Optical Engineering

SPIEDigitalLibrary.org/oe

# Rapid cascade condition assessment of ductwork via robot vision

Yongxiong Wang
Jianbo Su

Optical Engineering 51(2), 027201 (February 2012)

# Rapid cascade condition assessment of ductwork via robot vision

**Yongxiong Wang**
Shanghai Jiao Tong University
Department of Automation and Ministry of Education of China
Key Laboratory of System Control and Information Processing
Shanghai, China
and
Jinggangshan University
School of Electronics Information Engineering
No. 28 Xueyuan Road, Qingyuan District, Ji'an City
Jiangxi, China
E-mail: wyxiong@sjtu.edu.cn

**Jianbo Su**
Shanghai Jiao Tong University
Department of Automation and Ministry of Education of China
Key Laboratory of System Control and Information Processing
No. 800 Dongchuan Road, Minhang District
Shanghai, China

**Abstract.** Automatic assessment of condition in ductwork is very desirable in applications. Presented is a visual condition diagnosis approach, which is capable of processing images rapidly and achieving high accuracy rates. A hierarchical coarse-to-fine image segmentation method is employed. False alarms could thus be progressively eliminated, which is robust in strongly noisy conditions. The simple classifiers combined in a cascade quickly classify the detected images and discard the uninterested (non-object) images, leaving more computation power on promising object-like regions. The features of each simple classifier are selected based on the Bhattacharyya distance. The cascade can be viewed as an object-specific focus-of-attention mechanism. Experimental results validate the effectiveness and rapidity of the proposed assessment method.
 [DOI: 10.1117/1.OE.51.2.027201]

Subject terms: robot vision; image segmentation; cascade classification; duct assessment.

Paper 111079 received Sep. 6, 2011; revised manuscript received Dec. 5, 2011; accepted for publication Dec. 6, 2011; published online Feb. 23, 2012.

## 1 Introduction

Air quality in indoor public places, such as office and commercial buildings, is very important to human health. A heating, ventilation and air-conditioning (HVAC) system of a building, which have already been used for many years without cleaning, may become a source of contaminants introduced into occupied spaces. The National Air Duct Cleaners Association (NADCA) of America has recommended inspection duration of 1 to 2 years for HVAC systems.[1] Ductwork, the principal component of HVAC, should be evaluated. This is currently realized by mobile closed circuit television (CCTV) systems or human inspectors.[2–5] Although the HVAC images could be obtained automatically by CCTV, the degree of cleanliness and anomaly detection and classification of thousands of miles of duct are still operated manually online or off-line. This introduces the issue of subjectivity, fatigue, and high-cost. Thus, a reliable automatic detection and recognition system for duct defects is highly desirable for HVAC ductwork.

While the pipeline robot goes through the duct, videos or digital images are gathered, saved in storage system, and shown on screen. Anomaly detection and condition assessment of pipeline, also known in some literature as underground pipe detection, has been a high concern in a wide variety of applications.[6–16] To overcome the limitations diagnosed by human interpretation on CCTV inspection images, the main alternative automated technologies applicable to pipeline inspection are based on vision, laser,[9] and ultrasonic.[10] Obviously, the methods based on robot vision provide very rich visualized condition information. These visual data can be used by inspectors during and after pipe inspections and can be easily fused with other sensor data. Therefore, vision-based technology has become the most popular technology for pipeline inspection.[11,12]

Moselhi and Shehab-Eldden[13] employ support vector machines (SVM) to detect pipe defects based on geometrical features. Sinha and Fieguth[14] use mathematical morphology methodology for segmentation of pipe defects including crack, hole, joint, collapse pipe, and lateral, and perform neuro-fuzzy network, to classify the pipe defects. Yang and Su attempt to apply radial basis networks[15] to develop a diagnostic system based on shape features. And then they employ SVM to detect the sewer defects based on textural features instead of morphological or geometrical features.[16] Guo et al. employ a change detection-based approach to automatic defect detection in sewer pipelines via videos and images.[11] Whereas, in analyzing HVAC duct images, we need to consider the inherent noise and irregularly corroded areas as well as the wide range of ductwork background patterns, such as metal texture and rust of the duct. Because of rust and long-term settled dust, color of backgrounds may be similar to detected objects. Some defects may be camouflaged in the backgrounds of corroded areas or settled dust. Therefore, the performance of image segmentation is not perfect in a noise environment and the detection rate to some defects, such as crack and hole, are not satisfactory.[12–15] Due to natural shape irregularities and the complex imaging environment, there are not the specific textural features or shape features of duct defects for automatic defect survey. The segmented objects have to be further processed and classified using more features, such as shape, textural features, and

0091-3286/2012/$25.00 © 2012 SPIE

so on. Moreover, the speed of the above classifiers is slow and their efficiency is not high because the classifiers need more computation power in undesired processed images. These systems cannot be used in real-time applications.

Because of the specialties of HVAC ductwork and the different objects of condition assessment, the existing automated inspection systems for other pipelines are not completely suitable to detect HVAC ductwork which is commonly made of galvanized steel sheet. Evaluating contamination degree for the HVAC ducts, besides the defects, is one of the primary assessment purposes. According to the standard of NADCA and Institute of Inspection, Cleaning, and Restoration Certification S520 (IICRC S520),[1] HVAC ductwork must be inspected for contamination degree and physical defects. So HVAC inspection should recognize six items: 1. clean duct, 2. hole, 3. crack or joint, 4. heavy dust, 5. debris, 6. rust.

This paper develops an automatic hierarchical method to segment and extract features from defects. The goal of this phase is to develop an automated segmentation. The whole segmentation course can be considered as a process of coarse-to-fine segmentation, in which false alarms of the segmented image are progressively eliminated in strong noise. The process mainly includes three steps: 1. coarse image segmentation based on global adaptive thresholding, 2. fine edge segmentation using level set approach with local gray value, and 3. shape analysis and objects separation using parameterized mathematical morphology.

To classify the severity of contaminants and defects, this paper proposes a method for combining successively simple classifiers in a cascade structure which dramatically increases the speed of the detector by focusing attention on promising regions of the image. The method allows duct detection or assessment with high detection rates of a few false positives, strong robustness in dynamical environments, and high processing speed (real-time).

There are a large number of methods[17–19] which have used cascades (or ensembles) of (boosted) classifiers for object detection. A cascade of boosted classifiers which is introduced by Viola[19] has the capability of fast face detection with high detection rates. A framework combining boosting, nested cascade classification, and bootstrap training is introduced by Verschae.[17] Their detection and classification systems are applied to face detection, eyes detection, and gender classification systems with high accuracy, robustness, processing speed, and training speed. Tuzel et al. presents an algorithm to detect pedestrians utilizing a cascade of boosting classifiers.[18] Cascade classifiers[19] consist of several layers of classifiers for obtaining fast processing speed and high accuracy, such that it will increase the complexity of classifiers when proceeding in the cascade. This is possible because there are much more non-objects than objects windows in any given image, and most of the non-object windows are quite different from the object windows.[17] Therefore, they can be easily discarded.

Our motivation for using cascade structure is similar. In duct detection, most of the ducts are non-object images (clean duct) which are easily recognized by simple classifiers and are discarded in the first stages (short processing time), while more complex processing is only reserved for these promising images (object candidates) by several cascade stages. The complexity of the classifiers (i.e., number of features in each classifier) generically increases when proceeding in the cascade. Note that the average processing time of an image is almost completely determined by the expected processing time of non-object images.[17] Moreover, after discarding most of the non-object images, it greatly relieved the imbalance class problem which may be very harmful for performance of classification in difficult to learn tasks.[20]

The most discriminating features are selected via the maximal Bhattacharyya distance between the classes.[21,22] And for multiclass problems, the classes corresponding to the maximal Bhattacharyya distance are first classified in a cascade structure.

The rest of this paper is organized as follows. In Sec. 2, the detailed methodology is given to segment the defect candidates and extract the features. In Sec. 3, the cascade classification is presented. Selection of the parameters, experimental results, comparative analysis, and comparison with single classifiers are depicted in Sec. 4, followed by conclusions and future work in Sec. 5.

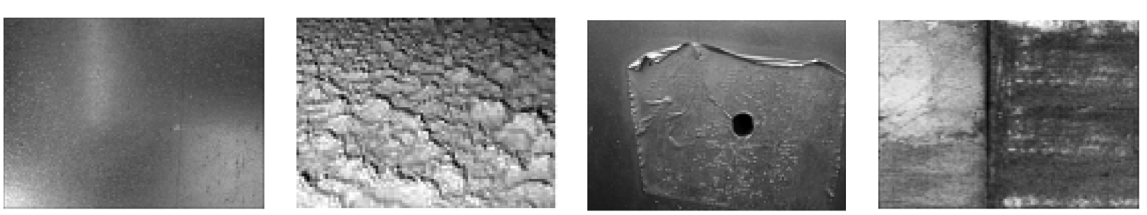
## 2 Image Segmentation and Features Extraction

The typical duct images are shown in Fig. 1. There are several factors, for example, corroded areas, various debris, mold and other microbial contamination, and a mass of settled duct,[1] which further complicate the segmentation of duct images and the detection of defects from duct images.

### 2.1 *Whole Framework of Hierarchical Segmentation*

The hierarchical segmentation method from duct images is illustrated in Fig. 2. It is composed of following three steps:

Step 1 *Coarse image segmentation:* The image preprocessing is employed to enhance the contrast of duct image between the duct background and the defect features. It can improve the understanding of the characteristics of air duct. The coarse image segmentation with global information is applied to obtain possible defect candidates. The object of this step is to acquire the duct defects candidates in the case of missing alarms as low as possible.



(a) (b) (c) (d)

**Fig. 1** Typical HVAC duct images with different objects: (a) clean surface, (b) heavy settled dust, (c) hole, and (d) joint with noise background.

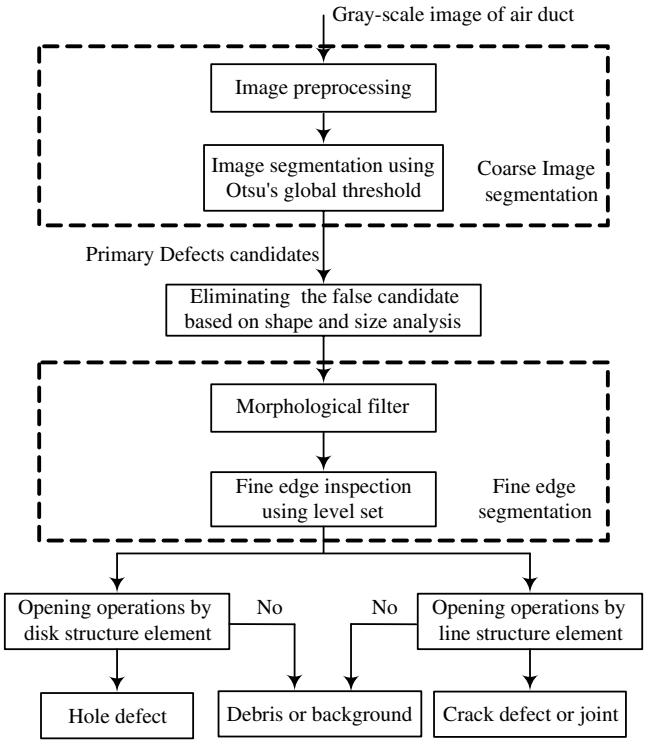



**Fig. 2** Flowchart of hierarchical automated duct images segmentation.

Step 2 *Fine edge segmentation:* Because the duct defects generally have their respective geometrical shapes and sizes, the sizes and shapes of the defects can be used to eliminate obvious false alarms and false defect candidates. The mathematical morphological operation is used to remove the noise and acquire the closed edge curves which are taken as the initialization value for the fine edge segmentation. Then, the level set is used to obtain fine edge with local information.

Step 3 *Objects separation:* Then, the anomalies are separated by the morphological operators with different structuring element, such that it is very easy to recognize the category of duct defects for classification.

### 2.2 *Coarse Image Segmentation*

Because of the long-term settled dust, it is essential to adjust image intensity values to improve the visibility of the images, as shown in Fig. 1. It is assumed that cracks, joints, and holes are the darkest objects in the image and the contrast of dark pixels is enhanced from the background. Then we apply a median filter to each duct image. The window size for the median filter is $5 \times 5$. This is determined based on the width of the crack, joint lines, and the size of the holes. The window size experimentally selected is large enough to erode the small features and at the same time small enough to be computationally fast. Therefore, a preprocessing step of enhancement of contrast is carried out before application of the proposed segmentation algorithm that works better in the image.

Because the cracks, joints, and holes are obvious in the duct image, the Otsu's technique is applied to coarsely segment the air duct image. The Otsu's technique,[19] an effective global adaptive thresholding method based on discriminant analysis, determines the optimal thresholds for the gray-level image segmentation by maximizing between-class variance and minimizing within-class variance.

### 2.3 *Fine Edge Segmentation*

After the coarse segmentation, a region-filling algorithm[20] is adopted to eliminate the holes within the interested segmented regions. The closed edges of the interested segmented areas for the following image process are obtained. Then, the morphological open and close operators with a three-pixel-diameter disk structure element are performed to filter very thin edges and random noise regions, which may be brought about by small debris and the color of the background.

Generally, Otsu's method is known as one of the best methods in automatically selecting threshold. However, when the difference of gray value between the background and the detected objects in the image is not obvious, the false alarms and edges error usually are brought about to the segmented image, and the method is sensitive to noise and size of objects.

Some edge localization errors or false alarms will be produced during the coarse image segmentation, as shown in Fig. 3(c). To remove the false alarms and extract the accurate sharp features from the segmented regions, it is necessary to refine every region's contour with the local gray quality. We adopt a level set based on the Chan–Vese model (CV)[23] to refine the curve of corresponding subimages which are cut from the original image and gain the piecewise smooth edge. CV model, which is the active contours to detect objects in a given image, has the ability of detecting smooth boundaries, scale adaptation, automatic change of topology, and robustness with respect to noise. The approach is based on techniques of curve evolution, Mumford–Shah functional for segmentation, and level sets. The optimal partition problem is solved by minimizing the following energy function:[23]

$$F(C, c_I, c_O) = \text{Area}[\text{inside}(C)] + \int_{\text{inside}(C)} [u(x,y) - c_I]^2 \mathrm{d}x\mathrm{d}y + \int_{\text{outside}(C)} [u(x,y) - c_O]^2 \mathrm{d}x\mathrm{d}y, \quad (1)$$

where $u(x, y)$ is the gray value of image pixel $(x, y)$, $C_I$ is the average value of the gray inside the contour $C$, $c_O$ is the average value of the brightness outside the contour $C$. Generally, the level set of CV model needs to manually set the initial curve.[23] In this paper, the closed edge acquired from preceding coarse image segmentation by Otsu's method is set as the initial curve $C$ of the level set, such that the automatic setting initial values are achieved.

### 2.4 *Extracting the Shape Feature and Eliminating the False Alarms Based on Simple Shape and Size Analysis*

Because the duct defects generally have their respective geometrical shapes and sizes, the shapes and sizes of the

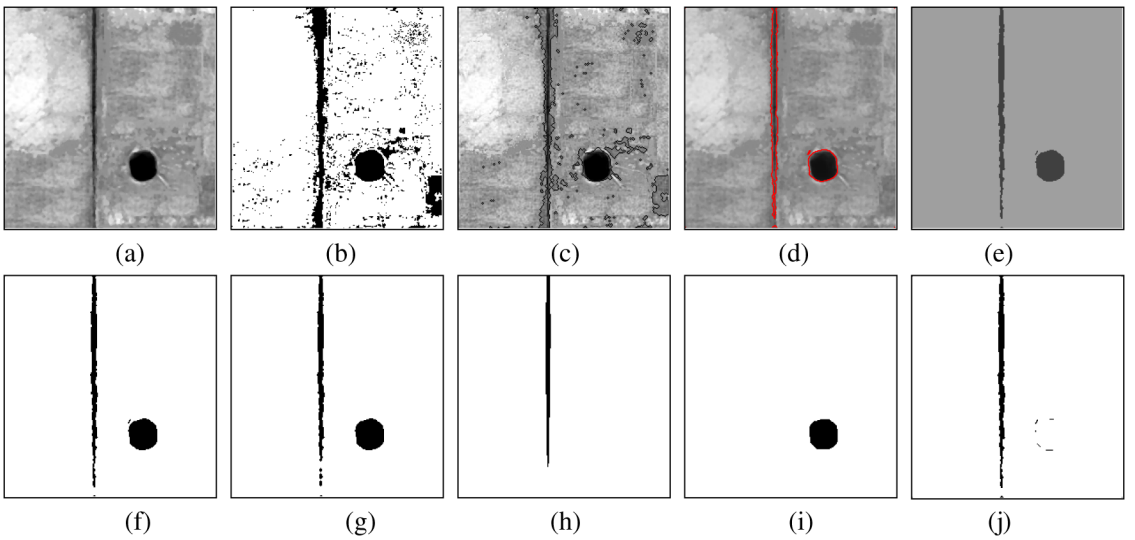



**Fig. 3** Intermediate results with two classes' defects of one typical image sample. (a) Original image. (b) Otsu' segmented image. (c) The image with mixed coarse edge. (d) The fine segmented image with mixed red curves of CV level set. (e) The fine segmented gray-value image by level set. (f) The binary image with a little noise after morphological opening using a disk structure element ($r = 1$ mm). (g) The binary image without noise after morphological opening using a disk structure element ($r = 2$ mm). (h) The image after morphological operator using a linear structure element ($l = 15$ mm). (i) The image after morphological operator using a disk structure element ($r = 6$ mm). (j) The result of (g)–(i).

defects can be used to eliminate obvious false alarms and false defect candidates. After the first step of image segmentation, the binary image can easily be obtained. However, due to the background of corroded areas or debris, there may still be some false alarms in the segmented duct image. The cracks and joints have characteristic length, width, range, and shapes which are long and thin, and the hole has characteristic area and shape approximating that of an ellipse. Thus, to reduce computation, we apply simple shape and size analysis to remove the false alarms before the second step of fine segmentation.

In this work, we adopt the following simple scalar region descriptors to represent other shape features besides the area, width, and length:

1. *Compactness.* Compactness is a popular shape description characteristic independent of linear transformations given by[24]

$$\text{Compactness} = \frac{(\text{region\_border\_length})^2}{\text{area}}. \tag{2}$$

   The most compact region in a Euclidean space is a circle. Using the outer boundary, compactness assumes values in the interval $(16, \infty)$.[24] The compactness of the hole equals approximately that of the circle, and the compactness of other shape objects, such as joints and cracks, is clearly larger than the hole.
2. *Eccentricity.* The simplest eccentricity characteristic[24] is the ratio of the major to the minor axes of an object. Obviously, the linear shape objects of joints and cracks have the largest eccentricity in the various shapes. The eccentricity of the hole is only slighted larger than 1.
3. *Convexness.* Let $u$ represent a set of contour points obtained from level set and its convex hull [denoted as $C(u)$]. The convexity measure is defined as $C_v(u) = \text{Area}(u)/\text{Area}[C(u)]$.

### 2.5 *Segmenting Independently Cracks, Joints, and Holes Using the Morphology with Line and Disk Structuring Element*

In this section, we aim to segment independently crack, joint, and hole based on morphology[20] so that the geometric features can be easily extracted and some bigger background noise removed. The morphological opening completely removes regions of an object that cannot contain the structuring element and retains the regions which can contain the structuring element. Sinha and Fieguth[25] put forward parameterized morphological operations to isolate crack, joint, and hole from noise background, however the horizontal structuring element they used can only isolate horizontal crack or joint. In this paper, we use two parameterized structuring elements, which are a disk structuring element of varying radius $r$, and a line structuring element of varying length $l$, angle $\theta$, and fixed width $w = 1$, to isolate the objects of arbitrary-angle crack, joint, and hole.

The opening of set $A$ by structuring element $B(x)$, denoted $A \circ B(x)$,[20] is defined as

1. For disk structuring element with varying radius $r$,

$$\phi(r) = A \circ B(r) = [A \Theta B(r)] \oplus B(r). \tag{3}$$

2. For line structuring element of varying length $l$, angle $\theta$,

$$\phi(l, \theta) = A \circ B(l, \theta) = [A \Theta B(l, \theta)] \oplus B(l, \theta). \tag{4}$$

$$\phi_{\text{open}}(l) = \underset{i=1,\cdots,18}{\text{Max}}[\phi(l, \theta_i)]. \tag{5}$$

where angle $\theta_i$ is oriented at every 10 deg from 0 deg to 180 deg. The element length $l$ and radius $r$ are based upon the range of cracks length and holes size that are of interest to the duct community. A more detailed discussion about parameters $l$ and $r$ will follow in Sec. 4.

### 2.6 *Extraction of Surface Roughness and Texture Features*

This section describes how the roughness of surface and texture features are extracted.

#### 2.6.1 *Roughness of surface*

Considering most ducts are made of galvanized steel sheets, the introduction of a critical feature, surface roughness of duct, is regarded as one of the important features to assess the clean duct.

Surface roughness is widely used in industry and generally quantifies the smoothness of a surface finish or surface texture. The arithmetic mean deviation of profile $R_a$ is universally recognized and the most used international parameter of roughness. Luk[26] proposes an approach of measuring the surface roughness based on machine vision. The roughness is defined as the ratio between the standard deviation of the image gray-scale and the root mean square height of the gray-level distribution along the row axis and column axis:

$$R_a = \left( \frac{\frac{1}{N-1}\sum_{i=0}^{l-2} n_i (1-\mu)^2}{\frac{1}{N}\sum_{i=0}^{l-1} n_i^2} \right)^{1/2}, \tag{6}$$

where

$$N = \sum_{i=0}^{l-1} n_i, \qquad \mu = \frac{1}{N}\sum_{i=0}^{l-1} i \cdot n_i, \tag{7}$$

$n_i$ is the number of pixels whose gray-scale are $I$, and $l$ is the gray-scale number.

From Eq. (6), it can be found that the roughness is the statistical feature for characterizing diverse gray scale in the image. It expresses distribution status of the different intensity pixel in a macroscopic perspective. Comparing to gray-scale co-occurrence matrices, the roughness is more intuitive and concise. The roughness roughly reflects the distribution status of the dust, debris, and rust within the duct.

For each cross-section of the 3-D surface texture, the Luk's formula to calculate $R_a$ of surface texture generates two aggregates of roughness of each profile scanning along the row axis and column axis, respectively. That means that it needs two $R_a$s to represent the roughness of surface texture. More reasonably, in this paper, we used a modified formula to achieve only one roughness average for full representation of surface texture roughness. The modified method is based on polar-coordinate scanning. Instead of scanning along $x$, $y$ coordinates, it now performs a circular scan in polar coordinates. The center of the image acts as the original point of the polar coordinates. We calculate the aggregate of roughness $R_a$ of each rotationally scanned profile according to the rotation angle, $\theta$. Equation (6) changes to

$$R_a|_{a=\theta} = \left[ \frac{\frac{1}{N-1}\sum_{i=0}^{254} n_i (i-\mu)^2}{\frac{1}{N}\sum_{i=0}^{255} n_i^2} \right]^{1/2}, \tag{8}$$

where $\theta$ is the rotation angle, $n_i$ denotes the number of pixels whose gray scale are $i$ at $\theta$ angle. While scanning 360 deg, the surface roughness of the whole image can be obtained. In Zaklit et al.,[27] the results are verified to be same as Luk's method, whereas the roughness only needs to be calculated once.

#### 2.6.2 *Texture feature*

To decrease the time of extracting texture feature, the texture feature is extracted from the defect candidate subimages obtained by first coarse segmentation, such that the computation time can be substantially decreased.

There are some efficient approaches to extraction of textural features. Wavelet Transform (WT) and co-occurrence matrices are used to extract the co-occurrence features from defective textile fabrics[28] and are powerful in detecting defects. Yang[16] effectively inspects the sewer defects based on WT and co-occurrence matrices. In this paper, the hybrid usages of wavelet transform and gray-level co-occurrence matrices is considered as an effective solution for the texture analysis of duct conditions.

*2.6.2.1 Wavelet transform.* WT has multi-resolution technique in the frequency domain, so its sub-band decomposition of tree structure is appropriate for detection of local signal varieties on images. When digital images are to be viewed or processed at multiple resolutions, the Discrete Wavelet Transform (DWT) is considered as an efficient mathematical tool of choice.[20] The effectiveness of performing the texture analysis using DWT has been shown.[28] DWT provides powerful insight to the spatial and frequency characteristics of an image. The Fast Wavelet Transform (FWT) is defined via the series expansions:[20]

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\varphi(2x-n), \tag{9}$$

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\varphi(2x-n), \tag{10}$$

where mother wavelet and scaling function, $\varphi(x)$ and $\psi(x)$, are computed by convolving $\varphi(2x-n)$ with the time-reversed scaling and wavelet vectors, $h_\varphi(n)$ and $h_\psi(n)$. The original function, $\varphi(2x-n)$, is split into a low-pass filter (approximation component) corresponding to $\varphi(x)$, and a high-pass filter (detail component) corresponding to $\psi(x)$.

Due to the separability of wavelet transform, a two-dimensional DWT can be represented as a number of one-dimensional transforms. Through the decomposition of 2D-DWT, which is implemented by consecutive low-pass (L) and high-pass (H) filtering through one-dimensional convolution, the CCTV duct image can be divided into an approximation image (LL) by a 2-D scaling function, $\varphi(x,y)$, shown as: $\varphi(x,y) = \varphi(x)\varphi(y)$, the kernels can be represented as three separable 2-D wavelets $\psi^H(x,y) = \psi(x)\varphi(y)$, $\psi^V(x,y) = \varphi(x)\psi(y)$, $\psi^D(x,y) = \psi(x)\psi(y)$, where $\psi^H(x,y)$, $\psi^V(x,y)$, and $\psi^D(x,y)$ are called horizontal (HL), vertical (LH), and diagonal wavelets (HH), respectively. At each recurring step of decomposition, the approximation image is split into the next level of approximation and detail images until no more valuable information is obtained.

*2.6.2.2 Co-occurrence matrices.* Co-occurrence matrixes describe statistical properties of patterns that jointly occur in the neighborhood of each pixel. Co-occurrence matrixes characterize to capture spatial dependence of image gray-level value. They are regarded as an effective approach for pattern classification,[16,28] and have been widely used in texture classification or texture segmentation.

The gray level co-occurrence matrix $p_{i,j}(\theta, s)$ is defined as the probability of pixel $i$ and $j$ with a specified distance of $s$ and direction of $\theta$ occurring in the image. The size of co-occurrence matrix depends on the range of the gray-level values of the CCTV image. Generally, there are 14 types of co-occurrence features derived from co-occurrence matrices useful for pattern classification. However, the relationship between the co-occurrence features is not absolutely independent so that the redundant features would reduce the efficiency of pattern classification. Through a discriminant analysis, the most independent co-occurrence features with correlation coefficients of less than 0.5 were found as entropy (Entro), correlation (Corre), and cluster tendency (Clust),[16] which were broadly used to describe the textures of objects on images. In this paper, the three features, which are defined as follows, have been used to describe the texture of sub-window in HVAC duct images:

$$\text{Entro} = -\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij} \cdot \log P_{ij}, \tag{11}$$

$$\text{Corre} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(i \cdot j)P_{ij} - \mu_x\mu_y}{\sigma_x\sigma_y}, \tag{12}$$

$$\text{Clust} = \sum_{i=1}^{n}\sum_{j=1}^{n}(i - \mu_x + j - \mu_y)^2 \cdot P_{ij}, \tag{13}$$

where

$$\mathrm{P}_{ij} = \frac{M_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}}, \quad \mu_x = \sum_{i=1}^{n}\sum_{j=1}^{n} i \cdot P_{ij}, \quad \mu_y = \sum_{i=1}^{n}\sum_{j=1}^{n} j \cdot P_{ij}, \tag{14}$$

$$\sigma_x = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(i - \mu_x)^2 \cdot P_{ij}}, \quad \sigma_y = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(i - \mu_y)^2 \cdot P_{ij}}. \tag{15}$$

Co-occurrence features are affected by the distance and the direction between the two positions mainly, and the gray scale quantization levels are not important in general cases. The distance and direction between the two adjacent pixels have inherent discipline. So in this paper, $s = 1$ and $\theta = 0$ deg, 45 deg, 90 deg, and 135 deg are chosen.

## 3 Cascade Classification

This section describes the basic SVM-based classifier we adopt and an algorithm for constructing a cascade of classifiers which step-by-step achieves condition diagnosis of HVAC ducts while radically reducing computation time. The insight is that the priority of classification is arranged in a cascade structure according to the Bhattacharyya distance which is often used as a class separability measure.[22] We construct simple cascade classifiers and are therefore more efficient.

### 3.1 *Whole Architecture of Cascade Classification*

The architecture of cascade classification is illustrated in Fig. 4. There are three simple classifiers in this cascade. To rapidly diagnose the duct defects from large numbers of HVAC images, we first classify clean and unclean duct based on binary SVM since the Bhattacharyya distance between them is maximal. As a result, many non-object images can be discarded as soon as possible before more complex classifiers are called upon to achieve more difficult classification and gain low false-positive rates.

The second-level classification is used to inspect the duct defects, such as hole, crack, and joint. In this stage, to reduce computation, the false candidates are eliminated based on shape and size analysis. Then the multi-class SVM-based classification is employed to classify the defect classes of HVAC duct.

The third-level classification is performed to assess duct condition, such as the degree of rust, dust, and even mold. Considering the diversities of features, it is necessary to divide them into more subclasses for classification. This approach aims at finding multiple boundaries between these subclasses and obtaining a better performance in duct condition assessment.

The structure of the cascade reflects the fact that an overwhelming majority of all images are negative samples and not necessary to further process, i.e., clean duct. As such, the cascade attempts to reject as many negative samples as possible at the first stage. While a positive sample will trigger the evaluation of every classifier in the cascade, this is a rare event. And parallel processing can be adopted, such that the computation time can radically be reduced.

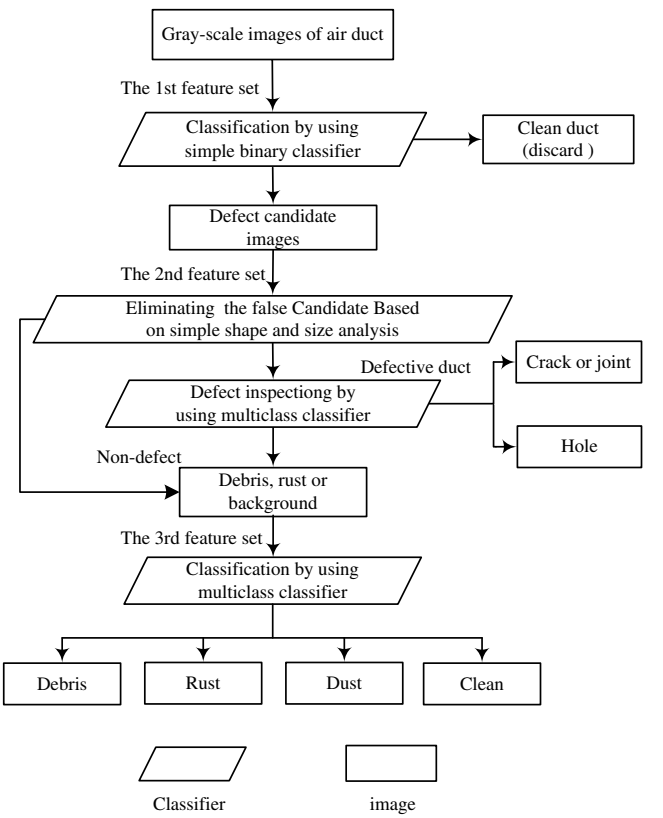



**Fig. 4** Flowchart of cascade classification.

### 3.2 *Selection of Features Based on Bhattacharyya Distance and the Classifiers*

#### 3.2.1 *Feature selection based on Bhattacharyya distance in training*

The numbers of the extracted features from all samples are enormous. Feature selection can reduce measurement costs, shorten computational time, and improve classification accuracy.

In statistics, the Bhattacharyya distance measures the similarity of two discrete probability distributions. It has been used as a class separability measure for feature selection and is known to provide the upper and lower bounds of the Bayes error.[21,22] Thus, we propose to select the features based on the Bhattacharyya distance as an optimization criterion assuming normal distributions. For multiclass problems, we present a simple extension. The Bhattacharyya distance of the $j$-th feature between the 1st and 2nd class is defined as

$$B_{12}^{j} = \frac{1}{8}(\mu_2^j - \mu_1^j)^T \left[\frac{\Sigma_1^j + \Sigma_2^j}{2}\right]^{-1} (\mu_2^j - \mu_1^j) + \frac{1}{2}\ln \frac{|(\Sigma_1^j + \Sigma_2^j)/2|}{\sqrt{|\Sigma_1^j||\Sigma_2^j|}}, \quad (16)$$

where the superscript $j$ denotes the $j$'th feature, the subscript indicates the number of classes, $\mu_i^j$ and $\Sigma_i^j$ are the mean vector and covariance matrix of class $i$, respectively.

In a multiclass situation, the average value of the Bhattacharyya distance, over all classes, is used to compute. The average value of the Bhattacharyya distance is given as

$$\bar{B}_i^j = \frac{1}{N}\sum_{k=1}^{N} B_{ik}^j, \quad (17)$$

where $N$ denotes the number of classes. Features are then ranked in order of descending values of the Bhattacharyya distance. For each classifier, the $l$ features corresponding to the $l$ maximum values of the Bhattacharyya distance are then selected and value of $l$ determined according to upper and lower bounds of the error.[21] The features are gradually extracted as needed in test. This approach does not take into account existing correlations between features.

#### 3.2.2 *Binary and multi-class classifiers*

The basic classifier which we use in this paper is SVM because it has the capability of providing classification probabilities aside from categories and high performance in many applications.[22,29]

The binary classification is used to solve the two-class recognition problem. In this classification procedure, the classes corresponding to the maximal Bhattacharyya distance are first classified in a cascade structure. Based on the Bhattacharyya distance, the 1st feature subset is selected which is used in the first-level classifier out of the feature space of all samples. The 2nd feature subset for the second-level classifier is selected from the remaining samples after the first-level classification, and so on. A multi-class classification strategy is general in many real-world recognition problems by finding multiple boundaries between these subclasses. The class of defect candidates includes many subclasses which include hole, crack, and joint. The non-defective subclass is partitioned into debris, dust, rust, and background. The multi-class SVMs are used to classify the subclasses.

#### 3.2.3 *Complexity analyses of the classifiers in per layer*

To obtain an optimal cascade classifier, we need to handle the tradeoff between the true positive rate (TPR), the false positive rate (FPR) and the number of features in each layer. To rapidly discard many non-object images, a simple classifier is first used to classify clean and unclean duct because of the maximal Bhattacharyya distance between the two classes in the first layer. Thus the first-layer classifier is the simplest one in the cascade. The error of the Gaussian maximum likelihood (GML) classifier is inversely proportional to the Bhattacharyya distance of features between two classes and can be empirically expressed as[21]

$$\varepsilon = 40.219 - 70.019B + 63.578B^2 - 32.766B^3 + 8.7172B^4 - 0.91875B^5, \quad (18)$$

where $B$ is the Bhattacharyya distance of features between two classes. To address the tradeoff between the TPR and FPR, the classifiers in the cascade are generically more complex when proceeding in the cascade as the Bhattacharyya distance between two classes is small, and more features are used in the classifier.

## 4 Experiment

### 4.1 *Experiment System Setup*

We fabricate an active infrared-CCD condition diagnosis system of HVAC duct by using the pocket hardware.[4] The mechanics of surveying the ducts are similar to the traditional CCTV sewer inspection. The pan and tilt CCTV camera and the schematic diagram of inspection principle of the ducts are shown in Fig. 5. The videos or digital images are gathered by the sensors fixed at a tractor robot going through the duct. The major benefit of our inspection system is that the flattened images acquired from pan camera greatly simplify automated computer-based analyses, and the operator will acquire much higher quality image data to make cleaning solutions and restoration decisions.[4]

We have acquired hundreds of HVAC duct images from several buildings in Shanghai, China. The size of the duct images are $870 \times 834$ pixels. The collected 321 duct images

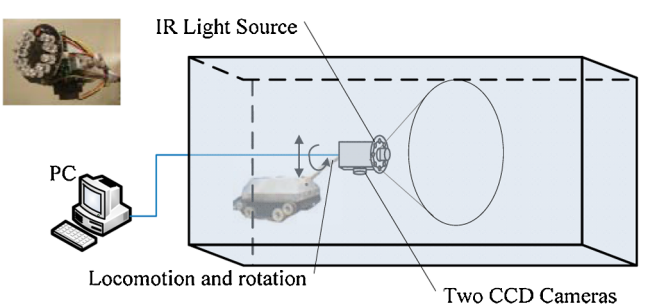



**Fig. 5** Schematic diagram of inspection principle and the inspection sensor with pan and tilt cameras.

**Table 1** Sub-image samples of each class.

| Sub-images | Clean ducts | Light contamination | Heavy contamination | Rust | Debris | Joint and crack |
|---|---|---|---|---|---|---|
| Number of the samples | 293 | 282 | 138 | 52 | 43 | 58 |

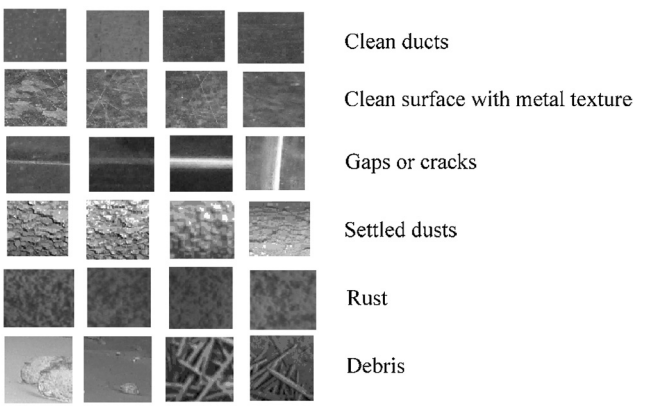



**Fig. 6** Typical sub-images of HVAC duct.

include 128 samples of clean ducts, 96 samples of light contamination, 35 samples of heavy contamination, 17 samples of all kinds of debris, 24 samples of rust, 7 samples of hole, and 14 samples of crack and joint.

From these original images, we extract 866 sub-images which include some typical samples, such as various clean ducts, light contamination surface, heavy contamination, and so on (listed in Table 1). The class "hole" does not include sub-images. The typical sub-images are shown in Fig. 6. The images of every class and subclass are divided into two even portions. One portion is randomly selected for training and the other is used for testing.

### 4.2 *Parameter Selection of the Length and Radius Structuring Element*

While performing a morphological opening on duct images, the area of regions will decrease as the radius of the disk structuring element increases, and the area of regions will be zero when the regions cannot contain the structuring element. Therefore, the selections of the size of the structuring element are essential to achieve the optimum performance of this segmentation step. We experimentally obtain the appropriate parameters value of length $l$ and radius $r$, shown in Figs. 7 and 8. These figures plot the relationships between the values of the structuring element and the average normalized areas of objects in each class (crack, hole, and joint). The normalized area[25] is defined as the ratio of the number of dark pixels after opening operation to that of dark pixels in image $I$

$$A_L = k_1 \frac{\sum_{i=0}[I \circ B(l)]}{\sum_{i=0} N(I_L)}, \qquad A_D = k_2 \frac{\sum_{i=0}[I \circ B(r)]}{\sum_{i=0} N(I_D)}, \tag{19}$$

where $N$ is the number of dark pixels in image $I$ after binary thresholding; $I_L$, $I_D$ are the idealized, prototype images of the perfect crack and hole; $k_1$, $k_2$ is constant parameter; $B$ is the structuring element.

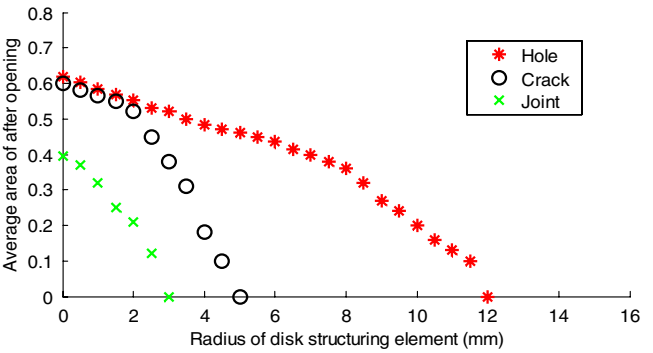



**Fig. 7** The average areas of each class after morphological opening with different adii of disk structuring elements.

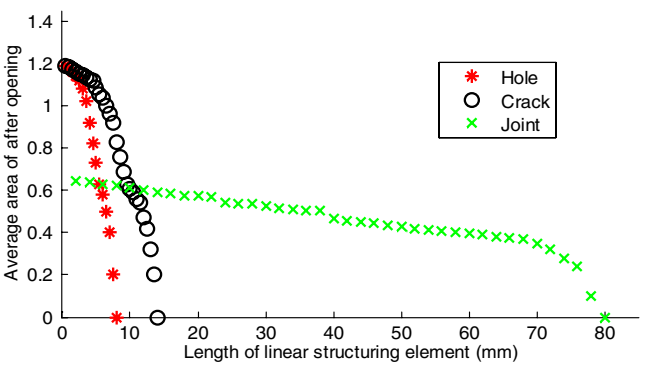



**Fig. 8** The average area of each class after morphological opening with different lengths of linear structuring elements.

From Fig. 7, the radii of disk structural elements, which are 3 mm, 5 mm, and 12 mm, respectively, corresponding to separate the joints, cracks, and holes clearly satisfies the cut-off thresholds and are selected as the optimum parameters that consistently provide good detection in all types of duct images. Similarly, the corresponding cut-off thresholds for lengths of linear structural elements are 8 mm, 14 mm, and 80 mm for segmenting the joint, crack, and hole, as shown in Fig. 8. For example, when the radius of disk structuring element in morphological opening is equal to 4 mm, the joints will be removed, holes and cracks will be preserved.

### 4.3 *Experimental Results*

#### 4.3.1 *Experimental results of hierarchical segmentation*

The proposed hierarchical segmentation method works very well for duct images containing only one class of object. Figure 3 represents the processing procedure of our

hierarchical segmentation method step-by-step. In real HVAC ductwork, duct images may contain two or more defect objects, such as containing a joint and a hole, as shown in Fig. 3(a). To segmentation and classification of such images, a similar morphological approach is applied to separate the two classes. The image (j) can be obtained by image (g) subtracting the image (i) in Fig. 3. Figure 3 illustrates the procedure of separating the two classes through coarse image segmentation, fine image segmentation, and successive morphological operation in turn.

In the segmentation experiment, to assess the performance of the proposed hierarchical segmentation method, two performance evaluations, *Accuracy* and *Completeness*, are defined as follows:

$$\text{Accuracy} = \frac{\text{number\_of\_correctly\_segmented\_defects}}{\text{number\_of\_real\_defects}} \times 100\%,$$

$$\text{Completeness} = 1 - \frac{\sum \left| I_{\text{Seg}} - I_{\text{Truth}} \right|}{\sum N} \times 100\%,$$

where $I_{\text{Seg}}$ is the number of object pixels in the segmented images, $I_{\text{Truth}}$ is the number of object pixels in the ground-truth image (i.e., the number of object pixels of reference), and $N$ is the number of pixels in the images $I$.

Based on the performance index of *Accuracy* and *Completeness*, we compare the automated hierarchical segmentation method to other segmentation methods, which are Canny edge approach,[30] the Smallest Univalue Segment Assimilating Nucleus (SUSAN) segmentation, Otsu and morphology segmentetion.[25] The results are listed in Table 2. The segmentation thresholdings are manually chosen for every duct image in Canny and SUSAN methods; the proposed method, Otsu and morphology segmentation, is based on automatic thresholding. The proposed method is effective and accurate for segmenting and classifying duct images as seen in Fig. 3. The experimental results show that the proposed method achieves a very small proportion of false detection and performs better than other approaches in Table 2.

#### 4.3.2 *Experimental results of cascade classification method*

To evaluate the performance of the proposed cascade classification method, a performance evaluation is used, defined as follows:

$$\text{Accuracy Rate} = \frac{\text{Number correctly detected in the class}}{\text{Number of each class}} \times 100\%$$

$$\text{FPR} = \frac{\text{Number of False Positives}}{\text{Number of True Negatives} + \text{Number of False Positives}} \times 100\%$$

$$\text{TPR} = \frac{\text{Number of True Positives}}{\text{Number of True Positives} + \text{Number of False Negatives}} \times 100\%$$

*4.3.2.1 The first-layer classifier.* In this experiment, the maximal Bhattacharyya distance is between the 1st class and others, i.e., the clean ducts are most easily classified. That may be because most ducts are made of galvanized steel sheets with smooth and glossy surfaces for clean ducts. Thus the ducts are first divided into two classes: clean and non-clean ducts, including defect candidates and others. The first-layer classifier eliminates a large number of clean samples with very little processing. The experimental results are shown in Table 3.

To simplify the design of the classifier and decrease the computation time, we select minimum numbers of features based on the Bhattacharyya distance within the set maximum allowed TPR and the minimum allowed FPR in each layer. In our experiment, two features, roughness of surface and a wavelet-based feature corresponding to two maximum Bhattacharyya distances are selected for the first-layer SVM.

**Table 3** Results from first-layer classifier for two classes.

| Classes | Non-clean | Clean |
|---|---|---|
| Accuracy_Rate (%) | 99.8 | 88.6 |
| FPR (%) | 1.8 | 9.2 |
| TPR (%) | 98.1 | 95.6 |

**Table 2** Comparison of accuracy and completeness for image segmentation.

| | Hole (%) | | Joint (%) | | Crack (%) | | Both joint and hole (%) | |
|---|---|---|---|---|---|---|---|---|
| Segmentation method | Accuracy | Completeness | Accuracy | Completeness | Accuracy | Completeness | Accuracy | Completeness |
| Canny edge | 76.6 | 75.3 | 74.2 | 72.3 | 72.6 | 70.9 | 75.2 | 73.8 |
| SUSAN | 83.3 | 81.6 | 76.3 | 74.6 | 76.3 | 72.2 | 79.2 | 75.3 |
| Otsu and morphology | 62.1 | 60.8 | 52.1 | 51.2 | 52.4 | 48.7 | 56.4 | 54.6 |
| Proposed method | 91.8 | 90.5 | 90.1 | 88.6 | 90.2 | 85.3 | 90.1 | 82.1 |

**Table 4** The classification results of second-layer classifier for three classes.

| Classes | Hole | Crack or joint | Non-defect |
|---|---|---|---|
| Accuracy_Rate (%) | 91.5 | 90.8 | 92.6 |
| FPR (%) | 2.6 | 2.0 | 5.2 |
| TPR (%) | 82.1 | 86.3 | 86.2 |

**Table 5** The classification results of last-layer classifier for four classes.

| Sub-classes | Clean | Dust | Debris | Rust |
|---|---|---|---|---|
| Accuracy_Rate (%) | 94.9 | 92 | 88.6 | 89.8 |
| FPR (%) | 3.9 | 2.8 | 2.3 | 4.5 |
| TPR (%) | 84.2 | 87.3 | 89.2 | 85.3 |

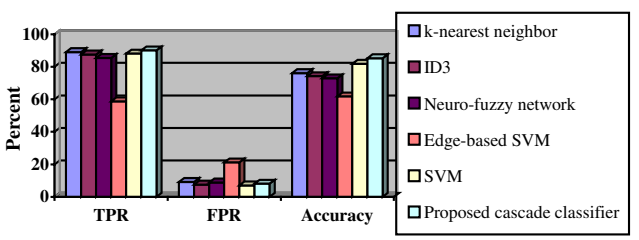



**Fig. 9** The total classification performances of the *k*-nearest neighbor, ID3, neuro-fuzzy network, edge-based SVM, SVM and the proposed cascade classifier for the "crack or joint" class.

*4.3.2.2 The second-layer classifier.* In order to accelerate computer speed, the simple shape and size analysis are carried out to discard the evident false positives before the second-layer classifier. For example, the cracks are long and thin and have a limited area. The very large or very small background and debris can be eliminated in the proper thresholds. After empirical analysis to cracks shown in Figs. 7 and 8, the threshold of the ratio of length and width is set between 4.23 mm and 9.55 mm in this experiment. The joints and holes have some similar characteristics. The joint's threshold of the ratio of length and width is larger than 9.55 mm. The hole's is below 1.58 mm.

After eliminating false defects candidates, the second-layer classifier is used to classify the defects of duct. From the experimental results shown in Table 4, the classifier based on SVM can efficiently identify the duct defects, including two subclasses: hole, crack or joint. From the result of experiment, five features, including two wavelet-based features and three low-dimensional geometric features, such as compactness, eccentricity, and convexness are chosen in this classifier. It can be seen that the classifier of this layer is little more complex than that in the first layer.

*4.3.2.3 The last-layer classifier.* After the second-layer classification, the remainders of duct images are considered to include four subclasses: clean, debris, rust, and dust. The classifier based on SVM is directly adopted to classify the four subclass ducts. From the experimental results shown in Table 5, the classifier based on SVM can efficiently identify the subclass. In this layer, five features, including two wavelet-based features, two co-occurrence features, and the roughness of surface, are chosen in this classifier.

The dimensionality of selected features in this layer is larger than that in the second layer. From the result of experiment, it may be concluded that the complexity, i.e., numbers of used features, in the three classifiers increase as proceeding in the cascade. It may reflect the idea or method of classification from simple to complex.

*4.3.2.4 Total performances of duct condition assessment.* To compare the proposed cascade classifier, the common pattern classification techniques, including *k*-nearest neighbor, ID3, neuro-fuzzy network methods,[14] edge-based SVM (only using shape features), and single SVM are used to classify the classes of duct condition based on the same method of selecting features. The *k* value of the *k*-nearest neighbor is set to 10. Figure 9 illustrates the total performance of the *k*-nearest neighbor, the ID3, the neuro-fuzzy network, and the proposed cascade classification methods for the "crack or joint" class. The performance of the other classes is similar

**Table 6** Comparison of total classification results for all classes.

| Methods | Accuracy_Rate (%) | | | | | | Average test time per an image[a] |
|---|---|---|---|---|---|---|---|
| | Clean | Hole | Crack or joint | Dust | Debris | Rust | |
| $\kappa$-nearest neighbor | 95.1 | 70.9 | 76.1 | 81.5 | 81.5 | 80.1 | 0.89 sec |
| ID3 | 95.8 | 72.2 | 74.2 | 78.9 | 80.1 | 76.8 | 0.88 sec |
| Neuro-fuzzy network | 96.5 | 71.3 | 72.9 | 81.6 | 79.9 | 78.9 | 0.98 sec |
| Edge-based SVM | 72.1 | 53.5 | 61.8 | 43.2 | 45.8 | 58.2 | 0.78 sec |
| SVM | 96.9 | 72.8 | 78.8 | 86.2 | 82.6 | 81.7 | 0.99 sec |
| Proposed cascade classification | 97.3 | 82.8 | 85.2 | 88.9 | 86.3 | 84.2 | 0.26 sec |

[a]The test time is acquired in a computation with 2.60 G Hz Pentium G620.

to the "crack or joint" class. Figure 9 shows that the proposed cascading method has better performance than the including *k*-nearest neighbor, ID3, neuro-fuzzy network, edge-based SVM, and SVM methods in terms of TPR, FPR, and accuracy.

Comparisons of accuracy from the classification methods are shown in Table 6. The average test time per image is listed in Table 6. Table 6 shows that the performance using edge-based SVM is worst in all classifiers. The experiments demonstrate that single classifiers do not work properly, especially edge-based classifiers. The experimental results show the proposed cascade classifier has good accuracy and high processing speed. The accuracy of rare classes, such as "hole" and "crack or joint", has improved obviously. It can be concluded that subclass division and cascade classification strategies are helpful for achieving a good classification performance.

## 5 Conclusions

This paper presented a novel cascade approach for rapid condition diagnosis of HVAC duct based on robot vision. The total procedure can be considered as a sequential process of focus of attention to objects.

The hierarchical segmentation method for duct image is used to obtain coarse-to-fine defect candidates and simple shape analysis is employed to eliminate obvious false alarms. The segmentation results using level sets of CV model are closer to the actual shape of defects because of the usage of local gray and optimization. The proposed hierarchical segmentation or understanding method is an effective method for segment the complex or multiple-goal images.

The most discriminating features, which have the maximal Bhattacharyya distance between the classes, are selected while the classification error can approximately be estimated. The cascade classification combines a series of simple classifiers in a cascade which can quickly reject the uninteresting images and leave more computation power on recognizing object-like images and regions. The most uninteresting images are discarded and no further processing is employed. Computation efficiency of the whole system is radically improved, which is essential to real-time applications. In duct detection, the structure of the cascaded detection has high detection rates with a few false positives, strong robustness in dynamic environments, high processing speed (real-time), and can improve the performance of classification with imbalanced data distribution.

Although our approach has shown promising results, some issues to enhance the performance of inspecting duct defects still remain. Future work will focus on better image preprocessing, more effective feature extraction, and selection strategies.

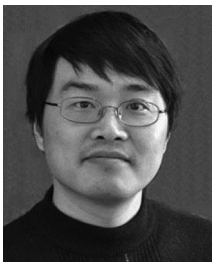
### *Acknowledgments*

This research was supported by the National Natural Science Foundation of China (Key Project No. 60935001).

### *References*

1. B. Lundquist et al., "Assessment, cleaning and restoration of HVAC systems," http://powervac.ca/pdf/ACR_2006.pdf (2006).
2. Biovac System Inc., "Robotic inspection, cleaning and spraying system," http://www.biovacsystem.com/index.php?option=com_content&view=article&id=73&Itemid=82&lang=en.
3. Y. Wang and J. Zhang, "Autonomous air duct cleaning robot system," in *Proc. IEEE Midwest Symp. Circuit.*, pp. 510–513, IEEE, Puerto Rico (2006).
4. Y. Wang and J. Su, "Movability of the tracked pipeline robot based on hierarchical fuzzy control," *High Technol. Lett.* **17**(2), 166–172 (2011).
5. Z. Zhang et al., "Inspection and sampling pipeline robot," http://www.eslon.com.cn/qx.htm.
6. O. Duran, K. Althoefer, and L. D. Seneviratne, "State of the art in sensor technologies for sewer inspection," *IEEE Sens. J.* **2**(2), 73–81 (2002).
7. R. Kirkham, P. D. Kearney, and K. J. Rogers, "PIRAT—a system for quantitative sewer pipe assessment," *Int. J. Robot. Res.* **19**(11), 1033–1053 (2000).
8. CERF, "Evaluation of SSET: the sewer scanner & evaluation technology," Report #40551, Civil Engineering and Research Foundation (CERF), Washington, DC (March 2001).
9. O. Duran, K. Althoefer, and L. D. Seneviratne, "Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network," *IEEE Trans. Autom. Sci. Eng.* **4**(1), 118–126 (2007).
10. F. Gomez, K. Althoefer, and L. D. Seneviratne, "An ultrasonic profiling method for sewer inspection," in *Proc. IEEE Int. Conf. Robot. Automation*, pp. 4858–4863, IEEE, New Orleans, LA (2004).
11. W. Guo, L. Soibelman, and J. H. Garrett Jr., "Automated defect detection for sewer pipeline inspection and condition assessment," *Automat. Constr.* **18**(5), 587–596 (2009).
12. J. Mashford et al., "A morphological approach to pipe image interpretation based on segmentation by support vector machine," *Automat. Constr.* **19**(7), 875–883 (2010).
13. O. Moselhi and T. Shehab-Eldden, "Automated detection of defects in underground sewer and water pipes," *Automat. Constr.* **8**(5), 581–588 (1999).
14. S. K. Sinha and P. W. Fieguth, "Neuro-fuzzy network for the classification of buried pipe defects," *Automat. Constr.* **15**(1), 73–84 (2006).
15. M. Yang and T. Su, "Segmenting ideal morphologies of sewer pipe defects on CCTV images for automated diagnosis," *Expert Syst. Appl.* **36**(2), 3562–3573 (2009).
16. M. Yang and T. Su, "Automated diagnosis of sewer pipe defects based on machine learning approaches," *Expert Syst. Appl.* **35**(3), 1327–1337 (2008).
17. R. Verschae, J. Ruiz-del-Solar, and M. Correa, "A unified learning framework for object detection and classification using nested cascades of boosted classifiers," *Mach. Vision Appl.* **19**(2), 85–103 (2008).
18. O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on riemannian manifolds," *IEEE Trans. Pattern Anal.* **30**(10), 1713–1727 (2008).
19. H. Yan, "Unified formulation of a class of image thresholding techniques," *Pattern Recogn.* **29**(12), 2025–2032 (1996).
20. R. C. Gonzalez and R. E. Woods, *Digital Image Process.*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ (2002).
21. E. Choi and C. Lee, "Feature extraction based on the Bhattacharyya distance," *Pattern Recogn.* **36**(8), 1703–1709 (2003).
22. S. Theodoridis and K. Koutroumbas, *Pattern Recogn.*, 4th ed., Elsevier Inc., London (2009).
23. F. T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
24. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 2nd ed., Brooks/Cole, Pacific Grove, CA (2002).
25. S. K. Sinha and P. W. Fieguth, "Segmentation of buried concrete pipe images," *Automat. Constr.* **15**(1), 47–57 (2006).
26. F. Luk, V. Huynh, and W. North, "Measurement of surface roughness by a machine vision system," *J. Phys. E* **22**(12), 977–980 (1989).
27. J. Zaklit et al., "Quantitatively characterizing automotive interior surfaces using an optical TIR-based texture sensor," in *Proc. IEEE Int. Conf. Robot. and Biomimetics*, pp. 1721–1726, IEEE, Guilin, China (2009).
28. A. Latif-Amet, A. Ertüzün, and A. Eric, "An efficient method for texture defect detection: sub-band domain co-occurrence matrices," *Image Vis. Comput.* **18**(6—7), 543–553 (2000).
29. V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York (2000).
30. J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal.* **8**(6), 679–714 (1986).
31. P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.* **57**(2), 137–154 (2004).

**Yongxiong Wang** received his BS degree from Harbin Engineering University in 1991 and MS degree from Shanghai JiaoTong University in 2006. He is currently a PhD candidate in Department of Automation, Shanghai JiaoTong University. His research interests include the intelligent robot and robot vision.



**Jianbo Su** received his BS degree in automatic control from Shanghai Jiao Tong University, Shanghai, China, in 1989, and MS degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1992, and PhD degree in 1995 from Southeast University, Nanjing, China. From 1995 to 1997, he was a postdoctoral research fellow at the Institute of Automation, Chinese Academy of Sciences. In 2000, he joined the Department of Automation, Shanghai Jiao Tong University, where he is currently a professor. He is an IEEE Senior Member, a member of the Technical Committee of Networked Robots, IEEE Robotics and Automation Society, and an associate editor of IEEE Transaction on System, Man, and Cybernetics (Part B). His research interests include intelligent robotics, network robots, multi-sensor data fusion, machine learning and human-machine interaction, and multi-robot coordination.