USING MULTI-SCALE AUTO CONVOLUTION MOMENTS TO GET IMAGE AFFINE INVARIANT FEATURES

Fengwen Zhai*
School of Electronic and Information Engineering
Lanzhou Jiaotong University,
Lanzhou, Gansu, 730070
China
930843420@qq.com

Jianwu Dang
School of Electronic and Information Engineering
Lanzhou Jiaotong University,
Lanzhou, Gansu, 730070
China

Yangping Wang
School of Electronic and Information Engineering
Lanzhou Jiaotong University,
Lanzhou, Gansu, 730070
China

Jing Jin
School of Electronic and Information Engineering
Lanzhou Jiaotong University,
Lanzhou, Gansu, 730070
China

Abstract. This paper includes two important works. First of all, the complete mathematical proof procedure of Multi-Scale Auto convolution was summarized and the simplified geometric proof the procedure was proposed. Secondly, Multi-Scale Auto convolution moments were adopted to describe images' maximally stable extremal regions to get affine invariant features of images. In the second job, the Multi-Scale Auto convolution moments of the image features were calculated on each feature's MSER region to form image features' descriptors, and then the image feature matching was performed. In order to verify the validity of the second job, the proposed algorithm were compared with the SIFT algorithm and MSER_SURE algorithm. Simulation experiments show that, for affine transformed images, the feature matching accuracy of the second job is much higher than the classical SIFT algorithm and the MSER_SURE algorithm, which indicates that using Multi-Scale Auto convolution moments on the MSER regions could get effective affine invariant image features.

Keywords: multi-scale auto convolution, maximally stable extremal regions, feature matching, affine invariant, pattern recognition.

∗. Corresponding author

1. Introduction

Images provide a variety of information of colors, intensities, textures, edges and shapes. To understand and recognize the images, it is necessary to acquire invariant features of digital images. The image features mainly fall into three categories: region-based features, texture-based features, and point features. The point features are widely used in the field of image retrieval, target recognition and tracking, image three-dimensional reconstruction, image super-resolution reconstruction, image mosaic and image registration. David Lowe proposed the scale-invariant feature transform (SIFT) algorithm in 1999 and further improved in 2004. SIFT features remain invariant to rotation, scaling, and brightness changes, and maintain a certain degree of stability against affine transformation and noises [1]. Morel et al. proposed the ASIFT algorithm in 2009[2]. ASIFT uses the discrete image shooting angle changes to simulate the continuous shooting angle changes. The matching effect of ASIFT greatly exceeds the SIFT algorithm in affine invariance aspect, but the time efficiency is lower than SIFT. In 2011, Ethan Rublee et al. proposed ORB algorithm[3] on ICCV. The ORB algorithm is based on the FAST algorithm and the BRIEF algorithm, but the ORB algorithm could not get scale invariant and affine invariant features. Image moments could keep some invariant characteristics. A lot of researchers have concentrated on the study of adopting moments to describe image features [4, 5, 6]. But the before moments could only get scale or rotational or transformational invariances while affine invariance is also important for real applications[7, 8, 9, 10]. In 2002, Heikkil J and Esa et al. proposed Multi-scale Auto convolution (MSA)[11] moments which are affine invariant when applied to pattern recognition. And in the later years, Esa et al. polished the MSA moments algorithm from different aspects [12, 13, 14, 15]. Due to the strong affine invariance of MSA moments, there are more and more researchers applied MSA moments into image processing. In 2014, Shao C et al. improved the MSA moments classification ability by taking the cosine value of the angle between two most prominent points within image feature's neighborhood as the feature's probability density value[16]. In 2015, Li et al. proposed a feature point matching method based on the geometric relation constraint and MSA moments [17]. In the same year, Zhang et al. Combined the MSA moments with the SIFT algorithm [18] and improved the match accuracy of the SIFT algorithm. However, it was found that the proof process of MSA moment calculation method is complicated and has a lot of jumping in the reference[12], which is inconvenient for readers to understand and to simulation. At the same time, MSA Moment has strict mathematical features, if the MSA moments were introduced into feature point matching, the neighborhood window similar to SIFT algorithm and ORB algorithm can not be used directly. In view of the above two problems, firstly, this paper summarizes the complete proof of the MSA moments and then proposes a more easy-to-understand calculation method through spatial geometry analysis. Secondly, the regions of the Maximally Stable Extremal

Regions(MSER)[19] were selected as the simple background neighborhoods of the image features, so that the MSA moments can be effectively applied to the description of the image features, which was defined as MSER_MSA for convenient. Thirdly, the MSER_MSA algorithm was compared with the SIFT algorithm and the MSER_SURF[20] algorithm to prove the effectiveness of the MSER_MSA algorithm for large-scale affine transformed image pairs.

The remaining part of this paper is structured as follows. The second part introduces the basic principle of the MSA moments and the MSER algorithm. The third part summarizes the complete proof procedure and proposes the easy-to-understand geometric proof of the MSA moments. The fourth part introduces the MSER_MSA algorithm to detail. The fifth part compares the proposed MSER_MSA algorithm with the SIFT and MSER_SURF algorithms. At last, the conclusion part concludes with a summary of the effectiveness of the proposed MSER_MSA algorithm and the next work direction.

2. MSA Moments and MSER algorithm

In 2005, Esa et al. Proposed MSA moments. To understand MSA moments, references [12] gives three definitions.

Definition 1. $\mathbf{X}$ is the base point, and the $\mathbf{\Gamma}$ is the affine transform matrix, $\mathbf{X}'$ corresponds to $\mathbf{X}$

$$\mathbf{X}' = \mathbf{\Gamma}(\mathbf{X}) = \mathbf{T}\mathbf{X} + \mathbf{t} \tag{1}$$

Among them $\mathbf{t}, \mathbf{x} \in \mathbb{R}^2$, and T is a non-singular real matrix, the inverse transform recorded as

$$\mathbf{Y}' = \Gamma^{-1}(\mathbf{Y}) = \mathbf{T}^{-1}\mathbf{Y} - \mathbf{T}^{-1}\mathbf{t} \tag{2}$$

Definition 2. Let $f(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$, and $f(\mathbf{x}) \geq 0$ is a grayscale function defined on a two-dimensional grayscale image. Applying the affine transformation to the image, a new image grayscale function can be obtained.

$$f'(\mathbf{x}) = f(\Gamma^{-1}(\mathbf{x})) = f(\mathbf{T^{-1}x} - \mathbf{T^{-1}t}) \tag{3}$$

Definition 3. For the feature $\mathbf{A}$ extracted from the image, the feature $\mathbf{A}$ is called the affine invariant feature when $f \circ (\mathbf{A})$ and $f' \circ (\mathbf{A})$ is the same.

Randomly take three non-collinear three points $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ in the image as the basis for the two arbitrary variables α and β, then random variable $\mathbf{u}$ could be gotten form Eq. (4)

$$\mathbf{u} = \alpha(\mathbf{x_1} - \mathbf{x_0}) + \beta(\mathbf{x_2} - \mathbf{x_0}) + \mathbf{x_0} \tag{4}$$

To transform the point $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ using $\Gamma = \Gamma\{\mathbf{T}, \mathbf{t}\}$, then $\mathbf{x}_0{}' = \mathbf{T}\mathbf{x}_0 + \mathbf{t}$, $\mathbf{x_1}' = \mathbf{T}\mathbf{x_1} + \mathbf{t}$, $\mathbf{x}_2{}' = \mathbf{T}\mathbf{x}_2 + \mathbf{t}$. Assume $\mathbf{u}' = \alpha(\mathbf{x}_1{}' - \mathbf{x}_0{}') + \beta(\mathbf{x}_2{}' - \mathbf{x}_0{}') + \mathbf{x}_0{}'$,

then

$$
\begin{aligned}
\mathbf{u}' =& \alpha(\mathbf{x}_1{}' - \mathbf{x}_0{}') + \beta(\mathbf{x}_2{}' - \mathbf{x}_0{}') + \mathbf{x}_0{}' \\
=& \alpha((\mathbf{T}\mathbf{x}_1 + \mathbf{t}) - (\mathbf{T}\mathbf{x}_0 + \mathbf{t})) + \beta((\mathbf{T}\mathbf{x}_2 + \mathbf{t}) - (\mathbf{T}\mathbf{x}_0 + \mathbf{t})) + \mathbf{T}\mathbf{x}_0 + \mathbf{t} \\
=& \alpha(\mathbf{T}\mathbf{x}_1 - \mathbf{T}\mathbf{x}_0) + \beta(\mathbf{T}\mathbf{x}_2 - \mathbf{T}\mathbf{x}_0) + \mathbf{T}\mathbf{x}_0 + \mathbf{t} \\
=& \mathbf{T}\mathbf{u} + \mathbf{t}
\end{aligned} \tag{5}
$$

From Eq.(5), to transform the basis point $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ by $\Gamma = \Gamma\{\mathbf{T}, \mathbf{t}\}$, then we can obtain three new base point $\mathbf{x}_0{}'$, $\mathbf{x}_1{}'$, $\mathbf{x}_2{}'$. The same affine transformation $\Gamma = \Gamma\{\mathbf{T}, \mathbf{t}\}$ is performed on point $\mathbf{u}$ to get a new point $\mathbf{u}'$. The linear relationship between $\mathbf{u}'$ and $\mathbf{x}_0{}'$, $\mathbf{x}_1{}'$, $\mathbf{x}_2{}'$ and the linear relationship between $\mathbf{u}$ and$\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ remain unchanged. And from Eq. (3), Eq.(6) can be obtained.

$$
f'(\mathbf{u}') = f(\Gamma^{-1}(\mathbf{u}')) = f(\mathbf{T^{-1}}\mathbf{u}' - \mathbf{T^{-1}}\mathbf{t}) = f(\mathbf{T^{-1}}(\mathbf{T}\mathbf{u} + \mathbf{t}) - \mathbf{T^{-1}}\mathbf{t})) = f(\mathbf{u}) \tag{6}
$$

In summary, we can obtain an affine invariant feature $f(\mathbf{U})$, where $\mathbf{u}$ is a random variable of U. The MSA moment is the mathematical expectation of $f(\mathbf{U})$.

$$
F(\alpha, \beta) = E[f(U_{\alpha,\beta})] \tag{7}
$$

The reference[12] gives the calculation formula of $F(\alpha, \beta)$ as equation (8)

$$
F(\alpha, \beta) = \left[\int_{\mathbb{R}^2} \hat{f}(-\xi)\hat{f}(\alpha\xi)\hat{f}(\beta\xi)\hat{f}(\gamma\xi)d\xi \right] / \hat{f}(0)^3 \tag{8}
$$

2.1 The basic principle of MSER algorithm

MSER's full name is Maximally Stable Extremal Regions, which was put forward by J. Matatas in 2002[19]. MSER algorithm uses three definitions: Region, Extremal Region and Maximally Stable Extremal Region.

Definition 4. The region is a part of an image, for any two points **A** and **B**, **A** can reach **B** along the adjacent pixel within this part, that is, there is a connected path between any two points.

Definition 5. The definition of Extreme region is related to the grayscale threshold. After the grayscale threshold is set, the image is divided into two part according to the threshold, the pixels smaller than the threshold are set to black, the pixels larger than the threshold are set to white. The region in the image is made up of black pixels. And the region can become the extreme region if no more pixels can be found to enlarge the current region.

For example, threshold segmentation is performed on an image. Then select a series of grayscale thresholds. For example, we take $i = 0, 1, \cdots, 255$. When the threshold is i, the gray value of each pixel in the image is compared with the threshold. Points below this threshold are "black spots", otherwise, "white

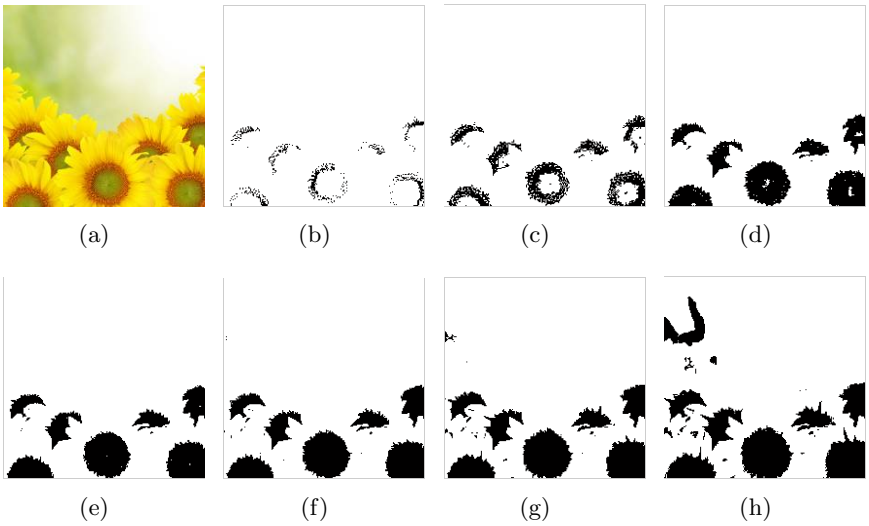


Figure 1: Extreme Region illustration figures. (a) is the original image. (b),(c),(d),(e),(f),(g),(h) are the Extreme Regions that appear in the image when the grayscale thresholds are set to 80, 100, 120, 140, 160, 180, 200.

spots". This will get a picture of a frame. When all the images are arranged in order of threshold, you can see: The first picture is all white (all pixels have a threshold of no less than 0). Then the black spots appear gradually on the image, and some black spots gradually fuse together to form a small area to become the connected areas. The last picture is all black as shown in Fig.1. In these figures, all the black connected areas are Extreme Regions. Even an isolated black spot is also an Extreme Region. The determination of Extreme Region is closely related to this gray threshold. It is unable to determine a region or point is Extreme Region or not when leaving this grayscale threshold.

Definition 6. During the process of obtaining Extreme Region, some Extreme Regions increase slightly with the increasing of the set gray threshold. Such a group of Extreme Regions is a nested relationship from small to large. We use $Q_1, Q_2, \cdots, Q_{i-1}, Q_i$ to represent this series of mutually exclusive Extreme Region sequences. The condition of Q_i^* could be Maximally Stable Extremal Region if and only if the minimum value of $q(i)$ is achieved when $i = i^*$.

$$q(i) = \frac{|Q_{i+\xi} - Q_{i-\xi}|}{|Q_i|} \tag{9}$$

Because Q is a collection of pixels, the absolute value represents the cardinality of the collection. We can think of it as the area of this Extreme Region. As can be seen from the above explanation, the maximum stable Extreme Region

is the Maximally Stable Extremal Region when the gray threshold changes. So we find all the Extreme Regions and then use the formula (9) to determine the final Maximally Stable Extremal Regions which is abbreviated as MSER.

The MSER regions can be processed as simple background image fragments. Esa et al. have experimentally proved that the MSA moments should be applied to image pattern recognition with simple backgrounds.

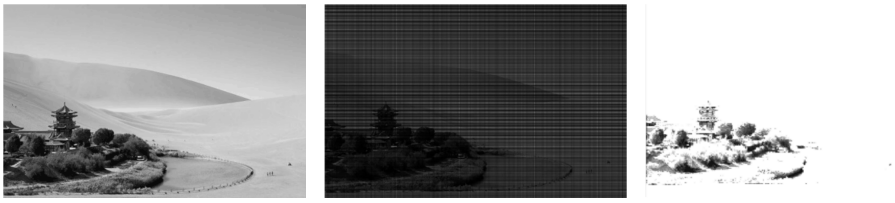
3. The MSA moment calculation method

3.1 The summarized proof for MSA moment calculation

The proof process of MSA moment calculation includes convolution calculation. As we all know that convolution calculation in the time domain is very time-consuming. It has been tested that the convolution of three $5 * 5$ matrices requires nearly 1 minute to run in our experimental environment. Therefore, Esa et al. converts the convolution computation of time domain into frequency domain product[21], and gives the proof process of MSA moment. However, the proof process of MSA moment calculation is complicated and inconsistent in the reference [12], which is inconvenient for readers to understand and to simulate. Firstly, this paper summarizes the detailed proof process of the MSA moment calculation method as follows.

$$
\begin{aligned}
(1)\ F(\alpha,\beta) &= \int_{D^2} f(\mathbf{u})(p_\alpha * p_\beta * p_\gamma)(u)d\mathbf{u} \\
(2)\ &= \int_{D^2} \overline{\overline{f(\mathbf{u})}}(p_\alpha * p_\beta * p_\gamma)(\mathbf{u})d\mathbf{u} \\
(3)\ &= \int_{D^2} \overline{\mathbb{F}(\overline{f(\mathbf{u})})}\mathbb{F}((p_\alpha * p_\beta * p_\gamma)(\mathbf{u}))d\xi \\
(4)\ &= \int_{D^2} \overline{\mathbb{F}(\overline{f(\mathbf{u})})}\mathbb{F}(p_\alpha)\mathbb{F}(p_\beta)\mathbb{F}(p_\gamma)d\xi \\
(5)\ &= \int_{D^2} \overline{\mathbb{F}(\overline{f(\mathbf{u})})}\frac{1}{\alpha\beta\gamma\|f_{R^2}\|_{L^1}{}^3}\mathbb{F}(f(\frac{\mathbf{u}}{\alpha}))\mathbb{F}(f(\frac{\mathbf{u}}{\beta}))\mathbb{F}(f(\frac{\mathbf{u}}{\gamma}))d\xi \\
(6)\ &= \frac{1}{\alpha\beta\gamma\|f_{R^2}\|_{L^1}{}^3}\int_{D^2} \overline{\mathbb{F}(\overline{f(\mathbf{u})})}\mathbb{F}(f(\frac{\mathbf{u}}{\alpha}))\mathbb{F}(f(\frac{\mathbf{u}}{\beta}))\mathbb{F}(f(\frac{\mathbf{u}}{\gamma}))d\xi \\
(7)\ &= \frac{\alpha\beta\gamma}{\alpha\beta\gamma\|f_{R^2}\|_{L^1}{}^3}\int_{D^2} F(-\xi)F(\alpha\xi))F(\beta\xi)F(\gamma\xi)d\xi \\
(8)\ &= \frac{1}{\|f_{R^2}\|_{L^1}{}^3}\int_{D^2} F(-\xi)F(\alpha\xi))F(\beta\xi)F(\gamma\xi)d\xi \\
(9)\ &= \frac{1}{F(0)^3}\int_{D^2} F(-\xi)F(\alpha\xi))F(\beta\xi)F(\gamma\xi)d\xi
\end{aligned}
$$

Among them, D^2 is the integral domain of variable $\mathbf{u}$, and $\overline{f(x)}$ is the conjugate of $f(x)$. Step (2) to (3) is based on the Plancherel formula $\int_{R^2} f(x)\overline{g(x)} = \int_{R^2} \mathbb{F}(f(x))\overline{\mathbb{F}(g(x))}$. The Fourier transform convolution theorem was used



(a) Original image $\mathbf{I}$ (b) $\alpha\mathbf{I}$ image, $\alpha = 2$ (c) $\alpha\mathbf{I}$ image, $\alpha = 0.5$

Figure 2: The comparison figure between images $\mathbf{I}$ and $\alpha\mathbf{I}$.

from step (3) to step (4). step (5) is calculated by substituting $p_\alpha(u)$, $p_\beta(u)$ and $p_\gamma(u)$ into the formula, where $p_\alpha(u) = f(\frac{u}{\alpha})/(\|f_{R^2}\|_{L^1} * \alpha)$. step (6) to step (7) is based on the similarity principle of the Fourier transform. Step (8) to step (9) is based on $\|f_{R^2}\|_{L^1} = F(0)$.

3.2 Geometric Proof of MSA Moment Calculation Method

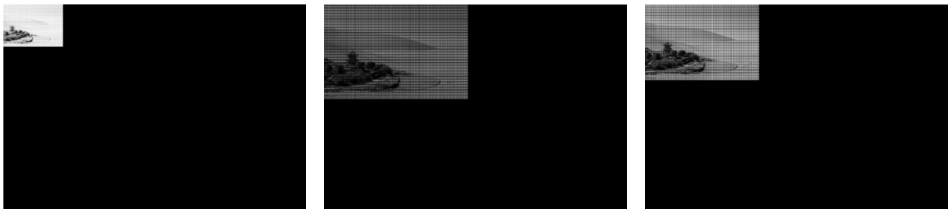
In Eq. (10), the proof of MSA moment calculation method is very rigorous. The functions and variables in the mathematical derivation from (5) to (9) have no specific physical meanings, which is hard for reader to understand. After detailed reasoning and simulating, this paper proofed the MSA moment calculation method directly through geometric analysis. The proof begins at step(4) in Eq.(10), as shown in Eq. (11)

$$F(\alpha,\beta) = \int_{D^2} \overline{\mathbb{F}(\overline{f(\mathbf{u})})}\mathbb{F}(p_\alpha)\mathbb{F}(p_\beta)\mathbb{F}(p_\gamma)d\xi \tag{10}$$

Firstly, the variables and functions in the Eq.(11) are described here. D^2 is the integral field of variable $\mathbf{u}$, $\mathbf{u} = \alpha\mathbf{x_1} + \beta\mathbf{x_2} + \gamma\mathbf{x_0}$, let the size of image I be $m \times n$, then $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2 \in [1:m, 1:n]$, so D^2 is $[1:(|\alpha|+|\beta|+|\gamma|)\times m, 1:(|\alpha|+|\beta|+|\gamma|)\times n]$. Where p_α represents the probability density of $\alpha\mathbf{x_1}$, p_β represents the probability density of $\beta\mathbf{x}_2$, p_γ represents the probability density of $\gamma\mathbf{x}_0$. The range or the integral field of $\alpha\mathbf{x_1}$, $\beta\mathbf{x}_2$, $\gamma\mathbf{x}_0$ is D^2, $f(\mathbf{x})$ represents the gray value of the point $\mathbf{x}$ in the image $\mathbf{I}$.

Secondly, the proposed geometric proof is given below. $\alpha\mathbf{I}$ represents α times expanded image of $\mathbf{I}$. Let $f_\alpha(\mathbf{z})$ represent the gray value of the point $\mathbf{z}$ in image $\alpha\mathbf{I}$. When $|\alpha| > 1$, $\mathbf{I}$ increases to $\alpha\mathbf{I}$ by inserting 0. When $|\alpha| < 1$, by taking the "enrichment" method, $\mathbf{I}$ will be shrunk to α times of image $\mathbf{I}$. The "enrichment" means to sum up the pixels' gray value within a small region. So $\|f_\alpha(\mathbf{z})\|_{L^1} = \|f(\mathbf{x})\|_{L^1}$. Fig. 2 shows the relationship between $\mathbf{I}$ and $\alpha\mathbf{I}$.

The following demonstrates that the density of variable $\alpha\mathbf{x_1}$ on image $\alpha\mathbf{I}$ is equal to the density on D^2. Because $\mathbf{x_1} \in [1:m, 1:n]$, so $\alpha\mathbf{x_1} \in [\alpha:\alpha m, \alpha:\alpha n]$. Even if the integration domain of $\alpha\mathbf{x_1}$ is D^2, $\alpha\mathbf{x_1}$ can only take values in



(a) Gray value of $\alpha\mathbf{x_1}$ on D^2 (b) Gray value of $\beta\mathbf{x_2}$ on D^2 (c) Gray value of $\gamma\mathbf{x_3}$ on D^2

Figure 3: the intensities of $\alpha\mathbf{x_1}$, $\beta\mathbf{x_2}$, $\gamma\mathbf{x_0}$ with D^2, $\alpha = 0.7$, $\beta = 1.5$.

sub-domains in $[\alpha : \alpha m, \alpha : \alpha n]$, as shown in Fig.3, and equation (12).

$$p_{\alpha\mathbf{X_1}}^{D_2}(\mathbf{u}) = \begin{cases} p_{\alpha\mathbf{X_1}}{}^{\alpha\mathbf{I}}(\mathbf{u}), & \mathbf{u} \in \alpha\mathbf{I} \\ 0, & \mathbf{u} \in D^2 \end{cases} \tag{11}$$

In the equation (12), $p_{\alpha\mathbf{X_1}}^{D_2}(\mathbf{u}_1)$ represents the density of the variable $\alpha\mathbf{x_1}$ on the integral domain D^2, and $p_{\alpha\mathbf{X_1}}^{\alpha\mathbf{I}}(\mathbf{z})$ represents the density of the variable $\alpha\mathbf{x_1}$ on the integral domain $\alpha\mathbf{I}$. Because $\alpha\mathbf{x_1}$, $\beta\mathbf{x_2}$, $\gamma\mathbf{x_0}$ are independent to each other, so $\int_{D^2} p_{\alpha\mathbf{X_1}}{}^{D^2}(\mathbf{u}_1) = \int_{\alpha\mathbf{I}} p_{\alpha\mathbf{X_1}}{}^{\alpha\mathbf{I}}(\mathbf{z}_1) = 1$. Similarly, $\int_{D^2} p_{\beta\mathbf{X_2}}{}^{D^2}(\mathbf{u}_2) = \int_{\beta\mathbf{I}} p_{\beta\mathbf{X2}}{}^{\beta\mathbf{I}}(\mathbf{z}_2) = 1$, $\int_{D^2} p_{\gamma\mathbf{X_0}}{}^{D^2}(\mathbf{u}_0) = \int_{\gamma\mathbf{I}} p_{\gamma\mathbf{X_0}}{}^{\gamma\mathbf{I}}(\mathbf{z}_0) = 1$. In the equation (11), p_α represents $p_{\alpha\mathbf{X_1}}^{D_2}(\mathbf{u}_1)$, p_β represents $p_{\beta\mathbf{X_2}}^{D^2}(\mathbf{u}_2)$, p_γ represents $p_{\gamma\mathbf{X_0}}^{D^2}(\mathbf{z}_0)$. During the actual calculation, $\mathbb{F}(p_\alpha)$ can be calculated according to equation (12). The specific calculation process is to stretch the original image $\mathbf{I}$ to $\alpha\mathbf{I}$, then perform Fourier transform in the way of zero-padding, as shown in Eq.(13).

$$\mathbb{F}(p_\alpha) = \mathbb{F}(p_{\alpha\mathbf{X_1}}^{\alpha\mathbf{I}}(\mathbf{z}), (|\alpha| + |\beta| + |\gamma|) * m, (|\alpha| + |\beta| + |\gamma|) * n) \tag{12}$$

In Eq. (13), $p_{\alpha\mathbf{X_1}}^{\alpha\mathbf{I}}(\mathbf{z}) = f_\alpha(\mathbf{z})/\|f_\alpha(\mathbf{z})\|_{L^1}$. The calculation methods of $\mathbb{F}(p_\beta)$, $\mathbb{F}(p_\gamma)$ are similar to $\mathbb{F}(p_\alpha)$

Through the above geometric analysis, the calculation of the MSA moment can be directly carried out according to Eq.(11). So, the deduction from step (5) to step (9) of the Eq.(10) could be omitted. The method presented in this paper is more intuitive than the mathematical proof method in [8], and is easy to understand and to realize. At the same time, the simulation results show that the calculation results of the proposed method are consistent with the results of the reference [12].

4. The MSA moments of characteristic points are computed on MSER region

In order to apply the scale invariance and affine invariance of MSA moments into the description of image features, MSER method here is used as feature point neighborhood extraction method, and the 31 orders' MSA moments are calculated as the feature descriptors of the most stable extremal regions. The

31 orders were recommended by reference [12]. For conveniently expressing, the method proposed in this paper was named as MSER_MSA, and the MSER_MSA algorithm was designed as follows:

Algorithm 1 MSER_MSA.

Input:

Reference image **I**, floating image **J**;

Output:

Matching feature pairs;

1: To perform Gauss smoothing on image **I** and image **J**;
2: The most stable extremal regions of **I** and **J** are extracted respectively, and the external ellipses are obtained;
3: The centroids of the external ellipses are taken as the feature points of the image;
4: The 31 orders' MSA moments of each feature point in its ellipse neighborhood are calculated and used as the feature descriptor of each feature point;
5: By comparing the feature descriptors, the features of image **I** and image **J** are matched.

5. Experiment

The experimental simulation platform of this paper is configured as follows: Windows XP operating system, 3.30GHz CPU, 3.0G memory, MATLAB 2015. In the experiments, the proposed MSER_MSA algorithm was compared with the SIFT algorithm and the classical MSER_SURF method. The floating images in Fig.4 were obtained by share transforming the reference images. The floating images in Fig.5 were obtained by share and scale transforming the reference images.

As can be seen that from Fig.4 and Fig.5, SIFT algorithm's matching effect of affine transformation is poor, and the MSER_SURF algorithm has better matching effect for the image with smaller affine transformation scale, but the matching effect of MSER_SURF is sensitive to scale changes, and can not obtain correct matching for the images with relatively large affine changes and scale changes. The MSER_MSA performed stably for scale change and affine change, and the correct matching number is only slightly less for image pairs with large scale and affine changes.

Fig.6 is the matching accuracy comparison chart of SIFT, MSER_SURF and MSER_MSA. The image samples came from 4 images. Each image generates a group of samples. The first one in each group is the reference image, and the rest 3 samples within each group are obtained by affine transforming the reference image, and the transforming scales are progressively increased in each group. From Fig.6, it can be seen that the accuracy rates of SIFT,

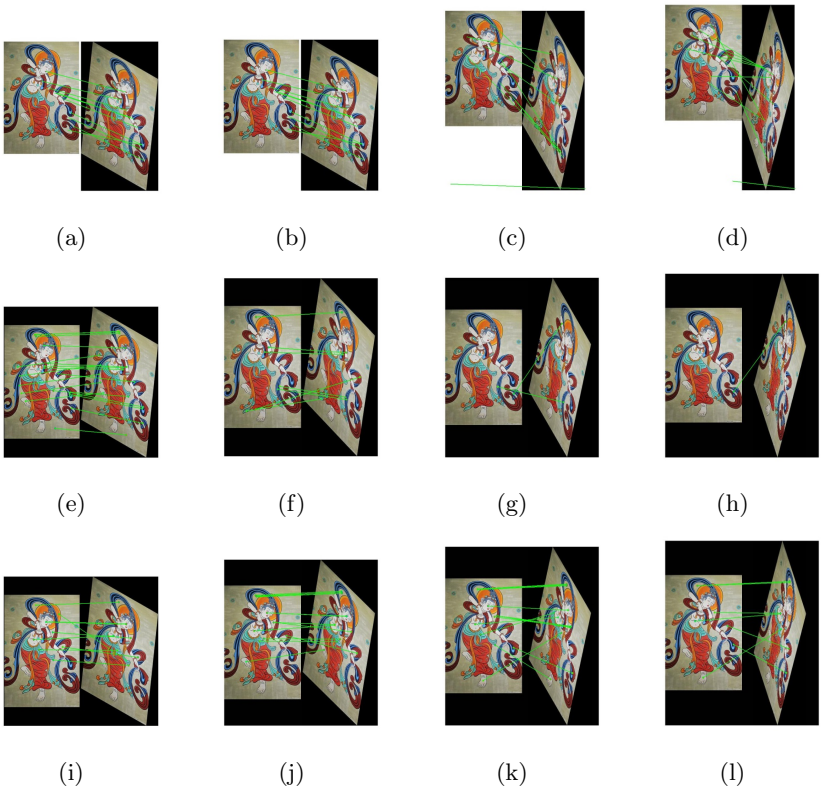


Figure 4: The feature matching results of SIFT, MSER_SURF and MSER_MSA of images with share transforming. (a),(b),(c) and (d) are the matching results of SIFT. (e),(f),(g) and (h) are the matching results of MSER_SURF. (i),(j),(k) and (l) are the matching results of MSER_MSA.

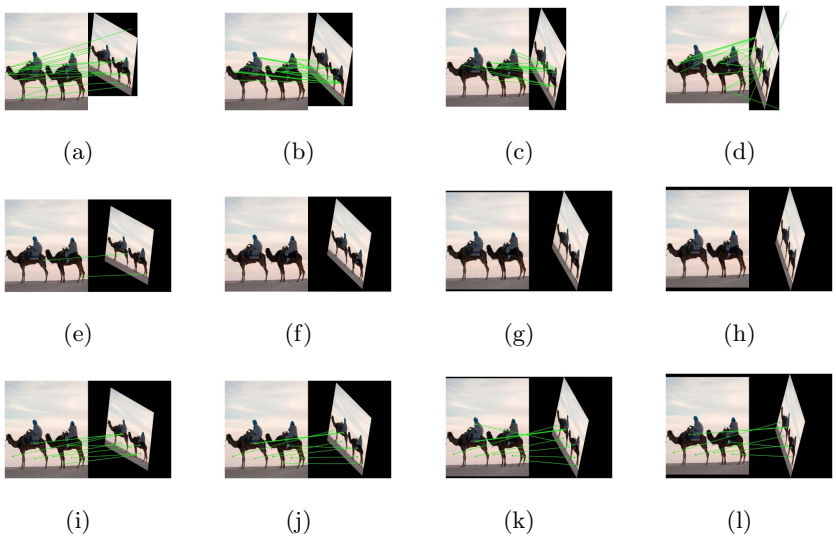


Figure 5: The feature matching results of SIFT, MSER_SURF and MSER_MSA of images with share and scale transforming. (a),(b),(c) and (d) are the matching results of SIFT. (e),(f),(g) and (h) are the matching results of MSER_SURF. (i),(j),(k) and (l) are the matching results of MSER_MSA.

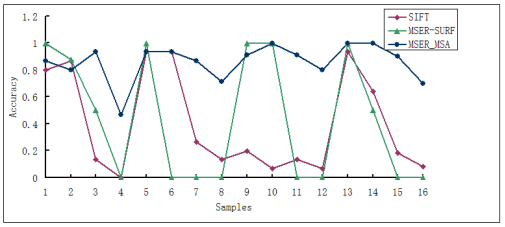


Figure 6: Accuracy comparison chart of SIFT, MSER_SURF and MSER_MSA

MSER_SURF and MSER_MSA are not far-off, but for large scaled transforming image pairs, the accuracy rates of SIFT and MSER_SURF are significantly lower than MSER_MSA.

Since the MSA algorithm involves Fourier transforms, the time efficiency is slightly lower. At the same time, the MSA algorithm relies on the segmentation effect of MSER algorithm. So the matching effect of artificial experimental images is better than that of realistic images.

6. Conclusion

MSA moments have scale invariant and affine invariant characteristics, and have achieved good recognition results in the field of image pattern recognition. However, there is not much research on the application of MSA moments in filed of image feature describing. After many experiments and analysis, it was found that if we want to apply MSA moments to the description of image features, the simple background neighborhoods of each feature points must be obtained firstly. Therefore, this article combines the MSER method with the MSA moments. The maximum stable extreme regions are taken as the neighborhoods of image feature points. And the 31 orders' MSA moments of each feature point are calculated as the feature descriptor of each feature point. Experimental results show that the proposed MSER_MSA algorithm can make full use of the invariant characteristics of MSA moments. However, it was also found that the MSA moments are sensitive for images with larger illumination changes. At the same time, the time efficiency of the MSA moment is low. Therefore, the next work of this study will focus on solving the illumination sensitivity of MSA moments and improving the time efficiency of MSA moments.

References

[1] Zhai F, Dang J, Wang Y, et al., *Application of Shape Context and Belief Propagation in Removing SIFT Mismatches*, Journal of Computer-Aided Design & Computer Graphics, 28(2016), 443-449.

[2] Morel J M, Yu G., *ASIFT: A New Framework for Fully Affine Invariant Image Comparison*, Siam Journal on Imaging Sciences, 2 (2011), 438-469.

[3] Rublee E, Rabaud V, Konolige K, et al., *ORB: An efficient alternative to SIFT or SURF*, IEEE International Conference on Computer Vision, IEEE, 2012, 2564-2571.

[4] MingKuei Hu, *Visual pattern recognition by moment invariants*, Information Theory Ire Transactions on, 8 (1962), 179-187.

[5] Zhai Feng-Wen, Muhammad Asim Azim, WANG Yang-PING et al., *Amendment of Zernike moment fast conputation*, Journal of Jilin University(Engineering and Technology Edition), 44 (2014):1860-1866.

[6] Dae San Kim, Seung-Hwan Yang, *A recursive formula for power moments of 2-dimensional kloosterman sums associated with general linear groups*, Italian journal of Pure and Applied Mathematics, 34 (2015), 7-16.

[7] Huang Bo, Zhao Ji-Yin, Zheng Rui-Rui et al., *Affine invariant pattern recognition based on multi-scale autoconvolution normalized histograms*, CTA Electronic Sinica, 39 (2011), 64-69.

[8] Huang Bo, Zhao Xiao-Hui, Chao Xin-Qin et al., *Histogram affine invariants extraction method based on two-scale auto-convolution*, Journal on Communications, 32 (2012), 140-146.

[9] Yang Jian-Wei, Li Pei-Yao, *Affine Invariant Feature Extraction Based on Fractional Order Moment*, Acta Automatica Sinica, 41 (2015), 2147-2154.

[10] Miao Jun, Chun Jun, Zhang Gui-Meri, *An Affine Invariant Line Descriptor and Line Matching*, Acta Electronic Sinica, 43 (2015), 2505-2512.

[11] Heikkil J., *Multi-Scale Autoconvolution for Affine Invariant Pattern Recognition*, 16Th International Conference on Pattern Recognition, IEEE Computer Society, 2002, 10119.

[12] Rahtu E, Salo M, Heikkila J., *Affine invariant pattern recognition using multiscale autoconvolution*, IEEE Transactions on Pattern Analysis & Machine Intelligence, 27(2005), 908-18.

[13] Kannala J, Rahtu E, Heikkila J., *Affine registration with multi-scale autoconvolution*, IEEE International Conference on Image Processing, IEEE, 2005, 1064-7.

[14] Rahtu E, Salo M, Heikkilä J., *Multiscale Autoconvolution Histograms for Affine Invariant Pattern Recognition*, British Machine Vision Conference 2006, Edinburgh, Uk, September. DBLP, 2006:1059-1068.

[15] Rahtu E, Salo M, Heikkil J, et al., *Generalized affine moment invariants for object recognition*, International Conference on Pattern Recognition, IEEE Computer Society, 2006, 634-637.

[16] Shao C, Ding Q, Luo H., *An improved multi-scale auto convolution transform*, Proceedings of SPIE - The International Society for Optical Engineering, 9301 (2014), 120-127.

[17] Li Hui-Hui, Hua Li, Yang Ning et al., *Multi-target association algorithm for remote sensing images based on MSA features and simulated annealing optimization*, Journal of Jilin University (Engineering and Technology Edition), 45 (2015), 1353-1359.

[18] Zhang Xiongmei,Yi Zhaoxiang, Cai Xingfu et al., *SAR Image Registration Method Based on Improved SIFT*, Computer Engineering, 41(2015), 223-226.

[19] Donoser M, Bischof H., *Efficient Maximally Stable Extremal Region (MSER) Tracking*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 2006, 553-560.

[20] Tao L, Jing X, Sun S, et al., *Combining SURF with MSER for image matching*, IEEE International Conference on Granular Computing, IEEE, 2014, 286-290.

[21] B.G. Sidharth, *A new integral transform*, Italian Journal of Pure and Applied Mathematics, 37 (2017), 15-18.

Accepted: 26.02.2018