Wavelet Entropy-based Feature Extraction for Crack Detection in Sewer Pipes

M. Browne, M. Dorn, R. Ouellette, T. Christaller and S. Shiry
GMD-Japan Research Laboratory,
Collaboration Center,
2-1 Hibikino, Wakamatsu-ku,
Kitakyushu-city
Contact e-mail: matthew.browne@gmd.gr.jp
keywords: wavelet transforms, shannon entropy, crack detection, sewer pipes

Abstract

This paper describes the use of wavelet entropy as a feature extractor for robust classification of cracks in sewer pipe structures. Video image data was acquired using an infrared camera in an experimental sewer pipe setup. Video frames were partitioned into $64 * 64$ pixel sections. 1,885 'crack' and 1,675 'clean' (non-crack) sections were manually classified. Each section was transformed to a combined space-frequency representation using a Haar wavelet transform. The concentration of wavelet coefficient distribution energy in 15 orthogonal wavelet subspaces was estimated using Shannon entropy, and the extracted features were used as inputs to a logistic regression model. An average validation set classification rate of 88.31% was obtained over 10 runs.

1. Introduction

Sewer systems are prone to damage in the pipewalls due to aging, traffic and chemical reactions, through which inflow (e.g. rainwater, groundwater, etc.) seeps into the pipesystems. Regional city reports [1], state that this inflow amounts to approximately 30% of the total flow. As well as the inflow of groundwater into the sewer pipes, outflow from damaged systems occurs. Available statistics for Germany ([6], [9] and [16]) show that several million liters of waste and disposable water contaminates surrounding environment each year.

Preceding repair, diagnostic measures need to be taken in order to identify sewer pipe damage. These diagnostic measures serve as assessment tools to identify the nature, degree, and location of sewer-pipe damage. Currently, inspection is undertaken using cable-tethered robots with onboard video camera systems. An operator remotely controls the movement of the robot and video system. By this means of video-supported visual inspection, any noted damages or abnormalities are recorded in the video stream. The reliability of such a system relies on the experience and subjectivity of the operator and is thus prone to human error. Consequently, effective autonomous online techniques to identify and extract objects of interest (e.g. cracks, etc.) from sensor data are of immediate importance.

Carino [5] gives a detailed overview of crack detection strategies such as infrared thermography, stress-wave propagation methods and ground-penetrating radar. However, detection strategies tend to be developed with specific pipe materials in mind. Widely used materials such as concrete or clay (see Table 1 [1]) have characteristically heterogeneous compositions, making application of simple fault detection methods problematic. Widely available and successful techniques used on steel pipes as that described by G.R. Stavroulakis et al. [19] cannot be applied to PVC or concrete pipes due to the electrical non-conductance of the material. Given feature detection methods of appropriate sophistication and sensitivity, lower cost and general-purpose systems such as video cameras could play an important role in sewer-pipe fault detection.

Table 1 Proportions of Pipe Materials used in Kitakyushu City, Japan

Clay	PVC	Hume	Concrete	Cast Iron
53%	19%	18%	7%	3%

The robust detection of cracks and other faults in sewer pipes from sensor data is a significant challenge. Bernatzki et al. [4] introduced a method for detecting small cracks in oil and gas pipelines. Raw ultrasonic data was transformed to a time-frequency representation using the wavelet transform. Edges were detected using the real part of wavelet coefficients. Artificial neural networks were used for classification.

Yoshimura et al. [22] described the application of a neural network-based inverse analysis method and the finite element method to the identification of cracks in solid objects with laser ultrasonics. They used an Error Propagation Coefficient to assess the accuracy of the neuro-based method for crack identification. Further, they are able to identify a surface defect with an estimated error of 2.4% - 12.0%, being able to verify the depths with an accuracy of 0.6% - 4.1%.

For the analysis of high dimensional spatially distributed data, wavelets may provide a useful feature detection method. Mojsilovic et al. in [12] used Haar wavelets for decomposition and classification of myocardial tissue images. Gunatilake et al. [3] introduce a mobile robot platform that delivers live imagery for remote aircraft surface inspection. The crack detection algorithm is modeled closely on the widely practiced test for detection of cracks using directional lighting. They apply a two-step multi-scale edge detection where a region of interest (ROI) is first decomposed into different resolutions by successive smoothing, followed by edge detection at each resolution. Wavelet-based filters are used for the projection of the ROI to different resolutions and estimation of intensity variation for multi-scale edge determination.

The present paper presents a novel approach to crack detection, which combines wavelet transforms with Shannon entropy as a feature extraction method. Wavelet-based methods appear to be well suited to the fundamental tasks of dimensionality reduction and feature extraction for image processing [13], [14]. The wavelet transform decomposes regions of video data into a combined space-frequency representation. The Shannon entropy measure is used to parameterize the concentration of energy distributions of wavelet coefficients at various scales. Ferraro et al. [7], [8] discuss the use of an entropy-based information measure in order to parameterize essential image features. The approach of using global statistical attributes of image data as feature detectors was also used by Antonie et al. [2], who applied the first 4 moments to subsections of image data to obtain classification attributes.

We suggest that images comprising faults are characterized by a distinctive energy distribution at certain spatial frequencies. This information may be passed to a discriminate function for classification of the data. Exploratory analyses in our laboratory indicated that features extracted by methods that were sensitive to translation of the input data (PCA, raw wavelet coefficients) resulted in poor classifier performance. The proposed method involves consideration of the entropy of overall wavelet coefficient distributions (comprising a complete set of translated basis functions), and is therefore translation independent*. Translation independent feature extraction is suitable for the present application because crack-like space-frequency structures are of interest, regardless of orientation or translation.

Because analysis filters based on Haar wavelets are extremely short (length 2), and due to the recursive downsampling method of the wavelet transform, the image transformation matrix used in the present study is extremely sparse. By parameterizing wavelet activity at each scale with a single Shannon entropy measure, the dimensionality of the input data is very low. Thus, through consideration of the global Haar wavelet energy density at each scale, the dimensionality of the data may be drastically reduced in a computationally efficient manner. The computational efficiency of this approach makes it suitable in online applications for the robust classification of cracks in sewer pipe systems.

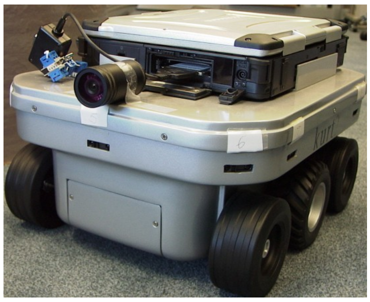

Fig.1 Autonomous KURT2 platform with infrared camera

2. Methology

2..1 Data collection

Video data was collected from a KURT2 mobile robot platform [10] (see figure 1), navigating through an experimental sewer pipe assembly. Grayscale video data (8 bit) was digitally sampled at a spatial resolution of $320*240$ pixels at a temporal resolution of 30 frames per second. Artificial sewer pipe cracks were created to simulate actual pipe faults as closely as possible.

Each video frame was split into fourteen $64*64$ subsections arranged around the periphery of the visual field. Figure 2 displays an example of a video frame, and the locations of the sections. In all, 3,560 sections were manually classified into 1,885 'crack' sections, and 1,675 'clean' (non-crack) sections.

2..2 Analysis

Sections were first filtered using an averaging filter in order to remove gross low-frequency lighting effects. An $8*8$ matrix containing elements with values of $1/64$ was convolved with each section, and the resulting low-pass output subtracted. Haar wavelet filters [20] were used:

$$h = \begin{pmatrix} -0.25 & 0.25 \end{pmatrix} \quad (1)$$
$$g = \begin{pmatrix} 0.25 & 0.25 \end{pmatrix} \quad (2)$$

A single iteration of the two-dimensional wavelet transform involves application of these filters and down-sampling to both the rows and columns of a

*Note that translation independence is used in the weak sense: at scale j the decimated wavelet transform is translation dependent for translations $j \ mod \ (2^j) \neq 0$. Translation invariant feature extraction is used to describe the property that spatial locality information is removed by considering the distribution of wavelet coefficients at each scale globally.

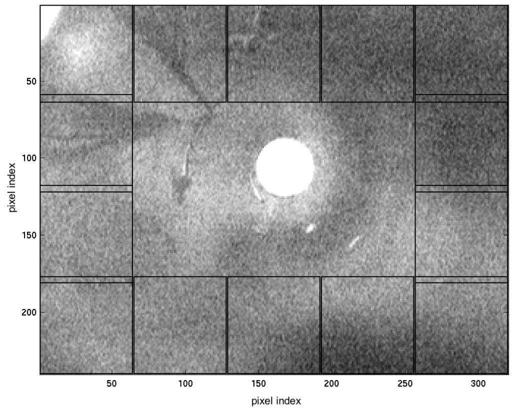

Fig.2 $64*64$ pixel sections drawn from the periphery of each $320*240$ pixel video frame. Note that some overlap was introduced to ensure that the size of the resulting sections were a factor of two

matrix, creating low-pass GG, vertical GH, horizontal HG, and diagonal HH image representations. Figure 3 displays example 'crack' and 'clean' sections, and their corresponding Haar transform.

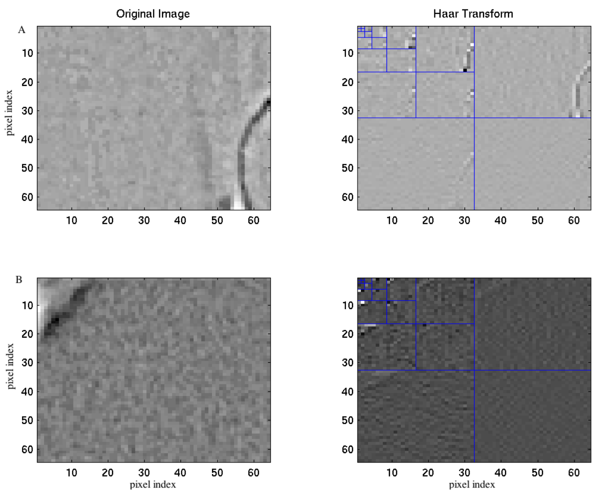

Fig.3 Examples of Haar wavelet transform on sections A (damaged) and B (undamaged). Note spurious lighting effect in top left hand corner of 'clean' image

The wavelet coefficient at scale j and translation k may be described via the scalar product, where $\phi(x)$ is the scaling function, $\psi(x)$ the wavelet function, and $f(x)$ the analyzed signal [18]:

$$c_j^G(k) = \frac{1}{2^j}\left\langle f(x), \phi\left(\frac{x-k}{2^j}\right)\right\rangle \quad (3)$$

$$c_j^H(k) = \frac{1}{2^j}\left\langle f(x), \psi\left(\frac{x-k}{2^j}\right)\right\rangle \quad (4)$$

Alternatively, the transform may be described as a recursive discrete convolution of a row or column of an image with the analysis filters h and g

$$c_j^G[k] = \sum_l g[l]\ c_{j-1}[k + 2^{j-1}l] \quad (5)$$

$$c_j^H[k] = \sum_l h[l]\ c_{j-1}[k + 2^{j-1}l] \quad (6)$$

For the standard wavelet transform, this process is implemented recursively on the low-pass GG output only. However, in the present study, the Haar coefficient matrix C was more efficiently generated by the prior creation of a sparse Haar wavelet matrix $\mathbf{W}$, and subsequent matrix multiplication with the entire image dataset $\mathbf{X}$:

$$\mathbf{C} = \mathbf{W^T X} \quad (7)$$

This implements an orthogonal rotation of the basis vectors of the analyzed data. The original pixel-coordinate space is transformed to a combined space-frequency representation. For the present $64*64$ wavelet transform, the recursion was implemented to a maximum level of $\log_2(64) = 6$. However, the 6^{th} level of the transform, which consisted of only a single coefficient (the mean of the data), was not included in further analyses. For each subspace, Haar wavelet coefficients w_i were squared and normalized to create energy probability distributions:

$$p_i = \frac{c_i^2}{\sum_n c_n^2} \quad (8)$$

where $\{c_i\}$ represents the set of wavelet coefficients over a number of translations k, given a particular scale j, and orientation $o \in \{HH, GH, GH\}$

For each wavelet scale, the sorted series p_i may be considered as an inverse empirical cumulative distribution function (ECDF). The degree of disorder or unpredictability of energy in each Haar wavelet subspace was calculated using Shannon entropy [17]:

$$\chi = \sum_i p_i \log_2(p_i) \quad (9)$$

Figure 4 shows an example of sorted wavelet coefficients at each of the first five wavelet scales for the "clean" and "cracked" images in figure 2. Note that the y-axis is plotted using a logarithmic scale. It may be seen that at certain scales the wavelet energy distributions differ in terms of the concentration of energy.

This procedure yielded 15 descriptor variables for each case, which quantified the concentration of energy of the Haar wavelet distribution for 3 orientations and 5 (frequency) levels. The total dataset was randomly split 10 times into model (90%) and validation (10%) subsets.

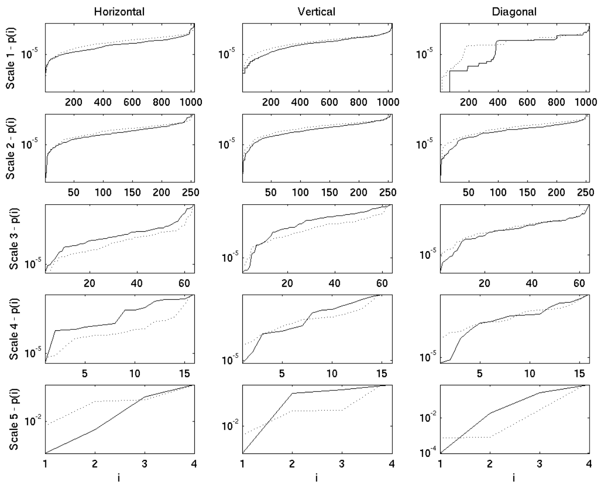

Fig.4 Cumulative normalized wavelet energy distributions for three orientations and five spatial scales. The dotted line represents coefficients associated with the 'clean' image; the solid line corresponds to the 'crack' image (from figure 2).

Logistic linear regression is a rigorous method for performing linear prediction on a dichotomous dependent variable [11]. Logistic linear regression models

$$y = \frac{\exp(\alpha + \beta_1\chi_1 + \cdots + \beta_n\chi_n)}{1 + \exp(\alpha + \beta_1\chi_1 + \cdots + \beta_n\chi_n)} \quad (10)$$

consist of a standard linear model, consisting of the sum of products of beta-weights β_n with features χ_n, plus a constant α. Logistic regression involves the application of the exponential function to this linear model, 'squeezing' Models were optimized using maximum-likelihood estimation using each of the model data sets, and classification performance assessed on the validation sets.

3. Results and Discussion

The combined Haar wavelet and Shannon entropy approach was applied to crack identification using video stream data acquired by an onboard infrared camera system.

Mean classification performance over 10 validation sets was 88.31%, with a minimum classification rate of 87.64% and a maximum classification of 89.61%.

When assessing these results, the challenging nature of the data set should be considered. Many 'clean' cases consisted of spurious lighting effects, dirt or other features, with structures quite similar to genuine cracks.

The present study represents an initial investigation of the utility of using the Shannon entropy of wavelet distributions as a feature extractor. Other aspects of the research (such as the classifier system) were deliberately kept as simple as possible. However, there exist many promising extensions of this feature extraction / classifier system. For example, the wavelet packet transform, which transforms data using a far more comprehensive range of space-frequency analysis functions, is expected to extract more information of interest. Preliminary inspection of the Shannon entropy measures indicates that the classification problem may be non-linear to some extent, indicating that use of an artificial neural network classifier may provide better results. The current experiment used Shannon entropy to parameterize the wavelet ECDF. However, other forms of discriminate information may be extracted from the ECDF using polynomial curve fitting or possibly PCA. For example, the polynomial regression coefficients of the ECDF can be used as well as / in place of Shannon entropy in order to better characterize the wavelet distribution. Finally, more complex structures may be extracted by creating new ECDF distributions based on combinations of spatially related wavelets. For example, creation of a new ECDF based on the geometric mean of triplets of vertically adjacent coefficients (for a vertical GH sub-bands) would capture the distribution of long vertical lines in an image.

We suggest that parameterization of the wavelet ECDF via Shannon entropy or other measures represent a useful feature extraction method for automated image classification. A useful property of this approach is the extraction of translation-invariant space - frequency image features. This experiment indicates that these features appear to be effective for detection of cracks in sewer pipes.

References

[1] Material received by Sewer System Administration Dept., City of Kitakyshu, Japan

[2] Maria-Luiza Antonie, Osmar R. Zäiane, Alexandru Coman, "Application of Data Mining Techniques for Medical Image Classification", Proceedings of the Second International Workshop on Multimedia Data Mining, San Francisco, USA, August 26, 2001

[3] Gunatilake, P., M. W. Siegel, A. G. Jordan, G. W. Podnar, "Image understanding algorithms for remote visual inspection of aircraft surfaces", Proceedings of the SPIE Conference on Machine Vision Applications in Industrial Inspection V, San Jose, February 1997, SPIE v. 3029, pp. 2-13.

[4] Bernatzki, A., Eppler, W., Gemmeke, H., "Interpretation of Neural Networks for Classification Tasks", Proceedings of EUFIT 1996, Aachen, Germany

[5] Nicholas J.Carino, "Concrete Technology - Past, Present and Future", ACI SP 144-30, American Concrete Institute, Detroit, MI, 1994, pp. 623-678

[6] Dürbeck, W., "Stand der Eigenkontrollverordnung in den Ländern der Bundesrepublik Deutschland aus der Sicht der Industrie", Dokumentation 4. Int. Kongress Leitungsbau, Hamburg, 16.-20.10.1994, Hamburg Germany, 1994, pp. 153-166

[7] M. Ferraro, G. Boccignone, T. Caelli, "On the Representation of Image Structures via Scale-Space Entropy Conditions" , IEEE Transactions on Pattern Analysis and Machine Intelligence , vol.21, no. 11 , 1999, pp. 1199 - 1203

[8] M. Ferraro, G. Boccignone, T. Caelli, "Entropy-based Representation of Image Information", Pattern Recognition Letters, vol., no. , 2002

[9] Keding, M. Riesen, S. van Esch, "Der Zustand der öffentlichen Kanalisation in der Bundesrepublik Deutschland. Ergebnisse der ATV-Umfrage 1990", Korrespondenz Abwasser, 37 (1990), Nr. 10, pp. 1148-1153

[10] F. Kirchner and J. Hertzberg, "A Prototype Study of an Autonomous Robot Platform for Sewerage System Maintenance", Autonomous Robots 4(4), 1997, pp. 319-331

[11] Menard, Scott., "Applied Logistic Regression Analysis", Sage Publications, Thousand Oaks, CA., 1995

[12] A. Mojsilovic, M. Popovic, A. Neskovic, A. Popovic, "Wavelet image extension for analysis and classification of infarcted myocardial tissue", IEEE Trans. Biomedical Engineering, vol. 44, no. 9, September 1997

[13] F. Murtagh and J.-L. Starck, "Image processing through multiscaleanalysis and measurement noise modeling", Statistics and Computing, vol. 10, 2000, pp. 95-103

[14] F. Murtagh, J.-L. Starck and M. Berry (Univ. Tennessee), "Overcoming the curse of dimensionality in clustering by means of the wavelet transform", The Computer Journal, vol. 43, 2000, pp. 107-120

[15] Lucas Paletta, Erich Rome, "Reinforcement Learning of Object Detection Strategies", SIRS 2000, 8th International Symposium on Intelligent Robotic Systems, The University of Reading, England, July 18-20, 2000.

[16] Rudolph, U., Wellnitz, J., "Zustand und Sanierungsbedarf der Abwasserkanäle in den neuen Bundesländern", Korrespondenz Abwasser, 38 (1991), Nr. 12, pp. 1625-1630

[17] Shannon, C.E. and Weaver, W. A., Mathematical Theory of Communication, Urbana, Illinois, University of Illinois Press, 1963

[18] Starck J., Murtagh F., Bijaoui A., Image Processing and Data Analysis: The Multiscale Approach. Cambridge University Press, 1998, Cambridge (GB)

[19] G.E. Stavroulakis, H. Antes, "Nondestructive elostatic identification of unilateral cracks through BEM and neural networks", Computational Mechanics 20, Springer Verlag 1997, pp. 439-451

[20] Strang, G. "Wavelet Transforms Versus Fourier Transforms.", Bull. Amer. Math. Soc. 28, 1993, pp. 288-305

[21] Shigeru Takaya, Gabriel Preda, Kazuyuki Demachi, Tetsuya Uchimoto and Kenzo Miya, "Electromagnetic Nondestructive Evaluation of Nuclear Fusion Reactor First Wall"

[22] S. Yoshimura, G. Yagawa, A. Oishi and K. Yamada, "Neural Network Based Inverse Analysis for Defect Identification with Laser Ultrasonics", Key Engeneering Materials vols. 145-149, Trans Tech Publications, Switzerland 1998, pp. 443-452