Surface Defects Detection of Steel Plate Based on Visual Attention Mechanism

1^{st}:Fang Guo
School of Control Science and Engineering
University of Jinan
Jinan, Shandong
Email:18366104849@163.com

2^{nd}:Jianyu Zhao
School of Control Science and Engineering
University of Jinan
Jinan, Shandong
Email: cse_zjy@ujn.edu.cn

3^{rd}:Ping Jiang*
School of Control Science and Engineering
University of Jinan
Jinan, Shandong
Email: cse_jiangp@ujn.edu.cn

***Abstract*—For the image of present surface defects in steel plates, a surface defect detection method based on human visual attention mechanism is proposed. Based on the physiological structure and function of human visual system, this paper establishes a steel surface defect detection model via Gaussian filter, Gabor filter and the calculation such as cross-scale reduction, cross-scale addition and normalization, from which we can get the feature maps, the saliency map and so on. The experimental results show that the method can not only detect the defect area accurately, but also can meet the requirements of online real-time detection.**

I. Introduction

Steel plate as the main form of steel industry, has become indispensable raw materials of automobile manufacturing, shipbuilding, aerospace and daily necessities. The plate defects detection plays an important role in the steel quality testing. The existing steel surface defects detection methods are mainly two types. One is the staff to distinguish between discrimination, after the body temperature cooling to the standard temperature [1]. The second is taking monitor images by the camera or video, and then measured by eyes [2]. These two methods belong to the manual detection mode, which consume much time and labor. There is great necessary to research newly steel surface defects detection method for improving the steel quality. In this paper, a detection algorithm based on visual attention mechanism is proposed.

Visual attention mechanism is essentially a biological mechanism, which makes people choose a few important scenes of interest from the complex scenes as the focus of attention (FOA) preferentially [3] . In recent years, there are a large number of in-depth, interesting research and application of visual saliency detection. In simple terms, visual saliency drives the perception of human beings and make people pay more attention to some salient objects. The mechanism of visual attention is a very important issue in biological visual information processing. Humans can quickly find the area of interest from a complex scene and easily understand the scene. The visual attention mechanism is used to guide the human eyes to focus on significant areas in mass data and assign resources to prioritize important areas. The purpose of this paper is to use the mechanism of visual attention to achieve robust detection of sheet defects. Our focus is to make vision inspection systems detect defects quickly, accurately, and effectively.

This work was supported by the National Natural Science Foundation (NNSF) of China under Grant 61601197 and the Research Fund for the Taishan Scholar Project of Shandong Province of China.

*Ping Jiang is the corresponding author.

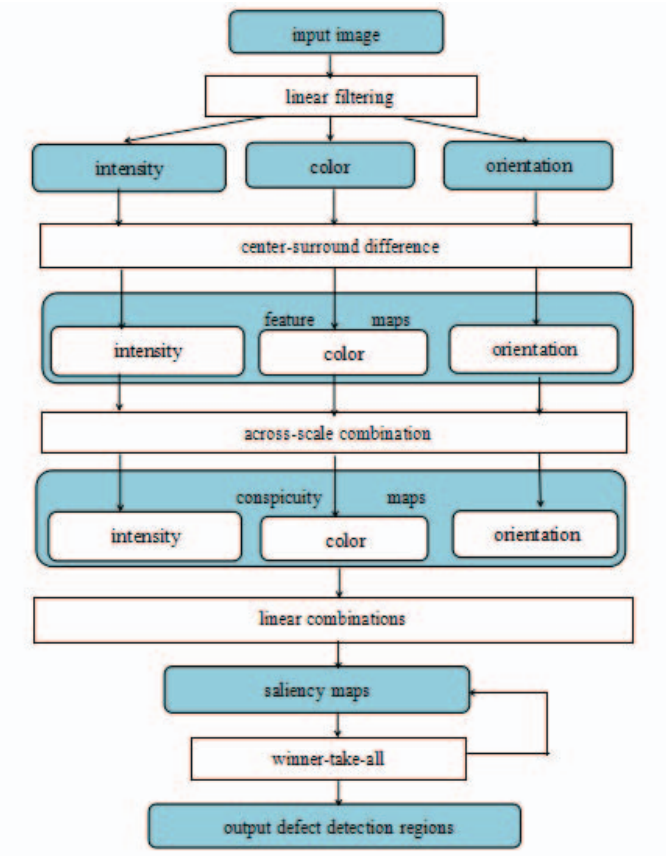
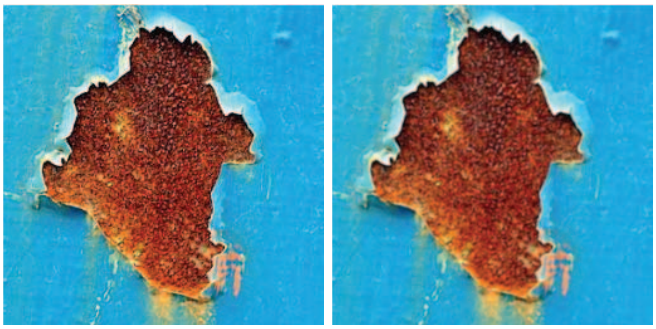


Fig. 1. Architecture of the BU model

II. The Description of Proposed Method

The visual attention mechanism is divided into two categories. One is the bottom-up approach, which is driven by



(a) before (b) after

Fig. 2. Gaussian filter function

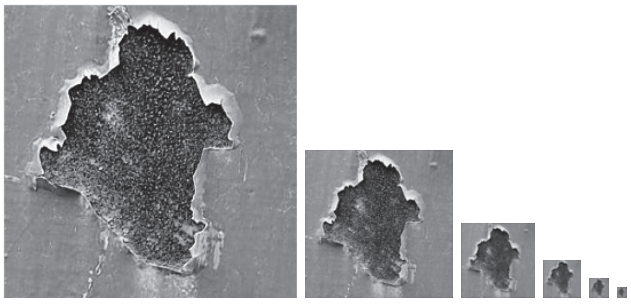


Fig. 3. the Gaussian pyramid of the intensity

data (scene-dependent). The other is the top-down approach, which is driven by task (task-dependent) [4]. In this paper, the first bottom-up approach is adopted. Initially, the low-level visual features such as colors, intensity and orientations are extracted from the input image. Furthermore, the feature maps are generated by using center-surround differences [5]. Additionally, the sub-saliency maps are combined by fusing the across-scale feature maps. At last the saliency map is acquired by combined all the sub-saliency maps and the saliency area is selected via winner-take-all mechanism [6].

A. Bottom-up Model

The model used in this paper is based on the BD model proposed by Itti in 1998 [7], in which the low-level features are extracted on several spatial scales by using the central perimeter filter structure of the biology. Three characteristic significance descriptions are generated by the combination of the feature maps, which are intensity, colors and orientations. And the saliency map shows the linear combination of these three characteristic significance descriptions, as shown in Fig1. The following is a detailed introduction to this model [8].

B. Calculation of the Gaussian Pyramid

The input image usually is an RGB image, which contains red (r), green (g), blue (b) three color channels, through their changes and superposition between each other to get a variety of colors [9]. Each color channel is divided into 256 levels of brightness, and 0 represents the darkest while 255 represents the brightest. All colors can be expressed as a mixture of three color channels in different proportions and intensities.

After inputting an image, we first need to do is linear filtering, and we use the Gaussian filter function because it is similar to the human eye model. And the Gaussian function see equation (1).

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{1}$$

Where x represents the abscissa of any point of the image, y represents the ordinate of any point of the image, and is the variance. And the initial image and the filtered image are as shown in Fig2.

After the image filtered, the Gaussian pyramid transformation is carried out. Through the method of line interpolation, we get a set of images with scale 0 to scale 5. We denote the scale as n, and the ratio of the scale n image to the original image is $1:2^n$. That means the scale 0 image is the same seize as the original image, and the scale 5 image is 32 times smaller than the original image. The smaller image reflects the background information of the original image, while the larger scale image represents the center information. We calculate the following three sets of features of the Gaussian pyramid:

$$GL(n) = (r+g+b)/3 \tag{2}$$

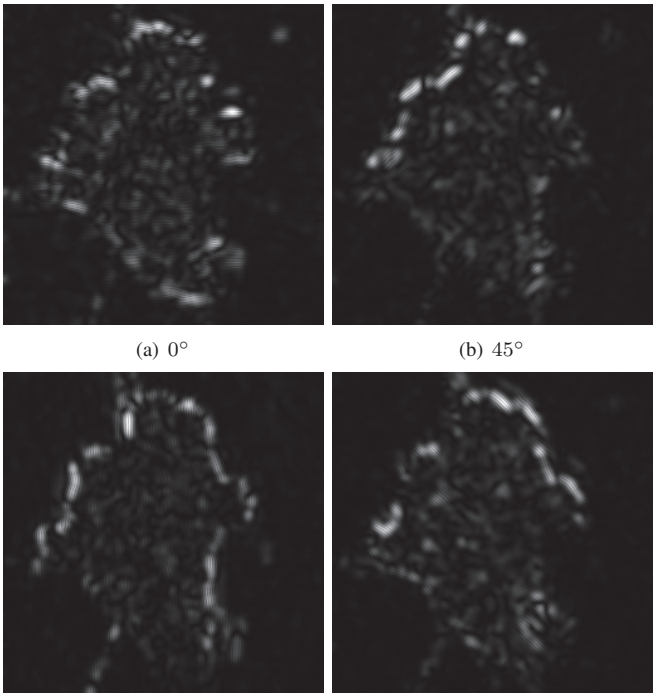
$$\begin{cases} GL(n) = (r+g+b)/3 \\ GC_R(n) = r-(g+b)/2 \\ GC_G(n) = g-(r+b)/2 \\ GC_B(n) = b-(r+g)/2 \\ GC_Y(n) = (r+g)/2-(|r-g|/2)-b \end{cases} \tag{3}$$

$$GO(n,\theta) = e^{\frac{-x'^2+r^2y'^2}{2\sigma^2}} \cos(2\pi\frac{x'}{\lambda}+\psi) \tag{4}$$

Equation (2) gets one group of Gaussian pyramid images about intensity-$GL(n)$, and Fig3. shows the Gaussian pyramid of the intensity; equation (3) get four groups about colors including red-$GC_R(n)$, green-$GC_G(n)$, blue-$GC_B(n)$ and yellow-$GC_B(n)$; Equation(4) is a Gabor function with good imitation of biological properties, where $x' = x\cos(\theta) - y\sin(\theta)$,$y' = -x\sin(\theta) - y\cos(\theta)$. We can get four groups images about orientation when $\theta \in \{0^o, 45^o, 90^o, 135^o\}$, the scale0 images of orientations different angles are shown in Fig4., (a) for 0^o,(b) for 45^o,(c) for 90^o and (d) for 135^o.

C. Get the Feature Maps

In order to obtain the feature maps of the local center and the surrounding background information, we operate the center-surround difference between the larger scale image and the smaller image, denoted as $\ominus$. The specific algorithm is to linearly interpolate the smaller scale image representing the surrounding background information so that it has the same size as the larger scale image representing the central information, and then perform the point-to-point subtraction operation. In this case, c represents the image scale of the center information and$c \in \{0,1\}$. while s represents the image of the surrounding information and $s = s+\delta, \delta \in \{2,3,4\}$. Therefore, there are six result images of each channel after



(a) 0° (b) 45°

(c) 90° (d) 135°

Fig. 4. Gabor filter function

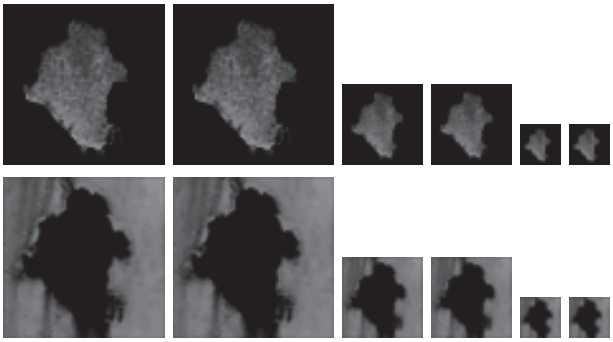


Fig. 5. Feature maps of colors

center-surround difference, denote $\{0 \ominus 2, 0 \ominus 3, 0 \ominus 4, 1 \ominus 3, 1 \ominus 4, 1 \ominus 5\}$.

$$FL(c,s) = |GL(c) \ominus GL(s)| \tag{5}$$

$$FC_{RG} = |(GC_R(c) - GC_G(c)) \ominus (GC_G(s) - GC_R(s))| \tag{6}$$

$$FC_{BY} = |(GC_B(c) - GC_Y(c)) \ominus (GC_Y(s) - GC_B(s))| \tag{7}$$

$$FO(c,s,\theta) = |GO(s,\theta) \ominus GO(s,\theta)| \tag{8}$$

Where $FL(c,s)$ denotes 6 feature maps of intensity; FC_{RG} and FC_{BY} denotes 12 feature maps of colors, which are shown in Fig5; and $FO(c,s,\theta)$ denotes 24 feature maps of orientations.

D. Get the Conspicuity Maps

Then,we will combine the comparative maps of all scales about every respective feature.However, if the comparative maps are combined directly, the salient targets could be covered by noise or by less-salient because the saliency is strong only on a part of the maps. Thus, normalization operation (defined $\aleph$) is carried out before combination for every comparative map. $\aleph$ is calculated as follows.

1) N: ormalize the values to a certain class $[\ 0...N\ ]$in the map;

2) F: ind N which are the global maximum in the comparative map and compute $(\bar{n})$ which are the average of all its other local maxima;

3) M: ultiply the entire image by $(N - \bar{n})^2$.

$$CL = \oplus N(FL(c,s)) \tag{9}$$

$$CC = \oplus[N(FC_{RG}(c,s)) + N(FC_{BY}(c,s))] \tag{10}$$

$$CO = \sum N[\oplus N(FO(c,s))] \tag{11}$$

Thus, three conspicuity maps can be obtained using (9) for intensity, (10) for color and (11) for orientation, as the Fig6 show.

(a) intensity (b) color (c) orientation

Fig. 6. Gaussian filter function

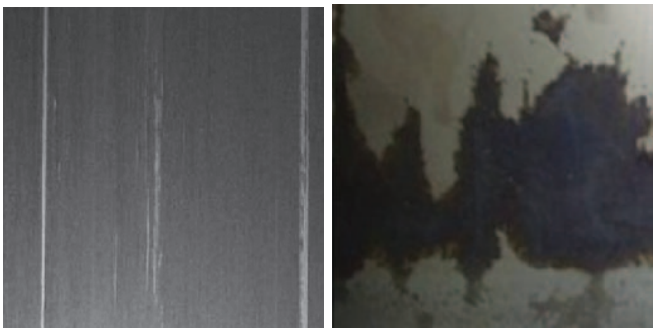
Fig. 7. Saliency Map

E. Get the Saliency Map

The saliency map are gained by combining all the conspicuity maps. It is noteworthy before combining that we will normalize the conspicuity maps by using $\aleph$ operation. Use Equation (12) to calculate the saliency map.

$$S = \frac{1}{3}(N(CL) + N(CC) + N(CO)) \tag{12}$$

As the Fig7 show, the saliency map calculates the saliency of a visual object based on the object area and its neighborhood, and takes the features of all scales into account, thus indicating the saliency of the image feature.



(a) (b)

(c) (d)

Fig. 8. Original images

(a) (b)

(c) (d)

Fig. 9. The saliency maps

III. Experiment Results and Analysis

In order to prove that our proposed method is effective, we conducted several tests on different plates with different defects, which appear with texture structures in different background. The original images as the Fig8 show. And the Fig9 is the relative saliency maps, which are gained by using BD model.

For different types of steel surface defects, such as shape, size, location, color and others, this way are all used in. From the shape of the defect, 8(a) is a vertical line, 8(b) is the irregular large area, 8(c) is the horizontal line,and 8(d) is irregular small area. It can be seen from Fig9 that this method can accurately find the location of defects, access to significant areas of information, and the method is not affected by the type, size and location and other factors of defects, size and location and other factors. However, there are some unsatisfactory parts. Comparing the defective regions identified by the computer and by the human eyes, we can see from Fig9 that the boundary line is not very smooth and like mosaic block due to linear interpolation, especially in the 9(c). In general, the BD model can achieve the requirements for detecting defective regions.

IV. Conclusion

In this paper, a detection method based on visual attention mechanism is put forward. This method makes use of the saliency map to quickly detect different types of defects on the surface of the steel plate to meet the real-time on-line inspection requirements and improve the detection efficiency. However, a large number of feature operations lead to slow algorithm, which means that enhancing the speed of the algorithm will be the next direction of research. We moreover hope we can get more information about the surface defects of the steel plate, such as the shape, size, area and even the types of defects, which is also the issue we will study in depth.

References

[1] A. Mohammed Al-Anezi, A. Tariq Al-Ghamdi, L. Waleed Al-Otaibi and M. Saad Al-Muaili, *Manufacturing, testing, and operational techniques to prevent sour service damages.* Elsevier Inc,2016.

[2] Xunwei Lin. *The Steel Surface Defects Region Detection Based On Visual Attention Mechanism.* Wuhan University of Science and Technology, 2013, 4: 3-26.

[3] Qian S, Chen D. *Discrete Gabor transform* IEEE Transactions on Signal Processing, 1993, 41(7): 2429-2438.

[4] JUKKA IIVARINEN, et al. *A Defect Detection Scheme for Web Surface Inspection.* International Journal of Pattern Recognition and Artificial Intelligence,2000, 14(6): 735-755.

[5] Koch C, Ullman S. *Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry.* Human Neurobiology, 1985, 4(4): 219.

[6] Li Fu. *Study on Interested Object Extraction in Video Image Based on Biological Visual Attention Mechanism.* China University of Petroleum, 2011: 11-22.

[7] L. Itti, C. Koch, E. Niebur. *A Model of Saliency-Based Visual Attention for Rapid Scene Analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.

[8] Jiang P, Gao T. *Paper Defects Detection via Visual Attention Mechanism.* Technical Committee on Control Theory, Chinese Association of Automation, 2011: 5852-5856.

[9] Milanese R. *Detecting salient regions in an image: from biological evidence to computer implementation.* University of Geneva, 1993.