An Improved SIFT Algorithm based on Multi-scale Homomorphic Filtering

Xiaoqing Dong[1] and Gengzhong Zheng[2]

[1]School of Physics and Electronic Engineering, Hanshan Normal University, Chaozhou, 521041, China
[2]School of Computer Information Engineering, Hanshan Normal University, Chaozhou, 521041, China
dxqzq110@163.com, 360313290@qq.com

Abstract

For the poor matching performance of traditional SIFT based algorithms under uneven illumination environments, a novel improved method based on multi-scale homomorphic filtering, named "multi-scale homomorphic filtering"+SIFT (MSHF+SIFT), is proposed in this paper. According to the good adaptability of the difference of Gaussian based homomorphic filtering algorithm for the illumination variation, in this paper, we use this algorithm to improve SIFT algorithm. In the proposed MSHF+SIFT algorithm, first, we extend the difference of Gaussian based homomorphic filtering algorithm to multi-scale space, named multi-scale homomorphic filtering(MSHF), and build a MSHF based pyramid. Second, use the MSHF based pyramid to replace the difference of Gaussian based pyramid of SIFT algorithm, so that the MSHF algorithm could be fused with SIFT algorithm effectively. And then, use original feature extraction and description method of SIFT to generate feature points. Finally, use distance-radio method for feature matching. Experimental results show that the proposed MSHF+SIFT algorithm has better robustness than traditional algorithms under various illumination variation environments.

Keywords: SIFT; Illumination Variation; Homomorphic Filtering; Multi-Scale Space

1. Introduction

In the process of large scale repetitive industrial production, some of the dangerous working environments which are not suitable for manual work or the situation that artificial vision is difficult to meet the requirements, machine vision is usually used to replace artificial vision. For example, in the production of chinaware, using machine vision to classify different types of chinaware can greatly reduce labor costs and improve production efficiency.

The machine vision is commonly used in some of the poor working environments. One of the main characteristics is that the light condition is relatively poor such that it cannot provide enough brightness and uniform illumination. This problem seriously affects the recognition rate and detection accuracy of the machine equipment, and become a difficult problem that restricts the applications of machine vision. So this paper attempts to start with the detection and recognition algorithm to improve the accuracy and adaptability of machine vision in the condition of poor light condition.

In recent years, algorithm based on local feature has become a research hotspot in the fields of computer vision, image recognition and image matching [1-5]. SIFT (Scale invariant feature transform) [6] algorithm is the more robust local feature algorithm, which has good characteristics of rotation and scale invariant, and has certain robustness for illumination variation, that has better matching performance than other local feature algorithms [7].Researchers proposed some improved algorithms for SIFT which can be summarized as follows:

1).In the aspect of recognition rate, Yurong Chen et.al[8] proposed an innovative Contourlet-SIFT algorithm, which could reduce the matching time and improve the matching robustness of the local similar regions by applying contourlet transformation to the feature and its neighborhoods, and only selecting the features of the top 1% smaller distance to be matched. Luigi Cinque et.al [9] improved the matching accuracy of SIFT algorithm by constraining matching points for their spatial position.

2).In the aspect of matching efficiency, Ke Y et.al[10] proposed the PCA-SIFT algorithm, which has good matching efficiency for the greatly reduced feature data by using PCA algorithm to reduce the dimensionality of the descriptor.

3).In the aspect of affine invariant, Kaiyang Liao et.al[11] used normalized elliptic field to solve the affine invariant. And Morel, J.M et.al[12] proposed the ASIFT algorithm which simulate latitude and longitude for achieving a complete affine invariant.

In summary, the above methods improve the performance of matching efficiency, recognition rate and affine invariant for SIFT algorithm, and achieve good results. However, the recognition performance of these algorithms including the SIFT are not good under the illumination variation environment. They calculate the descriptor using the local gradient information of key points, and normalize the length of the extracted feature vectors, so that they can eliminate the influence of the uniform illumination variation. But when the illumination changes greatly, it may produce different degrees of light and dark areas on the surface of matching target. And these illumination variations cannot be eliminated by normalizing the length of feature vectors, which would greatly influence the extraction of features. To the best of our knowledge, there is no improved algorithm of SIFT under the condition of uneven illumination, so we will try to improve the performance of the SIFT algorithm under the uneven light environment and apply it to the machine vision in the light of uneven environment to improve the adaptability of the equipment.

Usually case, the spectrum of the uniform light variation is mainly located in low frequency band, and the spectrum frequency of the matching target which has a certain structure and edge may be higher, so the traditional homomorphic filtering methods filter light components by high pass filter [13,14].But Tan X Y et.al[5] indicated that some significant informations both exist in the high and low frequency band, and completely eliminating the low-frequency components and retaining the high frequency components are not appropriate. Also, the high frequency components may be mixed with some noise. Based on this, fan et.al [14,15] proposed a homomorphic filter based illumination normalization algorithm, which uses difference of Gaussian filter (DoG) [16] instead of high pass filter to obtain the required bandpass filter, and experimental results shown that this method effectively reduces the influence of illumination variation and enhances the image edges and details.

Based on the good adaptability of the improved homomorphic filtering algorithm[14,15] for the illumination variation, in this paper, we extend it to multi-scale space [17-21], and put forward a difference of Gaussian based multi-scale homomorphic filtering algorithm, named multi-scale homomorphic filtering (MSHF) and construct a pyramid based on MSHF. At the same time, because the gray dynamic range of the image is small after multi-scale homomorphic filtering, we apply gray enhancement to the image for a better classification.

Since the proposed MSHF and the SIFT algorithm both use the difference of Gaussian filter to construct multi-scale space, we can fuse these two algorithms, *i.e.* MSHF instead of DoG to construct multi-scale space. Then, we propose a MSHF+SIFT method, which uses MSHF to construct pyramid and the feature extraction and feature descriptor of SIFT to generate feature, to eliminate the influence of large scale illumination changes meanwhile retaining good qualities of SIFT algorithm.

Our contributions in this paper are as follows:

1).We extend homomorphic filtering algorithm to multi-scale space, and put forward a

difference of Gaussian based multi-scale homomorphic filtering algorithm, named multi-scale homomorphic filtering (MSHF).

2).We propose a novel method named MSHF+SIFT, using MSHF to construct pyramid , and the feature extraction and feature descriptor of SIFT to generate features.

3).Experiments prove that the proposed algorithm effectively extract more stable feature points and achieve higher matching rate in uneven illumination environments.

The rest of this paper is organized as follows: THE homomorphic filtering and SIFT algorithms are introduced in section 2. In section 3, we present the proposed MSHF and MSHF+SIFT algorithms. The experimental results and analysis are provided in section 4. Finally, a conclusion is drawn in section 5.

2. Related Works

2.1. Homomorphic Filtering

2.1.1. Traditional Homomorphic Filtering

The reference [13] proposed a simplified Lambertian illumination model, and eliminated the influence of illumination change according to this model. The model is as follow:

$$f(x, y) = i(x, y) \cdot r(x, y) \tag{1}$$

Where $f(x, y)$, $r(x, y)$ and $i(x, y)$ denote the target image, reflection function and illumination function(outside light source), respectively. From the equation, we know that we can eliminate the influence of illumination variation by the elimination of $i(x, y)$. Figure 1 shows the commonly used process of homomorphic filtering [13]:

$f(x,y)$ → ln → DFT → $H(u,v)$ → $(DFT)^{-1}$ → exp → $g(x,y)$
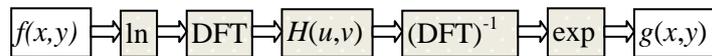
Figure 1. Flow Chart of the Traditional Homomorphic Filtering

As shown in Figure 1, the traditional homomorphic filtering includes the following steps:

1). Transform the multiplicative relation between the reflection function and the illumination function into additive relation by applying Logarithmic operation on both sides of Equation (1).

2). Transform the functions from spatial domain to frequency domain by using Fourier transformation. Then eliminate the low frequency light component $i(x, y)$ by using frequency domain enhancement function $H(u, v)$ to multiply both sides of the equation.

3).Finally, using inverse Fourier transformation to transform the image to the spatial domain, we can obtain the target image which has been eliminated by illumination variation.

2.1.2. DoG based Homomorphic Filtering

Frequency domain enhancement function $H(u, v)$ [14] based on difference of Gaussian filtering can be denoted as following:

$$H(u,v) = (\gamma_H - \gamma_L)[1 - e^{-D^2(u,v)/2D_0^2}] + \gamma_L \tag{2}$$

Where D_0 is the cut-off frequency, $D(u, v)$ is the distance from point (u, v) to the origin of frequency plane, the values of γ_H and γ_L determine the slope of the function $H(u,v)$ shown as Figure 2.
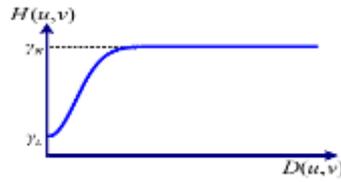
Figure 2. Profile of the Homomorphic Filtering

From Figure 2, we can know that γ_L is related to the low frequency suppression, and the smaller the γ_L, the greater the low frequency suppression. At the same time, high frequency components are in proportion to the γ_H. So, selecting the appropriate γ_H and γ_L, we can eliminate the influence of light variation, meanwhile enhancing the edge structure information of the target object.

Figure 3 shows the images processed by Homomorphic filtering. It can be observed from original image (a) that the images are with large illumination variation, and the last three image are too dark and with uneven brightness. After the treatment of Homomorphic filtering shown as image (b) of Figure (3), we can see that the phenomenon of dark and uneven has been effectively improved .However, the gray dynamic range is a little small.

(a) Original images

(b) Images after Homomorphic filtering
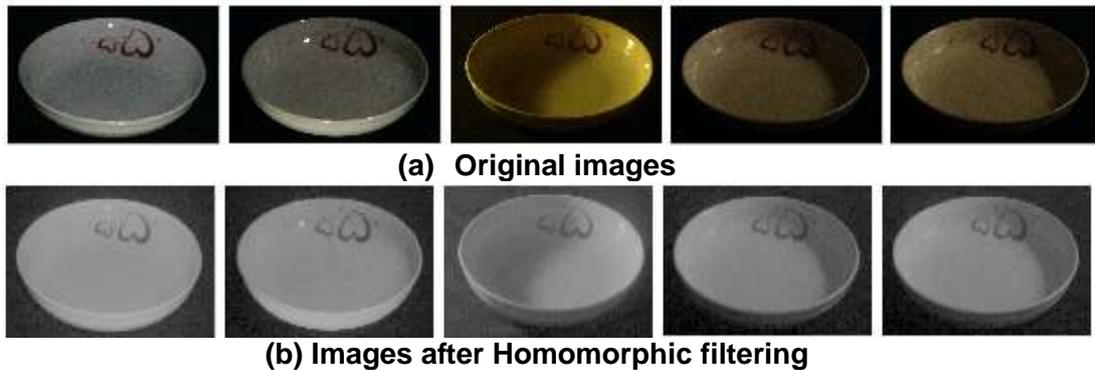
Figure 3. Images Processed by Homomorphic Filtering

2.2. SIFT Algorithm

SIFT algorithm mainly includes four key steps: construction of scale space, extraction feature point, generation of feature descriptor, matching feature points[6].

2.2.1. Construction of DoG based Scale Space

SIFT algorithm uses Gauss convolution kernel to construct the scale space *L*, defined as follow:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

Where $G(x, y, \sigma)$ is the scale variable Gauss function, and different scales are corresponding to different features. In order to reduce the computation time, SIFT uses the difference of Gaussian scale space, which is defined as follow:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{4}$$

Equation (4) executes a convolution between the input image and the Gauss difference kernel of the adjacent scale. The following figure is a gauss image pyramid with *O* group (octave), and each group has *S* layers (interval / 1evel).
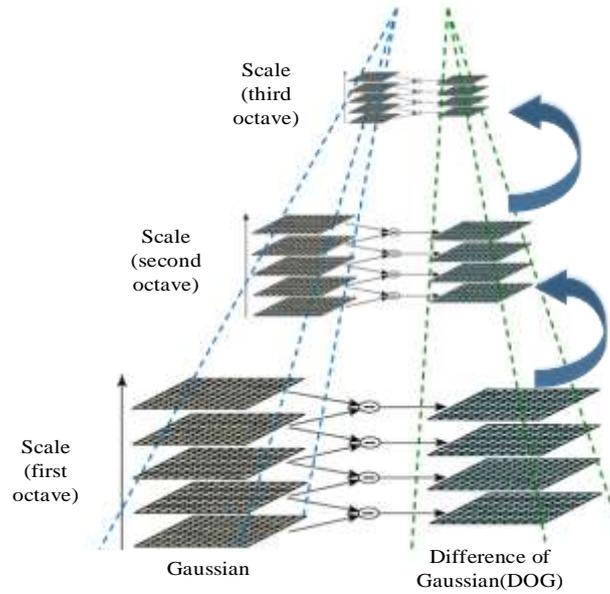
Figure 4. Gauss Image Pyramid

2.2.2. Extraction of Feature Point

The feature extraction of SIFT mainly has 3 steps, *i.e.* the detection of the extreme points in scale space, the location of the extreme points and orientation of the key points.

1).First, the extreme points should be detected both in the scale space and image spatial domain, *i.e.* the detecting point should be compared with all the adjacent points, as shown in Figure 5, a total of 26 adjacent points.
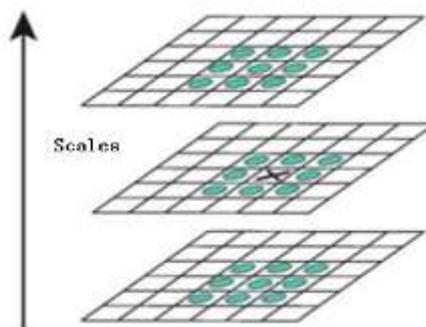
Figure 5. Detection of Extreme Point in Scale Space

2).Second, after detecting the extreme points, use fitting surface of quadratic Taylor expansion of the DoG function $D(x, y, \sigma)$ to determine the exact position and scale of the extreme points. Quadratic Taylor expansion of $D(x, y, \sigma)$ is as follow:

$$D(x) = D + \frac{\partial D^{\mathrm{T}}}{\partial x} x + \frac{1}{2} x^{\mathrm{T}} \frac{\partial^2 D}{\partial x^2} x \tag{5}$$

3).Finally, calculate the gradient magnitude and gradient direction of each key point, respectively:

$$m(x, y) = \sqrt{A^2 + B^2} \tag{6}$$

$$\theta(x, y) = \arctan(\frac{B}{A}) \tag{7}$$

where

$$A = L(x+1, y) - L(x-1, y) \tag{8}$$
$$B = L(x, y+1) - L(x, y-1) \tag{9}$$

After obtaining the gradient magnitude and direction, the key point will have the characteristics of rotation invariance.

2.2.3. Generation of the Feature Descriptor

In scale space, take a 16 x 16 area with the key point as its center, and divide this area into four 4*4 blocks which have 16 small cells. Then generate a seed point by calculating the gradient vector histogram and accumulate the gradient direction of each cell, shown as Figure 6.Usually, we use a key point with 16 seed points, each of which has 8 directions, to generate a 128 dimensional feature descriptor.
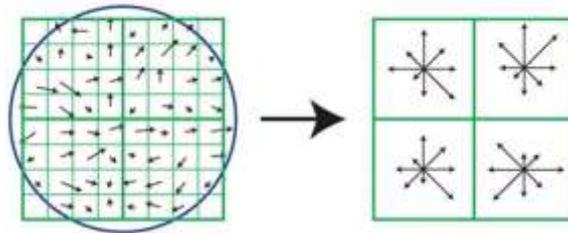
Figure 6. Generation of Feature Descriptor

2.2.4. Feature Point Matching

SIFT algorithm uses the Euclidean space to measure the distance of the key points, and the distance-radio criterion [6], shown as Eq(10), is as similarity criteria for the key points.

$$\begin{cases} d_{min} / d_{n\text{-}min} < \text{threshold} & \text{success} \\ & else \end{cases} \tag{10}$$

Where d_{min} and $d_{n\text{-}min}$ denote the first and second nearest distances. As shown in Equation (10), it would be true if the value of $d_{min} / d_{n\text{-}min}$ is less than a certain threshold. It means that the distance of the nearest two key points is much smaller than the second nearest, *i.e.* the two key points which match successfully are very similar while the distinctions among the target point and the other key points are quite big.

3. The Proposed Algorithm

Based on the characteristics of difference of Gaussian based homomorphic filter, we try to extend it to multi-scale space, and propose difference of Gaussian based multi-scale homomorphic filtering algorithm(MSHF). Then use the MSHF to construct homomorphism filter pyramid, and fuse it with the SIFT algorithm, *i.e.* using the MSHF multi-scale space instead of the DoG multi-scale space. And put forward an improved SIFT method based on the homomorphic filtering pyramid, named MSHF+SIFT. In this paper, we expect that the proposed algorithm could effectively eliminate the influence of most of the illumination variation and enhance the robustness of SIFT algorithm under illumination variation environment by using the MSHF multi-scale space.

The proposed method mainly consists of two parts. The first part is construction of MSHF multi-scale space, and the second part is the fusion of MSHF algorithm and SIFT algorithm.

3.1. Construction of MSHF Multi-scale Space

3.1.1. The Principle of MSHF Multi-scale Space

From the Equation (2), we know that the cut-off frequency D_0 denote the factor of image scale space. When the cut-off frequency is small, there would pass more low frequency components, and the frequency band is wide. When the cut-off frequency D_0 is big, the frequency band is narrow and only the high frequency component can pass. So, using some different cut-off frequencies, we can form the MSHF multi-scale space which is denoted as follow:

$$H\left(u,v\right) = (\gamma_H - \gamma_L)[1 - e^{-D^2(u,v)/2D(o,s)^2}] + \gamma_L \tag{11}$$

Where different scales corresponding to different values of $D(o, s)$ which can be denoted as follow:

$$D(o,s) = D_0 2^{o+s/S}, o \in [0,\dots,O-1], s \in [0,\dots,S-1] \tag{12}$$

Where D_0 is cut-off frequency of the reference layer, O is coordinate of the octave (Group), s is the coordinate of sub-level (layer). Shown as the above Equation (12), the distance of each octave is S, and each layer in the same octave is separated by a constant $2^{1/S}$. The selections of specific parameters of the MSHF multi-scale space are as section 3.1.2.

3.1.2. The Determination of Parameters for Multi-scale Space

(1). The Determination of Cut-off Frequency $D(o,s)$

From the Equation (12), we know that the Cut-off frequency is determined by three parameters: O (octave or group coordinates), S (sub-level or layer coordinates) and D_0 (the cut-off frequency of reference layer).

The key of matching successfully is the number of the matched points, which is increasing with the increase of S. So, in theory, selecting the large scale for sampling can increase the matching success rate. But, it may take more matching time. In this paper, the sampling scale is $S=3$ for each group, and the group number O is 2. Actually, due to the evaluation of extreme value, we need to compare with the two adjacent scales, and in fact it has 5 layers for S.

(a). Original Images

(b) Processed Images
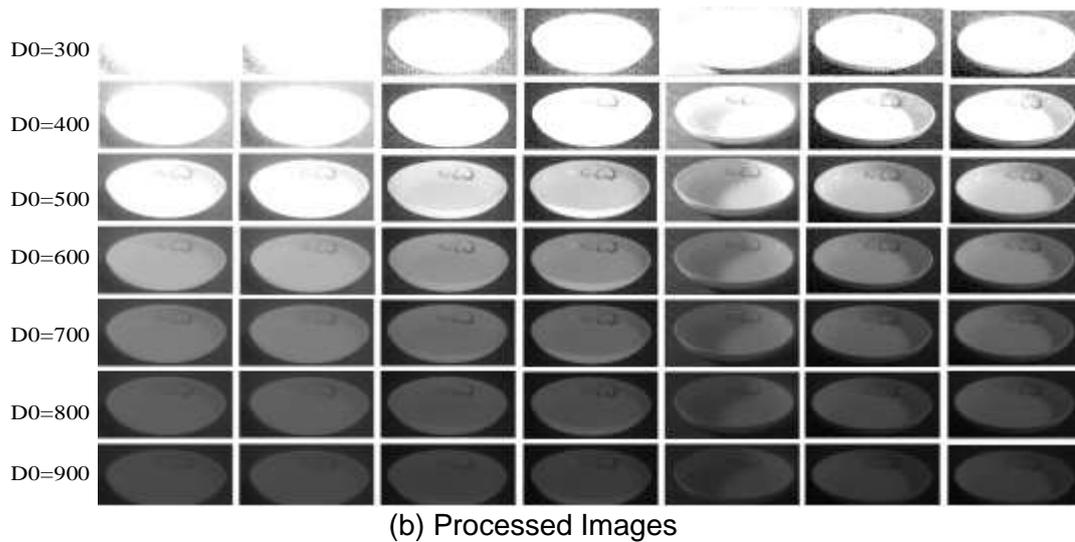
Figure 7. Processed Results of Different Cut-off Frequencies

Figure 7 shows the processed results after applying the homomorphic filtering algorithm with different cut-off frequencies of different reference layers under different illumination conditions. Where γ_H and γ_L are 0.5 and 0, respectively. It can be observed from the figure that the gray dynamic range is too small. And the matching objects are relatively bright for $D_0<600$. While $D_0>600$, the entire images are a little dark. So, in this paper, we select the $D_0=600$ as the cut-off frequency of reference layer because of its appropriate brightness and contrast.

(2) The Determination of Parameter γ_H

The slope of the function $H(u,v)$ is controlled by the parameters γ_H and γ_L, and appropriate γ_H and γ_L can eliminate the influence of illumination variation. From Figure 2, we know that，the smaller the γ_L ,the greater the low frequency suppression. At the same time, it is more able to enhance the high frequency components of the image when increase the γ_H . However, we should know that the value of γ_H is not the bigger the better, because part of the noise is mixed in the high frequency.

For selecting an appreciate value of γ_H , let $D_0=600$ and $\gamma_L=0$, and vary the value of γ_H .It can be observed from Figure 8 that the dynamic range of gray level of the images is small and the images are dark when $\gamma_H<0.5$. When $\gamma_H>0.5$, the image contrast gets better, but many high-frequency noises exist. And we obtain the best result when $\gamma_H=0.5$.

(a). Original Images

rh=0.2

rh=0.3

rh=0.4

rh=0.5

rh=0.6

rh=0.7

(b). Processed Images
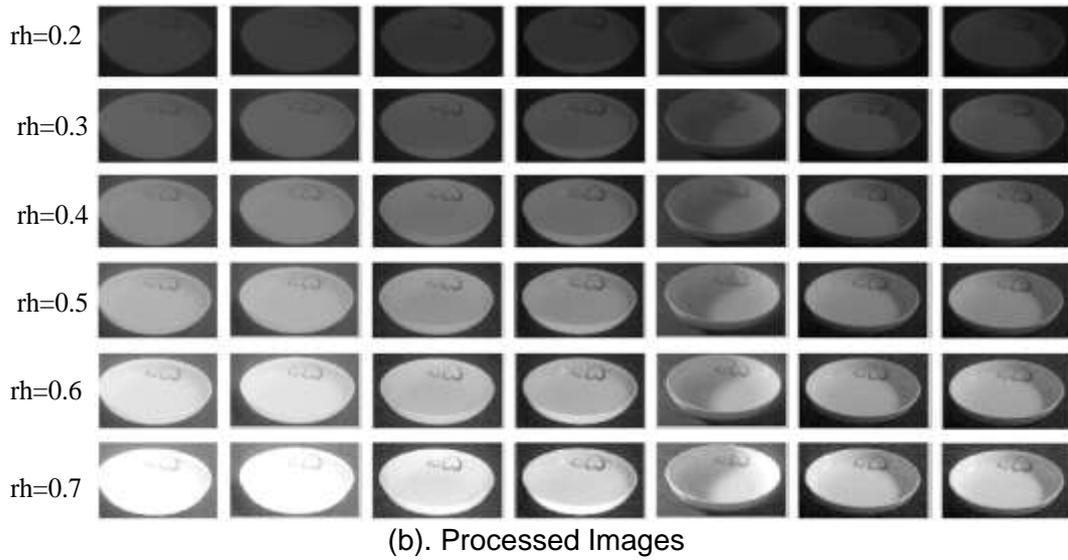
Figure 8. Processed Results of Different γ_H

(3). The Determination of Parameter γ_L

Similar to the selection of γ_H, let D_0=600, $\gamma_H = 0.5$, and vary the value of γ_l for obtaining the appreciate γ_l. We can observe from Figure 9 that we can have the best result when $\gamma_l = 0$.

(a). Original Images

rl=0

rl=0.1

rl=0.2

(b). Processed Images
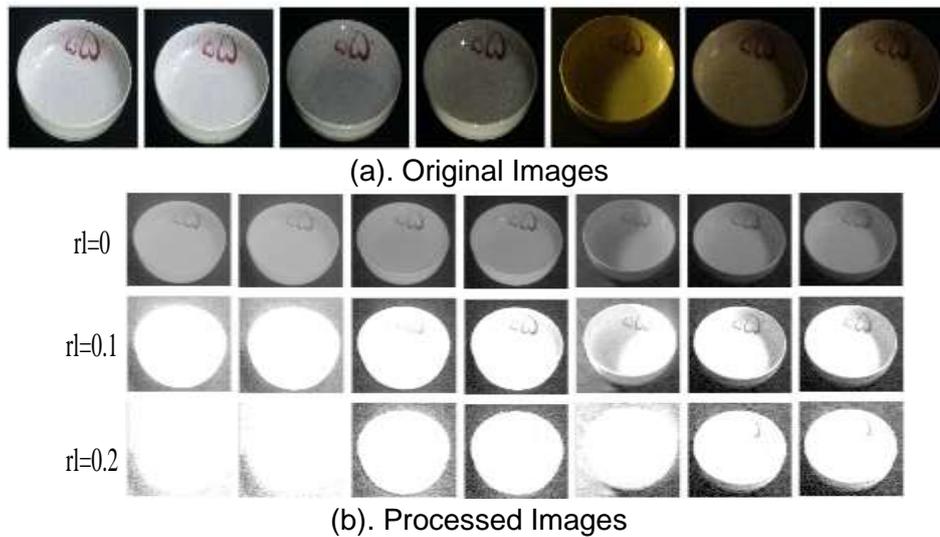
Figure 9. Processed Results of Different γ_l

From the above experimental results, we finally select D_0=600, $\gamma_H = 0.5$, $\gamma_l = 0$ as the optimal parameters to construct MSHF multi-scale space in this paper.

3.1.3. Gray Enhancement

Through the above experiments, we can select the optimal parameters for the multi-scale space. We can see that the problems of dark and uneven have been effectively improved, and the light distribution is more even after the treatment of Homomorphic filtering. However, just shown as Figure 10, we can see that the images of Figure 10.(a)

have poor contrast ratio, and their gray ranges are just a little small shown as gray histogram of Figure 10.(b), and mainly concentrated in the low gray area.

(a).Images Processed by Homomorphic Filtering
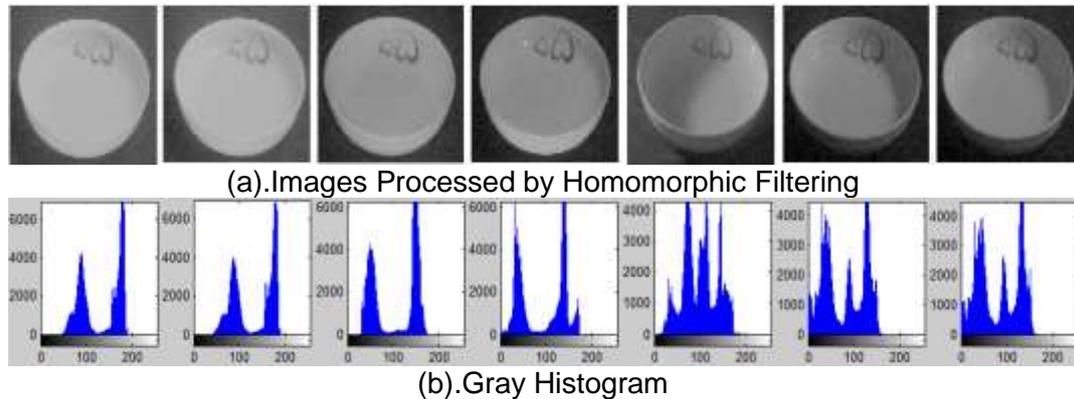
(b).Gray Histogram

Figure 10. Gray Histogram of Images Processed by Homomorphic Filtering

In order to enhance the contrast of the images, we use the following method to extend gray scale of the images.

Step1. Get the gray histogram H of the image.

Step2. Extract the main gray range R of gray histogram H. Just shown as Figure 11, the main gray range R is from 40 to 190.
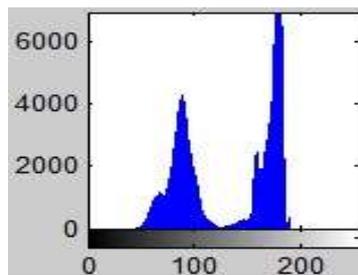
Figure 11. Gray Histogram

Step 3.Extend the gray range R to a greater range shown as Figure 12.
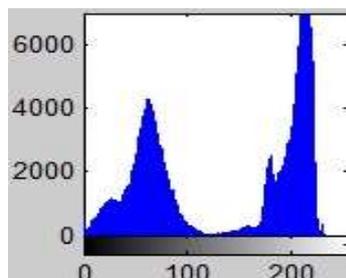
Figure 12. Gray Enhancement

By using the above gray enhancement algorithm, we can observed from Figure 13 that the contrasts of the images are greatly improved and the images are more clear, and their gray range increased which mainly distribute from 0 to 255.

(a).Images Processed by Gray Enhancement

(b).Gray Histogram after Gray Enhancement
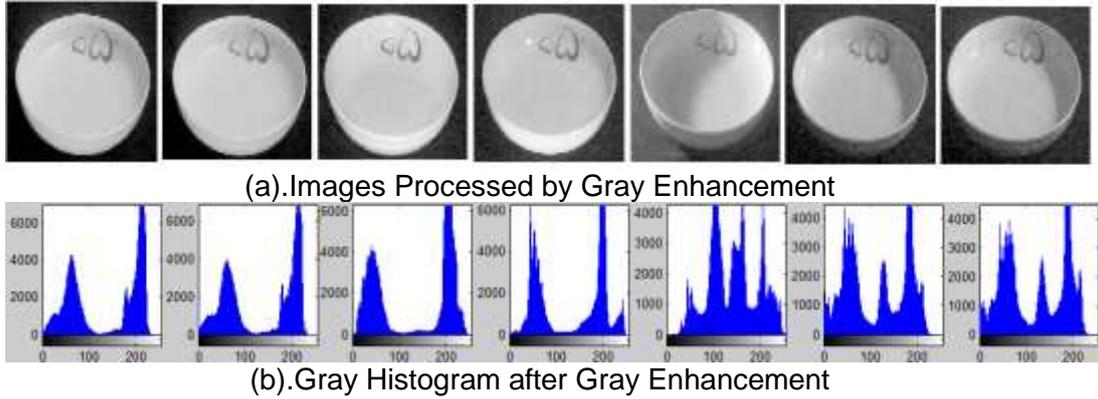
Figure 13. Images and Gray Histogram after Gray Enhancement

3.1.4. MSHF Pyramid

Figure 14 shows a DoG based MSHF Pyramid, which has 4 groups (octave), each group with 5 layers (1evel / interval), and the next group is obtained by down sampling to the last set. The MSHF Pyramid is given as Figure 14.
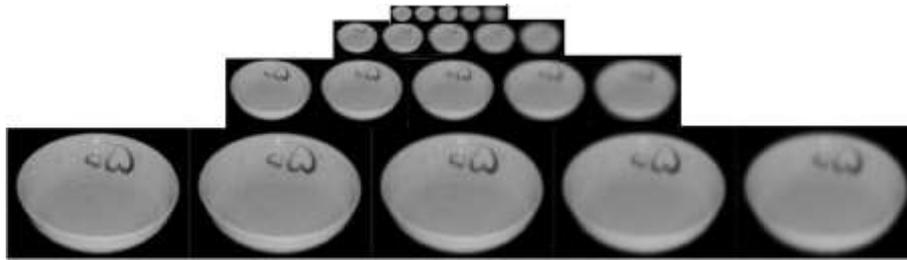
Figure 14. MSHF based Pyramid

3.2. Fusion of MSHF and SIFT (MSHF+SIFT)

In this paper, we use the MSHF based multi-scale space to replace the original scale space, and effectively fuse with SIFT algorithm. We use the feature extraction method of SIFT shown as section 2.2 to extract feature in the multi-scale space, including the detection of extreme point, determination of location and scale of the extreme point and calculation of the gradient magnitude and direction for the key points.

Similarly, we still use the generation method of feature descriptor of SIFT to generate feature descriptor for the MSHF+SIFT method. In the scale space, take a 16 x 16 area with the key point as its center, and divide this area into four 4*4 blocks which have 16 small cells. Then generate a seed point for calculating the gradient vector histogram and accumulate the gradient direction for each cell. Usually, we use a key point with 16 seed points, each of which has 8 directions, to generate a 128 dimensional feature descriptor. Figure 15 shows the extracted feature points.
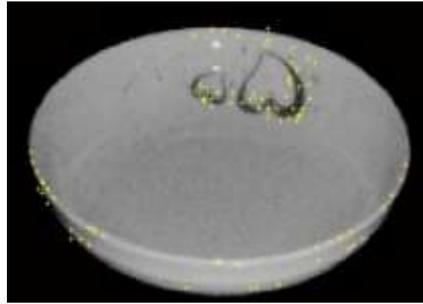
Figure 15. Feature Extraction

Finally, the proposed method use LOWE's distance-radio criteria as the Matching method.

4. Experimental Results and Analysis

In order to test the performance of the algorithms under illumination variations environment, we use some chinaware under varying illumination environment as the testing sample to compare the proposed algorithm MSHF+SIFT with SIFT, Contourlet-SIFT, PCA-SIFT and ASIFT algorithms. The parameters for the proposed algorithm are described in section 3.1.2, and the parameters of the other algorithms are the same as the corresponding references[6,8,10,12]. We used MATLAB2010 for all of the experiments and performed on a PC with Intel(R) Dual(TM)i5-4210U CPU, 4GB memory and windows 10 operating system.

4.1. Introduction of the Testing Samples

The matching and sorting of chinaware, which are usually finished by workers, are important parts of a chinaware production line. We attempt to replace the artificial realization by machine vision. Due to the relatively poor working environment of chinaware production, it is difficult to provide a good lighting environment so that the recognition software is required to have a good adaptability to light environment. We select 9 kinds of typical chinaware as the testing samples, each with 7 images under different illumination, and 63 images in total. The following images are some of the testing samples:

Figure 16. Chinaware under Different Illumination Environments

Figure 16 shows part of the samples, which are under different illumination environments. It can be observed from the figure that, some of the images are with good light condition, moderate brightness and uniform light. Some of the images are with

uneven illumination, whose left side or right side appears some shadows. Some are too dark that it is very difficult to see clearly. We expect that the proposed algorithm can have good recognition performance under these harsh environments.

4.2. Performance Comparison of Related Algorithms

In this paper, we fuse the different of Gaussian based homomorphic filtering algorithm and the SIFT algorithm, and expect that the good performance of homomorphic filtering algorithm for uneven illumination could improve the light adaptability for SIFT algorithm. In the following section, we conduct the following experiments to test the extraction of the stable feature points and recognition rates of the related algorithms.

4.2.1. Experiments of Feature Point Extraction

In these experiments, use 9 different kinds of chinaware, each with 7 images and 63 images in total, as the testing samples to test the performance of the extraction of the stable feature points for the related algorithms.

In the first experiment, we compare the related algorithms with 9 kinds of chinaware, which are all under 7 different illumination environments. The numbers of feature points are the average numbers which are extracted under these 7 different illumination environments.

As shown in Table 1, the numbers of extracted feature points are different depending on the size or shape of the chinaware. Generally speaking, chinaware with large size and complex structure is generally able to extract more feature points. Also, we can observe from Table 1 that, the SIFT, PCA-SIFT and Contourlet-SIFT algorithms extract the same number of key feature points, since these three algorithms use the same feature point extraction method. The differences among them are that, the PCA-SIFT algorithm applies PCA to the extracted feature points to reduce the dimensionality and data quantity. The Contourlet-SIFT applies contourlet transformation to the extracted feature points in the scale space, then use contourlet global texture to reduce the number of matching points. ASIFT algorithm transform the image by using simulation of all possible affine deformation, but the collected chinaware samples in the experiments are almost from the same angle and have no deformation so that the ASIFT cannot extract more feature points than the previous three algorithms. The proposed method extract maximum feature points in all kinds of chinaware, because it overcomes the influence of uneven illumination by using DoG based homomorphic filtering and enhances contrast of the image by using the gray enhancement algorithm.

Table 1. The Numbers of Key Feature Points Extracted from Different Kinds of Chinaware

Chinaware / Algorithm	1	2	3	4	5	6	7	8	9
SIFT	278	260	159	201	134	107	98	103	96
Contourlet-SIFT	278	260	159	201	134	107	98	103	96
PCA-SIFT	278	260	159	201	134	107	98	103	96
ASIFT	276	261	165	194	150	103	100	100	99
MSHF+SIFT	324	287	180	211	165	124	118	143	142

In the second experiment of feature point extraction, we compare the key point extraction performance of the related algorithms under different illumination environments. There are 7 different illumination environments, and 9 different types of chinaware in each environment. The numbers in Table 2 are the average numbers of the

feature points which are extracted from these 9 different types of chinaware.

Table 2 shows the number of extracted feature points under 7 different illumination environments. The environment 1 denotes the ideal illumination environment. The environments of 2,3 and 4 denote the little dark environment. The environments of 5 and 6 are the environments with light source in the left or right side. The environment 7 is the very dark environment and also with light source in the left or right side. It can be observed from Table 2 that, all algorithms extract maximum stable feature points under the ideal light environment.As mentioned in the first experiment, the SIFT, PCA-SIFT, and Contourlet-SIFT use the same method of feature point extraction, so they have the same number of extracted key feature points. The maximum number and the minimum number extracted under the ideal light environments and the worst light environment are 221 and 84, respectively. So the difference between the maximum number and the minimum number is up to 137. While the difference of the ASIFT algorithm is 133. Comparing to the proposed method in this paper, whose maximum number and minimum number extracted under the ideal light environments and the worst light environment are 229 and 144 respectively, the difference is only 85. It means that the differences of numbers of feature points extracted by the proposed method under various illumination conditions are not large, while the other algorithms could only extract enough feature points under the good light environments. So the proposed method has better robustness to illumination variation than the other algorithms.

Table 2. The Number of Key Feature Points Extracted under Different Illumination Environments

Illumination environments / Algorithms	1	2	3	4	5	6	7
SIFT	221	201	195	198	106	110	84
Contourlet-SIFT	221	201	195	198	106	110	84
PCA-SIFT	221	201	195	198	106	110	84
ASIFT	223	196	199	194	105	112	90
MSHF+SIFT	229	211	209	210	154	158	144

4.2.2. Experiment of Recognition Rate

To compare the recognition performance of the related algorithms under different light environments, we designed two experiments.

In the first experiment, in order to test the recognition performance of the related algorithms under little dark environment, the images under ideal light environment is chosen as the training sample, and the images from illumination environments of 2,3 and 4 are chosen as the testing samples, so there are 9 training samples and 27 testing samples in total. Similar to the first experiment, in order to test the comprehensive recognition ability of the related algorithms under a variety of different light environments, we also chose the images under ideal light environment as the training sample for the second experiment, while the images from the other 6 illumination environments are as testing samples, so there are 9 training samples and 54 testing samples in total.

(1).Sub-experiment 1

As shown in Table 3, in addition to the PCA-SIFT algorithm, the other algorithms in Sub-experiment 1 achieve higher recognition rates that have reached more than 80%. Especially the proposed method, only 2 of the 54 testing samples are wrong and the recognition rate is up to 96.3%. It means that, in the little dark environments in which light changes uniformly, most of the related methods have good recognition performance,

since they eliminate the uniform illumination changes to a certain extent by normalizing the length of extracted feature vectors. The recognition rate of PCA-SIFT algorithm is slightly lower than the other methods, because in order to reduce the amount of data, it uses PCA to reduce the dimension of extracted features that leads to the decline of feature description accuracy.

Table 3. Recognition Rates of the Related Algorithms

Algorithms	Recognition rates
SIFT	85.2%
Contourlet-SIFT	88.9%
PCA-SIFT	74.1%
ASIFT	85.2%
MSHF+SIFT	96.3%

(2).Sub-experiment 2

In the Sub-experiment 2, we test the comprehensive recognition ability of the related algorithms under a variety of different light environments, and results are shown in Table 4.We can observe from Table 4 that the proposed method still has a relatively high recognition rate of 90.7%, while the recognition rates of the other algorithms decrease rapidly, especially PCA-SIFT algorithm whose recognition rate is only 48.1%. This indicates that, when the images with uneven light are added to the testing sample database, the recognition rates of the traditional algorithms drop rapidly, *i.e.* the traditional algorithms have poor adaptability to the uneven illumination. Because when the illumination angle changes, some light and dark areas would exist on the surface of the object. Such illumination changes cannot be eliminated by the normalization of the feature vector length, so the recognition rates of the traditional algorithms are decreased dramatically. And the proposed method could eliminate the influence of uneven illumination by using DoG based homomorphic filtering so that it could still maintain a high recognition rate.

This experiment shows that the proposed algorithm in this paper has better robustness to the uneven illumination than the traditional algorithms

Table 4. Recognition Rates of the Related Algorithms

Algorithms	Recognition rates
SIFT	59.3%
Contourlet-SIFT	70.4%
PCA-SIFT	48.1%
ASIFT	61.1%
MSHF+SIFT	90.7%

5. Conclusions

In order to solve the problem that the traditional SIFT algorithms has poor matching performance under uneven illumination environment, a novel improved MSHF+SIFT method based on improved homomorphic filtering is proposed in this paper. The main works of this paper are as follows:

First, extend homomorphic filtering algorithm to multi-scale space, and put forward a difference of Gaussian based multi-scale homomorphic filtering algorithm, named

multi-scale homomorphic filter (MSHF). Then use gray enhancement algorithm to enhance the gray dynamic range and contrast of the image, and obtain relatively good pretreatment effect.

Second, effectively fuse the MSHF method and SIFT algorithm, *i.e.* using MSHF to construct pyramid, and using feature extraction and feature descriptor of SIFT to generate features, and propose a novel method named MSHF+SIFT which could obviously improve the matching performance.

Finally, experiments prove that the proposed algorithm effectively extract more stable feature points and achieve higher matching rate in uneven illumination environments.

Acknowledgments

References

[1]   L. Liu, M. Yu and L. Shao, "Unsupervised Local Feature Hashing for Image Similarity Search", IEEE Transactions on Cybernetics, **(2015)**.

[2]   S. U. Yu, S. G. Shan, X. L. Chen and W. Gao, "Integration of Global and Local Feature for Face Recognition", Journal of Software, vol. 21, no. 8, **(2010)**, pp. 1849-1862.

[3]   [W. Tao and K. Sun, "Robust Point Sets Matching by Fusing Feature and Spatial Information Using Nonuniform Gaussian Mixture Models", IEEE Transactions on Image Processing, vol. 24, no. 11, **(2015)**, pp. 3754-67.

[4]   R. Weng, J. Lu and Y. P. Tan, "Robust Point Set Matching for Partial Face Recognition", IEEE Transactions on Image Processing, vol. 25, no. 3, **(2016)**, pp. 1-1.

[5]   X. Tan and B. Triggs, "Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions", IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, vol. 19, no. 6, **(2010)**, pp. 1635-50.

[6]   D. G. Lowe, "Distinctive Image Features from Scale-Invariant Key-points", International Journal of Computer Vision, vol. 60, no. 2, **(2004)**, pp. 91-110.

[7]   K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 27, no. 10, **(2005)**, pp. 1615-30.

[8]   S. R. Chen, "Contourlet-SIFT Feature Matching Algorithm", Dianzi Yu Xinxi Xuebao/journal of Electronics & Information Technology, vol. 35, no. 5, **(2013)**, pp. 1215-1221.

[9]   L. Cinque, G. Iovane, M. Manzo and E. Sangineto, "Face recognition using SIFT features and a region-based ranking", Journal of Discrete Mathematical Sciences & Cryptography, vol. 13, no. 2, **(2013)**, pp. 153-170.

[10] Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors", IEEE Computer Society Conference on Computer Vision & Pattern Recognition IEEE, **(2004)**, pp. 506-513.

[11] K. Liao, G. Liu and Y. Hui. "An improvement to the SIFT descriptor for image representation and matching", Pattern Recognition Letters, vol. 34, no. 11, **(2013)**, pp. 1211-1220.

[12] J. Michel Morel and G. Yu. "ASIFT: A New Framework for Fully Affine Invariant Image Comparison", Siam Journal on Imaging Sciences, vol. 2, no. 2, **(2009)**, pp. 438-469.

[13] Y.J. Zhan, "Image engineering-Image processing (Second Edition)", Tsinghua university press, Beijing, **(2006)**.

[14] C. N. Fan, "Research on illumination processing algorithms for face recognition", Nanjing University, **(2011)**.

[15] C.N. Fan and F. Y. Zhang, "Homomorphic filtering based illumination normalization method for face recognition", Pattern Recognition Letters, vol. 32, no. 10, **(2011)**, pp. 1468-1479.

[16] D. Catarious, A. H. Baydush and F. C. Jr., "Characterization of difference of Gaussian filters in the detection of mammographic regions", Medical Physics, vol. 33, no. 33, **(2006)**, pp. 4104-14.

[17] M. A. Leibold, M. Holyoak, N. Mouquet, Amarasekare, P. Chase, J. M. and M. F. Hoopes, "The Metacommunity Concept: A Framework for Multi-Scale Community Ecology", Ecology Letters, vol. 7, no. 7, **(2004)**, pp. 601-613.

[18] Z, Farbman, R. Fattal and D. Lischinski, "Edge-preserving decompositions for multi-scale tone and detail manipulation", Acm Transactions on Graphics, vol. 27, no. 3, **(2008)**, pp. 15-19.

[19] Moghaddam, R. Farrahi and M. Cheriet. "A multi-scale framework for adaptive binarization of degraded document images", Pattern Recognition, vol. 43, no. 6, **(2010)**, pp. 2186-2198.

[20] K. Sunkavalli, M. K. Johnson, W. Matusik and H. Pfister, "Multi-scale image harmonization", Acm Transactions on Graphics, vol. 29, no. 4, **(2010)**, pp. 4.

[21] G. Cui, "Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition", Optics Communications, vol. 341, no. 341, **(2015)**, pp. 199-209.