Gaussian Affine Feature Detector

Xiaopeng Xu, Xiaochun Zhang

Abstract—A new method is proposed to get image features' geometric information. Using Gaussian as an input signal, a theoretical optimal solution to calculate feature's affine shape is proposed. Based on analytic result of a feature model, the method is different from conventional iterative approaches. From the model, feature's parameters such as position, orientation, background luminance, contrast, area and aspect ratio can be extracted.

Tested with synthesized and benchmark data, the method achieves or outperforms existing approaches in term of accuracy, speed and stability. The method can detect small, long or thin objects precisely, and works well under general conditions, such as for low contrast, blurred or noisy images.

Index Terms—LoG, DoG, differential geometry, Hessian, Harris, affine, Fourier, Laplacian.

1 INTRODUCTION

Detecting two dimension signals is more difficult than one dimension ones and many heuristic algorithms are proposed to deal with it. However, following appearance of scale space theory [12], [7], many effective feature detectors come into being [5].

Originally, scale space theory is proposed by physicists, and developed by computer scientists. It is studied thoroughly from view point of vision and mathematics, and a consistent way to find new detector had been built [2]. Many successful feature detectors are built upon scale space, including [3], [15].

An additional dimension is introduced in scale space, namely scale dimension. In order to get image's information, such as affine shape parameters, some methods [9], [6], [4] iteratively search in scale dimension. They are based on fix point theory: they will finally get a solution if there is one. In practice, however, these methods have several drawbacks, including,

- waste lots of candidate features;
- very slow;
- get abundant duplicated or false features.

To overcome these drawbacks, we propose a new method based on analytic solution. It achieves or outperforms iterative methods with much less computation resources.

Some feature detectors [10], [8] are ideal for noise free images, but are incapable of blurred or noisy images. The proposed detector is more robust with similar performance.

Features' position, orientation, background luminance, contrast, area and aspect ratio can also be extracted from images. Until recently, information such as background luminance contrast are not commonly used in feature extraction. Others including area, orientation and aspect ratio are studied extensively, but with a limited accuracy.

In this paper, a feature model is proposed, and the above mentioned parameters will be calculated.

- *X. Xu and X. Zhang are with Nanjing University of Science and Technology, China.*

2 GAUSSIAN AFFINE SHAPE

In this section, firstly a feature model is proposed, and then analytic result is derived based on the model to get various parameters.

2.1 Feature Model

From a view point of systematology, images, feature extractor and features correspond to input, system, and output. We need build a system that can transform input to output. In another words, image is system input and feature parameters are output. In order to study behavior of the system, we need define input signals. As it is not possible to build an all-purpose feature extractor, we will concentrate on some specific image signals. Since (two dimensional) Gaussian has nice analytic properties and simple form, it is chosen as input signal. As to be shown later, Gaussian based model will filter out high frequency signal, hence ideal for noisy images.

Based on above mentioned idea, image surface is modeled as Gaussian function, as shown in Fig. 1. In this way, image feature parameters are related to Gaussian parameters, including orientation, long and short radii, baseline height and contrast. Traditionally, baseline height and contrast are not considered in feature extraction, they are included for completeness. The signal can be defined as Equ. 1. Parameters and model variables are listed in Tab. 1.

$$f(x) = c \; e^{-\frac{1}{2}\left(\vec{x}-\vec{\mu}\right)^T \Sigma^{-1} \left(\vec{x}-\vec{\mu}\right)} + d \qquad (1)$$

Before continuing, it is helpful to clarify a fact, that is, rotating an image will not affect our discussion. This fact greatly simplifies our deduction. It is proved in Appendix. A.

Known the fact, two-dimensional axis-aligned Gaussian will be used as input signal, as shown in Equ. 2. For this function, we need get value of α, β, c and d.

$$\mathrm{I}(x,y;\alpha,\beta,c,d) = c \; e^{-\frac{1}{2}\left(\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}\right)} + d \qquad (2)$$
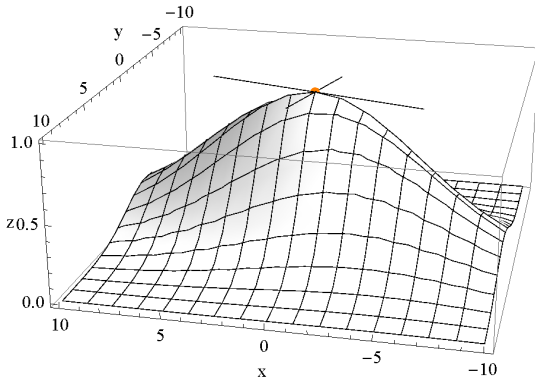
Fig. 1. Model of input signal, whose baseline height is 0 and contrast, is 1.

TABLE 1
Factors of model

Factors	Parameters
Contrast	c
Baseline height	d
Long radius	β
Short radius	α
Nominal radius	$\sqrt{\alpha\beta}$
Aspect ratio	$\frac{\beta}{\alpha}$
Orientation	θ
LoG detected scale	σ

2.2 Feature Detection

Before computing parameters, we need detect feature's position. There exists many feature detectors, we need choose the one that has good performance and solid mathematical foundation. It will be chosen from rotation invariant differential operator family. As defined in Equ. 3, LoG detector is a good candidate, it is very stable, has fast implementation, and is Gaussian based. The last one is the most important reason, because input signal is also a Gaussian, and they may have close relation.

$$\text{LoG} = \nabla \cdot \nabla G, G = \frac{e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^2} \quad (3)$$

$$\partial_x (f_1 * f_2) = f_1 * \partial_x f_2$$
$$= \partial_x f_1 * f_2 \quad (4)$$

Using Equ. 5, which is based on Equ. 4, we can define LoG operation on image function, as shown in Equ. 6.

$$\partial_{x,x}(f_1 * f_2) = \partial_{x,x}f_1 * f_2 = f_1 * \partial_{x,x}f_2$$
$$\partial_{x,y}(f_1 * f_2) = \partial_{x,y}f_1 * f_2 = f_1 * \partial_{x,y}f_2$$
$$\partial_{y,y}(f_1 * f_2) = \partial_{y,y}f_1 * f_2 = f_1 * \partial_{y,y}f_2 \quad (5)$$

$$\text{LoG} * I = (\partial_{x,x}G + \partial_{y,y}G) * I$$
$$= \partial_{x,x}G * I + \partial_{y,y}G * I$$
$$= \nabla \cdot \nabla(G * I) \quad (6)$$

As shown in Equ. 7, convolving I with G is another Gaussian, which is called (Gaussian) scale space. For zero shifted I, its Laplacian will get extreme at origin. Normalizing this value will get normalized Laplacian of Gaussian operation upon I, which is basis of some feature extractors.

$$G * I = \frac{c\alpha\beta}{\sqrt{\sigma^2+\alpha^2}\sqrt{\sigma^2+\beta^2}} e^{-\frac{1}{2}\left(\frac{x^2}{(\alpha^2+\sigma^2)}+\frac{y^2}{(\beta^2+\sigma^2)}\right)} + d \quad (7)$$

Applying LoG to image to get extreme points, and with information provided by $G * I$, we need get radii (standard deviations) of original input (Gaussian) function I.

2.3 Parameters Calculation

As shown before, image I can be considered as a surface in three-dimensional space. From differential geometry, we know its hessian matrix directly relates to principal curvatures and principal directions, and for Gaussian function, principal curvatures connect with its standard deviations. In one word, eigenvalues of hessian matrix relate to radii and eigenvectors relate to directions. We also know that two principal directions are perpendicular to one another. Based on these facts, we will derive formulas for parameters.

Obviously, convolving I with isotropic Gaussian will not change principal direction. For extreme point, we can use $G * I$'s principal direction as I's principal direction. For the case of axis-aligned Gaussian, we already know principal directions, otherwise, compute eigenvectors.

Remaining question is, giving information of $G * I$, how to get I's radii α and β, its contrast c and baseline height d.

Here, we will exploit a fact, that LoG can detect Gaussian at one and only one scale. In another word, every α and β pair must produce one and only one σ, as shown in Equ. 8. If analytic form of f is determined, we can recover α and β from σ.

$$\sigma = f(\alpha, \beta) \quad (8)$$

For any input image, $\frac{\beta}{\alpha}$ is fixed. LoG will detect extreme point in a fixed scale σ. Let us denote $k = \frac{\beta}{\alpha}$, and $h = \frac{\alpha}{\sigma}$.
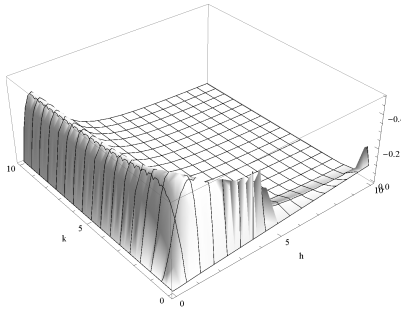
Apply normalized LoG operator to I, and substitute $\beta = \alpha k$ and $\alpha = \sigma h$, and let $x = 0$, $y = 0$, we get Equ. 9.

$$\sigma^2 \nabla \cdot \nabla(G*I)_{x=0,y=0} = -\frac{ch^2 k\left(2 + h^2\left(1 + k^2\right)\right)}{\left(\left(1 + h^2\right)\left(1 + h^2 k^2\right)\right)^{3/2}} \quad (9)$$

Let c be constance 1 and draw this expression in Fig. 2. It is clearly shown that for $k > 1$, extreme of $LoG * I$ is located on a smooth ridge.

For a fixed k, at extreme point, the formula's one order derivative will be zero. After some calculation, we can get Equ. 10.

$$-4 - 2h^2\left(1 + k^2\right) - h^4\left(1 - 6k^2 + k^4\right) + 2h^6\left(k^2 + k^4\right) == 0 \quad (10)$$

Fig. 2. k and h constrained

Fig. 3. Relation of H and ratio of eigenvalues

To solve this equation, let $k^2 = K$ and $h^2 = H$, and we get two order equation Equ. 11.

$$-4-2H-H^2+\left(-2H+6H^2+2H^3\right)K+\left(-H^2+2H^3\right)K^2 \quad (11)$$

It is easy to solve, as Equ. 12.

$$K_1 = \frac{1-3H-H^2-(1+H)\sqrt{-3+H(6+H)}}{H(-1+2H)}$$
$$K_2 = \frac{1-3H-H^2+(1+H)\sqrt{-3+H(6+H)}}{H(-1+2H)} \quad (12)$$

Known constraint of K and H, we need more information to get their values. As mentioned above, eigenvalues of hessian matrix relate to radii closely. We calculate hessian matrix over scale space, as shown in Equ. 13.

$$\begin{pmatrix} \partial_{x,x} & \partial_{x,y} \\ \partial_{x,y} & \partial_{y,y} \end{pmatrix}(G*I) \quad (13)$$

As before, we calculate eigenvalues of this matrix, and let $x = 0$, $y = 0$. Since our discuss based Gaussian, we can get analytic solution of two eigenvalues, shown in Equ. 14.

$$e_1 = -\frac{ch^2k}{(1+h^2)^{3/2}\sqrt{1+h^2k^2}\sigma^2}$$
$$e_2 = -\frac{ch^2k}{\sqrt{1+h^2}(1+h^2k^2)^{3/2}\sigma^2} \quad (14)$$

These two eigenvalues have complicated form, but their ratio is simpler, shown in Equ. 15.
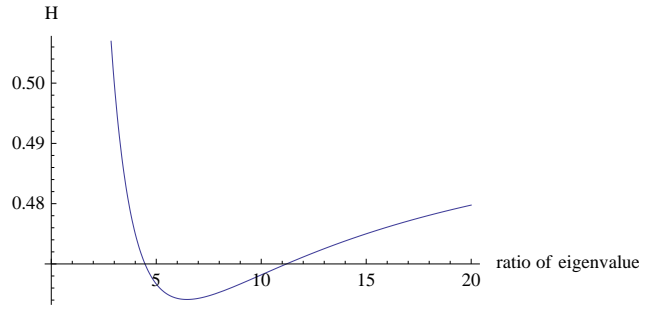
$$r = \frac{e_1}{e_2} = \frac{1+HK}{1+H} \quad (15)$$

From Equ. 15, we can solve for K, shown in Equ. 16.

$$K = \frac{-1+r+Hr}{H} \quad (16)$$

Combined Equ. 12 and Equ. 16, we can solve for H, result is Equ. 17.

$$H = \frac{3+r^2}{2r(1+r)} \quad (17)$$

We draw this relation in Fig. 3, which shows detecting scale tends to be constancy as shape gets elongate. Simply put, elongating a shape contributes little to its detecting scale.

Got H, k, h, α and β will be solved directly.

c and d can also be solved in analytic form. Equ. 9 is used to get c. Because d is constant component of scale space, it will disappear by differential operation; therefore can only be solved in scale space itself. Let $x = 0$ and $y = 0$ in $G*I$, we will get Equ. 18, so d can be solved upon extreme point of scale space.

$$G*I_{x=0,y=0} = d + \frac{c}{\sqrt{1+\frac{1}{H}}\sqrt{1+\frac{1}{HK}}} \quad (18)$$

Until now, we have calculated all parameters of the zero shifted and axis aligned Gaussian. Because axis can be shifted or rotated, our discussion will be applied to Gaussian of any position or rotation. We will summary the steps of our algorithm.

- Detect extreme point in normalized LoG space, and get its σ.
- compute hessian matrix of extreme point in corresponding scale space
- compute eigenvectors as principal directions of the point.
- compute eigenvalues, let absolute larger one divide smaller one, and represented as r
- use Equ. 17 to compute H, Equ. 16 to compute K, and use $\alpha = \sqrt{H}\sigma, \beta = \sqrt{K}\alpha$ to compute other parameters, use Equ. 9 and Equ. 18 to solve for contrast and baseline height.

2.4 Data Transformation

Until now, signal's radii and angle are extracted. In order to comparing with other methods' results, we depend on some publicly available tools. Therefore radii and angle need to be transformed to a common form, such as symmetric positive definite matrix, as shown in Equ. 19.
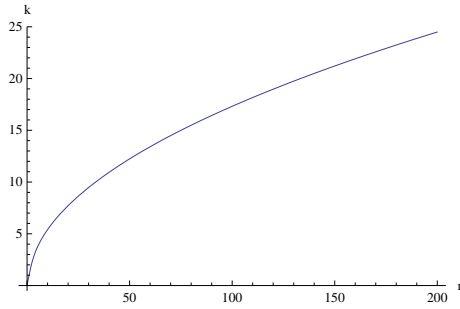
$$\begin{pmatrix} x & y \\ y & z \end{pmatrix} \quad (19)$$

If let θ be signal's orientation, and $t = \arctan(\theta)$, in a similar way as before, we get Equ. 20.

$$x = \frac{\beta+t^2\alpha}{1+t^2}$$
$$y = \frac{t(\alpha-\beta)}{1+t^2}$$
$$z = \frac{\alpha+t^2\beta}{1+t^2} \quad (20)$$

Fig. 4. relation of k and r

3 IMPLEMENTATION DETAILS

In this section, some important implementation details are outlined.

3.1 Approximation and Adjustment

As shown in Equ. 21, LoG can be implemented by DoG, and together with pyramid algorithm, which makes proposed method ready for application. We use similar DoG pyramid as Lowe's. Extremum of DoG should be adjusted by a constant multiplier, for its value is used to compute c and d.

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2 \nabla^2 G \qquad (21)$$

3.2 Removal of False Features

Tested with synthesized data, we found one common problem among several (affine) feature detectors, that is, for a single Gaussian signal, often there are several features detected out. Some of them have similar radii and orientations, located around true position, as shown in Fig. 5 and Fig. 9. Others are false features arisen from noise, as shown in Fig. 9 and Fig. 10.

In practice, we found a large part of false features coming from sampling and digitization process, that is to say, they are small sized, low contrast features. True features seldom have such properties. Therefore features with small value of c, α and β are considered as noises.

3.3 Detector Threshold

Like SIFT, we uses ratio of principal curvatures (ratio of hessian's eigenvalues, or r in our method) to remove points on valley or ridge. To accept more features, the ratio needs to be refined. Combining Equ. 17 and Equ. 16, with $K = k^2$, we have Equ. 22.

$$k = \sqrt{\frac{r + 3r^3}{3 + r^2}} \qquad (22)$$

We have drawn relation of k and r in Fig. 4. For aspect ratio k to be as high as 40, r need at least to be 535 theoretically. The r in Equ. 23 is threshold of features.

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(e_1 + e_2)^2}{e_1 e_2} = \frac{(re_2 + e_2)^2}{re_2^2} = \frac{(r+1)^2}{r} \qquad (23)$$

TABLE 2
Test Image condition

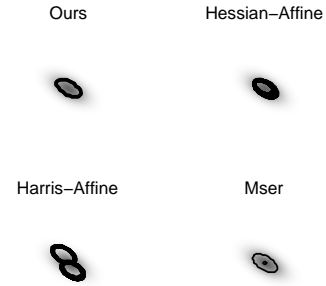| Parameter	Range
c	$[-255, 255]$
d	$[0, 255]$
$\sqrt{\alpha\beta}$	$[5, 40]$
$\frac{\beta}{\alpha}$	$[1, 30]$
θ	$\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

Fig. 5. Typical results for an isotropic Gaussian

4 EXPERIMENTS

In order to evaluate performance of our method, we firstly test it with synthesized data. In this way, we will know true parameters and therefor can compare them with calculated ones. We will compare results of our method and others, including Harris-Affine, Hessian-Affine and Mser. Only common parameters such as orientation, long and short scale can be compared, because contrast and base height are unique provided by our method. Nevertheless, we will show the results alone.

Gaussian will be used as test image. Image size is 256x256, and gray scale level is 256. Our method can detect a large range of parameters, and Tab. 2 lists parameters used in experiments.

4.1 Results of ideal signals

As demostrated in Fig. 5, Hessian-Affine and Harris-Affine tend to detect duplicated features. Fig. 6 and Fig. 7 show, for noise free Gaussian signal, Mser has highest accuracy for detecting position and aspect ratio. Our method achieves similar results as MSer. Compared with Harris-Affine, Hessian-Affine gets better results. Both Mser and our method can detect signals of high aspect ratio, but Hessian-Affine and Harris-Affine are limited to low aspect ratio signals.

Our method is to compute original parameters from blurred output image. For very long and thin shapes, our method may slightly underestimate true aspect ratio, as shown in Fig .7.

As shown in Fig. 8, our method and Mser achieve highest accuracy for detecting short radii. However, in addition to true signals, Mser often finds small concentric signals.
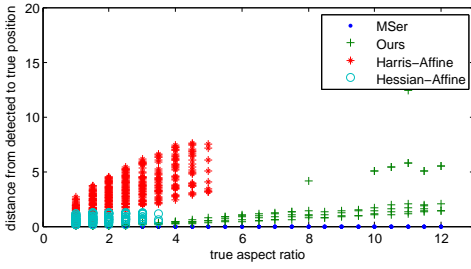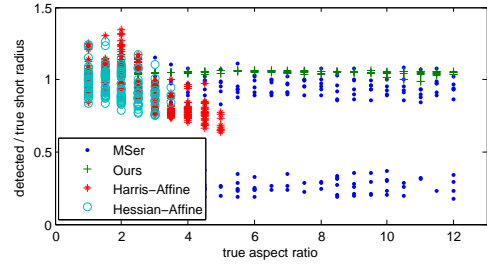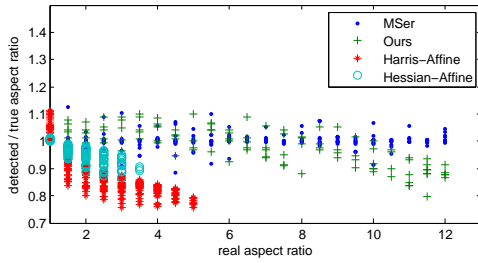
Fig. 6. Position inaccuracy

Fig. 7. Aspect ratio accuracy

In conclusion, for ideal noise free Gaussian, Mser get best results, and ours is similar to that of Mser. Hessian-Affine and Harris-Affine are not as stable as Mser and ours.

4.2 Results of noisy signals

Fig. 9 is a typical noisy image, and Mser is the most sensitive to noise. Even a small amount of noise can impact Mser seriously. Fig. 10 is distance of true and detected points. It is difficult for Mser to differentiate noises from true signals. Therefore we only compare Hessian-Affine, Harris-Affine and ours for noisy images.

As shown in Fig. 10, Fig. 11, Fig. 11, our method performs well when other methods reach their limits.

Using Mikolajczyk's evaluation images and toolbox, we get repeatability in Fig. 13. For these noisy free images, Mser get highest accuracy, and Hessian-Affine, Harris-Affine and our methods have similar results. Our 1 and 2 are results of different thresholds.

Fig. 14 is detecting results of graffiti under different view angle. Compared with Hessian-Affine and Harris-Affine, Mser and ours detect fewer features. It seems that the former two detect many redundant features. Compared with ours, Mser tends to detect many small features.

5 CONCLUSION

In this paper, we have proposed a new feature detector. Compared with other methods, it is very stable, accurate
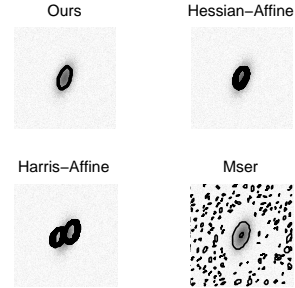

Fig. 8. Short radius accuracy

Fig. 9. Typical results for noisy Gaussian

and quick. Tested with Gaussian, for ideal noisy free signal, our method produces one of the best results, and for noisy signal, it outperforms others significantly. The proposed method can also extracts parameters unavailable for other methods, such as contrast and baseline height.

Test with benchmark images, the method get similar repeatability as Harris-Affine and Hessian-Affine.

APPENDIX A
PROOF OF ROTATION INVARIANT FOR IMAGE SURFACE

Let F be Fourier operator, and f be an input function; Fourier transform is shown in Equ. 24.

$$F \circ f(x) = \int_{\infty}^{\infty} f(x)e^{-2\pi i\langle x,\xi\rangle} dx = \hat{f}(\xi) \quad (24)$$

If input function rotates in x space, and let $\xi = Rx$, its Fourier transform will also rotate same angle, as shown in
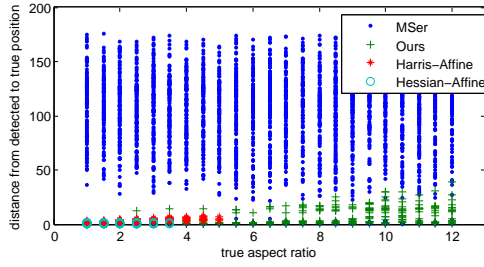
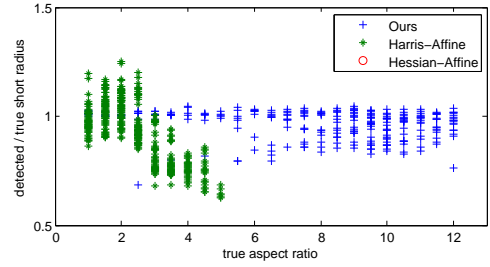Fig. 10. Position inaccuracy for noisy Gaussian

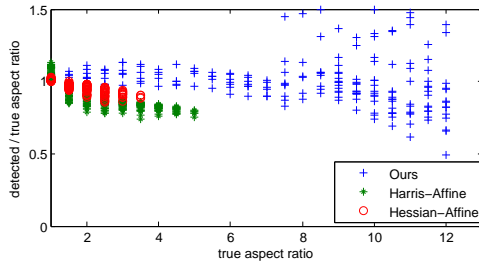Fig. 12. Short radius accuracy for noisy Gaussian
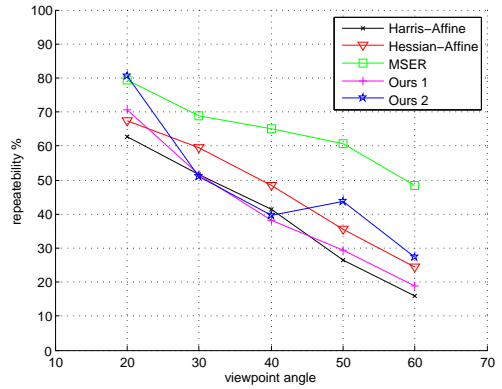

Fig. 11. Aspect ratio accuracy for noisy Gaussian

Fig. 13. Repeatebility comparasion. Our 1 and 2 have different threshold.

output's geometrical property will not change on rotating input and system.

ACKNOWLEDGMENTS

Equ. 25.

$$F \circ f(Rx) = \int_{\infty}^{\infty} f(Rx)e^{-2\pi i \langle x, \xi \rangle} dx$$
$$= \int_{\infty}^{\infty} f(y)e^{-2\pi i \langle R^T y, \xi \rangle} dy$$
$$= \int_{\infty}^{\infty} f(y)e^{-2\pi i (R^T y)^T \xi} dy$$
$$= \int_{\infty}^{\infty} f(y)e^{-2\pi i y^T R\xi} dy$$
$$= \int_{\infty}^{\infty} f(y)e^{-2\pi i \langle y, R\xi \rangle} dy$$
$$= \hat{f}(R\xi) \quad (25)$$

Convolution in space domain can be implemented by multiplication in ξ domain, as shown in Equ. 26.

$$f(x) * g(x) \leftrightarrow \hat{f}(\xi)\hat{g}(\xi) = \hat{h}(\xi) \leftrightarrow h(x) \quad (26)$$

$$f(Rx) * g(Rx) \leftrightarrow \hat{f}(R\xi)\hat{g}(R\xi) = \hat{h}(R\xi) \leftrightarrow h(Rx) \quad (27)$$

Using Equ. 24, Equ. 25 and Equ. 26, we can get Equ. 27. It means if input and system are rotated with same angle, the output will also rotate the same angle. In one word,

REFERENCES

[1] J.J. Koenderink and A.J. Doorn, "Generic Neighborhood Operators," IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, No. 6. (1992), pp. 597-605.
[2] B.M.H Romeny, *Front-end vision and multi-scale image analysis.* Berlin, Germany.: Springer Verlag, 2003.
[3] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, No. 2. (1 November 2004), pp. 91-110.
[4] K. Mikolajczyk and C. Schmid, "Scale and Affine Invariant Interest Point Detectors," International Journal of Computer Vision, Vol. 60, No. 1. (1 October 2004), pp. 63-86.
[5] T. Lindeberg, "Feature Detection with Automatic Scale Selection" International Journal of Computer Vision, Vol. 30, No. 2. (1 November 1998), pp. 79-116.
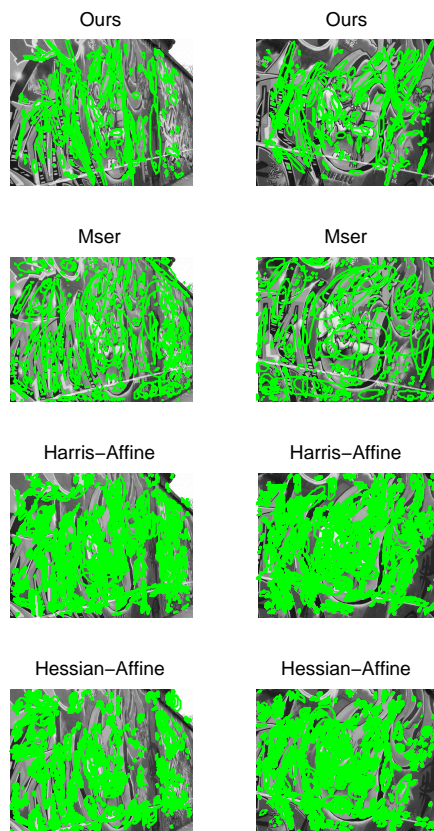
Fig. 14. Graffiti of different view angles.

[6] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 10. (October 2005), pp. 1615-1630.

[7] J.J. Koenderink "The structure of images," Biological Cybernetics, Vol. 50, No. 5, pp. 363-370-370, Aug 1984, doi:10.1007/BF00336961.

[8] T. Tuytelaars and Luc Van Gool, "Matching Widely Separated Views Based on Affine Invariant Regions," Int. J. Comput. Vision, Vol. 59, No. 1. (August 2004), pp. 61-85.

[9] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir and L. Van Gool, "A comparison of affine region detectors," International Journal of Computer Vision, Vol. 65, No. 1. (13 November 2005), pp. 43-72.

[10] J. Matas, O. Chum, U. Martin and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," In Proceedings of the British Machine Vision Conference, Vol. 1 (2002), pp. 384-393.

[11] X. Xu and J. Yang, "Directional SIFT – An Improved Method Using Elliptical Gaussian Pyramid," Chinese Conference on Pattern Recognition (CCPR '10), pp. 1 - 5, 2010, doi:10.1109/CCPR.2010.5659135.

[12] A.P. Witkin, "Scale-Space Filtering," In 8th Int. Joint Conf. Artificial Intelligence, Vol. 2, pp. 1019-1022, Aug 1983.

[13] A. Vedaldi and B. Fulkerson, "VLFeat: An Open and Portable Library of Computer Vision Algorithms," http://www.vlfeat.org/

[14] T. Kadir, A. Zisserman and M. Brady, "An Affine Invariant Salient Region Detector," In Computer Vision - ECCV 2004, pp. 228C241.

[15] D.G. Lowe, "Object Recognition from Local Scale-Invariant Features," IEEE International Conference on In Computer Vision, 1999, Vol. 2, (06 August 1999), pp. 1150-1157.