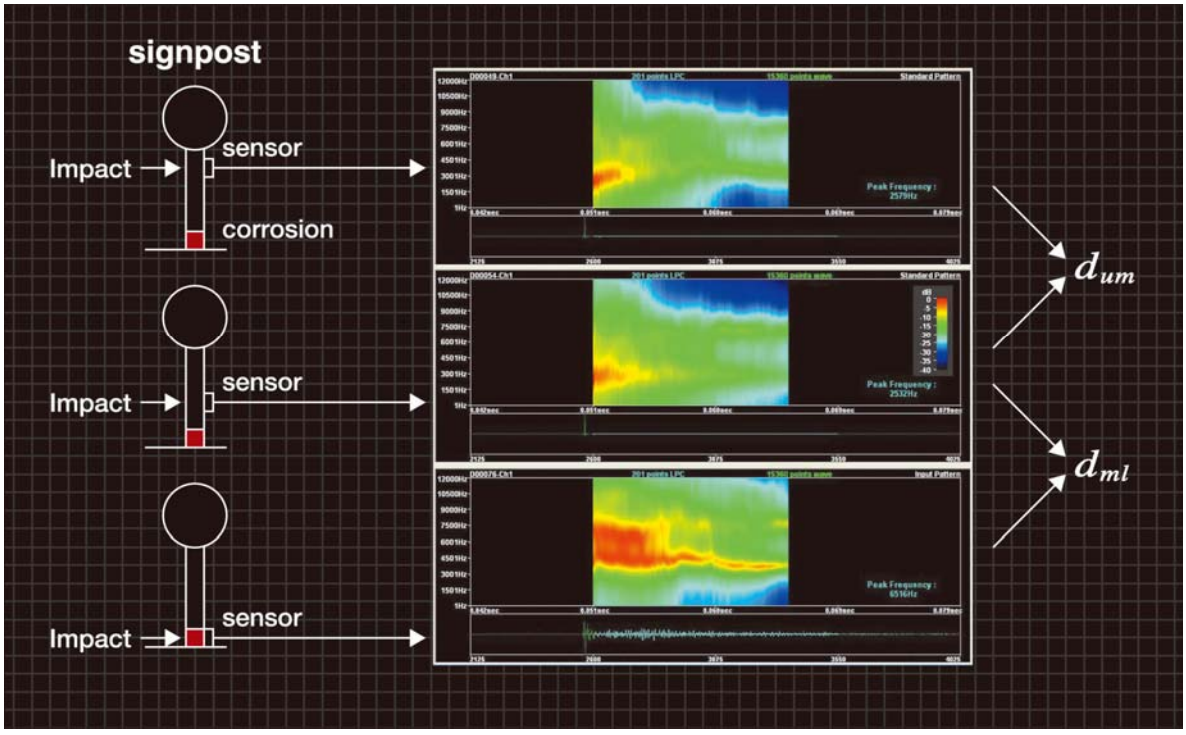

Figure 1. Sonograms of impact vibrational response of signpost

3.1. Euclidean distance and cosine similarity – Conventional similarity scales

Conventional similarity scales Euclidean distance and cosine similarity compare the patterns using one-to-one mapping. The result of the one-to-one mapping is that input patterns with different shapes may have the same distance from the standard pattern when the sonograms have the "difference" and "wobble".

The upper diagram of Figure 2 shows an example of the "difference" where the standard pattern has two peaks in the sonogram, and input patterns 1, 2, and 3 have a different position on the first peak. Note that the standard and input patterns have the same volume. As shown in the bar graph at the bottom left of Figure 2, the Euclidean distances and cosine similarities e_1, e_2, and e_3 have the relationship of $e_1=e_2=e_3$ between the standard pattern and each of input patterns 1, 2, and 3. Therefore, input patterns 1, 2, and 3 cannot be distinguished.

The upper diagram of Figure 3 shows an example of the "wobble" where the standard pattern has a flat sonogram, input patterns 4 and 5 have the "wobble" on the flat sonogram, and input pattern 6 has a single peak. However, each pattern is assumed to have variable α in the relationship shown in Figure 3. Therefore, the standard and input patterns always have the same volume. As shown in the bar graph at the bottom left of Figure 3, the Euclidean distances and cosine similarities e_4, e_5, and e_6 have the relationship of $e_4=e_5=e_6$ between the standard pattern and each of input patterns 4, 5, and 6. Therefore, input patterns 4, 5, and 6 cannot be distinguished.

3.2. One-dimensional Geometric Distance

As an improvement, we have developed a new similarity scale called the Geometric Distance [16]. A similarity scale is a concept that should intuitively concur with the human concept of similarity in hearing and sight. First we need to develop a mathematical model for the similarity scale so that we can perform numerical processing by computation. In the Geometric Distance, a mathematical model of the similarity scale is proposed to improve the shortcomings that are found in the Euclidean distance, cosine similarity and others. A mathematical model incorporating the following two characteristics is used.

<1> The distance metric must show good immunity to noise.

<2> The distance metric must increase monotonically when a difference increases between peaks of the standard and input patterns.

The bar graphs at the bottom right of Figures 2 and 3 express the mathematical model by figures. Following on from above, a new algorithm based on one-to-many point mapping is proposed to realize the mathematical model. This section describes the 1-d GD algorithm.

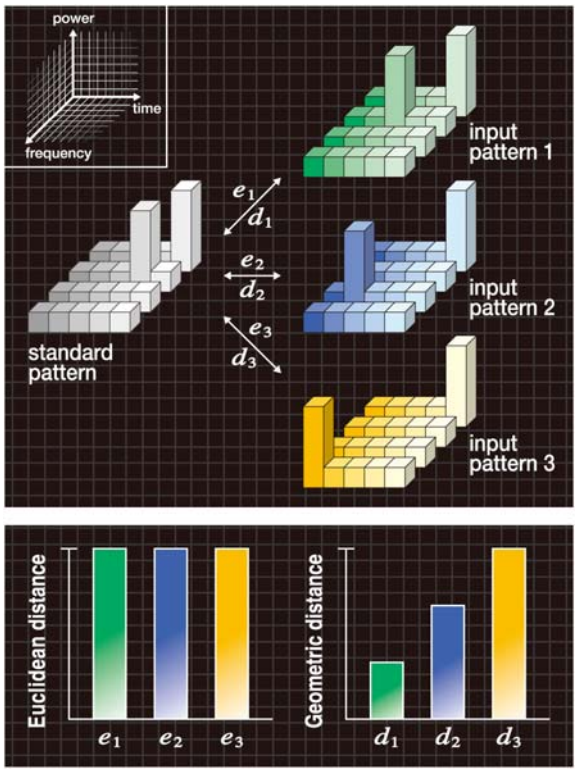

Figure 2. Typical example of "difference"

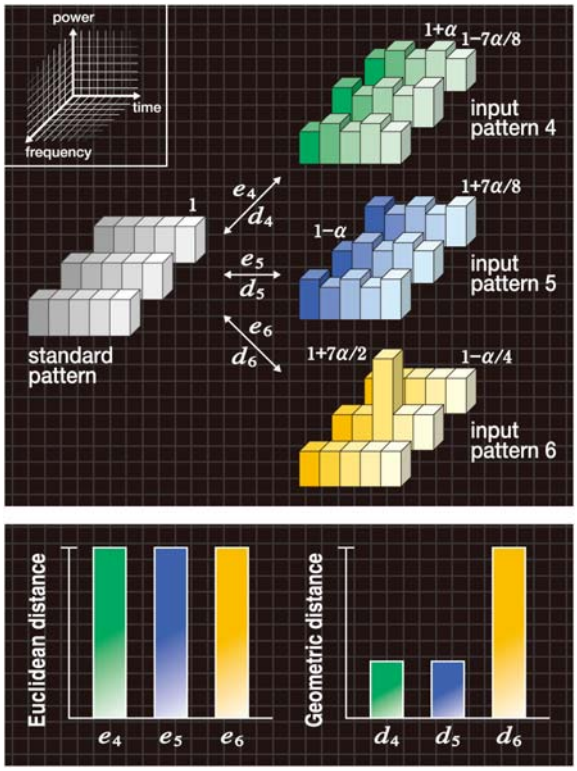

Figure 3. Typical example of "wobble"

Figures 4(a)-(e) respectively show typical examples of the standard and input patterns that have been created using the momentary power spectrum (frequency-power) of standard and input sounds. Note that the power spectrum is generated from the output of filter bank with the m frequency bands. The i-th power spectrum values (where, $i = 1, 2, \ldots, m$) are divided by their total energy, so that normalized power spectra s_i and x_i have been calculated. At this moment, the standard and input patterns have the same area size. Moreover, Figures 4(a)-(e) respectively show reference patterns that have the initial shape r_i of a normal distribution.

With the 1-d GD algorithm, a difference in shapes between standard and input patterns is replaced by the shape change of the reference pattern using the following equation.

$$r_i \longleftarrow r_i + (x_i - s_i) \qquad (i = 1, 2, 3, \cdots, m) \tag{1}$$

Next, we explain Eq. (1) using Figure 4.

●Figure 4(a) gives an example of the case where standard pattern and input pattern have the same shape. Because values r_i of Eq. (1) do not change during this time, the reference pattern shown in Figure 4(a) does not change in the shape from the normal distribution.

●Figures 4(b)-(d) respectively show examples exhibiting a small, medium, and large "difference" of peaks between the standard and input patterns. If Eq. (1) is represented by the shapes, as shown in Figures 4(b)-(d), value r_i decreases at peak position i of each standard pattern. At the same time, value r_i increases at peak position i of each input pattern.

●Figure 4(e) typically shows the standard pattern having a flat shape and the input pattern where a "wobble" occurs in the flat shape. Because values r_i increase and decrease alternatively in Eq. (1) during this time, the reference pattern shown in Figure 4(e) has a small shape change from the normal distribution.

For the reference pattern whose shape has changed by Eq. (1), the magnitude of shape change is numerically evaluated as the variable of moment ratio. The moment ratio of the reference pattern can be calculated using the following equation.

$$A = \frac{\left\{\sum_{i=1}^{m} r_i\right\} \cdot \left\{\sum_{i=1}^{m} (L_i)^4 \cdot r_i\right\}}{\left\{\sum_{i=1}^{m} (L_i)^2 \cdot r_i\right\}^2} - 3 \tag{2}$$

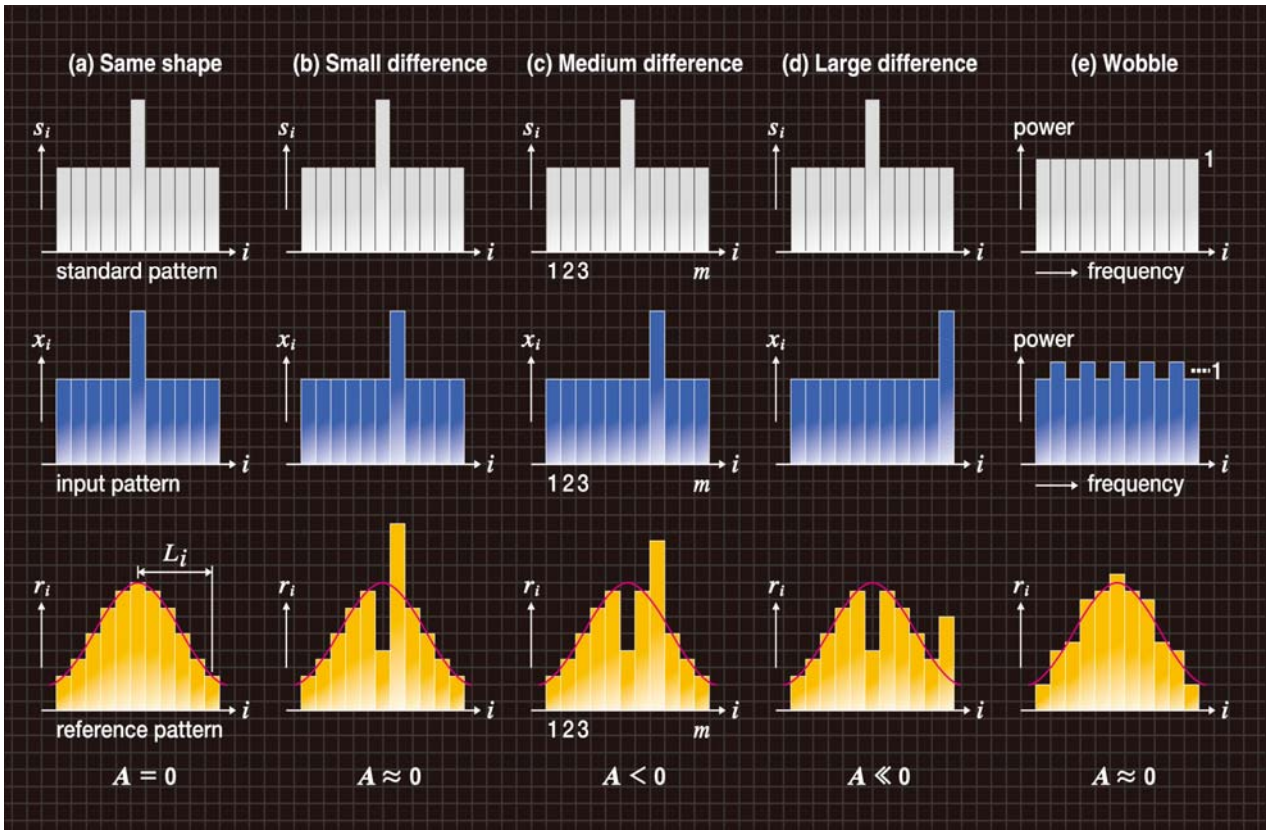

Figure 4. Shape changes of reference patterns

Where, L_i (i = 1, 2, … , m) is a deviation from the center axis of the normal distribution as shown in the reference pattern of Figure 4(a). The moment ratio A is derived from the kurtosis from a statistical analysis. If the shape of the reference pattern follows the normal distribution, then $A=0$. If it has peakedness relative to the normal distribution, then $A>0$. Alternatively, if it has flatness relative to the normal distribution, then $A<0$. Figures 4(a)-(e) show how A varies with r_i.

●In Figure 4(a), the values r_i do not change. The moment ratio becomes $A=0$.

●In Figure 4(b), the position i of the decreased r_i and that of the increased r_i are close. Because the effect of an increase and a decrease is canceled out, the moment ratio becomes $A\approx 0$.

●In Figure 4(d), because the shape of reference pattern has flatness relative to the normal distribution, the moment ratio becomes $A\ll 0$.

●In Figure 4(c), because the shape of the reference pattern is an intermediate state between (b) and (d), the moment ratio becomes $A<0$.

●In Figure 4(e), the reference pattern has small shape change from the normal distribution, and the moment ratio becomes $A\approx 0$.

From Figures 4(a)-(d), we can understand that value $|A|$ increases monotonically according to the increase of the "difference" between peaks of the standard and input patterns. Also, from Figure 4(e), it is clear that $A\approx 0$ for the "wobble".

As shown in Figure 4, we have determined the moment ratio A by assuming that the center axis of the normal distribution locates at the center of standard and input patterns. Next, as shown in Figure 5, we determine the amount of moment ratio A_j for each j in the case where the center axis of the normal distribution moves to any component position j (where, j = 1, 2, … , m) of the standard and input patterns. Using the m parts of the moment ratios A_j that we have obtained in Figure 5, we can calculate the difference in shapes between standard and input patterns by the following equation and we define it as the "One-dimensional Geometric distance d".

$$d = \sqrt{\sum_{j=1}^{m}(A_j)^2} \qquad (3)$$

In this method, when a "difference" occurs between peaks of the standard and input patterns with a "wobble" due to noise or other non-linearity, the "wobble" is absorbed and the distance metric increases monotonically according to the increase of the "difference". From the above description, we could verify that the 1-d GD algorithm matches the characteristics <1> and <2> of the mathematical model. In the actual 1-d GD algorithm, we create a pair of reference patterns that have the initial shape of the normal distribution, because Eq. (2) cannot be defined if the value r_i is negative [16].

3.3. Two-dimensional Geometric Distance – New similarity scale

The 1-d GD algorithm is expanded to the 2-d GD algorithm. Figures 6 and 7 respectively show stylized examples of the standard and input patterns that have been created using the sonogram (time-frequency-power) of standard and input sounds. The (i_1, i_2)-th power spectrum values (where, time axis i_1 = 1, 2, … , m_1; frequency axis i_2 = 1, 2, … , m_2) are divided by their total energy, so that normalized power spectra $s_{i_1 i_2}$ and $x_{i_1 i_2}$ have been calculated. At this moment, the standard and input patterns have the same volume size. Moreover, Figures 6 and 7 respectively show reference patterns that have the initial shape $r_{i_1 i_2}$ of a two-dimensional normal distribution.

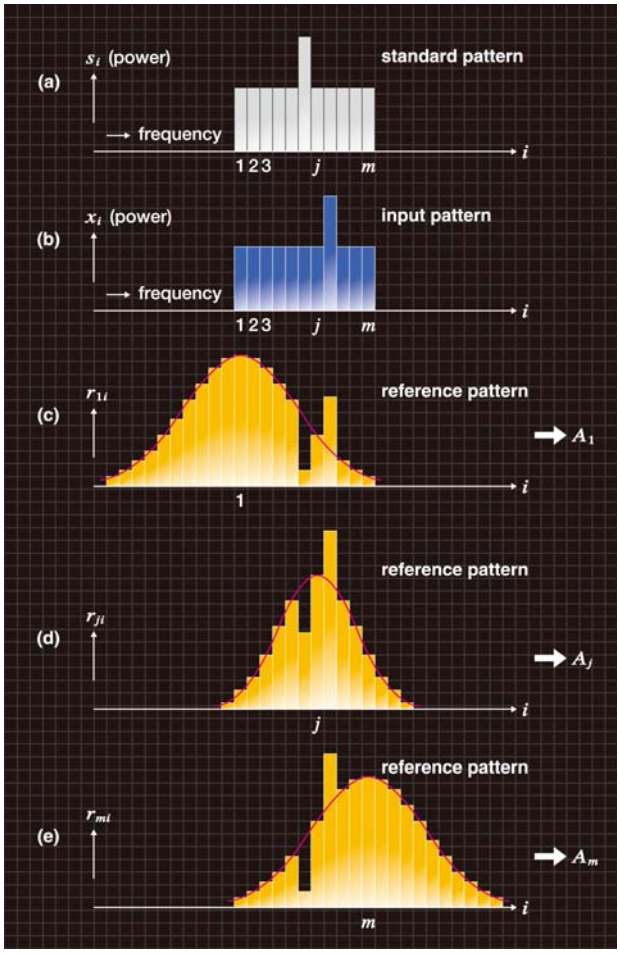

Figure 5. Movement of reference pattern

Figure 6 shows an example exhibiting a "difference" of peaks between the standard and input patterns. Figure 7 shows the standard pattern having a flat shape and the input pattern where a "wobble" occurs in the flat shape. We suppose that the standard and input patterns have the same volume. With the 2-d GD algorithm, a difference in shapes between standard and input patterns is replaced by the shape change of the reference pattern using the following equation

$$r_{i_1 i_2} \longleftarrow r_{i_1 i_2} + (x_{i_1 i_2} - s_{i_1 i_2}) \quad \begin{matrix}(i_1 = 1, 2, 3, \cdots, m_1) \\ (i_2 = 1, 2, 3, \cdots, m_2)\end{matrix} \tag{4}$$

In Figure 6, the value of the reference pattern decreases at peak position of the standard pattern. At the same time, the value of the reference pattern increases at peak position of the input pattern. In Figure 7, because the values of the reference pattern increase and decrease alternatively, the reference pattern has a small shape change from the two-dimensional normal distribution.

For the reference pattern whose shape has changed by Eq. (4), the magnitude of shape change is numerically evaluated as the variable of moment ratio using the following equation.

$$A = \frac{\left\{\sum_{i_1=1}^{m_1}\sum_{i_2=1}^{m_2} r_{i_1 i_2}\right\}\cdot\left\{\sum_{i_1=1}^{m_1}\sum_{i_2=1}^{m_2} (L_{i_1 i_2})^4 \cdot r_{i_1 i_2}\right\}}{\left\{\sum_{i_1=1}^{m_1}\sum_{i_2=1}^{m_2} (L_{i_1 i_2})^2 \cdot r_{i_1 i_2}\right\}^2} - 3 \tag{5}$$

However, the deviation L_i shown in Figure 4(a) and Eq. (2) is replaced by a deviation $L_{i_1 i_2}$ shown in Figure 6.

Next, as shown in Figure 8, we determine the amount of moment ratio $A_{j_1 j_2}$ for each (j_1, j_2) in the case where the center axis of the two-dimensional normal distribution moves to various positions (j_1, j_2) relative to the standard and input patterns. Using these moment ratios $A_{j_1 j_2}$, we can calculate the difference in shapes between standard and input patterns by the following equation and we define it as the "Two-dimensional Geometric Distance d".

$$d = \sqrt{\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2} (A_{j_1 j_2})^2} \tag{6}$$

Also, we can verify that the 2-d GD algorithm matches the characteristics <1> and <2> of the mathematical model.

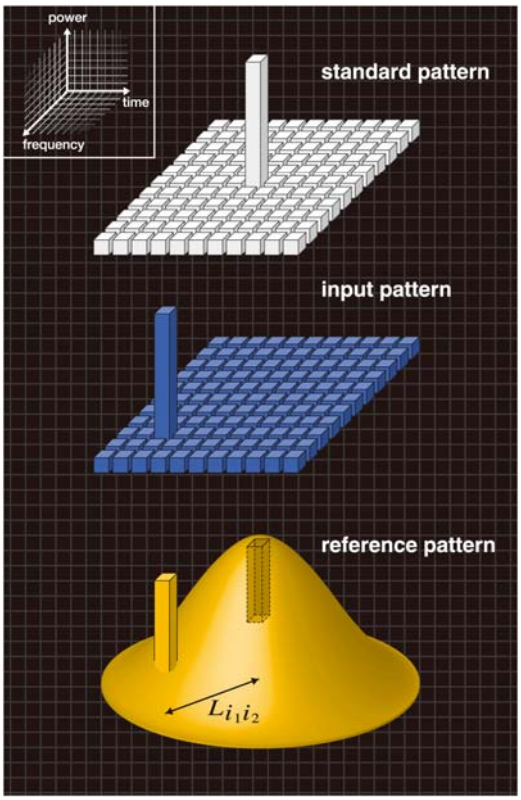

Figure 6. Shape change of reference pattern (difference)

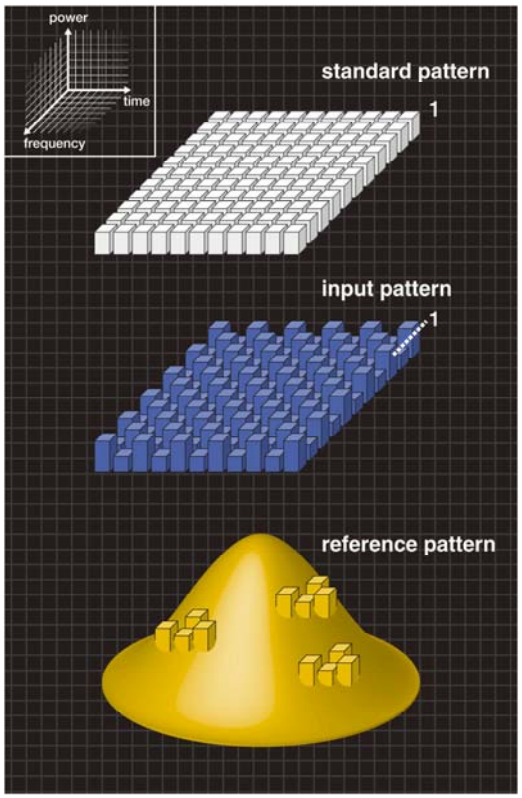

Figure 7. Shape change of reference pattern (wobble)

4. Recognition of abnormal impact vibrational responses

In our system, we impact the signpost 5 times on each position shown in Figure 1 using wooden hammer and measure the vibrational waves. Then, the 5x3 pieces of the sonograms are extracted from each vibrational waveform. As shown in Figure 1, we calculate the value d_{um} of the 2-d GD between the sonogram obtained on the upper position and the sonogram obtained on the intermediate position. The value d_{um} expresses the difference in shapes between the sonograms of the normal part of signpost.

The 5x5 pieces of the values d_{um} are calculated for each combination of the 5 sonograms obtained on the upper position and the 5 sonograms obtained on the intermediate position. We represent them as "sample 1". Similarly, the 5x5 pieces of the values d_{ml} are calculated for the intermediate position and the bottom of the signpost. We represent them as "sample 2". Figure 9 shows an example of distributions of samples 1 and 2. If the signpost is eroded at the bottom as shown in Figure 1, then the values d_{ml} are large. Therefore, the distribution of sample 2 is greater than that of sample 1 as shown in Figure 9.

We perform Wilcoxon rank-sum test with the following processing procedure to detect the difference between samples 1 and 2 automatically.

(Step 1) *Null hypothesis* H_0 : The distributions of samples 1 and 2 are the same.

Alternative hypothesis H_1 : The distribution of sample 2 is greater than that of sample 1.

(Step 2) We calculate a statistic W of Wilcoxon rank-sum test.

(Step 3) If the statistic W is in the critical region with a significant level of 10%, then we reject the hypothesis H_0 and determine that the signpost is eroded at the bottom.

We have measured the impact vibrations of the five normal signposts and the five abnormal signposts eroded at the bottom, and performed Wilcoxon test. As the result of the experiments, the software has recognized correctly in all cases. We have verified the effectiveness of the proposed method.

5. Conclusions and future work

We have proposed a new method for detecting the abnormal impact vibrational responses of the signposts. The software extracts the sonogram from the impact vibrational wave using the LPC spectrum analysis, and matches the images of the sonogram between a standard and an input impact vibrations using the 2-d GD. Then, the software distinguishes the abnormality of the input impact vibration using Wilcoxon test. As the result of the recognition experiments, we have verified the effectiveness of the proposed method.

Finally, we describe future work. We will continue the recognition experiments using the signposts with various types of corrosions and will verify the effectiveness of the proposed method. Furthermore, we will develop mobile devices for estimating the degree of the corrosion so that non-skilled workers can inspect the signposts on expressways.

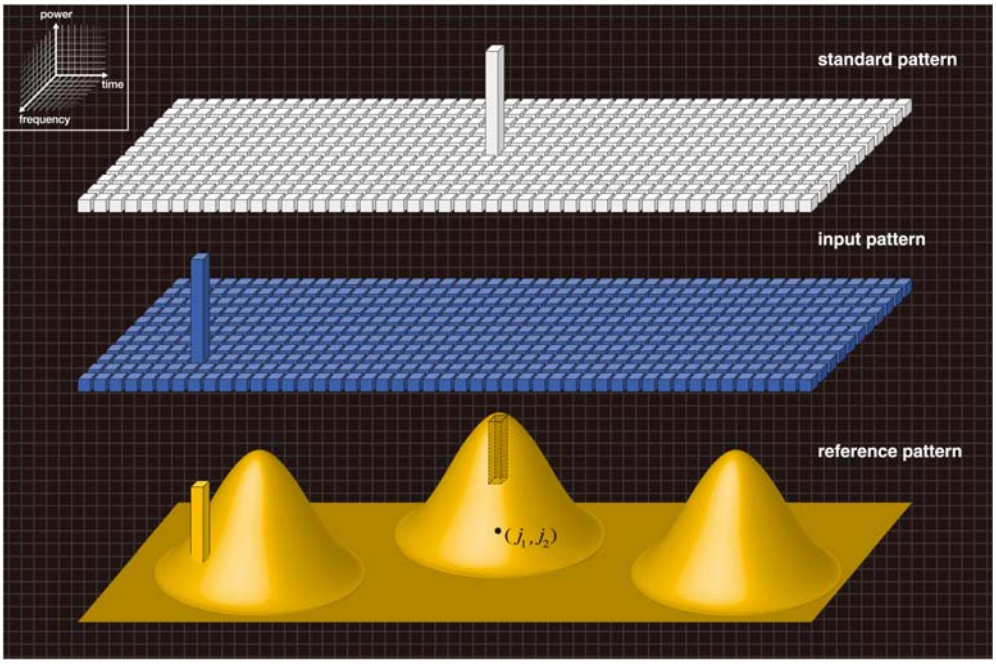

Figure 8. Movement of reference pattern

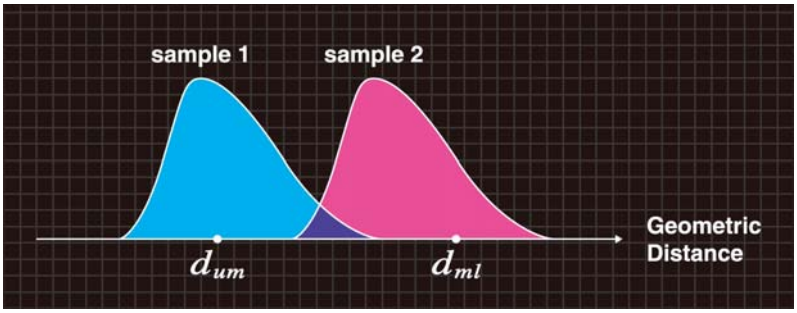

Figure 9. Distributions of two-dimensional geometric distances

Acknowledgements

This research has been partially supported by New Energy and Industrial Technology Development Organization (NEDO) of the Japanese Government under Grant No. 10HC7011, by Queensland Parks and Wildlife Service of the Australian Government under Coxen's Fig-parrot Recovery Plan, and funded by Mitsubishi Heavy Industries, Ltd. of Japan and Tokyo Gas Co., Ltd. of Japan.

References

[1] R.O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, second ed., Wiley, NewYork, 2000.

[2] K.K. Paliwal, "Effect of preemphasis on vowel recognition performance", Speech Communication, 3, pp.101-106, 1984.

[3] L.R. Rabiner and B.H. Juang, Fundamentals of Speech Recognition, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[4] F. Itakura and S. Saito, "An analysis-synthesis telephony based on maximum likelihood method", Proc. 6th Int. Congr. Acoustics, C-5-5, 1968.

[5] F. Itakura, "Minimum prediction residual principle applied to speech recognition", IEEE Trans. Acoust., Speech and Signal Processing, 23, pp.67-72, 1975.

[6] S. Furui, Digital Speech Processing, Synthesis, and Recognition (Electrical and Computer Engineering), Marcel Dekker, Inc., NewYork, 1989.

[7] K. Shikano and M. Sugiyama, "Evaluation of LPC spectral matching measures for spoken word recognition", Trans. IECE, 565-D, 5, pp.535-541 1982.

[8] D. Klatt, "Prediction of perceived phonetic distance from critical band spectra: A first step", Proc. ICASSP 82, 2, pp.1278-1281, 1982.

[9] D. Mansour and B.H. Juang, "A family of distortion measures based upon projection operation for robust speech recognition", IEEE Trans. Acoustics, Speech and Signal Processing, ASSP-37, 11, pp.1659-1671, 1989.

[10] N. Nocerino, F.K. Soong, L.R. Rabiner and D.H. Klatt, "Comparative study of several distortion measures for speech recognition", Speech Communication, 4, pp.317-331, 1985.

[11] S.-H. Cha and S.N. Srihari, "On measuring the distance between histograms", Pattern Recognition, 35, pp.1355-1370, 2002.

[12] J.-K. Kamarainen, V. Kyrki, J. Ilonen and H. Kälviäinen, "Improving similarity measures of histograms using smoothing projections", Pattern Recognition Lett., 24, pp.2009-2019, 2003.

[13] F.-D. Jou, K.-C. Fan and Y.-L. Chang, "Efficient matching of large-size histograms", Pattern Recognition Lett., 25, pp.277-286, 2004.

[14] F. Serratosa and A. Sanfeliu, "Signatures versus histograms: Definitions, distances and algorithms", Pattern Recognition, 39, pp.921-934, 2006.

[15] V.V. Strelkov, "A new similarity measure for histogram comparison and its application in time series analysis", Pattern Recognition Lett., 29, pp.1768-1774, 2008.

[16] M. Jinnai, S. Tsuge, S. Kuroiwa, F. Ren and M. Fukumi, "New similarity scale to measure the difference in like patterns with noise", International Journal of Advanced Intelligence, Volume 1, Number 1, pp. 59-88, November 2009.
http://aia-i.com/ijai/contents.html