



This is a repository copy of *Image fusion for remote sizing of hot high quality steel sections*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/176450/>

Version: Accepted Version

---

**Proceedings Paper:**

Lin, Y., Wang, P., Muroiwa, R. et al. (2 more authors) (2021) Image fusion for remote sizing of hot high quality steel sections. In: Jansen, T., Jensen, R., Mac Parthaláin, N. and Lin, C.-M., (eds.) *Advances in Computational Intelligence Systems : Contributions Presented at the 20th UK Workshop on Computational Intelligence*. 20th UK Workshop on Computational Intelligence (UKCI'2021), 08-10 Sep 2021, Aberystwyth, Wales, UK. *Advances in Intelligent Systems and Computing* (1409). Springer Nature , pp. 357-368. ISBN 9783030870935

[https://doi.org/10.1007/978-3-030-87094-2\\_31](https://doi.org/10.1007/978-3-030-87094-2_31)

---

This is a post-peer-review, pre-copyedit version of an article published in *Advances in Computational Intelligence Systems*. The final authenticated version is available online at: [https://doi.org/10.1007/978-3-030-87094-2\\_31](https://doi.org/10.1007/978-3-030-87094-2_31).

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Image Fusion for Remote Sizing of Hot High Quality Steel Sections

Y. Lin<sup>1</sup>, P. Wang<sup>2</sup>, R. Muroiwa<sup>3</sup>, S. Pike<sup>3</sup>, and L. Mihaylova<sup>1</sup>

<sup>1</sup> The University of Sheffield, Sheffield, UK,

<sup>2</sup> Manchester Metropolitan University, Manchester, UK,

<sup>3</sup> Liberty Speciality Steels, UK,

{ylin42, l.s.mihaylova}@sheffield.ac.uk, p.wang@mmu.ac.uk,  
{ree.muroiwa, simon.pike}@specialityuk.com.

**Abstract.** This paper proposes an adaptive method for dual camera based steel section sizing, where high accuracy measuring is challenging due to the lack of well pronounced image features. The proposed approach includes additional information from a sidewise positioned checkerboard and enables adaptive image registration. A thorough evaluation of the registration results based on the virtual checkerboard is presented. On the accomplishment of image registration, both a fast Fourier transform and a discrete wavelet transform are adopted for fusion of the registered images. A range of comparisons with various metrics is conducted to achieve the best fusion quality. The hot steel section sizing results show an accuracy that is in line with the rolling standards, i.e. in the tolerance range less than 1.5 *mm* error.

**Keywords:** Image Fusion, Manufacturing and Automation, Metrology, Computer Vision, Registration

## 1 Introduction

In the process of automatic rolling, the real-time and automated non-contact measurement of steel sections on producing line is an effective process control method. Such real-time sizing provides great convenience for product quality control and improvement and innovation of the manufacturing process.

In order to obtain the physical size of objects, many three dimensional (3D) measurement and estimation methods based on vision have been studied [1]. The 3D vision measurement methods can be broadly divided into three categories - time of flight methods, structured light methods and stereo vision methods.

The *time of flight method* [2] uses a particular time of flight camera. In addition to capturing the colour information of pixels, the camera also records the time from the pixel light source to the camera. Therefore, the time of flight camera needs an individual artificial light source. The time is obtained by calculating the phase difference of light, and then the propagation speed of light is used as the reference to calculate the depth map. The error of the time-of-flight method is affected by many aspects, such as colour and material of the object,

the distance to the camera and illumination of the environment. The depth error in  $5m$  is less than  $4.6mm$ .

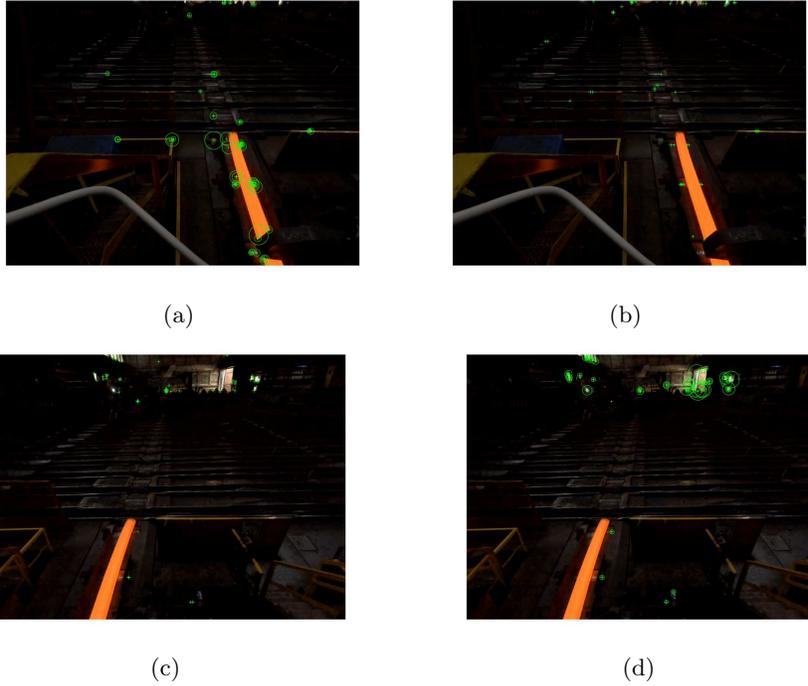
The *structured light method* [3, 4] relies on projecting the inferred light with specific structural characteristics onto the object, then collect the reflected structured light pattern by the infrared camera, and calculate the 3D information according to the principle of triangulation. Since the encoded structured light image or speckle spots are easily confused by intense natural light outdoor, the structured light scheme is not easy to use outdoor. When the object is far away from the camera, the intervals between light points projected on the object increases and lead to a decrease in accuracy. The structured light camera is also easily affected by the reflection of a smooth plane, such as a mirror.

*Stereo vision measurement systems* have also been proposed [5, 6]. In the analysis of two images taken by the binocular camera in the same scene, the parallax image is obtained by the stereo matching algorithm. Then the depth map and depth information are obtained by including geometry information. For the binocular stereo vision system, the camera calibration is the most crucial aspect that affects the measurement accuracy. The number of calibration images and the position of the calibration plate plays a significant role [7]. For instance, Zhou et al. [6] propose an approach for online diameter estimation of large hot forgings. A fast measurement method based on feature line reconstruction of stereo vision is developed to increase the calculation speed. Their measurement accuracy is considerable, but the method needs good lighting conditions and high-precision camera calibration.

Our previous work [8, 9] proposes a real-time measurement system based on monocular camera data. A fast structural random forest algorithm detects the steel bars' edges, and a regression algorithm extracts the edges in both optical and thermal videos. The steel sections' sizes are then calculated from the edge information and converted from the image plane to the physical plane. When the diameter of the steel section changes, the horizontal cross section of steel height to take the measurement. Although the depth of field can be predicted by monocular data to a certain extent through deep learning methods, this requires a large amount of data from the corresponding environment to train the system, and the accuracy of the final results is not enough for sizing tasks [10].

A binocular system has advantages and this is part of the novelty of this work. Depth information generated by the two cameras is used to improve the accuracy. However, image registration is needed. Unfortunately, there is no generic method that is suitable for all registration tasks. Each method is chosen based on different considerations, including the assumed geometric deformation types between images, the radiation deformation, noises, accuracy of registration, and the characteristics of application data.

The initial steps of the registration method start with feature detection. The commonly used automatic feature detection operators can extract the salient and unique features of the object such as closed boundary regions, edges, contours, intersections and corners. However, in the process of steel rolling, the illumination condition changes and high temperature condition of the steel section



**Fig. 1.** Feature points extracted by different methods. The top row images are captured by the left camera, and the bottom row images are captured by the right camera: (a) & (c) Results with SURF; (b) & (d) Features from FAST.

formation process are challenging to the autonomous computer vision methods. The feature points extracted by traditional methods are sparse, and this affects the accuracy of the results. In Fig. 1 the green points are the feature points detected by the Speeded-Up Robust Features (SURF) and Features from Accelerated Segment Test (FAST) algorithms [11, 12], respectively. We can see features on the steel sections are fairly sparse.

The input data is collected by two GoPro cameras installed at 2.5 m apart, and 5 m above the hot steel sections. Due to the need for high-precision vision based sizing at a long distance, a two camera solution is the choice that can provide accurate depth estimation. Two cameras are positioned on the left and right sides of the measured steel section respectively.

This paper proposes an adaptive image registration algorithm based on an external camera checkerboard. In this way, the spatial relationship between the measured object (hot steel sections) and the camera can be accurately estimated. Hence, this affords embedding the calibration board data directly in the sizing process and guarantees high image registration accuracy. After the image registration, the images from the left and right cameras are fused based on the fast Fourier transform (FFT) and discrete wavelet transform. A series of evaluation metric for the fusion results without reference image are used to analyze



**Fig. 2.** Input steel section images: (a) Left camera; (b) Right camera

---

**Algorithm 1** Image Registration and Evaluation

---

Phase 1: Image Registration

**Input:**  $I_L, I_R$

**Output:** The registered image  $I_{Rr}$

Extract the corner points  $C_L, C_R$  of checkerboards in  $I_L, I_R$

Calculate the geometric transformation  $T_{RL}$ , which transform  $C_R \rightarrow C_L$

Apply  $T_{RL}$  to  $I_R$ :  $T_{RL}(I_R) = I_{Rr}$

Phase 2: Registration Evaluation

**Input:**  $I_{BOI}, P_{re}$

**Output:** The Quality of Registration  $Q_R$

Find the middle point of checkerboard  $[x_{board} \ y_{board}]$

Create the testing area  $R_{test}$  with bounds  $y_{board} \pm P_{re}$

Create polygons  $P_{r,l}$  with Edges in  $R_{test}$

Calculate the percentage  $Q_R$  of overlapped area between  $P_r$  and  $P_l$

---

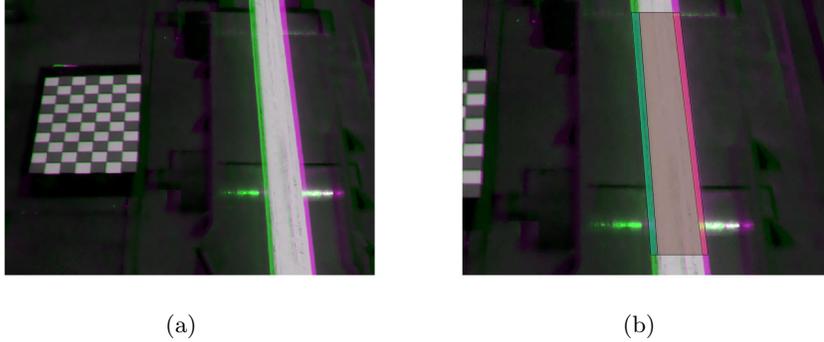
the quality of fusion. The remaining part of the paper is structured as follows. Section 2 describes how the virtual checkerboard is applied and the registration process. The results of sizing validation and evaluation of fusion results are given in Section 3. Section 4 summarises this work.

## 2 The Proposed Approach

To overcome the challenges caused by feature sparsity, we utilise the checkerboard used for camera calibration to provide external feature points. However, with only a few images of the checkerboard, an interpolation of the corner information can be done, which enables adaptive image registration and high accuracy image fusion.

### 2.1 Image Registration

To evaluate the image registration results, first the steel section shape is extracted by using edges in the images. Although deep learning methods [13]



**Fig. 3.** (a) Registered Images; (b) Polygons

have a big potential, they require large amounts of data for training and testing. Hence, this work adopts the structural random forest method [14] which does not require the availability of big data and can extract well edges in video frames. Compared with the traditional threshold based edge detection methods, the structural random forest algorithm can provide relatively stable adaptive results without setting parameters. On the effect of suppressing local maximum, structured random forest algorithm also has better results. Compared with the edge detection method based on deep learning, the structural random forest algorithm is easier to train, lighter to deploy.

After obtaining the steel section edges, the section part close to the checkerboard for evaluation is extracted. Next, the positive and negative  $P_{re}$  pixels above and below the middle point of checkerboards are selected. These points surround the testing area  $R_{test}$ . The polygons  $P_{l,r}$  are composed of the upper and lower bounds of the detection area and the steel section's edges in the detection area. After  $P_{l,r}$  are created, calculate the percentage of overlapped area between two polygons, which is the quality  $Q_R$  of the registration. The higher  $Q_R$  is, the better the registration quality is.

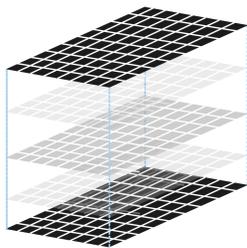
Fig. 3(a) shows the image pairs of the original image taken from the left and right cameras' registered images. The green part is from the left camera, and the magenta part is from the right camera. The regions with grayscale colour are overlapped area with the same intensity from two cameras. Fig. 3(b) are the detected polygons for steel sections from left and registered right images. The overlapping area between two polygons divided by the polygon area from the left image gives the quality of registration index  $Q_R$  which is given as follows

$$Q_R = a \left( \frac{P_r \cap P_l}{P_l} \right) \quad (1)$$

Here the coefficient  $a$  is positive (+1) when the centroid of  $P_l$  is situated on the left of the centroid of  $P_r$  and vice versa. For the purposes of image registration, the height of the checkerboard in the image needs to be changed. So the data is coming from different heights of the calibration plate. Checkerboard data at

the lowest and highest position are collected, and the data in the middle are generated by interpolation method. Taking images with a camera is essentially a projection process that projects the real three-dimensional object to the two-dimensional image plane. The projection process is linear. The middle point between checkerboards with different heights in the real-world is also in the middle in the image plane and this is represented by the equation

$$\mathbb{P} \frac{(x_1, y_1) + (x_2, y_2)}{2} = \frac{\mathbb{P}(x_1, y_1) + \mathbb{P}(x_2, y_2)}{2}. \quad (2)$$



**Fig. 4.** Virtual Checkerboard

Hence, the interpolation process can be realised by directly inserting data points between the checkerboard's corresponding points at different heights. These interpolated data will improve the accuracy of adaptive registration. The automatic update of the registration result is achieved by optimizing the registration quality value  $Q_R$ . According to the value of  $100 - |Q_R|$ , when  $Q_R$  is positive, the algorithm will choose the virtual checkerboard with higher height above the ground. When  $Q_R$  is negative, a lower virtual checkerboard is chosen, which make the virtual checkerboard at the same height with the horizontal cross section of the steel section. A larger value of  $100 - |Q_R|$  increase the searching interval between virtual checkerboards.

## 2.2 Image Fusion

After the images are registered, they are fused to enhance the information contained. The single view image taken by a single camera can not describe the target very well. If the images taken from different view can be fused together, the important information in the two images will be superimposed together, so that we can have a better understanding of the geometric shape of the target object. Image fusion based on the FFT and discrete wavelet are tested.

The FFT is an important image processing tool, which applies the discrete Fourier transform to the image and changes the image information from the spatial domain to the frequency domain. The left and right images are fused by combining the phase and magnitude maps of two images in the frequency

domain. The two dimensional (2D) discrete Fourier transformation for a image  $f[m, n]$  of size  $m \times n$  is defined as:

$$F_{k,l} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m, n] e^{-j2\pi(\frac{k}{M}m + \frac{l}{N}n)} \quad (3)$$

where  $F[k, l]$  can be decomposed into its amplitude  $\|F[k, l]\|$  and phase  $\angle F[k, l]$ . To fuse two images I1 and I2 in the frequency domain, the fusion rule is given as Algorithm 2. The discrete Fourier transform is implemented as a FFT. Im-

---

**Algorithm 2** FFT Fusion

---

**Input:**  $F_{I1}[k, l], F_{I2}[k, l]$

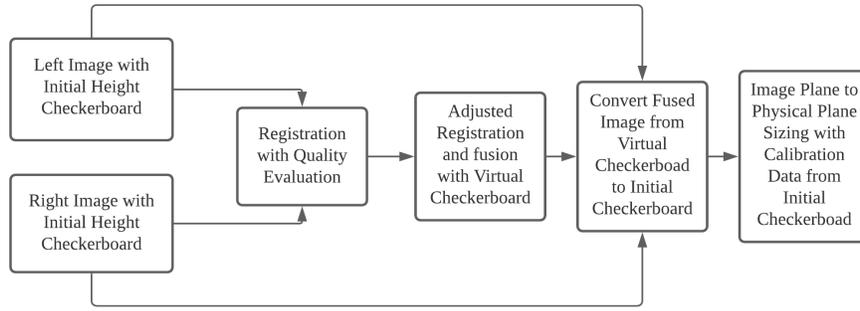
**Output:** The fused image  $I_F$

- 1: Calculate the magnitude  $\|F[k, l]\|$  and phase  $\angle F[k, l]$
  - 2: **for**  $k = 1, \dots, M$  **do**
  - 3:   **for**  $l = 1, \dots, N$  **do**
  - 4:     **if**  $\|F_{I1}[k, l]\| > \|F_{I2}[k, l]\|$  **then**
  - 5:        $\|F[k, l]\| = \|F_{I1}[k, l]\|$
  - 6:        $\angle F[k, l] = \angle F_{I1}[k, l]$
  - 7:     **else**
  - 8:        $\|F[k, l]\| = \|F_{I2}[k, l]\|$
  - 9:        $\angle F[k, l] = \angle F_{I2}[k, l]$
  - 10:    **end if**
  - 11:   **end for**
  - 12: **end for**
  - 13: Inverse FFT  $F[k, l] \rightarrow I_F$
- 

age fusion based on a discrete wavelet transform (DWT) is also a common method in image fusion. In essence, a two-dimensional DWT applies several one-dimensional discrete wavelet transforms based on the selected wavelet type in the horizontal and vertical directions of the image. The choice of wavelets affects the image fusion results. In this paper, the DWT is implemented with Daubechis wavelets. In the evaluation and validation process, a different number of wavelet coefficients are retained. The DWT method results are presented with 2, 4 and 16 coefficients and this is denoted as DWT2, DWT4 and DWT16, respectively. After the fusion of the images, edge extraction and image segmentation will be performed. The size information of the measured target can be estimated. By analyzing the contour of the measured object, the width of steel section can be intercepted. The mapping relationship between two-dimensional space and 3D space can be determined after calibrating the cameras. Through this mapping relationship, the size of the object in 3D space can be calculated.

### 3 Performance Validation and Evaluation

The whole process of sizing is shown in the flow chart of Fig. 5. The process starts with the video feeds from optical cameras monitoring the steel section from the left and right directions at the same time. At this time, the images have the initial checkerboard, which is used to calibrate the two cameras. It is not known whether the plane of the checkerboard is consistent with the plane of the steel section and how much difference there is. Then the next step is the registration of the two images with the aid of the virtual checkerboard. The registered images are with respect to the height of the steel section and it is generated by evaluating the registration quality in Section 2.1. After this registration process, the left and right images are fused and the steel section shape is extracted. The extracted steel section is next transformed from the image plane to the physical plane size by calculating the geometric transformation between the virtual checkerboard and the initial checkerboard.



**Fig. 5.** Flow Chart of Sizing Process

#### 3.1 Remote Sizing of Hot Steel Sections

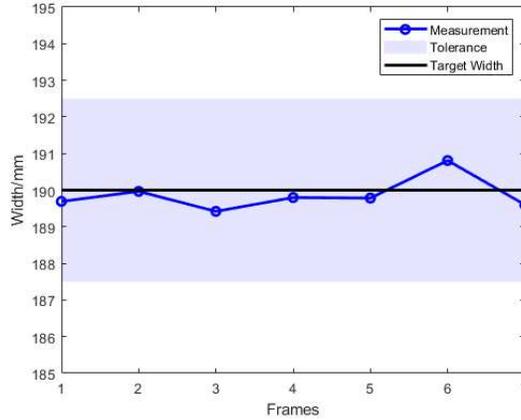
When the registration process is finished, with the camera internal and external parameters obtained from the calibration process and based on the approach from [9], the steel section size can be calculated and converted from the image plane to the physical plane. In our case, two hot rolling bar (HRB) edges are expected to be within the sliding window  $I_{H \times W}$ , with height  $H$  and width  $W$ . Therefore, we set up all  $n_i$  to be equal to  $n = 2$ . The  $\mathbf{x}_{ij}$  and  $\mathbf{y}_{ij}$  coordinates from the image plane are converted to coordinates in the physical plane through the transformation

$$\begin{bmatrix} x_{ij}^w \\ y_{ij}^w \\ z_{ij}^w \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \mathbf{K} \begin{bmatrix} x_{hj}^I \\ y_{hj}^I \\ 1 \end{bmatrix}, \quad (4)$$

where  $\mathbf{R}$  and  $\mathbf{T}$  are respectively the rotation and translation matrices, and  $\mathbf{K}$  is the intrinsic matrix of the camera parameters. These matrices are obtained via the calibration process. The coordinates  $x_{ij}^I \in \mathbf{x}_{ij}$  and  $y_{ij}^I \in \mathbf{y}_{ij}$  are from the image plane and  $[x_{ij}^w, y_{ij}^w, z_{ij}^w, 1]^T$  is the vector of corresponding coordinates in the physical plane. Given the vectors  $I_{i1} = [x_{i1}^I, y_{i1}^I]^T$  and  $I_{i2} = [x_{i2}^I, y_{i2}^I]^T$  on two HRB edges with  $x_{i1}^I = x_{i2}^I$ , the diameter  $l$  of the HRB is then calculated

$$l = \|P_1 - P_2\|_2, \quad (5)$$

with  $P_1 = [x_{i1}^w, y_{i1}^w]^T$  and  $P_2 = [x_{i2}^w, y_{i2}^w]^T$ , which are physical plane correspondences to  $I_{i1}$  and  $I_{i2}$ . Here  $\|\cdot\|_2$  denotes the Euclidean norm.



**Fig. 6.** Sizing Results for 7 Different Frames

### 3.2 Image Fusion Analysis Performance Metrics

Since we do not have a ground truth image for fusion performance evaluation, metrics [15] that do not require reference images are employed to assess the fusion results. These metrics are given below:

1) Information Entropy:  $H = -\sum P \log P$ , where  $P_i$  is the normalised histogram. The amount of information contained in the fused image can be measured by the information entropy. Images with more information have higher information entropy.

2) Standard Deviation:  $SD = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (f(i, j) - \bar{\mu})^2 / MN}$  where  $f$  is the pixel intensity,  $\bar{\mu}$  is the average intensity of pixels,  $M$  and  $N$  are the width and height of the image, respectively. The image standard deviation  $SD$  reflects the discrete degree of the image pixel brightness and mean value. The larger the standard deviation is, the more pronounced the contrast between light and dark areas is.

**Table 1.** Fusion Performance Evaluation

Metrics	Left Image	Right Image	FFT	DWT2	DWT4	DWT16
Information Entropy	4.67	4.36	<b>5.20</b>	4.42	4.45	4.53
Standard Deviation	32.13	32.32	30.98	30.58	30.26	29.36
Spatial Frequency	3.73	3.48	3.31	<b>4.62</b>	4.30	4.23
Average Gradient	4.70	4.33	<b>6.25</b>	6.22	5.75	5.78
Feature Mutual Information	-	-	<b>0.951</b>	0.947	0.950	0.947
Sum of the Correlation of Differences	-	-	<b>0.47</b>	0.32	0.28	0.23

**Table 2.** Efficiency evaluation of the algorithms

	FFT	DWT2	DWT4	DWT16
Time (ms)	1,826	6,717	6,778	7,158

3) Spatial Frequency:  $SF = \sqrt{(RF)^2 + (CF)^2}$  is defined based on the row frequency  $RF$  is and the column frequency  $CF$  and is calculated from the image row and column respectively. A higher frequency means better image quality.

4) Average Gradient:  $AG = \frac{1}{(H-1)(W-1)} \sum_x \sum_y \frac{G(x,y)}{\sqrt{2}}$  where  $H$  and  $W$  are height and width of the image, respectively and  $G$  is the gradient magnitude of the image. The first-order difference between the pixel value of a pixel and its adjacent pixels reflects the edge information of the pixel. The magnitude of this rate of change can be used to represent the clarity of the image.

5) Feature Mutual Information  $FMI = I_{FA} + I_{FB}$ , [16], where  $I_{FA}$  and  $I_{FB}$  denote respectively the mutual information between the individual image A or B and the fused image F. The FMI evaluates the dependency between input images and fused image. A larger FMI usually means a better fusion quality.

6) Sum of the Correlation of Differences  $SCD = r(D_1, S_1) + r(D_2, S_2)$ , [15], is the sum of two correlations, where  $D_i$  is the difference between the input image  $S_i, i = 1, 2$  and the fused image and  $r(\cdot)$  denotes the correlation.

### 3.3 Analysis of the Results

The efficiency of the proposed approach has been tested and validated with video data collected with two GoPro<sup>®</sup> Hero 7 Black cameras. The cameras are set to 2.7K mode and the shutter speed is 1/480s. The configuration of the test computer is an Intel(R) Core(TM) i7-9700K CPU and 32.0GB RAM. The software used is Matlab r2019a. The cameras are calibrated on the scene with a checkerboard of  $10 \times 7$  squares of size  $80 \times 80$ mm.

Fig. 6 presents the measurement results of seven images in a video sequence. The seven frames are separated by 100 frames, showing the measurement process of 700 frames. The target diameter of the hot rolling steel section is 190mm. Through our measurement, we can see that the accuracy of the estimated size of the rolling steel section is high.

Table 1 and Table 2 give the fusion performance evaluation and the time consumption of different methods. These performance measure values are obtained based on averaging results from 20 different video frames (from the left and right cameras). The results show that the FFT method gives better accuracy than the DWT method with respect to the fusion quality index and it also faster than the DWT. The DWT method results are given in Table 1.



**Fig. 7.** Fast Fourier transform/discrete wavelet transform fusion Image (Left/Right)

## 4 Conclusion

In this paper, an adaptive method is proposed for dual camera based high quality, high temperature steel section sizing. To cope with challenges such as the lack of strong image features on the sections and very few paired features from the two cameras, the proposed approach includes first cameras calibration. Next, the checkerboard information is extrapolated to the sizing plane. High resolution extrapolation is achieved by constraints imposed during the image registration process. The performance of the approach is evaluated on various real data with different metrics. The proposed approach achieves high precision sizing results that meet industrial requirements which is approximately  $1.5mm$  error (which corresponds to 0.79% error in percentages).

**Acknowledgements.** We are grateful to our sponsors through the Knowledge Exchange grant (Internet of Things for Overcome Barriers in the Steel Rolling Measurement Technology) with the Liberty Speciality Steels, the UK EPSRC IAA grant R/164995-11-1. We are grateful to the EPSRC for funding this work through EP/T013265/1 project NSF-EPSRC:ShiRAS Towards Safe and Reliable Autonomy in Sensor Driven Systems which is a joint grant with the USA Grant NSF ECCS 1903466. We also thank China NSFC (grant 61703387).

## References

1. S. Giancola, M. Valenti, and R. Sala, *A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies*. Springer Briefs in Computer Science. Springer, 2018.
2. Y. He, B. Liang, Y. Zou, J. He, and J. Yang, "Depth errors analysis and correction for time-of-flight (tof) cameras," *Sensors*, vol. 17, no. 1, p. 92, 2017.
3. S. Zhang, "High-speed 3d shape measurement with structured light methods: A review," *Optics and Lasers in Engineering*, vol. 106, pp. 119–131, 2018.
4. S. Lv, M. Jiang, C. Su, L. Zhang, F. Zhang, Q. Sui, and L. Jia, "Phase difference-3d coordinate mapping model of structural light imaging system based on extreme learning machine network," *IEEE Access*, vol. 8, pp. 68 974–68 981, 2020.
5. Y. M. Mustafah, R. Noor, H. Hasbi, and A. W. Azma, "Stereo vision images processing for real-time object distance and size measurements," in *Proc. of the 2012 ICCCE*. IEEE, 2012, pp. 659–663.
6. Y. Zhou, Y. Wu, and C. Luo, "A fast dimensional measurement method for large hot forgings based on line reconstruction," *The International Journal of Advanced Manufacturing Technology*, vol. 99, no. 5-8, pp. 1713–1724, 2018.
7. L. Yang, B. Wang, R. Zhang, H. Zhou, and R. Wang, "Analysis on location accuracy for the binocular stereo vision system," *IEEE Photonics Journal*, vol. 10, no. 1, pp. 1–16, 2017.
8. P. Wang, Y. Lin, R. Muroiwa, S. Pike, and L. Mihaylova, "A weighted variance approach for uncertainty quantification in high quality steel rolling," in *Proc. of the IEEE 23rd International Conf. on Information Fusion*. IEEE, 2020, pp. 1–7.
9. —, "Computer vision methods for automating high temperature steel section sizing in thermal images," in *Proc. of the Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2019, pp. 1–6.
10. Y. Luo, J. Ren, M. Lin, J. Pang, W. Sun, H. Li, and L. Lin, "Single view stereo matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 155–163.
11. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
12. E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.
13. Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects, in press," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2021.
14. P. Dollár and C. L. Zitnick, "Fast edge detection using structured forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1558–1570, 2014.
15. V. Aslantas and E. Bendes, "A new image quality metric for image fusion: The sum of the correlations of differences," *AEU-international Journal of Electronics and Communications*, vol. 69, no. 12, pp. 1890–1896, 2015.
16. M. B. A. Haghghat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Computers & Electrical Engineering*, vol. 37, no. 5, pp. 744–756, 2011.