AN IMAGE-MATCHING METHOD 
BASED ON THE CURVATURE OF COST CURVE 
FOR PRODUCING TUNNEL LINING PANORAMA

Cuong Nguyen KIM¹, Kei KAWAMURA², Masando SHIOZAKI³ and Amir TARIGHAT⁴

1Ph. D. Candidate, Graduate School of Science & Engineering, Yamaguchi University 
(Tokiwadai, Ube, Yamaguchi 755-8611, Japan) 
E-mail: v004wf@yamaguchi-u.ac.jp
2Member of JSCE, Assoc. Prof., Graduate School of Science & Technology for Innovation, Yamaguchi University 
(Tokiwadai, Ube, Yamaguchi 755-8611, Japan) 
E-mail: kay@yamaguchi-u.ac.jp
3Member of JSCE, Sumitomo Mitsui Construction Co., Ltd., Tokyo, Japan 
(Tsukuda, Chuo-ku, Tokyo 104-0051, Japan) 
E-mail: MasandoShiozaki@smcon.co.jp
4Assoc. Prof., Dept. of Civil Engineering, Shahid Rajaee Teacher Training University 
(Lavizan, Shabanloo St., Tehran 16788-15811, Iran) 
E-mail: tarighat@srttu.edu

Human inspection of defects of concrete tunnel lining is usually slow, laborious, and disruptive to traffic. This necessitates automated alternatives using sensors and computer-aided processing. The conventional image-matching methods only use the cost value of the pixel being processed based on similarity metric to estimate an image-matching location. To improve the image-matching efficiency, this paper proposes an image-matching method that relies on the curvatures of the cost curve at candidate matching points. This is followed by applying a median filter to mitigate the matching errors. Moreover, experimental results for an actual tunnel demonstrate that the curvature measurement can select the matching points accurately. The authors have developed an image-stitching software to create a high-resolution panorama of the tunnel lining surface for assisting in defect inspection.

Key Words: curvature metric, image matching, tunnel lining inspection, panorama

1. INTRODUCTION

Many transport infrastructure structures, including buildings, bridges, and tunnels, have been in use for over 50 years in Japan. If abnormalities in these structures (cracks, leakage, spalling, cavity, etc.) are not detected and properly treated, the structures may deteriorate or eventually be damaged.

This study focuses on defect inspection of tunnel lining. One of the most important tasks for preliminary tunnel inspection¹ is creating a layout map. To identify changes on the tunnel lining surface, visual analysis techniques are often used in practice². There are three main approaches to visualize abnormalities of the tunnel lining: manual drawing, image mosaicing approaches by hand camera, and tunnel scanning by multi-cameras.

The first approach is to draw a layout map of abnormality positions via visual inspection on the site. This approach is simple to perform at the expense of positioning accuracy and speed of inspection. The second approach³, ⁴ commonly uses a hand-held camera to capture images with an auxiliary positioning equipment and an illuminator. This approach yields higher inspecting speed, improved accuracy, and low cost. However, the main drawbacks are the long image acquisition time for large-size tunnels and disruption to traffic. At each image shooting step at a location, multi-images need to be acquired for a large scene. To capture the images of the full tunnel length, this step must be repeated at multiple locations. This image acquisition method results in complex geometric distortions because of many different camera positions. Hence a large number of raw original images have to be rectified before warping and stitching. This makes the approach less practical for inspecting large-size tunnels.

The third approach utilizes high-speed multi-
cameras mounted on an inspection vehicle, and an illumination device to acquire images continuously. This allows constructing high-resolution 2D and 3D profiles of the tunnel lining. By far, this approach yields the fastest inspecting speed and the highest accuracy compared to the other approaches. However, the inspection cost is high as the approach requires high-precision laser scanners and special equipment. As the inspection device shifts parallel to the imaging plane, the obtained images are simple geometric transformation (e.g., translation) from the reference image. Both the image mosaicing and the tunnel scanning approaches can create a layout panorama.

To reduce inspection cost, Ukai and Nagamine used a tunnel scanner with line sensor cameras, which does not require a special vehicle, to continuously scan images for detecting tunnel wall deformations. This image acquisition system can be mounted on a railcar running at 20km/h and achieving a record resolution of 0.5-1 mm/pixel. This image scanning system is used on railways, but it has not been applied to road tunnels. Besides, the image matching algorithm is not presented in detail.

In recent years, many researchers have considered automated crack detection on the tunnel lining surface based on auto inspection; e.g., see Yu et al. and Zhang et al. However, they did not expound on tunnel panorama generation to assist visual inspection.

This study utilizes the tunnel scanning approach to provide the layout panorama for defect inspection on the tunnel lining surface. The goals are to achieve reasonable accuracy, active image-acquisition time, and manageable inspection cost. For image matching, the authors use color pixel-intensity-based feature and the curvature measure to find the appropriate matching points. This is different from standard image matching techniques that make use of key-point (feature)-based image matching methods such as SIFT, SURF, which may fail when the tunnel lining images contain noises or do not have many discriminated features.

Previously, the authors proposed a Mobile Tunnel Inspection System (MOTIS) consisting of an image acquisition equipment and an image-stitching software for creating panoramic images. This software matches pixel brightness within the overlapped region. The initial optimal matching point is the position with maximum similarity value. The method is simple but effective in acquiring high-quality image data for inspection. More specifically, in the paper, the authors used a full search method (FS method) and a correction technique based on the parameters of image acquisition device and inspection speech. The average matching error rate of six cameras was 33% before the correction technique was applied. After correction, the average error rate was reduced to 2.7%. In the subsequent paper, the authors adopted a local search method (LS method) to check pixels around the estimated image-matching location, which was calculated using the parameters of the image acquisition device and inspection speech. As a result, the average error rate of image-matching locations was reduced to 3.45% before correction; and the error was eliminated after correction.

In this paper, the authors propose a search-strip method to extract candidate matching points from a search space based on the similarity metrics. Next, the precise matching point is selected from a curvature metric (CUR). This metric measures the sharpness of the peaks of the curve established from the matching point candidates. The matching accuracy is further improved by measuring the cost of Curvature function at the Nearest-neighbor Pixels and the pixel being processed (CNP). The experimental results demonstrate that CNP gives higher accuracy than CUR.

The remainder of the paper is organized as follows: Section 2 reviews the existing methods including the video acquisition system and image-stitching procedure. Section 3 proposes a new image-matching procedure, which consists of a similarity metric and its refinement. Section 4 reports the experimental results validating our proposed method against previous methods. Section 5 assesses the performance of the proposed search algorithm by evaluating the accuracy of image motion quantity (IMQ). The calculated IMQ is compared with the target IMQ based on the parameters of the imaging system and the manual image matching. The efficiency (time and accuracy) of this algorithm is compared with our previous published papers. The section concludes with a description of an image-stitching software that makes use of the available image-matching information to produce tunnel panoramic images. This is followed by a discussion in Section 6. Finally, Section 7 concludes the paper.

2. EXISTING METHOD

In our previous papers, the authors reported a video acquisition system and a procedure of the image stitching for the entire tunnel lining as follows:

(1) Video acquisition system

The video images of the entire tunnel lining surface are continuously scanned by a video acquisition system assembled on a car. Figure 1 (a) shows the system consisting of six digital video cameras (from V1 to V6), three illuminators attached to a steel
framework shaped to work within a half of the cross-section of the tunnel, and a car. This device is able to slide from the side to the top of the car so that the full tunnel lining surface can be captured by several passes through the tunnel as shown in Fig.1 (b).

Figure 1 (b) presents the designed six regions to take images of the full cross-section of the tunnel lining in forward and backward directions. Figure 1 (c) shows how to obtain an image in each move step of the system in the longitudinal direction of the tunnel. For example, in the tunnel scanning process, the image acquisition system takes six frames at six positions. Quality of the image data depends on many factors, such as distance between the tunnel lining and the inspection car, the illumination, and the specification of the video camera. This system is simple and does not require an exclusive vehicle. The video data are converted into the consecutive images with the specific overlapped region depending on the scanning speed of the video acquisition system. By design, the resulting image motion is only the translation, thus the rotation is not a concern. Therefore, the direct method, which makes use of available information in the image-matching location, is applied to stitch images.

(2) Image-stitching procedure

Figure 2 shows an image-stitching procedure for the entire tunnel lining, which involves three steps. Captured image data are first retrieved from the six video cameras, each of which contains a number of images depending on the length of the tunnel. In Step 1, the images for each camera are stitched into a panoramic image in the longitudinal direction of tunnel for each camera. In Step 2, these panoramic images obtained in Step 1 are stitched in the circumferential direction of tunnel for each region corresponding to each pass. Subsequently, in Step 3, the connected images of the six regions (R1-R6) in step 2 are stitched together to make a layout panorama, which consists of a full view of the entire tunnel lining surface.

In this paper, the automatic image-stitching method is used in Step 1. Manual methods are applied to Steps 2 and 3 because Steps 2 and 3 stitch images via the circumferential direction, and their accuracy depends on the accuracy of Step 1. Therefore, the accuracy of Step 1 plays the most important role. Step 1 is presented in detail in the following section.

3. PROPOSED METHOD

The key to image stitching is to find the correct matching point for a pair of consecutive images. To this end, the authors propose a new image-matching algorithm, as follows: First, an initial image-matching location (IML) is estimated using Sum Absolute Difference (SAD) and normalized cross-correlation (NCC) measurements. Next, the images are matched using the Curvature of the cost curve at the Nearest-neighbor Pixels as well as the pixel being pro-
cessed (CNP). Finally, the matching errors are corrected using the median filter.

(1) Image-matching procedure

Figure 3 illustrates the automatic matching process of two consecutive images in the tunnel longitudinal direction (X-axis). Here, image 1 is the referenced image, and image 2 is the registered image of the image-matching process. A search area is set in advance. Each movement step of the search point is with respect to the location, which enables image 2 to be shifted onto image 1 to find the appropriate image-matching location by measuring the similarity and sharpness in the brightness of all pixel pairs in the overlapped region of the two images.

Furthermore, to accelerate the search and measurement process, the search point and similarity measurement are skipped in the search area and overlapped region with predefined values, respectively. The search process for the matching point is summarized in Pseudo algorithm 1. The search algorithm is detailed as follows:

**a) Similarity metric**

In the previous papers\(^{(10), (11)}\), the authors used SSD cost function computed in Eq. (1) to find the correct matching point. However, this metric is expensive to compute. In this paper, Sum Absolute Difference (SAD) and Normalized Cross Correlation (NCC) denoting matching cost functions (C) are used to measure intuitive similarity of the color pixel-wise pair in terms of brightness in the overlapped region, as in Eqs. (2) and (3):

\[
SSD = \frac{\sum_{i=m}^{M} \sum_{j=n}^{N} [(I_1(i,j) - I_2(i,j))]^2}{(N-n+1) \times (M-m+1)}
\]

\[
SAD = \frac{\sum_{i=m}^{M} \sum_{j=n}^{N} |(I_1(i,j) - I_2(i,j))|}{(N-n+1) \times (M-m+1)}
\]

\[
NCC = \frac{\sum_{i=m}^{M} \sum_{j=n}^{N} [(I_1(i,j) - \bar{I}_1) \times (I_2(i,j) - \bar{I}_2)]}{\sqrt{\sum_{i=m}^{M} \sum_{j=n}^{N} I_1^2(i,j) \times \sum_{i=m}^{M} \sum_{j=n}^{N} I_2^2(i,j)}}
\]

In these equations, \(I_1\) and \(I_2\) are the intensities of pixels at coordinates \((i, j)\) of images 1 and 2 in the overlapped region, respectively; \(SAD\) is the sum of difference between the pixel values of images 1 and 2 for each color channel (R, G, and B).

However, the standard \(SAD\) function has poor performance because image data are attributable to radiometric distortion and noise\(^{(12)}\). For that, the authors propose a modified \(SAD\) function as follows: The numerator of Eq. (2) is divided by the area of the overlapping region (the total number of pixels) to normalize the overlapping region in each measure-
ment. Furthermore, \((m, n)\) and \((M, N)\) are the lower left and upper right coordinates (pixel) of the overlapped region of the image pairs, respectively. In this function, the higher similarity of the overlapped region of two images yields the smaller score of \(SAD\).

Conversely, the \(NCC\) metric is determined from the sum of the correlation between the pixel intensity values of images 1 and 2 in the overlapped region. The larger score of \(NCC\) metric corresponds to the higher similarity of the overlapped region of two images.

These similarity metrics have several advantages. These functions are relatively simple to use in finding the appropriate matching point\(^{13}\). Moreover, they provide dense matching for every image pair, thus they can stitch directly for all input dataset with reasonable accuracy without pre-processing and post-processing to get the features. However, the traditional similarity metrics are affected by artifacts such as non-uniform brightness, periodic structures, featureless, and noises. These factors result in image-matching error. Here, the authors introduce a metric to improve the accuracy of IML, as described below:

b) Proposed metric

To overcome these image-matching errors, the authors propose the search strip method and a metric including three steps. First, candidate matching points are determined by the similarity metrics in each search strip. For uniformity with \(NCC\), the cost value of the \(SAD\) function is converted to the negative value shown in the Eq. (4) because the cost values of \(SAD\) and \(NCC\) are opposite:

\[
P_{(t)} = \{P_1, P_2, ..., P_N\} = \text{Max} \left[ \frac{-C_{SAD}(k,l)}{C_{NCC}(k,l)} \right] \quad (4)
\]

where \(P_{(t)}\) is the cost value of the \(t^{th}\) candidate matching point at the \((k, l)\) coordinates in the \(t^{th}\) search strip; \(N\) is the total number of the candidate matching points in the search area; \(C_{SAD}(k,l)\) and \(C_{NCC}(k,l)\) are the cost value given by Eqs. (2) and (3), respectively.

The size of each search strip is \((11 \times 100)\) pixels. Each strip has a represented candidate of the image-matching location. The search point is skipped to 11 and 4 pixels via the horizontal and vertical directions, respectively, in the search strip and in the overlapped region.

Second, the set of these points \(\{P_t\}\) obtained from the first step forms a cost curve. The conventional similarity metrics measure the cost value of the pixel being processed as follows:

\[
MSM = \text{Max}\{P_t\}, t \in [1, N] \quad (5)
\]

where \(MSM\) is the Maximum value of the Similarity Measurement for all of the candidate matching points in the full search space.

The curvature of the cost curve is computed, as shown in Eq. (6). Finally, the \(CNP\) metric is calculated, as shown in Eq. (7).

\[
CUR_{(t)} = -2P_{(t)} + P_{(t-1)} + P_{(t+1)} \quad (6)
\]

\[
CNP_t = \frac{CUR_{(t-1)} + CUR_{(t)} + CUR_{(t+1)}}{3} \quad (7)
\]

In Eq. (6), \(P_{(t-1)}\) and \(P_{(t+1)}\) are the cost values to the left and right of the candidate matching point being processed in the \(t^{th}\) strip. If one of \(t-1\) and \(t+1\) positions is out of the search range, it is replaced with the nearest valid one. \(CUR_{(t)}\) is the curvature of the cost curve at the \((ti,l)\) coordinates.

The \(CUR\) metric shows the confidence of a match because of the rapid cost change near the position having the maximum cost\(^{14}\). The larger \(CUR\) means higher confidence. If the curvature of the cost curve at the pixel being processed is small, then the image-matching location will not be reliable due to the quite similar cost of three pixels. Therefore, the \(CNP\) metric checks the confidence of the candidate matching point by evaluating the curvature values of its neighbor candidates.

In Eq. (7) \(CNP_t\) is the average curvature value at the \(t^{th}\) candidate with respect to the two adjacent candidates. The higher \(CNP\) score means the more reliable matching location.

c) Refinement

After the proposed metrics are applied, some local bad matching points still remain. Further, the desired running speed of the inspection car is unchanged. Therefore, to smooth the IMQ graph, the median filter is adopted as follows:

\[
Q_i = \begin{cases} 
M_i & \text{if } |S_i - M_i| \leq T \\
S_i & \text{otherwise}
\end{cases} \quad (8)
\]

where

\[
S_i = \text{med}(M_{(i-k)}, ..., M_i, ..., M_{(i+k)}) \quad (9)
\]

Here, \(Q_i\) is IMQ in unit pixels after refinement, and \(M_i\) is IMQ before refinement. \(S_i\) is a median value obtained from Eq. (9), and \(T\) is a threshold value depending on the speed change of the inspection car.
The size of the median filter is $2k+1$ (where $k$ is an integer).

4. EXPERIMENTAL WORK

(1) Case study
The tunnel inspected in this experiment was an actual single-core circular tube in Yamaguchi Prefecture, Japan with length of 230 m, width of 10.25m, and height of 4.7m. The video digital cameras were SONY HDR-CX630V, with LED light 600lx and camera angle of view between 5°-6°. The distance between the cameras and the tunnel wall was 3m. The dataset acquired by each camera consisted of 1,558 images with a resolution of 1920 × 1080 pixels for each picture. A photographic laser distance meter was used to maintain a constant distance between the tunnel lining and the inspection car to ensure a constant resolution. The speed of the inspection car was maintained at $30 \pm 5$ km/h corresponding to the desired image motion quantity from 500 to 700 pixels.

Further, a prototype software written in C++ was developed to implement the full image stitching algorithm. The search range was set to large to ensure the search result of matching point objectively as follows:

$$0 \leq k \leq 810$$
$$-50 \leq l \leq 50$$

The search range via the horizontal axis designed from [0 810] pixels corresponded to the inspection car speed from [0 40] km/h with the frame rate and resolution of the captured images shown in Table 1.

The search range in [-50 50] via the vertical axis was on account of the vibration of the inspection vehicle. The parameters of the refinement step were chosen as follows: size of the median filter was 100, and the threshold value was 50. The purpose of the median filter application was to reduce local errors when comparing neighbor IMQs.

(2) IMQ results
Figures 5 and 6 show the IMQ results of the imaging data in camera V2. The horizontal axis is the number of consecutive images before and after matches, and the vertical axis presents the IMQ of two subsequent images. Figures 5(a) and 6(a) present the IMQ in the longitudinal direction of the tunnel called IMQx, and Figs.5(b) and 6(b) show the IMQ in the circumferential direction of the tunnel denoted IMQy. As can be seen in Fig.5 (a), the IMQx is distributed at two critical positions on the graph. There are positions close to 0 and 600 pixels. Some local locations have IMQx over 700 pixels, and some other locations have IMQx between 0 and 500 pixels. Therefore, it is reasonable to estimate the first matching point around 600 pixels. To extend the reliable image-matching range, the authors define the accuracy of IMQx based on the parameters of the imaging system shown in Table 1. As a result, the IMQx between the lower threshold value (500 pixels) and the upper threshold value (700 pixels) is considered as the correct-matching point.

Table 1 Parameter values of the imaging system.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed of shooting vehicle car ($u$)</td>
<td>[25-35] km/h</td>
</tr>
<tr>
<td>Frame rate ($v$)</td>
<td>60 frames/second</td>
</tr>
<tr>
<td>Resolution ($r$)</td>
<td>0.231 mm/pixel</td>
</tr>
<tr>
<td>Estimated value of image motion quantity: $E = u/(v \times r)$</td>
<td>[500 700] pixels</td>
</tr>
</tbody>
</table>
Accordingly, the error rate is calculated by dividing the number of the image-matching errors by the total of the input images. The error rate of the IMQx using the MSM metric for all images in camera V2 is about 45.38%. In that, the error rate above the upper threshold is about 1%, and the error rate below the lower threshold is 44.38% (see Table 2). As shown in the IMQx graph in Fig. 6 (a) using the CNP metric, most IMQs distribute around 600 pixels. The error rate is only 0.26%. It is demonstrated that the use of CNP metric improves the accuracy significantly.

Moreover, the authors define the permitted accuracy of the IMQy to be \( \pm 20 \) pixels (\( \pm 4.6 \) mm). This accuracy is relatively high due to vibration of the inspection car. In effect, Fig.5 (b) shows the IMQy around \( \pm 10 \) pixels, some local locations above \( \pm 20 \) pixels.

Meanwhile, Fig.6 (b) reflects that the IMQy resides in the reliable range and its fluctuation is about \( \pm 5 \) pixels. From the results of the obtained IMQx and IMQy, it can be concluded that the CNP metric gains the higher accuracy than that of the MSM metric. Moreover, the fluctuation of IMQy shown in Figs. 5(b) and 6(b) does not have influence on the quality of image stitching.

To demonstrate the effectiveness of the CNP metric, Fig.7 presents the relationship between IMQx and score deviation value of the MSM and CNP metrics using SAD cost function for four image-matching cases, which correspond to four original pairs shown in Fig.8.

Figures 7 (a), (b), and (c) show three cases of the inappropriate matches resulting from the MSM metric.

However, the results of the CNP metric obtain all correct matching positions. For detailed explanations, in Fig.7, the horizontal axis presents the IMQx, and the vertical axis shows the score deviation value defined by the following equation:

\[
\sigma_i = \frac{S_i - S_{\text{avg}}}{S_{\text{max}} - S_{\text{min}}}
\]

where \( \sigma_i \) and \( S_i \) are the \( i \)th score-deviation value and the \( i \)th score for a measurement of similarity as well as curvature; and \( S_{\text{avg}} \), \( S_{\text{max}} \), and \( S_{\text{min}} \) are average, maximum, and minimum scores of the metrics in the search range for each image pair, respectively.

In fact, Fig.7 (a) shows a case of a bad match at coordinates (198, -0.3) using the MSM metric. This case occurs due to featureless and non-uniform lighting contribution on frames 311 and 312 shown in Fig.8 (a). The matching point location of CNP metric is at the (627, 0.5) coordinates.

Figure 7 (b) shows an error of the MSM metric with the matching point at the (803,-0.6) coordinates. This case tends to be out of the search range. This case searches matching point of a pair of frames 524 and 525 shown in Fig.8 (b). Simultaneously, the matching point of CNP metric is at the (583, 0.7) coordinates. This point is a reliable matching point.
In addition, Fig. 7 (c) shows a common error of matching point at the (0,-0.6) coordinates using the MSM metric. This case occurs due to the influence of the featureless and the periodic structure in the pair of frames 216-217 shown in Fig. 8 (c).

Meanwhile, the location of the matching point of the CNP metric is at the (594, 0.43) coordinates. In Fig. 7(d), both MSM and CNP metrics expose the desired matches of the pair of frames 265 and 266 shown in Fig. 8 (d).

Moreover, Fig. 9 shows the graphs of IMQs using the CNP metric with SAD and NCC cost functions after refinement, called CNP(SAD) and CNP(NCC), respectively. By observation, the graph of IMQs using CNP(SAD) metric shown in Fig. 9 (b) is smoother than the one of CNP(NCC) metric shown in Fig. 9 (a).

For this reason, CNP(SAD) metric gets the sharper curvature of the pixel being processed compared with the neighbor pixels of the CNP(NCC) metric. In effect, the error rate of CNP(NCC) and CNP(SAD) metrics are completely eliminated after refinement technique is applied (see Table 3).

3) Accumulative IMQ results

This subsection shows the accumulative IMQ of six cameras in the region 1 (R1) of the video acquisition device. The accumulative IMQ results of R2 to R6 are not presented due to space limitation. Figures 10, 11, and 12 show the accumulative IMQ value for each camera. In these figures, the horizontal axis is the order numbers of cameras V1 to V6. The vertical axis is the accumulative IMQ in the tunnel longitudinal direction.

As shown in Fig. 10, in the comparison of accumulative IMQ (a.IMQ) of the SAD metric and NCC metric, the NCC metric is better than the SAD metric for all cameras. Although the SAD metric is normalized by the total number of pixels in the overlapped region, it is sensitive to the linear lighting change. Whereas the NCC metric is a normalized
metric, it is invariant to the linear change of the brightness.

Figures 11 and 12 show the results of the a.IMQ of CUR and CNP metrics using SAD and NCC cost functions after the refinement is applied. Their a.IMQ difference is small. The maximum difference of the a.IMQ is 9000 pixels belonging to camera V4, as shown in Fig. 11. It shows the effectiveness of the refinement technique leading to the stability of the a.IMQ.

Comparing the MSM, CUR, and CNP metrics, the results of a.IMQ using MSM are the lowest with the maximum a.IMQ of about 800,000 pixels in all cameras; whereas the maximum results of the a.IMQ using CUR and CNP metrics are 933861 and 923582 pixels, respectively.

5. PERFORMANCE EVALUATION

There are two criteria to evaluate the effectiveness of the proposed method. Those are (a) target IMQ based on the parameters of the video acquisition system and (b) ground-truth image matching.

(1) Evaluation of image matching error based on the parameters of the video acquisition system

The dataset used to test the image-matching error (IME) consists of six cameras in Region 1. The total number of images for evaluation is 9,348. The IMQ results of the MSM and CNP metrics are compared with the target IMQ based on the following equation:

\[ E = \frac{u}{v \times r} \]  

(12)

where \( u, v, \) and \( r \) are the parameters of the imaging system the values of which are given in Table 1; and \( E \) is the target IMQ.

The speed of the inspection car is set at about 30 km/h, with possible variation in the range of 25 to 35 km/h. From Eq. (12), this speed variance is equivalent to the reliable range of \( E \) of [500 700] pixel. The error rate is computed using the following equation (12):

\[ AME = \frac{1}{N} \sum_{i=1}^{N} \frac{IME_i \times 100%}{100} \]  

(13)

where \( N \) is the number of images in each camera; \( IME_i \) is \( i \)th image matching error; \( IMQ_i \) is image motion quantity at \( i \)th measurement; and \( AME \) is the accumulative error rate for each camera. This equation indicates that an IMQ value is out of range [500 700] pixels considered as an IME.

<table>
<thead>
<tr>
<th>Camera</th>
<th>Proposed method</th>
<th>Previous methods</th>
<th>Cost function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSM</td>
<td>CUR</td>
<td>CNP</td>
</tr>
<tr>
<td>V1</td>
<td>63.48</td>
<td>0.77</td>
<td>0.45</td>
</tr>
<tr>
<td>V2</td>
<td>45.38</td>
<td>0.77</td>
<td>0.26</td>
</tr>
<tr>
<td>V3</td>
<td>46.47</td>
<td>1.67</td>
<td>0.26</td>
</tr>
<tr>
<td>V4</td>
<td>33.18</td>
<td>1.99</td>
<td>0.83</td>
</tr>
<tr>
<td>V5</td>
<td>39.00</td>
<td>1.60</td>
<td>0.52</td>
</tr>
<tr>
<td>V6</td>
<td>37.66</td>
<td>5.80</td>
<td>0.77</td>
</tr>
<tr>
<td>V1</td>
<td>42.59</td>
<td>3.79</td>
<td>1.41</td>
</tr>
<tr>
<td>V2</td>
<td>29.63</td>
<td>5.65</td>
<td>1.16</td>
</tr>
<tr>
<td>V3</td>
<td>32.84</td>
<td>7.45</td>
<td>3.60</td>
</tr>
<tr>
<td>V4</td>
<td>20.85</td>
<td>6.55</td>
<td>2.57</td>
</tr>
<tr>
<td>V5</td>
<td>21.55</td>
<td>8.03</td>
<td>2.83</td>
</tr>
<tr>
<td>V6</td>
<td>41.94</td>
<td>21.16</td>
<td>7.59</td>
</tr>
</tbody>
</table>

Table 2. Criterion 1-based image matching error result (unit %) before refinement.

<table>
<thead>
<tr>
<th>Camera</th>
<th>Proposed method</th>
<th>Previous methods</th>
<th>Cost function</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSM</td>
<td>CUR</td>
<td>CNP</td>
</tr>
<tr>
<td>V1</td>
<td>82.92</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V2</td>
<td>31.45</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V3</td>
<td>43.77</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V4</td>
<td>16.55</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V5</td>
<td>20.47</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V6</td>
<td>9.69</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. Criterion 1-based image matching error results (unit %) after refinement.

<table>
<thead>
<tr>
<th>Segments</th>
<th>MSM</th>
<th>CUR before refinement</th>
<th>CNP before refinement</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1-1</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V1-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V1-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V1-1</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V1-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V1-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4. Criterion 2-based image matching error (unit %).

As a result, Table 2 reports the results of the AME based on the proposed method and our previous methods for six cameras before refinement. Among
the cost functions, the SAD metric gets the worst accuracy with an average AME of 44.20%. Meanwhile, the average AMEs of NCC and SSD metrics (in FS method) are 31.57 % and 32.87%, respectively.

Comparing the CUR, CNP, and LS metrics, the average AMEs of six cameras using CUR(SAD) and CNP(SAD) metrics are 2.10% and 0.51%, respectively. The average AME of CUR and CNP using NCC cost are 8.77% and 3.19%. The average AME of all metrics decreases as the refinement. The error rates of all metrics decrease dramatically except for the result of the FSM(SAD) function.

Table 6 shows the comparison of the length of tunnel based on a.IMGQ. The average accumulative IMQ (Avg.IMQ) values of six cameras using FN(SAD), CNP(SAD), and CNP(NCC) metrics are eliminated completely, while the AME of CUR(NCC) metric retains small errors.

Table 3 also shows that the AME of the LS method is eliminated completely, but the AME in the FS method is still retained with an average AME of six cameras of 2.7%. It demonstrates that the proposed method has improved the accuracy of the image-matching locations significantly.

(2) Evaluation of image matching error based on the ground-truth (G-T) image matching

To evaluate the accuracy of the proposed image-matching algorithm, the authors created the G-T image matching using manual method to compare the results of the proposed method in unit pixels. Each camera contained 1,558 images, and each image had a resolution of 1080 × 1920 pixels. Because the six cameras provided a large number of images, the authors did not test all images.

Therefore, from the original image data sets, three segmentations at the tunnel portal and in the middle of the tunnel are extracted from camera 1. The three segments of the camera V1 are denoted by V1-1, V1-2, and V1-3. Each segment has 40 images, and the total of three sections including 120 images are used as test samples.

$$AME = \frac{1}{N} \sum_{i=1}^{N} IME_i \times 100\%$$

$$IME_i = \begin{cases} 0 & \text{if } |T-IMQ| \leq \varepsilon \\ 1 & \text{otherwise} \end{cases}$$

where N, AME, IME, and IMQ are similar to the abovementioned definition. T is a manual matching of the ith G-T image in the vertical and horizontal directions; and \(\varepsilon\) is threshold value that is computed as ± 100 pixels equivalent to the permitted speed change, which is ± 5 km/h (Eq. 12).

Table 4 presents the results of AME measurement defined in Eq. (13). In the V1-1, MSM(NCC) metric has AME of 8%, and MSM(SAD) metric has AME of 15%. In other segments, the MSM metric has zero error. Furthermore, using CUR and CNP metrics, the AMEs are eliminated completely.

The curve graphs in Fig. 13 show the comparison between the G-T image matching and the image-matching results of CNP(SAD) and CNP(NCC) metrics. There is negligible difference between the graphs of the proposed metrics and the G-T. It can be concluded that both CNP(SAD) and CNP(NCC) metrics yield a high accuracy for the image-matching process of the tunnel lining surface.

(3) Evaluation of the length of tunnel based on accumulative IMQ

The actual tunnel length collected image data is 230 m on 0.231 mm/pixel resolution. The perfect accumulative IMQs will be around 1 million pixels. To evaluate the validation of the proposed method, the accumulative IMQs of six cameras need to be tested.

The average accumulative IMQ (Avg.IMQ) values
of six cameras using the MSM and CNP metrics are compared with the total number of pixels corresponded to the length of the entire tunnel. This difference is defined in the following equation:

$$\text{Diff} = \frac{[1,000,000 - \text{Avg.IMQ}]}{1,000,000} \times 100\%$$ (15)

Table 5 reflects that the largest Diff belongs to MSM(SAD) with 45.61%. However, the Diff of CNP(SAD) is 8.35%. The Diff of MSM(NCC) is 30.28% less than MSM(SAD). However, the Diff of CNP(NCC) is 8.31%. The experimental results express that CNP metric improves the accuracy of IMLs.

(4) Evaluation of the computational time

To test the computational time for each camera with other image-matching methods, the authors implemented a parallel running program for six cameras in each region to get the image-matching information based on MSM and CNP metrics. Table 6 reports the computational time for running the image-stitching software using the proposed method and the previous methods.

In the MSM method, the computational time of SAD cost function is shorter than the one of NCC cost function. The timing cost of the CNP metrics is 2 and 4 minutes after subtracting the time of the similarity measure of SAD and NCC, respectively. Comparing the computational time of FS and LS methods, the time of the proposed algorithm is the fastest. For this reason, the FS method searched all of points in the predefined area, and the LS method searched the matching point surrounding the initial estimated position (600 pixels). Meanwhile, the proposed search algorithm considers the curvature of the cost curve at the nearest neighbor pixels and the pixel being processed, which have local maximum MSM value to find the best image-matching point. Therefore, the algorithm reduces the number of search points.

(5) Creating panoramic images

Unfolded panoramas of tunnel lining are created by an image-stitching software based on the IMQ results of CNP metric. This software stitches consecutive images for each camera via longitudinal direction of the tube. The panoramic pictures of each camera for the entire tunnel lining include 22 segments, and the length of each segment is 10.5 m excluding overlapped parts at the two ends of the next segment.

For visualized comparison between MSM and CNP, Fig.14 shows the first three segments of the MSM and CNP metrics with SAD cost function. Using MSM metric to stitch images has many errors due to the featureless (red dash rectangle) and the periodic structures (green dash box) shown in Fig.14 (a). These errors are rectified and respectively shown from segment 1 to 3 using the CNP metric shown in Fig.14 (b).

Figure 15 shows a representation of one part of each camera in the longitudinal direction of the tunnel. Each mosaic compressed at a scale of 1 per 32 has 45600x1920 pixels based on stitching 65 RGB (1080x1920) resolution images. Figure 16 shows a represented result of the manual camera stitching via the circumferential direction. From this layout panorama with high resolution, inspectors can detect defects easily. For illustration, inspectors detect crack types (chalk lines) with predetermined maximum widths. That is an important purpose of producing the tunnel lining panorama.

6. DISCUSSION

Unlike the FS and LS methods, which only exploited the cost value of the pixel being processed, the proposed method extracted the candidate matching points based on the similarity metrics in each strip. Subsequently, the precise matching-point is selected based on the curvature value.

The MSM metric considered the global maximum value of the cost curve so that the results of the matching points had many errors due to the influence of the noises. The CUR metric focused on the candidates, which had the local maximum values of the cost curve. So far, the CNP metric has tested the stability of the candidates based on the mean curvature value.
The main advantage of the proposed method was to eliminate the artifacts of the neighbor pixels surrounding the candidate matching point in each search strip before measuring the sharpness at the peaks of the cost curve. Therefore, it improved the computational time and the matching accuracy compared with our previous studies.

As a result of this study, a large number of the raw images of tunnel lining surface were matched automatically with high accuracy. The average accuracy of the \( \text{CNP}(SAD) \) metric was 99.49% before refinement without depending on the initial estimated IMQ, and the computational time of the \( \text{CNP}(SAD) \) was 2 minutes in order to match 1558 image pairs.

7. CONCLUSIONS

The main contributions of the paper are:

1. Introducing an imaging equipment capable of scanning pictures of the entire tunnel. A total of six cameras, in which each camera produced 1558 images with 1290x1080-pixel resolutions, were mounted on a moving inspection car, which saved much time compared to the other conventional imaging systems.
2. Proposing a new image-matching method based on the curvature of the cost curve at the matching-point candidate locations. Consequently, the application of median filter rectified the error of inappropriate IMQ. The parameters of the imaging system were used to estimate the error rate of the initialized image matching. The use of image-stitching software enabled parallel running of six cameras, which significantly reduced computational time. The experimental results showed the high performance of the \( \text{CNP} \) metric compared with \( \text{MSM} \) and \( \text{CUR} \) metrics.

In addition, the image-matching results of the proposed method were improved significantly compared with those of our two previous methods and the G-T image matching.

In addition, this study found abnormalities in the tunnel lining surface without inspecting the shape of the tunnel. Therefore, the obtained images of the curved shape of the tunnel lining were manipulated to create a flat panorama for better visual inspection. The shaped 3D reconstruction of the tunnel-lining surface would be considered in the future research.

The paper observed several limitations. The proposed image-matching method relies exclusively on the color-pixel intensity. There is lack of feature descriptor-based methods such as SIFT and SURF. Moreover, the speed of the inspection car should be increased to eliminate disrupting traffic flow. These limitations will be addressed in future work.

ACKNOWLEDGMENT: This research is supported by JSPS KAKENHI Grant Number 15K0618.

Fig.15 The represented results of consecutive image stitching for each camera in the longitudinal direction.
REFERENCES


(Received September 15, 2017)