Accepted Manuscript

Infrared and Visible Images Registration with Adaptable Local-Global Feature Integration for Rail Inspection

Chaoqing Tang, Gui Yun Tian, Xiaotian Chen, Jianbo Wu, Kongjing Li, Hongying Meng

PII: S1350-4495(17)30281-5
DOI: https://doi.org/10.1016/j.infrared.2017.09.013
Reference: INFPHY 2384

To appear in: *Infrared Physics & Technology*

Received Date: 23 May 2017
Accepted Date: 16 September 2017

Please cite this article as: C. Tang, G.Y. Tian, X. Chen, J. Wu, K. Li, H. Meng, Infrared and Visible Images Registration with Adaptable Local-Global Feature Integration for Rail Inspection, *Infrared Physics & Technology* (2017), doi: https://doi.org/10.1016/j.infrared.2017.09.013

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Infrared and Visible Images Registration with Adaptable Local-Global Feature Integration for Rail Inspection

Chaoqing Tang[a], Gui Yun Tian[a,b,*], Xiaotian Chen[a], Jianbo Wu[c], Kongjing Li[a], Hongying Meng[d]

[a]*School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.*
[b]*School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*
[c]*School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065, China*
[d]*Department of Electronic and Computer Engineering, Brunel University London, Uxbridge UB8 3PH, U.K.*

Abstract

Active thermography provides infrared images that contain sub-surface defect information, while visible images only reveal surface information. Mapping infrared information to visible images offers more comprehensive visualization for decision-making in rail inspection. However, the common information for registration is limited due to different modalities in both local and global level. For example, rail track which has low temperature contrast reveals rich details in visible images, but turns blurry in the infrared counterparts. This paper proposes a registration algorithm called Edge-Guided Speeded-Up-Robust-Features (EG-SURF) to address this issue. Rather than sequentially integrating local and global information in matching stage which suffered from buckets effect, this algorithm adaptively integrates local and global information into a descriptor to gather more common information before matching. This adaptability consists of two facets, an adaptable weighting factor between local and global information, and an adaptable main direction accuracy. The local information is extracted using SURF while the global information is represented by shape context from

*Corresponding author
Email address: `g.y.tian@newcastle.ac.uk` (Gui Yun Tian)

Preprint submitted to INFRARED PHYSICS & TECHNOLOGY *September 13, 2017*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

edges. Meanwhile, in shape context generation process, edges are weighted according to local scale and decomposed into bins using a vector decomposition manner to provide more accurate descriptor. The proposed algorithm is qualitatively and quantitatively validated using eddy current pulsed thermography scene in the experiments. In comparison with other algorithms, better performance has been achieved.

Keywords: Rail inspection; infrared and visible image registration; local feature; global feature

---

1. Introduction

Rails are exposed to intense train loading and dynamic weather/geographic conditions, resulting in safety hazard like plastic deformation, wear, flaking, head checking, cracks, squats, corrugation etc. Some hazards also bring pollution to the environment. For example, rails with heavy corrugation not only cause train vibration but also make serious traffic noise pollution to the surroundings [1, 2]. Routine rail inspection is vital for rail operation. And for the benefit of both customers and railway companies, inspection techniques cannot bring serious disruption for transport line [3, 4]. To this end, inspection vehicles running on-line with on-board non-destructive testing and evaluation (NDT&E) techniques are cost-effective [5]. These techniques test the integrity and serviceability of rails based on different physical parameters. When equipped with cyber-enabled ability, these techniques play an important role in intelligent transport systems [5, 6].

Generally, the widely used NDT&E techniques in rail industry are from (a) visual, (b) electromagnetic, (c) vibration, and (d) thermal perspective. (a) From visual inspection perspective, on-board track inspection systems based on computer vision are widely applied [6, 7, 8, 9]. These visual inspections conduct fast assessment to the surface condition of rail tracks. Such visual inspection systems usually use high-speed cameras to search for broken clips in real-time with advanced pattern recognition algorithms. But visual inspection can only obtain

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

surface defect information. (b) From electromagnetic inspection perspective, eddy current inspection is one of several NDT methods that use the principal of electromagnetism. This method is sensitive to the surface and subsurface cracks, but the penetration depth is limited due to skin effect [10, 11]. Eddy currents are created through a process called electromagnetic induction. When alternating current is applied to conductors, such as copper coil, a magnetic field develops in and around the conductor. Other electrical conductors which locate in the proximity of this changing magnetic field will induce current. Variations in the electrical conductivity and magnetic permeability of the test object result in changes in eddy currents. Measuring these changes can test the presence of defects [12]. Besides eddy current, another method in this category which used in rail inspection is magnetic flux leakage (MFL) [10, 13]. Magnetic flux lines generated by a magnetizer are coupled into test objects simply by air. Any geometrical discontinuity or local anomalies are manifested as an abrupt change of magnetic permeability and force magnetic flux to leak out of the object in the poles of yoke in the air. Leakage magnetic field which contains information of defect can be collected by magnetic field sensors. For the advantages of its air coupling, MFL testing is suitable for automatic in-line and real time defect inspection. (c) Vibration-based techniques can test the surface and inner defect. They normally use piezoelectric transducers configured with different angles and positions to generate and receive vibration signals. The popular technique in this category is ultrasonic guided waves (UGW) [14, 15]. UGW employs mechanical stress waves that propagate along an elongated structure (e.g., rails and pipelines) while guided by its boundaries. The waves carry defect information when captured by sensors. This method can detect very long distance in a single test. Traditional ultrasonic inspection normally requires a coupling medium to promote the transfer of sound energy into the test specimen, which leads to low efficiency in rail inspection. Couplant free transduction can be achieved using electromagnetic acoustic transducers (EMATs) [16]. EMATs also make use of eddy current and the current is at the desired ultrasonic frequency. If a static magnetic field is present, these eddy currents will experience Lorentz

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

forces, thus eliminating couplant. But ultrasonic inspections are difficult to detect the defects in an acoustic shadow of rails and have low scan speed do to excitation complexity [17]. Another two techniques in rail inspection which do not need excitation are acoustic emission testing (AET) [17] and axle box acceleration (ABA) [18, 19]. Acoustic emission is transient elastic waves produced by a sudden redistribution of stress. Sources of AE vary from natural events like earthquakes and rock bursts to the initiation and growth of cracks, slip and dislocation movements etc. AE systems can only qualitatively gauge how much damage is contained in a structure [20]. ABA measures the vibrations of the wheel in the vehicle-track system, excited during the wheel-rail interaction, it can give an indication of an irregularity at the wheel-rail interface. ABA has the ability to measure the irregularities (usually short track defects) of the rail at line speeds. (d) From thermal inspection perspective, infrared/thermal NDT methods measures surface temperatures of rails. The use of thermal imaging systems allow thermal information to be very rapidly collected over a wide area and in a non-contact mode. This makes it promising in pantograph-catenary system inspection in high-speed train system [21]. Infrared information can be captured by bulky infrared cameras which offer high-resolution infrared images, or by small-size infrared sensors which connected to wireless sensor networks for long-term and remote monitoring [22, 23].

As discussed above, all these NDT techniques have its strength and weakness. To reduce the maintenance cost, integrating them together is necessary [4]. For example, EMATs integrates magnetic and ultrasonic testing to provide a non-contact solution. Eddy current pulsed thermography (ECPT), which integrates eddy current and thermal testing, can reveal surface and subsurface defect information quickly according to temperature distribution [10, 24]. To integrate the strength of visual inspection and thermal testing or even eddy current from the ECPT in rail inspection, this paper seeks for solution from computer vision perspective. Visible images from visual inspection and infrared images from thermal/ECPT testing are supposed to be fused.

Image fusion schemes can be classified into pixel-level, feature-level and

ACCEPTED MANUSCRIPT

decision-level fusion [25, 26]. The major difference among them is the sequence of information extraction. For pixel-level fusion, each image is combined in a pixel-by-pixel basis followed by the information extraction step. Many spare representation methods are in this category. A review of sparse representation-based multi-modality images fusion can be found in [27]. For feature-level fusion, the information is extracted separately from source images and then combined. As for decision-level fusion, the information is extracted from source images separately and then making decision on which content to be combined. This paper follows a feature-level fusion route. The general processing routines for feature-level methods are [28, 25] find features, extract features, match features, image transform and fusion. There are some widely used algorithms to find and extract features [29, 30], e.g., Harris corner detector, Scale-invariant feature transform (SIFT) [31], Speeded-Up-Robust-Features (SURF), and some improved version, e.g. GA-SIFT [32], among which SURF has faster speed than SIFT methods.

The above-mentioned methods extract features from pixels in local area (a sub-area around interest points), however, local information for infrared and visible images in active thermography rail inspection have obvious difference as illustrated in Figure 1(a) to Figure 1(d):

- Different contrast level. Thermal images only sensitive to temperature difference. The contrast is very low for small temperature difference. Unfortunately, this is the case for ECPT-based rail inspection. The contrast of visible image is much higher under sufficient illumination. This contrast difference results in little common interest point between them.
- Different intensity distribution. For example, the part with lowest intensity is area A for Figure 1(a), but it is area B for its counterpart. The difference in intensity distributions leads to different descriptors for corresponding interest point.
- Different defect visibility due to different modality. The enlarge part in Figure 1(d) shows an obvious scratch while the thermal counterpart not.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

This kind of complementary information is the motivation to fuse them, but brings difficulty for registration.

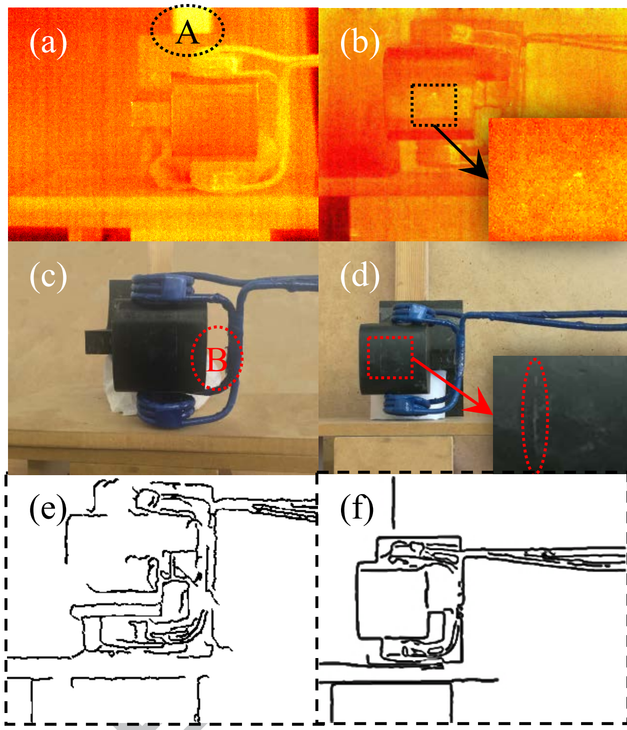

Figure 1: Rail track images. (a)∼(b) Infrared images; (c)∼(d) Visible images; (e) Canny edge for infrared image (b); (f) Canny edge for visible image (d). The infrared image is captured by a Flir SC650 infrared camera, and the visible image is captured by phone camera.

115 Given the highly polluted local information, consulting global information is an intuitive route. S. Raza et al. [33] use silhouette as global information for thermal and visible images registration of diseased plants, which is not applicable here because the low-contrast thermal images in rail inspection do not have robust silhouette. There are some literatures explore salient features [34, 35] for 120 fusion, but it is very difficult to find salient feature in rail inspection as illustrated in Figure 1. Large amount of literatures explore edges/lines as relatively robust global or local information. In consulting edges as local information, C. Aguilera et al. [36] propose EOH-SIFT which uses edge oriented histogram

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

(EOH) to characterize SIFT interest points. M. I. Patel et al. [37] use edge direction to characterize SURF interest points. In consulting edges as global information, J. Lpez et al. [38] use line segments in low-textured images registration. Shape context [39] is another category which making use of edge of the whole image rather than a sub-area around interest points. Descriptor is formed in a log-polar space which centres on each interest point. Y. Gu et al. [40] use shape context based on interest points that extracted using polynomial fitting. Only using edges as global information is suitable for scenes where there are robust edges, such as satellite remote sensing image where there are coastlines, rivers etc. There are no robust edges in rail inspection unfortunately. Figure 1(e) and Figure 1(f) are one set of Canny edges based on the preprocessed image of Figure 1(b) and Figure 1(d) respectively. These two edges images have much difference even extracting from preprocessed input images (The detail discussion of pre-process is given in section 3.1), which indicate that the performance of solo edge-based algorithms (no matter local or global) is poor in our case.

The shortcomings of solo local or global information based algorithms motivate us to consult methods that integrate local and global information. H. Jin et al. [41] propose a coarse-to-fine method for registration of multispectral images. Their method adopts SURF as local information and performs an initial matching in the first stage. In the second stage of matching in the reference, the whole edge image is divided into blocks. Each block of edges are represented using histogram of edge orientations. A similarity metric is used to refine the initial matching, a test point contributes most to the similarity metric if it has the same orientation as the edges. Y. Li et al. [42] also adopt a two stage matching process, an initial matching screens out some potential matching descriptor pairs, then followed by a second matching based on overlapped edge pixels. J. Han et al. [43] incorporate straight line as global feature and Harris corner detector as local information for infrared and visible image registration. All these three methods have shortcomings in our case. If divided entire edge image into small blocks, the edge in most blocks is different because the edge images have significant difference as shown in Figure 1(e) and Figure

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1(f). The local difference in edge image also brings trouble for overlapped edge pixels method and straight line method. This means edge image must take as a whole in our case, shape context can achieve such goal. Y. Gui et al. [44] propose a matching method using SURF and shape context. In their method, initial matching is based on SURF descriptor then shape context descriptor is used to refine the matching. However, the overall performance for all algorithms which sequentially integrate local and global information subjects to the matching quality in the first stage. Apparent difference in either local or global level will seriously decay the registration performance, which is a buckets effect.

This paper proposes an EG-SURF algorithm for infrared and visible image registration in rail inspection. Rather than sequentially integrating local and global information in matching stage which suffered from buckets effect, this algorithm integrates local and global information into a descriptor with adaptable weighting before matching. The adaptability consists of an adaptable weighting factor between local and global information and an adaptable main direction accuracy. The local information is extracted using SURF while the global information is represented by shape context from edges. Meanwhile, in shape context generation process, edges are weighted according to local scale and decomposed into bins using a vector decomposition manner to provide more accurate descriptor.

The rest of this paper is organized as follows: Section 2 gives a detail introduction to the proposed EG-SURF. The experimental validations are conducted in section 3 before the conclusions in the last section.

2. Proposed registration algorithm

This section proposes an algorithm called Edge-Guided Speeded-Up-Robust-Features (EG-SURF) to gather common information in local and global level. Figure 2 shows the overall diagram. This algorithm contains three major processes, i.e., SURF interest point detector, main direction assignment & global information extraction, descriptor construction. Compared to traditional SURF

ACCEPTED MANUSCRIPT

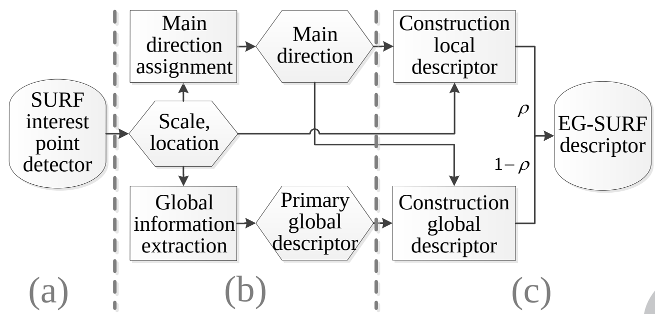

Figure 2: Diagram for EG-SURF. This algorithm contains three major process, i.e., (a) SURF interest point detector; (b) main direction assignment & global information extraction; (c) descriptor construction.

or shape context descriptor [40] or sequential combination [44], this algorithm combines local and global information into a descriptor before matching with adaptable weighting factor, which adapts to various inspection scene in rail inspection. The local information is extracted using SURF. Shape context from edges is used to represent global information, because edges and cracks in specimens show higher temperature contrast than other areas in active thermography. It is worth noting that the primary global descriptor is concurrently extracted with main direction assignment in our design, which improves efficiency. Meanwhile, the scale and location of SURF interest points are used to generate shape context descriptors, which improves the descriptor accuracy. Detail discussions are given in the following subsections.

2.1. SURF interest point detector

We use same interest point detector as SURF, which is based on the Hessian matrix. For a point (x, y) in an image I with size $a \times b$, the Hessian matrix $\mathrm{H}(\mathbf{x})$ in (x, y) at scale σ is:

$$\mathrm{H}(\mathbf{x}) = \begin{bmatrix} L_{xx}(\mathbf{x}) & L_{xy}(\mathbf{x}) \\ L_{xy}(\mathbf{x}) & L_{yy}(\mathbf{x}) \end{bmatrix} \tag{1}$$

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

where $\mathbf{x} = \begin{bmatrix} x & y & \sigma \end{bmatrix}^T$, $L_{xx}(\mathbf{x})$ is the convolution of the Gaussian second order derivative $\partial^2 g(0,\sigma)/\partial x \partial y$ with the image I in point (x, y), and similar for $L_{xy}(\mathbf{x})$ and $L_{yy}(\mathbf{x})$. Interest points are found by thresholding the determinant of Hessian matrix. To achieve scale invariant, SURF also making use of Laplacian of Gaussian (LoG) approximation D_{xx}, D_{xy}, D_{yy} with box filters. Thus, the determinant yields,

$$\det(H) \approx D_{xx} D_{yy} - (\alpha D_{xy})^2 \tag{2a}$$

$$\alpha = \frac{\|L_{xy}(1.2)\|_F \|D_{xx}(9)\|_F}{\|L_{xx}(1.2)\|_F \|D_{xy}(9)\|_F} \approx 0.9 \tag{2b}$$

where $\|\cdot\|_F$ is the Frobenius norm. The filter responses are normalized with respect to the mask size, $L \times L$, which is

$$L = 3\left(2^{o+1}(s+1)+1\right) \tag{3}$$

where $o, s \in \mathbb{N}^+$ are the octave number and layer number respectively. Even though α is dependent on the scale size σ (which relates to L and initial scale size (σ_0) as $\sigma = \sigma_0 L/9$), it turns out that in practice α can be approximated using a static constant of 0.9. Once located a set of potential interest points within each octave, 3D quadratic interpolation is necessary to get more accurate interest point using 2nd order Taylor series approximation:

$$\begin{aligned} \det(H)|_{\mathbf{x}} \approx \det(H)|_{\mathbf{x}_0} + \frac{\partial [\det(H)|_{\mathbf{x}}]^T}{\partial \mathbf{x}} \cdot \mathbf{x} \\ + \frac{1}{2}\mathbf{x}^T \cdot \frac{\partial^2 \det(H)|_{\mathbf{x}}}{\partial \mathbf{x}^2} \cdot \mathbf{x} \end{aligned} \tag{4}$$

The interpolation position and scale is obtained by differentiating the Eq. (4) and equating it to zero:

$$\mathbf{x}_{\max} = \left[\frac{\partial^2 \det(H)|_{\mathbf{x}}}{\partial \mathbf{x}^2}\right]^{-1} \left[\frac{\partial \det(H)|_{\mathbf{x}}}{\partial \mathbf{x}}\right] \tag{5}$$

By far, the location and size of each interest point are determined, we denote them as $\mathbf{p} = [\mathbf{x}\ \mathbf{y}]^{\mathbf{T}}$ and σ respectively.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

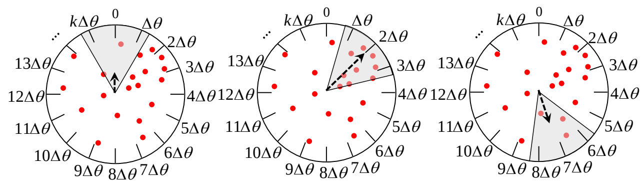

Figure 3: Graphical representation of main direction assignment. The rotation angle which has the greatest Haar response (denote as arrow) is the main direction of current interest point. Each red point stands for a unit of Haar response.

2.2. Main direction assignment & global information extraction

To ensure rotation invariant, every interest point are assigned a main direction, θ. This is achieved firstly by convolving pixels within radius 6σ in its neighborhood (denote as $N_{6\sigma}$) in x and y direction with Haar wavelet filters, F_H. The filter response of every pixels are weighted using Gaussian function with parameter 2σ according to their distance (denote as l) to the interest point position, $\mathbf{p}$. This process can be mathematically denoted as:

$$R_x = N_{6\sigma}^{(x)} * F_H \cdot G\left(l/6\sigma, 2\sigma\right) \tag{6a}$$

$$R_y = N_{6\sigma}^{(y)} * F_H \cdot G\left(l/6\sigma, 2\sigma\right) \tag{6b}$$

where $G\left(x, \lambda\right) = \exp\left(-x^2/2\lambda^2\right)$.

Then the weighted Haar responses are accumulated in a $\pi/3$ sector (denote as W_k) which rotates with certain step size, $\Delta\theta$:

$$m_{W_k} = \sum_{W_k} R_x + \sum_{W_k} R_y \tag{7}$$

where $2\pi/\Delta\theta \in \mathbb{N}^+$, $k \in \mathbb{N}$ and $k < 2\pi/\Delta\theta$. Figure 3 shows a graphical representation for Eq. (7). Those red spots stand for Haar responses, more spots covered by the sector means greater response and the longer of the arrow. The rotation angle which has the greatest amplitude is the main direction of current interest point:

$$\theta = k \cdot \Delta\theta \tag{8}$$

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

This equation indicates that the angle step size $\Delta\theta$ will influence the accuracy of main direction. Smaller step size leads to higher accuracy but improves computation complexity on the other hand. $\Delta\theta$ will also influence the length of final descriptor according to the shape context generation in the rest of this subsection.

Meanwhile, this paper introduce the shape context in main direction assignment stage to involve global information. The shape context is based on a binary edge image (denote as E) which should be prepared in advance. Let $r = \sqrt{a^2 + b^2}$, similar to the main direction assignment, a neighborhood on E with radius r (denote as N_r) is divided into bins in log-polar space. Note that this neighborhood is wider than the image size. The overflow area is regard as blank image. Then this neighborhood is weighted making use of the local feature size, because corresponding interest point pairs on thermal and visible images may have different size due to different scale. The edge within interest point size should have more weight. Originally, all edge pixels are regarded as 1 while non-edge pixels as 0. We recall Gaussian function and plot different λ values versus x in Figure 4. The function outputs decrease more rapidly with smaller λ. This property makes it suitable for our weighting requirement. So all edge in N_r are weighted using Gaussian function as shown in Eq. (9).

$$R_E = N_r \cdot G\left(l/r, 6\sigma\right) \tag{9}$$

The weighted edges (denote as R_E) are then used to calculate histogram as shown in Figure 5. Same step size as Figure 3 is used for histogram bins and $C \in \mathbb{N}^+$circles are used. The red spots here stand for edge pixels on E. All R_E are decomposed to its nearby bins in a vector decomposition manner as shown in the right side of Figure 5. This vector decomposition manner provides more accurate and robust shape context information than non-decompose methods, because the non-decompose method [40, 39] simply accounts the number of edge pixels in a bin. Finally, a $K = C \times 2\pi/\Delta\theta$ dimension vector $\mathbf{g}'$ (primary global descriptor) is obtained which contains global shape information.

ACCEPTED MANUSCRIPT

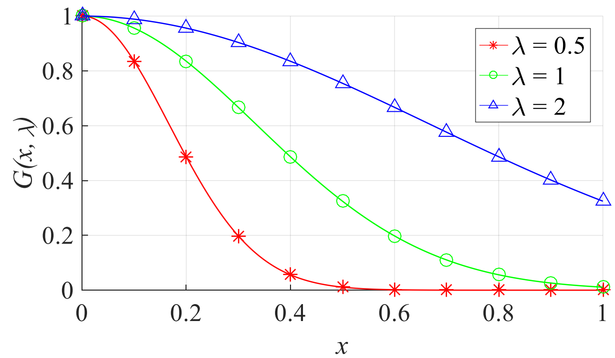

Figure 4: Gaussian function under different λ. The Gaussian function is defined as $G(x,\lambda) = \exp\left(-x^2/2\lambda^2\right)$. The scale of local information is used as the input to weight edges as Eq.(9).

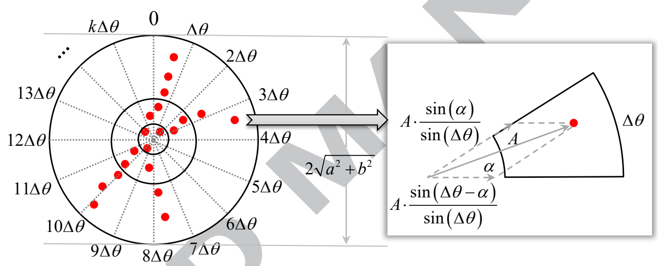

Figure 5: Graphical representation of shape context descriptor. All edges pixels (denote as red points) are decomposed into nearby bins using a vector decomposition manner as shown on the right side, where A denotes a weighted edge pixel in R_E.

It is worth noting that same angle step size as main direction assignment is adopted, so the histogram creation of these two processes can perform concurrently in implementation.

2.3. Descriptor construction

A primary global descriptor and a main direction are obtained for each interest point in previous subsection. These information are used to construct local descriptor and eliminate rotation influence for the primary global descriptor. Ultimately, the final descriptor is formed by proposing a weighting factor between the global descriptor and local descriptor. The details of these processes

ACCEPTED MANUSCRIPT

are introduced hereunder.

Based on a interest point location $\mathbf{p}$, a neighborhood on I with radius 20σ (denote as $N_{20\sigma}$) is evenly divided into 4×4 subareas (denote as N_S). Then, filtering every subarea uses Haar wavelet filter and weighting with Gaussian function in x and y direction:

$$R_S^{(x)} = N_S^{(x)} * F_H \cdot G\left(l/20\sigma, 2\sigma\right) \tag{10a}$$

$$R_S^{(y)} = N_S^{(y)} * F_H \cdot G\left(l/20\sigma, 2\sigma\right) \tag{10b}$$

The filter responses are rotated according to θ:

$$R'^{(x)}_S = -R_S^{(x)} \cdot \sin\left(\theta\right) + R_S^{(y)} \cdot \cos\left(\theta\right) \tag{11a}$$

$$R'^{(y)}_S = R_S^{(x)} \cdot \cos\left(\theta\right) + R_S^{(y)} \cdot \sin\left(\theta\right) \tag{11b}$$

Subsequently, counting $\sum R'^{(x)}_S$, $\sum \left|R'^{(x)}_S\right|$, $\sum R'^{(y)}_S$, $\sum \left|R'^{(y)}_S\right|$ obtains a $4 \times 4 \times 4 = 64$ vector to form the local information part descriptor (denote as $\mathbf{l}$).
230 For the global part (denote as $\mathbf{g}$), the only step is to shift $\mathbf{g}'$ cyclically for k times in order to eliminate rotation influence. We denote $\mathbf{l}$ and $\mathbf{g}$ for all interest points as $\mathbf{L}$ and $\mathbf{G}$, and all normalized to $0{\sim}1$.

Finally, concatenating $\mathbf{l}$ and $\mathbf{g}$ together to get a $K + 64$ dimension vector as the EG-SURF descriptor $\mathbf{d}$ using Eq. (12),

$$\mathbf{d} = \left[\begin{array}{cc} \rho\mathbf{l} & \left(1-\rho\right)\xi\mathbf{g} \end{array}\right] \tag{12}$$

where $0 \leq \rho \leq 1$ is a weighting factor used to adjust the weight of local information and global information. When $\rho = 1$, EG-SURF descriptor degrades to ordinary SURF descriptor. Likewise, only global information is used when $\rho = 0$. ξ is a normalization factor to guarantee same weight between $\mathbf{l}$ and $\mathbf{g}$, i.e. $E\left[\|\mathbf{l}\|_F\right] = E\left[\|\mathbf{g}\|_F\right]$, where $E\left[\cdot\right]$ is the mathematical expectation. One can empirically let $\xi = \|\mathbf{L}\|_F \Big/ \left(\sqrt{card\left(\mathbf{G}\right)}\|\mathbf{G}\|_F\right)$, where $card\left(\mathbf{G}\right)$ is the number of

ACCEPTED MANUSCRIPT

elements in $\mathbf{G}$. Given various of application scene, we assume each element of $\mathbf{l}$ (denote as L_i, $i \in \mathbb{N}^+$ and $i \leq 64$) and $\mathbf{g}$ (denote as G_j, $j \in \mathbb{N}^+$ and $j \leq K$) obey independent and identically distributed uniform distribution, L_i, $\mathrm{G}_j \sim \mathrm{U}(0,1)$, then we have:

$$\begin{aligned} E\left[\sum_{i=1}^{64} L_i^2\right] &= E\left[\sum_{i=1}^{K} (\xi G_i)^2\right] \\ &= \xi^2 E\left[\sum_{i=1}^{K} G_i^2\right] \end{aligned} \tag{13}$$

ξ can be easily obtained as $\xi = 8\big/\sqrt{K}$, which only decided by the length of $\mathbf{l}$ and $\mathbf{g}$, thus eliminating complex calculation.

As discussed in this section, a weighting factor ρ is proposed to make the weighting of local and global information adaptable. The solo SURF and shape context descriptors are given as two special cases in this algorithm. Besides this weighting factor, the angle step size $\Delta\theta$ and the circle number C are two critical parameters in this algorithm. The influence of the three parameters for fusion performance will be discussed in more details in the next section.

3. Experimental validation

Rail track health inspection is an important field in rail inspection industry, among which ECPT demonstrates good ability to reveal subsurface information. We build up a scene for ECPT as shown in Figure 6(b). In ECPT, coil is necessary to excite the specimen under test. So we place a coil near the rail track specimen but without imposing current. This is the initial setup for ECPT. Infrared images are captured by a Flir SC650 infrared camera with resolution of 640×480. RGB/visible images are captured by a phone which configured to the same resolution as the infrared camera. The infrared camera and RGB camera take photos from similar view angle but have different scale. More than 20 sets of images with resolution 640×480 are obtained, each set contains an infrared image and a corresponding visible image. Two of them are shown in

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure 1. The set which contains Figure 1(b) and Figure 1(d) is used to qualitatively validate the proposed algorithm, because this set is a representative one. Quantitative validation is based on the average performance of total datasets.

3.1. Qualitative validation

The thermal images are usually very blurry and have low contrast, which cause problem to edge extraction. The root mean square (RMS) contrast, defines as $\sqrt{\sum (I - \mathrm{E}[I])^2 \big/ ab}$ and $0 \leq I \leq 1$, is -74.3 dB for the thermal image, which is much lower than its visible counterparts with -33.7 dB. So, preprocessing is necessary. We adopt the Perona-Malik anisotropic diffusion [45] for the thermal image thus improving the RMS contrast to -27.1 dB. The processed result in Figure 6(c) shows higher contrast and better smooth compare to the raw image in Figure 6(a). The visible images are pre-processed by edge-preserving smoothing to remove local difference with thermal images as much as possible.

Then, Canny edge detection is used to extract the edge of both images because its superior performance [46]. The results are shown in Figure 1(e) and Figure 1(f). It is obvious that the two edge images have much difference, which brings challenge to solo edge-based registration methods.

The above preparations are used to validate the proposed EG-SURF algorithm. 38 and 44 interest points are found by thresholding the determinant of Hessian matrix for the processed infrared and visible images respectively as shown in Figure 7. Relaxing the thresholding obtains more interest points on the background than the track. This goes against the needs of rail track health diagnostics, because the inspection cameras are supposed to focus on tracks rather than the background. We set $\Delta\theta = 22.5°$ and $C = 4$, then all descriptors are in 128 dimension and the normalization factor $\xi = 1$. The weighting factor is set to be 0.5, thus the descriptors have same weight of local and global information. The descriptor numbers are 38 and 44 for infrared and visible image respectively.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

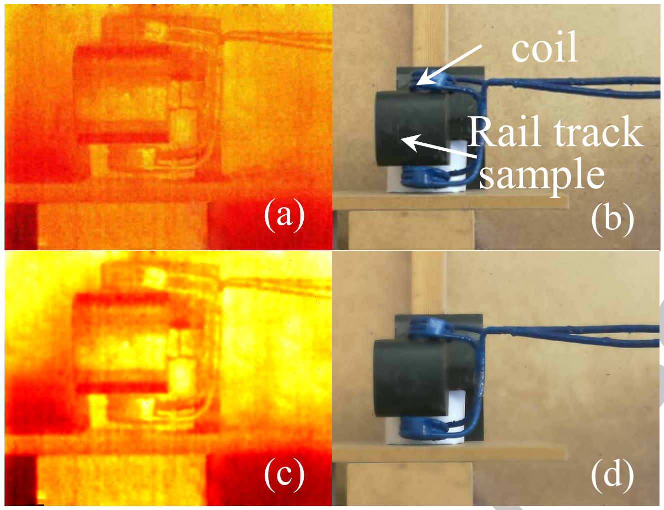

Figure 6: Preprocessed images of ECPT scene for (a) infrared image and (b) visible image. (c) Perona-Malik anisotropic diffusion for infrared image; (d) Edge-preserving smoothing for visible image.

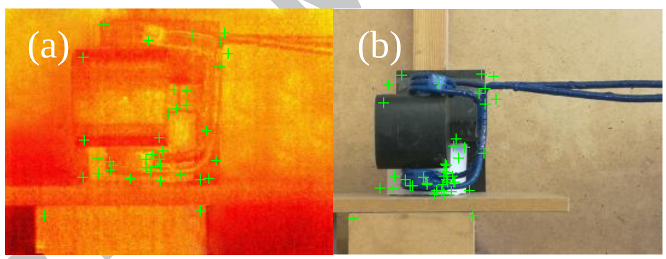

Figure 7: Interest points. (a) Infrared image; (b) Visible image.

Based on the descriptors, the Euclidean distance between every pair of descriptor from different images is calculated. Thresholding these distances obtains potential matching pairs. To ensure unique matching, only the pair with shortest distance for same interest point is reserved. The registration results are shown in Figure 8(a). 11 out of 12 pairs are correctly matched, which is more than 4 pairs in order to apply affine transformation to fuse them. Incorrect matches are common for all registration algorithms, one can adjust the matching parameters to reduce wrong matches. In contrast, random sample

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

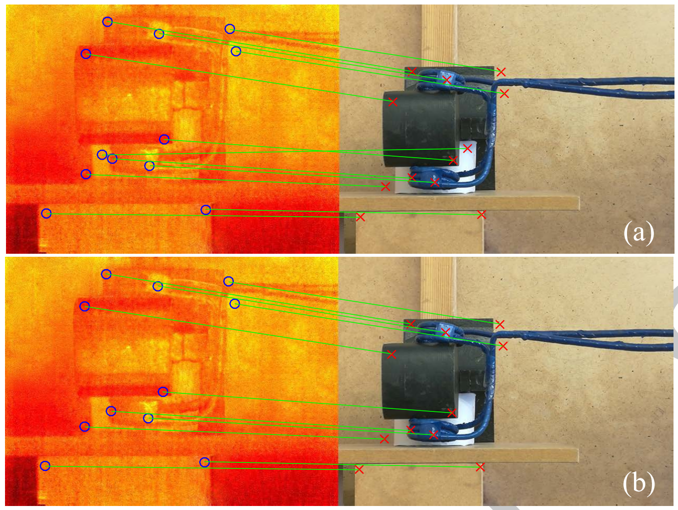

Figure 8: Registration results using (a) the proposed EG-SURF and (b) after additional RANSAC with $\Delta\theta = 22.5^\circ$, $C = 4$, $\rho = 0.5$. Our algorithm shows dominant percentage of correct match for infrared and visible rail track health images, the minor wrong match can be removed by RANSAC.

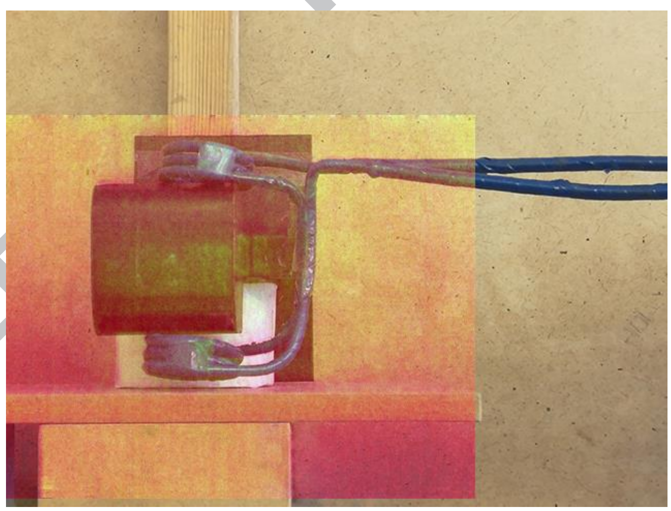

Figure 9: Fused image.

consensus (RANSAC) is a powerful technique to refine the matching without loss of generality [47, 48, 49]. In this technique, an initial affine transform ma-

ACCEPTED MANUSCRIPT

trix is formed by randomly choose several matched pairs. Then other matched pairs which fitting with the matrix are used to iteratively refine it. However, this technique is on the foundation that most of matched pairs are correct. Our proposed algorithm achieves this as shown in Figure 8(a). The matching results after RANSAC in Figure 8(b) already remove the wrong match pair. The fused image using this match results is shown in Figure 9, both visible and thermal information are presented in it. These results qualitatively validate the proposed algorithm.

3.2. Quantitative evaluation

To quantitatively evaluate the algorithm, the widely used *precision* [32, 50] is chosen as evaluation metric. We briefly introduce the definition here. Suppose a registration algorithm matches u pairs for two set of descriptors. If w pairs in u are correct, but $u - w$ pairs are actually wrong, the algorithm's *precision* is [51],

$$precision = w/u \tag{14}$$

The *precision* indicates how useful the search results are. In image registration, if *precision* is high enough, there is high probability to obtain correct transform matrix using RANSAC if the matching number is greater than 4, which already qualitatively illustrated in Figure 8. Note that all the following quantitative validations are based on the average results of datasets. Because different image set has different level of common local/global information, which leads to different optimal parameter settings. Automatically optimize these parameters is part of our future work.

Based on the above definition, the *precision* vs. distance threshold under different ρ is shown in Figure 10. Only *precision* when matching number greater than 4 are reserved, this is why there are sharp jumps in low distance threshold region in Figure 10 and Figure 13. Different ρ shows different precision, but all converges to a constant with distance threshold increasing, because the potential match pairs for two sets of descriptors are constant. As discussed previously, ρ is weighting factor between local information and global information. The

ACCEPTED MANUSCRIPT

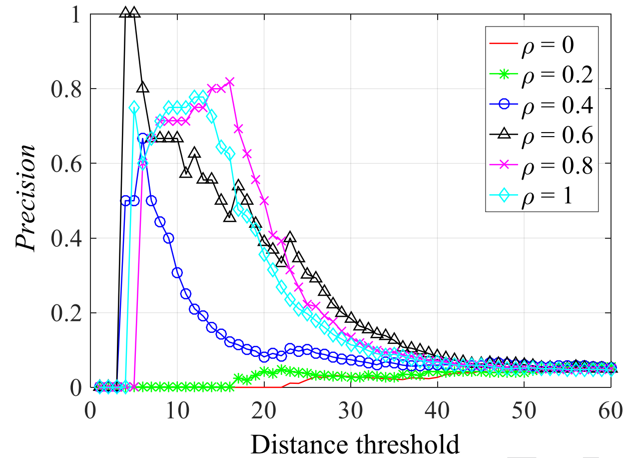

Figure 10: *Precision* vs. distance threshold under different ρ with $\Delta\theta = 22.5°$, $C = 4$.

curves under low ρ have the worst *precision*, which means relying too much on global information cannot get good matching results. Because the edges for thermal and visible rail track health images are so different. The high ρ region curves show better performance comparing with their low ρ counterparts, which means that the local information is a relatively reliable feature in rail track health images.

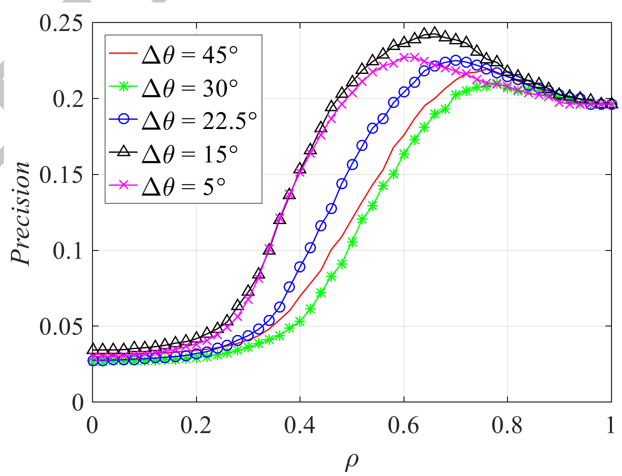

Figure 11: *Precision* vs. ρ under different $\Delta\theta$ with $C = 4$.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

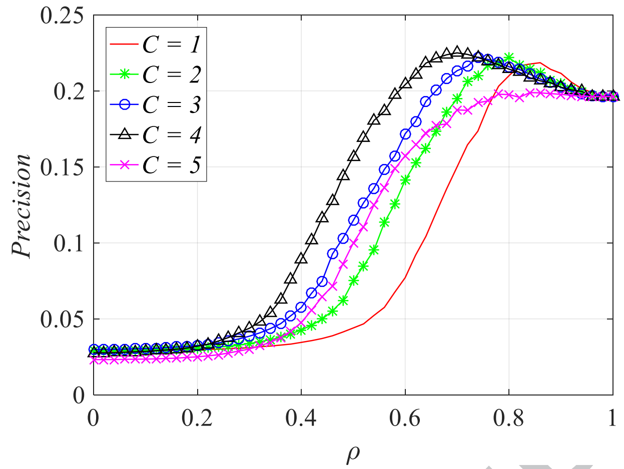

Figure 12: *Precision* vs. ρ under different C with $\Delta\theta = 22.5°$.

In order to further analyzing the influence of ρ and also investigating the influence of $\Delta\theta$, the average *precision* vs. ρ under different $\Delta\theta$ are plotted in Figure 11. Curves with different $\Delta\theta$ all start from small *precisions* in low ρ region and end at a constant, which means only rely on global information will lead to poor results. There is a constant at $\rho = 1$, because only local information is used, different $\Delta\theta$ do not have any influence. It is also obvious that every curve have a peak at certain weighting factor, which means properly combining local and global information outperforming solo local or global based methods. For thermal and visible rail track health images in ECPT, this weighting factor distributes between 0.6∼0.7 under 4 circles in log-polar space.

To investigate the influence of circle number, the average *precision* vs. ρ under different C is shown in Figure 12. These curves show similar trend as Figure 11, anther observation is that the peak shifting towards larger ρ for smaller C generally. Because smaller C has less bins leading to more global information in some sense, more weighting on local information will balance this effect.

The above discussions quantitatively show the influence of parameter ρ, C

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

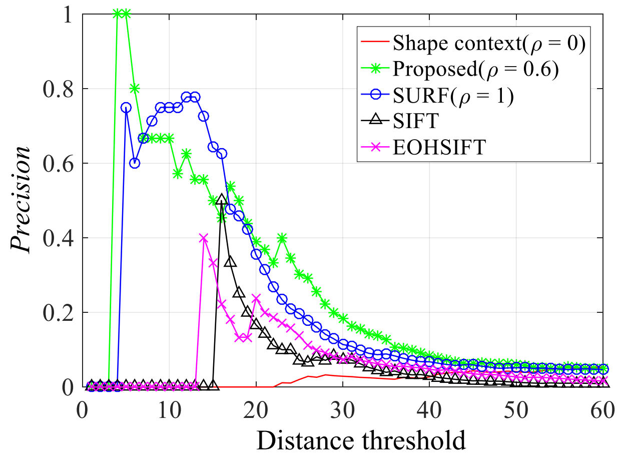

Figure 13: *Precision* vs. distance threshold for different algorithms.

and $\Delta\theta$. Trade-offs are necessary in practical application. Smaller $\Delta\theta$ and C increases both descriptor generation time and registration time, which increases in-situ inspection time as a results. If the computation ability of rail inspection system is powerful enough, this is not a vital issue. Besides, Smaller $\Delta\theta$ and C not necessarily leads to better precision as already discussed. Automatic optimization of all these parameters is part of our future work.

To further validate the performance of proposed algorithm, a horizontal comparison between different registration methods is plotted in Figure 13. Only *precision* when matching number greater than 4 are reserved. Same amount of interest points are obtained for different methods. The proposed method, which includes SURF and shape context as special case, shows best precision curves. The *precision* of SIFT, shape context and EOHSIFT all drops below 0.5 for all distance threshold, which means these methods almost unacceptable in registration of cross-spectrum rail track health images.

4. Conclusions & Future works

This paper proposes a registration algorithm called Edge-Guided SURF (EG-SURF) about feature-level fusion of infrared and visible images for rail

ACCEPTED MANUSCRIPT

inspection. Rather than sequentially integrating local and global information in matching stage which suffered from buckets effect, this algorithm adjustably integrates local and global information into a descriptor to gather more common information before matching. This adaptability reflected in an adaptable weighting factor between local and global information and an adaptable main direction accuracy. The local information is extracted using SURF while the global information is represented by shape context from edges. Meanwhile, in shape context generation process, edges are weighted according to local scale and decomposed into bins using a vector decomposition manner to provide more accurate descriptor. During the main direction assignment, a primary global descriptor is formed concurrently. This primary global descriptor only needs to cyclically shift for direction adjustment. This character makes the algorithms easy to implement. The experimental results using infrared and visible images of normal temperature rail tracks illustrate better performance than other state-of-the art algorithms.

Furthermore, this work paves the way for a 3D fusion model [52] of infrared and visible/RGB-D images. In this model, 2D thermal images need to correspond to 2D visible images which project from RGB-D image. After this process, every pixel in the 2D visible images have both RGB-D and temperature information. Thus 2D temperature distribution could be mapped to corresponding position in 3D model. This paper solves the 2D registration process, which laid the foundation for future investigation.

It should be noted that current weighting factor between local and global information (ρ), circle number of shape context (C), and angle step size ($\Delta\theta$) in main direction assignment are empirically set, which is a limitation of this method. Self-adaption ability of these parameter settings which fully consider the trade-off between fusion accuracy and in-situ inspection efficiency is part of our future work. We also plan to extract defects such as crack depth/width with 3D visualization.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

5. Acknowledgments

This work is supported by Engineering and Physical Sciences Research Council (EPSRC) [grant numbers EP/K503885/1]. Chaoqing Tang would also like to thank China Scholarship Council for funding his Ph.D study.

References

[1] X. Cui, G. Chen, J. Zhao, W. Yan, H. Ouyang, M. Zhu, Field investigation and numerical study of the rail corrugation caused by frictional self-excited vibration, Wear 376 (2017) 1919–1929. doi:http://dx.doi.org/10.1016/j.wear.2017.01.089.

[2] Z. Popovi, L. Lazarevi, L. Brajovi, M. Vilotijevi, The importance of rail inspections in the urban area-aspect of head checking rail defects, Procedia Engineering 117 (2015) 596–608.

[3] K. Odolinski, A. S. Smith, Assessing the cost impact of competitive tendering in rail infrastructure maintenance services: evidence from the swedish reforms (1999 to 2011), Journal of Transport Economics and Policy (JTEP) 50 (1) (2016) 93–112.

[4] M. H. B. Osman, S. Kaewunruen, A. Jack, J. Sussman, Need and opportunities for a 'Plan B' in rail track inspection schedules, Procedia Engineering 161 (2016) 264–268.

[5] F. Franklin, F. Nemtanu, P. F. Teixeira, Rail infrastructure, its and access charges, Research in Transportation Economics 41 (1) (2013) 31–42.

[6] Q. Li, Z. Shi, H. Zhang, Y. Tan, S. Ren, P. Dai, W. Li, A cyber-enabled visual inspection system for rail corrugation, Future Generation Computer Systems , In press.

[7] H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi, L. Chen, Automatic fastener classification and defect detection in vision-based railway inspection systems,

ACCEPTED MANUSCRIPT

IEEE Transactions on Instrumentation and Measurement 63 (4) (2014) 877–888. doi:10.1109/TIM.2013.2283741.

[8] S. Hajizadeh, A. Nnez, D. M. Tax, Semi-supervised rail defect detection from imbalanced image data, IFAC-PapersOnLine 49 (3) (2016) 78–83.

[9] E. Resendiz, J. M. Hart, N. Ahuja, Automated visual inspection of railroad tracks, IEEE Transactions on Intelligent Transportation Systems 14 (2) (2013) 751–760. doi:10.1109/TITS.2012.2236555.

[10] Y. Gao, G. Y. Tian, K. Li, J. Ji, P. Wang, H. Wang, Multiple cracks detection and visualization using magnetic flux leakage and eddy current pulsed thermography, Sensors and Actuators A: Physical 234 (2015) 269–281. doi:10.1016/j.sna.2015.09.011.

[11] X. Lu, Y. Li, M. Wu, J. Zuo, W. Hu, Rail temperature rise characteristics caused by linear eddy current brake of high-speed train, Journal of Traffic and Transportation Engineering (English Edition) 1 (6) (2014) 448–456. doi:http://dx.doi.org/10.1016/S2095-7564(15)30295-6.

[12] J. Garca-Martn, J. Gmez-Gil, E. Vzquez-Snchez, Non-destructive techniques based on eddy current testing, Sensors 11 (3) (2011) 2525–2565.

[13] P. Wang, L. Xiong, Y. Sun, H. Wang, G. Tian, Features extraction of sensor array based PMFL technology for detection of rail cracks, Measurement 47 (2014) 613–626. doi:http://dx.doi.org/10.1016/j.measurement.2013.09.047.

[14] S. Coccia, I. Bartoli, A. Marzani, F. Lanza di Scalea, S. Salamone, M. Fateh, Numerical and experimental study of guided waves for detection of defects in the rail head, NDT & E International 44 (1) (2011) 93–100. doi:http://dx.doi.org/10.1016/j.ndteint.2010.09.011.

[15] C. m. lee, J. L. Rose, Y. Cho, A guided wave approach to defect detection under shelling in rail, NDT & E International 42 (3) (2009) 174–180. doi:http://dx.doi.org/10.1016/j.ndteint.2008.09.013.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[16] P. A. Petcher, M. D. G. Potter, S. Dixon, A new electromagnetic acoustic transducer (EMAT) design for operation on rail, NDT & E International 65 (2014) 1–7. doi:http://dx.doi.org/10.1016/j.ndteint.2014.03.007.

[17] X. Zhang, N. Feng, Y. Wang, Y. Shen, An analysis of the simulated acoustic emission sources with different propagation distances, types and depths for rail defect detection, Applied Acoustics 86 (2014) 80–88. doi:http://dx.doi.org/10.1016/j.apacoust.2014.06.004.

[18] M. Molodova, Z. Li, A. Nez, R. Dollevoet, Validation of a finite element model for axle box acceleration at squats in the high frequency range, Computers & Structures 141 (2014) 84–93. doi:http://dx.doi.org/10.1016/j.compstruc.2014.05.005.

[19] M. Molodova, Z. Li, R. Dollevoet, Axle box acceleration: Measurement and simulation for detection of short track defects, Wear 271 (1) (2011) 349–356. doi:http://dx.doi.org/10.1016/j.wear.2010.10.003.

[20] W. Caesarendra, B. Kosasih, A. K. Tieu, H. Zhu, C. A. S. Moodie, Q. Zhu, Acoustic emission-based condition monitoring methods: Review and application for low speed slew bearing, Mechanical Systems and Signal Processing 72 (2016) 134–159. doi:http://dx.doi.org/10.1016/j.ymssp.2015.10.020.

[21] S. Gao, Y. Wang, Z. Liu, X. Mu, K. Huang, X. Song, Thermal distribution modeling and experimental verification of contact wire considering the lifting or dropping pantograph in electrified railway, IEEE Transactions on Transportation Electrification 2 (2) (2016) 256–265. doi:10.1109/TTE.2016.2558841.

[22] V. J. Hodge, S. O'Keefe, M. Weeks, A. Moulds, Wireless sensor networks for condition monitoring in the railway industry: A survey, IEEE Transactions on Intelligent Transportation Systems 16 (3) (2015) 1088–1106.

ACCEPTED MANUSCRIPT

[23] A. Soni, S. Robson, B. Gleeson, Structural monitoring for the rail industry using conventional survey, laser scanning and photogrammetry, Applied Geomatics 7 (2) (2015) 123–138. doi:10.1007/s12518-015-0156-1. URL http://dx.doi.org/10.1007/s12518-015-0156-1

[24] X. Li, B. Gao, W. L. Woo, G. Y. Tian, X. Qiu, L. Gu, Quantitative surface crack evaluation based on eddy current pulsed thermography, IEEE Sensors Journal 17 (2) (2017) 412–421. doi:10.1109/JSEN.2016.2625815.

[25] Y. Zhou, A. Mayyas, A. Qattawi, M. Omar, Feature-level and pixel-level fusion routines when coupled to infrared night-vision tracking scheme, Infrared Physics & Technology 53 (1) (2010) 43–49. doi:https://doi.org/10.1016/j.infrared.2009.08.011.

[26] X. Jin, Q. Jiang, S. Yao, D. Zhou, R. Nie, J. Hai, K. He, A survey of infrared and visual image fusion methods, Infrared Physics & Technology 85 (2017) 478–501. doi:http://dx.doi.org/10.1016/j.infrared.2017.07.010.

[27] Q. Zhang, Y. Liu, R. S. Blum, J. Han, D. Tao, Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review, Information Fusion 40 (2017) 57–75.

[28] C. A. Corneanu, M. O. Simn, J. F. Cohn, S. E. Guerrero, Survey on RGB, 3D, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (8) (2016) 1548–1568. doi:10.1109/TPAMI.2016.2515606.

[29] R. A. Guler, S. Tari, G. Unal, Landmarks inside the shape: Shape matching using image descriptors, Pattern Recognition 49 (2016) 79–88. doi:http://dx.doi.org/10.1016/j.patcog.2015.07.013.

[30] P. Loncomilla, J. Ruiz-del Solar, L. Martnez, Object recognition using local invariant features for robotic applications: A survey, Pattern Recognition

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

60 (2016) 499–514. doi:http://dx.doi.org/10.1016/j.patcog.2016.05.021.

[31] J. Dou, Q. Qin, Z. Tu, X. Peng, Y. Li, Infrared and visible image registration based on SIFT and sparse representation, in: Control and Decision Conference (CCDC), 2016 Chinese, IEEE, pp. 5420–5424.

[32] Y. Li, W. Liu, X. Li, Q. Huang, X. Li, GA-SIFT: A new scale invariant feature transform for multispectral image using geometric algebra, Information Sciences 281 (2014) 559–572. doi:http://dx.doi.org/10.1016/j.ins.2013.12.022.

[33] S. Raza, V. Sanchez, G. Prince, J. P. Clarkson, N. M. Rajpoot, Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain, Pattern Recognition 48 (7) (2015) 2119–2128. doi:http://dx.doi.org/10.1016/j.patcog.2015.01.027.

[34] J. Han, E. J. Pauwels, P. De Zeeuw, Fast saliency-aware multi-modality image fusion, Neurocomputing 111 (2013) 70–80.

[35] F. Wu, B. Wang, X. Yi, M. Li, J. Hao, H. Qin, H. Zhou, Visible and infrared image registration based on visual salient features, Journal of Electronic Imaging 24 (5) (2015) 053017–053017.

[36] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, R. Toledo, Multispectral image feature points, Sensors 12 (9) (2012) 12661–12672.

[37] M. I. Patel, V. K. Thakar, S. K. Shah, Image registration of satellite images with varying illumination level using HOG descriptor based SURF, Procedia Computer Science 93 (2016) 382–388.

[38] J. Lpez, R. Santos, X. R. Fdez-Vidal, X. M. Pardo, Two-view line matching algorithm based on context and appearance in low-textured images, Pattern Recognition 48 (7) (2015) 2164–2184. doi:http://dx.doi.org/10.1016/j.patcog.2014.11.018.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[39] L. Zhao, Q. Peng, B. Huang, Shape matching algorithm based on shape contexts, IET Computer Vision 9 (5) (2015) 681–690. doi:10.1049/iet-cvi.2014.0159.

[40] Y. Gu, K. Ren, P. Wang, G. Gu, Polynomial fitting-based shape matching algorithm for multi-sensors remote sensing images, Infrared Physics & Technology 76 (2016) 386–392.

[41] H. Jin, C. Fan, Y. Li, L. Xu, A novel coarse-to-fine method for registration of multispectral images, Infrared Physics & Technology 77 (2016) 219–225.

[42] Y. Li, J. Zou, J. Jing, H. Jin, H. Yu, Establish keypoint matches on multispectral images utilizing descriptor and global information over entire image, Infrared Physics & Technology 76 (2016) 1–10.

[43] J. Han, E. J. Pauwels, P. de Zeeuw, Visible and infrared image registration in man-made environments employing hybrid visual features, Pattern Recognition Letters 34 (1) (2013) 42–51. doi:http://dx.doi.org/10.1016/j.patrec.2012.03.022.

[44] Y. Gui, A. Su, J. Du, Point-pattern matching method using SURF and shape context, Optik-International Journal for Light and Electron Optics 124 (14) (2013) 1869–1873.

[45] N. Osia, T. Bourlai, A spectral independent approach for physiological and geometric based face recognition in the visible, middle-wave and long-wave infrared bands, Image and Vision Computing 32 (11) (2014) 847–859.

[46] Q. Xu, S. Varadarajan, C. Chakrabarti, L. J. Karam, A distributed canny edge detector: Algorithm and FPGA implementation, IEEE Transactions on Image Processing 23 (7) (2014) 2944–2960. doi:10.1109/TIP.2014.2311656.

[47] Z. Hossein-Nejad, M. Nasri, An adaptive image registration method based on SIFT features and RANSAC transform, Computers & Electrical En-

ACCEPTED MANUSCRIPT

gineering , In press. doi:http://dx.doi.org/10.1016/j.compeleceng.2016.11.034.

[48] M. Saval-Calvo, J. Azorin-Lopez, A. Fuster-Guillo, J. Garcia-Rodriguez, Three-dimensional planar model estimation using multi-constraint knowledge based on k-means and ransac, Applied Soft Computing 34 (2015) 572–586. doi:http://dx.doi.org/10.1016/j.asoc.2015.05.007.

[49] X. Wu, Q. Zhao, W. Bu, A SIFT-based contactless palmprint verification approach using iterative ransac and local palmprint descriptors, Pattern Recognition 47 (10) (2014) 3314–3326.

[50] Y. Ma, J. Han, J. Liu, Y. Zhang, L. Bai, A multi-scaled hierarchical structure model for multispectral image detection, Signal Processing: Image Communication 47 (2016) 193–206. doi:http://dx.doi.org/10.1016/j.image.2016.06.010.

[51] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (10) (2005) 1615–1630. doi:10.1109/TPAMI.2005.188.

[52] D. Gonzlez-Aguilera, P. Rodriguez-Gonzalvez, J. Armesto, S. Lagela, Novel approach to 3D thermography and energy efficiency evaluation, Energy and Buildings 54 (2012) 436–443. doi:http://dx.doi.org/10.1016/j.enbuild.2012.07.023.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights

1) Registration and fusion of an infrared and visible image registration algorithm for rail defect inspection.
2) Local and global information are integrated into a descriptor.
3) Adaptable weighting between local and global information.

Edge-Guided Speeded-Up-Robust-Features

4) (EG-SURF) to address this issue.
5) Local information is used to guide global information extraction.
6) The experimental results illustrate better performance than other algorithms.

ACCEPTED MANUSCRIPT