The Fourteenth International Symposium on Artificial Life and Robotics 2009 (AROB 14th '09),
B-Con Plaza, Beppu, Oita, Japan, February 5 - 7, 2009

Estimation of the optimal image resolution using the SIFT

M. Yoshioka, Y. Maeda and S. Omatu

Osaka Pref. Univ.
1-1 Gakuen-tyo Naka-ku Sakai-city, Osaka Japan 599-8531
(Tel : 81-72-254-9279)

Abstract

The image processing is used to check products in many factories. If we use down-sampled images, we can reduce the calculation time and the image noise. However, the accuracy of the detection also becomes low. The purpose of this paper is to estimate the optimal image resolution for the detection, keeping the accuracy of detection high. To achieve our purpose, we adopt the SIFT(Scale Invariant Feature Transform)as the criterion of the optimal image resolution. Finally we confirm that our proposed method is useful by the simulation.

Keywords: SIFT, Lanczos, Normalized-correlation, MSE.

I. INTRODUCTION

Checking products in the factory, the image includes blur, noise and so on. Using the original size image, we need long computation time and the accuracy of the detection becomes lower on account of blur and noise. If we use down-sampled image, we can reduce the calculation time, blur and noise. However, the accuracy of the detection also becomes low and there are a few researches to estimate the optimal resolution for the detection. So, our purpose is to propose new criterion that is independent of input images and determined only by the template image. Using down-sampled images, we need to use the feature that is scale invariant in order to achieve our purpose. We adopt the SIFT[1] feature and improve the SIFT. We have proposed new criterion by using the SIFT feature, the normalized correlation and the MSE. Finally we have confirmed that our proposed method is useful by the simulation.

II. IMPROVEMENT ON THE SIFT

We compute the SIFT feature of the template and the input image and use them for estimation of the

Fig.1. Using simple algorithm

$$
\begin{aligned}
G(2x, 2y) &= I(x, y) \\
G(2x+1, 2y) &= (I(x, y) + I(x+1, y))/2 \\
G(2x, 2y+1) &= (I(x, y) + I(x, y+1))/2 \\
G(2x+1, 2y+1) &= (I(x, y) + I(x+1, y) + \\
&\quad I(x, y+1) + I(x+1, y+1))/4
\end{aligned}
\tag{1}
$$

$$G(x, y): Output \quad I(x, y): Input$$

optimal image resolution and the detection. In the first stage of the SIFT, the image is up-sampled and down-sampled by using simple algorithms which do not interpolate in order to reduce the computation time. The down-sampling algorithm is the 4 neighbor averaging and the up-sampling algorithm is defined as the above equation.

We show the example using simple algorithm in Fig.1(double scale). The image in Fig.1 is not high quality and loses some features, because we use algorithms which do not interpolate. Thus, we need to change simple algorithm for advanced algorithm which interpolate not to lose features. We adopt three famous interpolation algorithms (Lanczos3[2], Bicubic, Bilinear). We experiment with the best combination of up-sampling and down-sampling algorithms on 9 following images. The original resolution of these images are VGA and we use down-sampled images from 160x120 to 80x60[pixel]. We compute corresponded keypoints between original resolution images and down-sampled images by using SIFT feature and show the result following Table1.

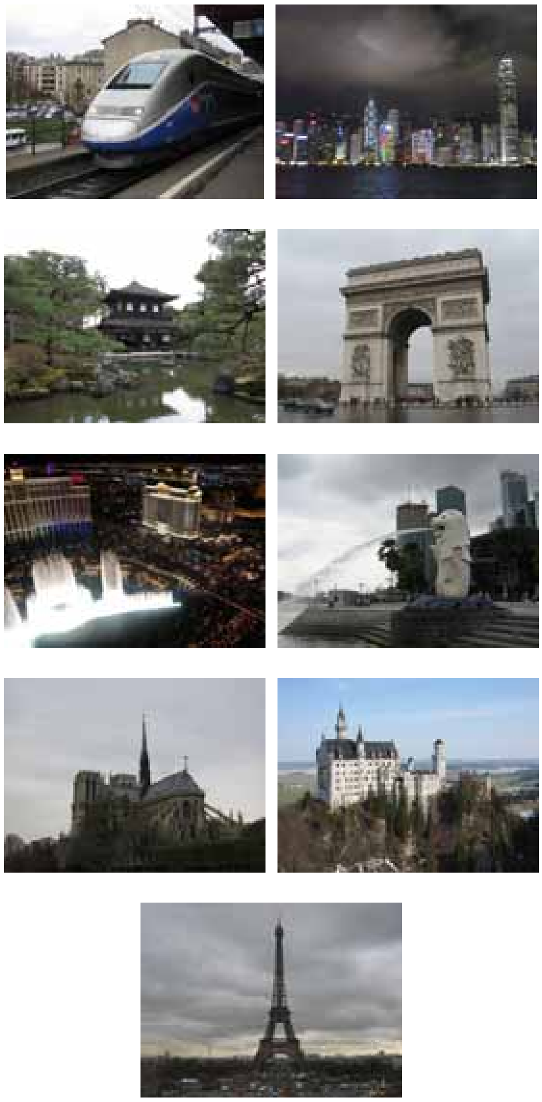

Fig.2. Using images

Table.1. Matching rate (%)

Up \ Down	Original	Bi linear	Bi cubic	Lanc zos3
Original	39.1	39.8	29.3	31.4
Bilinear	38.2	38.9	28.1	30.4
Bicubic	39.4	42.7	30.7	33.0
Lanczos3	39.5	42.9	32.9	34.5

As the results, the best combination is the bilinear as the up-sampling algorithm and the lanczos3 as the down-sampling algorithm and the matching rate is improved and 14% as Max and 3.8% as average are achieved. We use this improved SIFT in this paper.

As the resolution become lower, the number of SIFT keypoints and matching SIFT keypoints to original resolution changes.

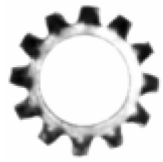

Fig.3. Template image

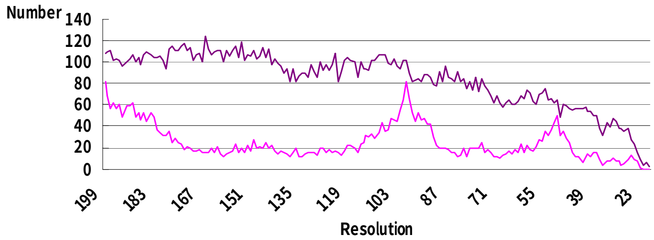

Fig.4. Number of keypoints and matching Keypoints of Fig.3

The original resolution of the template image is 200x200 [pixel]. The purple line shows the number of keypoints and the pink line shows the number of matching keypoints to ones in the original resolution in Fig.4.

The minimum resolution that has matching keypoints to ones in the original resolution is 23 x 23[pixel]. The input image of the SIFT is converted to double size, half size and …. of the original resolution in computing the SIFT feature. Thus, at 1/2 size and 1/4 size…. of the original resolution, there are many matching keypoints. From Fig3., purple and pink line do not show simple behavior. So, we need to use not only the SIFT feature but also other feature for proposing new criterion.

III. OUR PROPOSED METHOD

1. Normalized correlation and MSE

The SIFT feature is described by using neighbor gradients of the keypoint. So, the SIFT feature describes not global features of the image but local features of the image. Therefore we need to combine it with global features of the image to estimate the optimal image resolution. So, we adopt the normalized correlation and the MSE as the global feature. We compute them between original resolution of the template image and the down-sampled template image. There is a little difference in the normalized correlation between them and a large difference in the MSE between them.

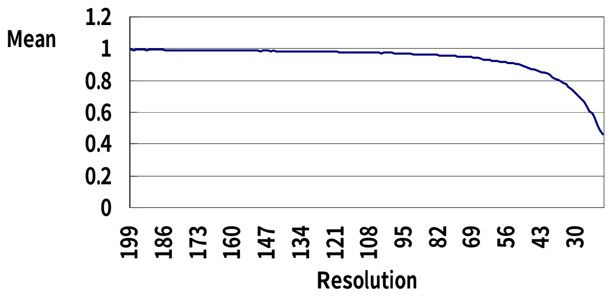

Fig.5. Mean between the NC and the 1-normalized MSE

So, we use the mean of the normalized correlation and the MSE. However, as the resolution becomes lower, the normalized correlation becomes lower and the MSE becomes higher. So, we reverse the MSE. We need to normalize the MSE because maximums of the MSE and the normalized correlation are not same. We normalize the MSE with the maximum of it. We compute the mean between the normalized correlation and (1 - normalized MSE). We use this mean as the global feature of the image. We show the value of mean between normalized correlation and (1 –normalized MSE) in Fig.5.

2. OUR CRITERION

In this section, we propose new criterion to estimate the optimal image resolution by using the SIFT feature, the normalized correlation and the MSE. We define the new fitness function:

$$F(x) = \left(0.9g(x) + 0.1\frac{x}{K}\right)\left(\frac{NC + (1 - normalizedMSE)}{2}\right) \quad (2)$$

$$g(x) = \left(1 - \frac{1}{1 + \left(\frac{x}{4}\right)^2}\right)$$

where x the number of matching keypoints to ones in the original resolution, K the number of keypoints and NC the value of the normalized correlation. The x / k means the matching rate, g(x) means the function of the number of matching keypoints and the mean of the normalized correlation and the reversed and normalized MSE means the criterion of the global features of the image. We show the behavior of g(x) in Fig.6. We use normalized value as the global features of image, so we use the normalized value as the local feature of the image. We adopt the function of the number of matching keypoints and matching rate as the local feature of the image because

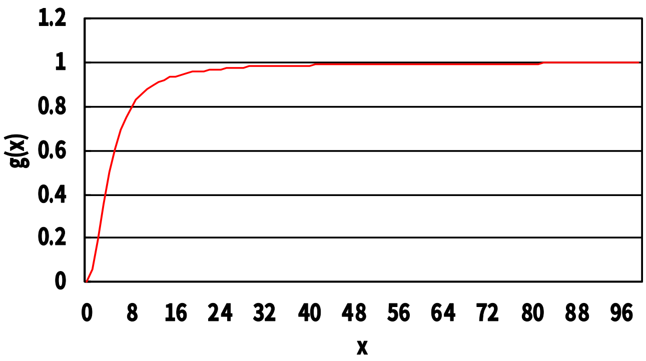

Fig6. g(x) the function of number of matching keypoints

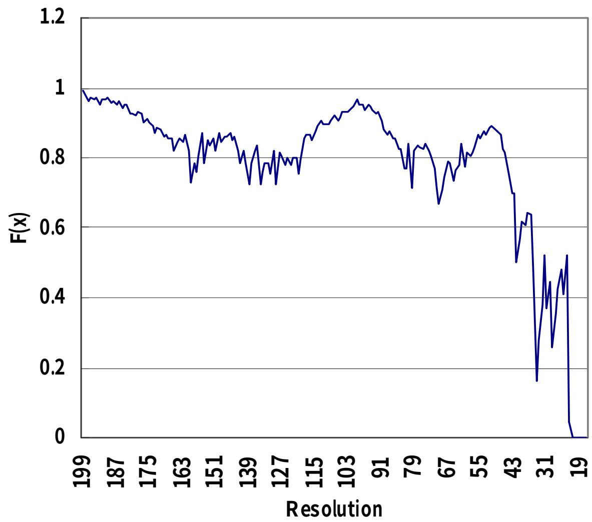

Fig.7 Fitness function of the optimal resolution

they mean how many local features the down-sampled image keeps. If there are two other resolution which have same matching keypoints, we select the resolution which has higher matching rate than another resolution. However the number of matching keypoints is more important than the matching rate. So, we weight the function of matching keypoints.

If the value of this fitness function is more than 0.4 at some resolution, that resolution is better. The minimum resolution where the value of our fitness function is more than 0.4 is the optimal resolution. We show the value of our fitness function with the template image in Fig.7. The behavior of our fitness function is not smooth. The half size, the quarter size and the 1/8 size images of the original resolution have more matching keypoints than other size images because the input image of the SIFT is converted to double size, half size, quarter size … of the original resolution in computing the SIFT feature. So, our fitness function has high value at 100, 50, 25.

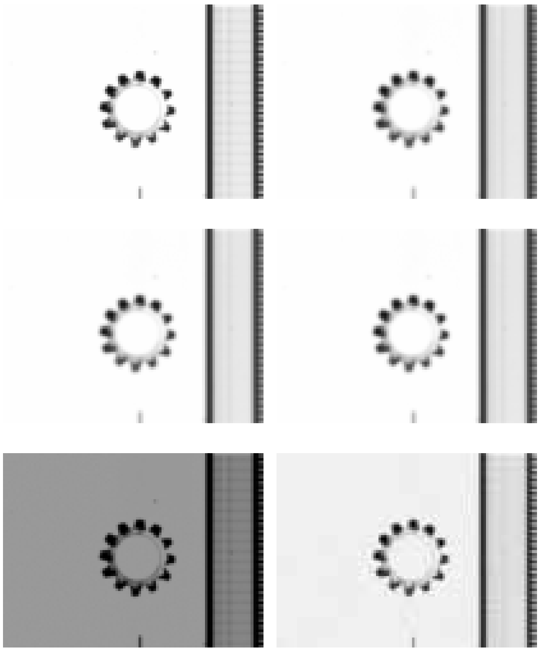

Fig.8. Input images

IV. SIMULATION

We show input images used for simulation in Fig.8. Original resolutions of them are VGA size. Checking products in factories, images include blur, noise and so on. So, we use not only the normal image but also three blur images, the noise image and the dark image as input . All input images are made from the normal image. Gaussian radiuses of three blur images are 5, 7 and 10. The dark image is 40 point darker than the normal image in brightness. The noise image is made by RGB diffusion and its parameters are 0.3 respectively. Based on our proposed criterion, the optimal image resolution is 24 x 24[pixel] in all resolution. So, down-sampling rate is 12%. The template image and input images are down-sampled to 12% and check whether we can detect or not. We show result in Table.2. In all input images, there are more than 5 matching keypoints, so we can detect the template from input images. We have confirmed that our proposed criterion is useful for not only the normal image but also blur, noise and dark input images. And we simulate with 11.5% down-sampled template and input images. There is no matching keypoints, so we can not detect the template from the noise image. Therefore, the down-sampling rate, at 12%, is the optimal and our proposed fitness function is effective. We use only one template image in this simulation. So, we need to simulate with more other template images and input images.

Table.2. Result of the detection (12%)

	Normal	Blur.5	Blur.7
Detection	○	○	○
	Blur.10	Dark	Noise
Detection	○	○	○

Table.3. Result of the detection (11.5%)

	Normal	Blur.5	Blur.7
Detection	○	○	○
	Blur.10	Dark	Noise
Detection	○	○	×

V. CONCLUSION

We have improved the SIFT feature by changing up-sampling and down-sampling algorithm and confirmed usefulness by the experiment. We have proposed the criterion that is independent of input images and robust. We have confirmed that our proposed method is useful by simulation. As the future work, we improve the SIFT feature and our criterion and use other image features such as edge.

REFERENCES

[1]David.G.Lowe, "Distinctive Image Feaures from Scale-Invariant Keypoints". International Journal of Computer Vision Vol.60,No.2,pp.91-110, 2004

[2]Ken.Turkowski. "Filters for Common Re-sampling Tasks" Graphic Gems I, Academic Press, pp. 147-165, 1990

[3]Mikio Takagi and Haruhisa Shimda, "Handbook of Image Analysis [Revised Edition]", University ofTokyo Press, 2004

[4]Philippe Thevanaz et al, "Image Interpolation and Resampling", Handbook f Medical Imaging, Processing and Analysis, I.N.Bankman, ED., Academic Press, San Diego CA, USA, pp. 393-420, 2000