Multi-view and multi-scale recognition of symmetric patterns

Dereje Teferi and Josef Bigun

Halmstad University. SE 30118 Halmstad, Sweden
{Dereje.Teferi, Josef.Bigun}@hh.se

Abstract. This paper suggests the use of symmetric patterns and their corresponding symmetry filters for pattern recognition in computer vision tasks involving multiple views and scales. Symmetry filters enable efficient computation of certain structure features as represented by the generalized structure tensor (GST). The properties of the complex moments to changes in scale and multiple views including in-depth rotation of the patterns and the presence of noise is investigated. Images of symmetric patterns captured using a low resolution low-cost CMOS camera, such as a phone camera or a web-cam, from as far as three meters are precisely localized and their spatial orientation is determined from the argument of the second order complex moment I_{20} without further computation.

1 Introduction

Feature extraction is a crucial research topic in computer vision and pattern recognition having numerous applications. Several feature extraction methods have been developed and published in the last few decades for general and/or specific purposes. Early methods such as Harris detector [3] use stereo matching and corner detection to find corner like singularities in local images whereas more recent algorithms use extraction of other features from gradient of images [4, 7] or orientation radiograms [5] with the intention of achieving invariance or resilience to certain adverse effects in vision, e.g. rotation, scale, view and noise level changes, to match against a database of image features.

In this paper, the strength of symmetric filters in localizing and detecting the orientation of known symmetric patterns such parabola, hyperbola, circle and spiral etc in varying scales and spatial and in-depth rotation is investigated. The design of the pattern via coordinate transformation by analytic functions and their detection by symmetry filters is discussed. These patterns are non-trivial and often do not occur in natural environments. Because they are non-trivial, they can be used as artificial markers to recognize certain points of interest in an image. Symmetry derivatives of Gaussians are used as filters to extract features from their second order moments that are able to localize as well as detect the local orientation of these special patterns simultaneously. Because of the ease of detection, these patterns are used for example in vehicle crash tests by using the known patterns as markers on artificial test driver for automatic tracking [2]

and in ffingerprint recognition by using the symmetry filters to detect core and delta points (minutia points) in fingerprints[6].

2 Symmetry features

Symmetry features are discriminative features that are capable of detecting local orientations in an image. The most notorious patterns that contain such features are lines (linear symmetry), that can be detected by eigen analysis of the ordinary 2D structure tensor. However, with some care even other patterns such as parabolic, circular or spiral (logarithmic), or hyperbolic shapes can be detected but by eigen analysis of the generalized structure tensor [1, 2] which is summarized below.

First, we revise the structure tensor S which enables to determine the dominant direction of ordinary line patterns (if any) and the fitting error through the analysis of its eigenvalues and their corresponding eigenvectors. S is computed as:

$$S = \begin{pmatrix} \iint (\omega_x)^2 |F|^2 & \iint (\omega_x \omega_y)|F|^2 \\ \iint (\omega_x \omega_y)|F|^2 & \iint (\omega_y)^2 |F|^2 \end{pmatrix} \tag{1}$$

$$= \begin{pmatrix} \iint (D_x f)^2 & \iint (D_x f)(D_y f) \\ \iint (D_x f)(D_y f) & \iint (D_y f)^2 \end{pmatrix} \tag{2}$$

Where $F = F(\omega_x, \omega_y)$ is the fourier transform of f and the eigenvectors k_{max}, k_{min} corresponding to the eigenvalues λ_{min}, λ_{max} represent the inertia extremes and the corresponding axes of inertia of the power spectrum $|F|^2$ respectively.

The second order complex moment I_{mn} of a function h, where m, n are non negative integers and $m + n = 2$ is calculated as,

$$I_{mn} = \iint (x + iy)^m (x - iy)^n h(x, y) dx dy \tag{3}$$

It turns out that I_{20} and I_{11} are related to the eigenvalues and eigenvectors of the structure tensor S as follows:

$$I_{20}|F|^2 = (\lambda_{max} - \lambda_{min}) e^{i2\varphi_{min}} \tag{4}$$

$$I_{11}|F|^2 = \lambda_{max} + \lambda_{min} \tag{5}$$

$$|I_{20}| = \lambda_{max} - \lambda_{min} \leq \lambda_{max} + \lambda_{min} = I_{11} \tag{6}$$

Here $\lambda_{max} \geq \lambda_{min} \geq 0$. If $\lambda_{min} = 0$ then $|I_{20}| = I_{11}$ which signifies the existence of a perfect linear symmetry which is also the unique occasion where the inequality in Eq. (6) is fulfilled with equality, i.e. $|I_{20}| = I_{11}$. Thus a measure of linear symmetry (LS) can be written as:

$$LS = \frac{|I_{20}|}{I_{11}} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} e^{i2\varphi_{min}} \tag{7}$$

In practice this is a normalization of I_{20} with I_{11}. The value of LS falls within $[0, 1]$ where $LS = 1$ for perfect linear symmetry and 0 for complete lack of linear symmetry (balanced directions or lack of direction).

The Generalized structure tensor (GST) is similar in its essence with the ordinary structure tensor but its target patterns are "lines" in curvilinear coordinates, ξ and η. For example, using $\xi(x, y) = log\sqrt{x^2 + y^2}$ and $\eta(x, y) = tan^{-1}(x, y)$ as coordinates, "oriented lines" in the log-polar coordinate system ($a\xi(x, y) + b\eta(x, y) = constant$), GST will simultaneously estimate evidence for presence of circles, spirals and parabolas etc. In GST, the I_{20} and I_{11} interpretations remain unchanged except that they are now with respect to lines in curvilinear coordinates, with the important restriction that the allowed curves for coordinate definitions must be drawn from harmonic curve family. [2] has shown that as the consequence of local orthogonality of ξ and η the complex moments I_{20} and I_{11} of the harmonic patterns can be computed in the cartesian coordinates system without the need for coordinate transformation as:

$$I_{20} = \iint e^{iarg((D_\xi - iD_\eta)\xi)^2} [D_x + iD_y f]^2 dxdy \tag{8}$$

$$I_{11} = \iint |(D_x + iD_y)f|^2 dxdy \tag{9}$$

where $\eta = \eta(x, y)$ and $\xi = \xi(x, y)$ represent a pair of harmonic coordinate transformations. Such pairs of harmonic transformations satisfy the following constraint: $\xi(x, y) = constant_1$ and $\eta(x, y) = constant_2$ are orthogonal to each other i.e. $D_x\xi = D_y\eta$ and $D_y\xi = -D_x\eta$.

Thus, the measure of linear symmetry in the harmonic coordinate system by the generalized structure tensor is in fact the analogue of the measure of linear symmetry by the ordinary structure tensor but in a cartesian coordinate system. The advantage is that we can use the same theoretical and practical machinery to detect the presence and quantify the orientation of for example parabolic symmetry (PS), circular symmetry (CS), hyperbolic symmetry (HS) drawn in cartesian coordinates depending on the analytic function $q(z)$ used to define the harmonic transformation. Some of these patterns are shown in Figure 1 where the iso-curves represent a line as $a\xi + b\eta = constant$ for predetermined ξ and η.

Harmonic transformation pairs can be readily obtained as the real and imaginary parts of (complex) analytic functions by restricting us further to $q(z)$ such that $\frac{dq}{dz} = z^{\frac{n}{2}}$. Thus we have,

$$q(z) = \begin{cases} \frac{1}{\frac{n}{2}+1} z^{\frac{n}{2}+1} \text{ if } n \neq -2 \\ \\ log(z), \quad \text{ if } n = -2 \end{cases} \tag{10}$$

Each of the curves generated by the real and imaginary parts of $q(z)$ can then be detected by symmetry filters Γ shown in the fourth row of Figure 1. The gray values and the superimposed arrows respectively show the magnitude and orientation of the filter that can be used for detection.
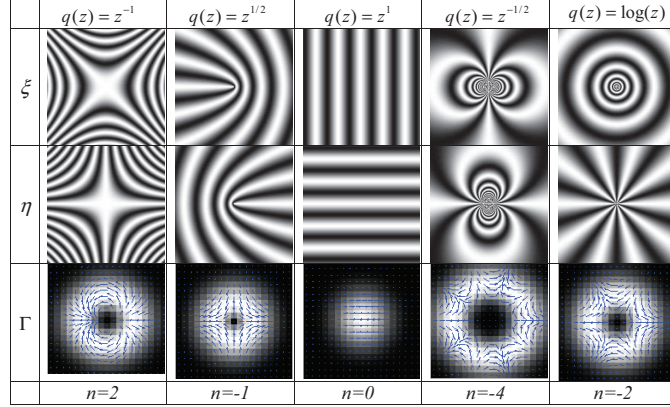
	$q(z) = z^{-1}$	$q(z) = z^{1/2}$	$q(z) = z^{1}$	$q(z) = z^{-1/2}$	$q(z) = \log(z)$
ξ					
η					
Γ					
	$n=2$	$n=-1$	$n=0$	$n=-4$	$n=-2$

Fig. 1. First row: Example harmonic function $q(z)$, second and third rows show the real and the imaginary parts ξ and η of the $q(z)$ where $z = x + iy$. The fourth row shows the filters that can be used to detect the patterns in row 2 and 3. The last row shows the order of symmetry

$$\Gamma^{\{n,\sigma^2\}} = \begin{cases} (D_x + iD_y)^n g & \text{if } n \geq 0 \\ \\ (D_x - iD_y)^{|n|} g & \text{if } n < 0 \end{cases} \tag{11}$$

Here $g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$ is the Gaussian and n is the order of symmetry. For $n = 0$, Γ is an ordinary Gaussian. Moreover, $(D_x + iD_y)$ and $\left(\frac{-1}{\sigma.^2}\right)^p (x + iy)^p$ behave identically when acting on and multiplied to, a Gaussian respectively [2, 1]. Due to this elegant property of Gaussians functions, the symmetry filters in the above equation can be rewritten as:

$$\Gamma^{\{n,\sigma^2\}} = \begin{cases} \left(-\frac{1}{\sigma^2}\right)^n (x + iy)^n g & \text{if } n \geq 0 \\ \\ \left(-\frac{1}{\sigma^2}\right)^{|n|} (x - iy)^{|n|} g & \text{if } n < 0 \end{cases} \tag{12}$$

3 In-depth (non-planar) rotation of symmetric patterns

Recognition of a pattern when rotated spatially in 3D is a challenging issue and requires resilient features. To test the strength of the symmetry filters in recognizing patterns viewed from different angles, we rotated the patterns geometrically using ray tracing as follows.

Suppose we are looking at the world plane W from point O through an image plane I in a pin-hole camera model as in Figure 2. Note that, if the image plane I is parallel to the world plane W, we would see a zoomed version of the world image depending on how far the image plane is from the world plane. When W is not parallel to I, then the image plane is a skewed and zoomed version of the world plane.
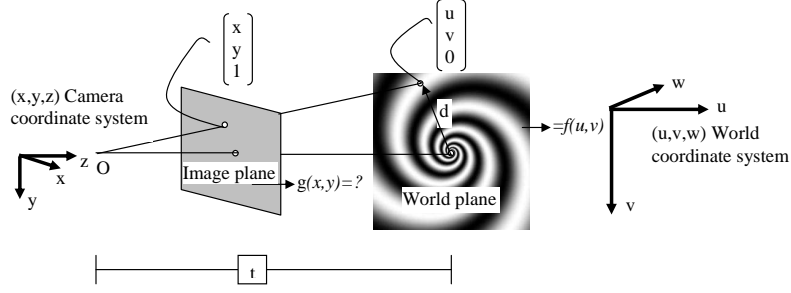
Fig. 2. Ray tracing for non-planar rotation

A point P represented in the world coordinates as d, transfers to the camera coordinates as $R(t + d)$ if both t and d are in world coordinates. Here R is a rotation matrix aligning the world coordinate axes with the camera coordinate axes and t is the translation vector aligning the origin of the world coordinate system to the origin of the camera coordinate system. The rotation matrix R of the world plane is the product of the rotation matrices around each axis R_x, R_y, and R_z relative to the world coordinates. As an example R_x is given as:

$$R_{x(\alpha)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\alpha) & -sin(\alpha) \\ 0 & sin(\alpha) & cos(\alpha) \end{pmatrix} \tag{13}$$

similarly R_y and R_z are defined and the overall rotation matrix R is given as:

$$R = R_{x(\alpha)} * R_{y(\beta)} * R_{z(\gamma)} \tag{14}$$

The normal n to the world plane is the 3rd row of the rotation matrix R expressed in the camera coordinates.

To find the distance vector from O to the world plane W, we can proceed in two ways as $L^T n$ and $t^T n$. Because both measure the same distance, they are equal, i.e. $L^T n = t^T n$

$$L = \tau \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \tau L_s \Rightarrow \tau L_s^T n = t^T n \tag{15}$$

where $L_s = (x, y, 1)^T$. Thus

$$\tau = \frac{t^T n}{L_s^T n} \tag{16}$$

$$\Rightarrow L = \left(\frac{t^T n}{L_s^T n} L_s \right) \tag{17}$$

$$d = R(L - t) \tag{18}$$

Accordingly, $g(x, y) = f(u, v)$, where $d = (u, v, 0)$. The last two rows of Figure 3 show the results of some of the symmetric patterns painted on the world plane but observed by the camera in the image plane.
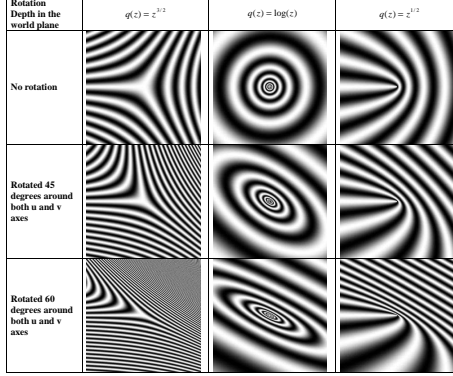

Rotation Depth in the world plane	$q(z) = z^{3/2}$	$q(z) = \log(z)$	$q(z) = z^{1/2}$
No rotation			
Rotated 45 degrees around both u and v axes			
Rotated 60 degrees around both u and v axes			

Fig. 3. Illustration of in-depth rotation of symmetric patterns

4 Experiment

4.1 Recognition of symmetric patterns using symmetry filters

We used the filter designed as in Eq. 12 to detect the family of patterns f generated by the analytic function $q(z)$, where

$$f = cos(k_1 \Re(q(z)) + k_2 \Im(q(z))) + 1 \tag{19}$$

Here $q(z)$ is given by Eq. 10 and $n \in -4, -3, -2, -1, 0, 1, 2$

The following steps are applied on the image to detect the pattern and its local orientation:

1. Compute the square of the derivative image h_k by convolving the image f with a symmetry filter of order 1, $\Gamma^{\{1, \sigma_1^2\}}$ and pixelwise squaring of the complex valued result as:
   $h_k = <\Gamma_k^{\{1, \sigma_1^2\}}, f_k>^2$. Here σ_1 controls the extension of the interpolation function, i.e. the size of the derivative filter Γ^{1, σ_1} that is modeling the expected noise in the image;
2. Compute I_{20} by convolving the complex image h_k of step 1 with the appropriate complex filters from Eq 12 according to their pattern family defined by n and by their expected spatial extension controlled by σ_2. That is:
   $I_{20} = <\Gamma_k^{\{m, \sigma_2^2\}}, h_k>$.
3. Compute the magnitude image I_{11} by convolving the magnitude of the complex image h_k with the magnitude of the symmetry filters from Eq 12 as:
   $I_{11} = <|\Gamma_k^{\{m, \sigma_2^2\}}|, |h_k|>$;

4. Compute the certainty image and detect the position and orientation of the symmetry pattern from its local maxima. The argument of I_{20} at locations characterized by high response of the certainty image, I_{11} yields the group orientation of the pattern;

The strength of the filters in detecting patterns and their rotated version is tested by applying the in-depth rotation of the symmetric patterns as discussed in the previous section. Figure 4 illustrates the detection results of circular and parabolic patterns rotated 45° around the x and y axes.
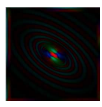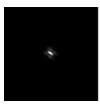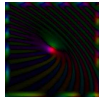
Original Image	Rotated Image $I_{(45,45)}$	Complex moment I_{20}	Detected pattern I_{20}/I_{11}

Fig. 4. Detection of symmetric patterns using symmetry derivatives of Gaussians on simulated rotated patterns

The color of the I_{20} image corresponding to the high response on the detected pattern (last column) indicates the spatial orientation of the symmetric pattern. The filters are also tested on real images captured with low-cost off the shelf CMOS camera. The result shows that symmetry filters detect these patterns from distance of up to 3 meters and in-depth rotation of up to 45 degrees, see Table 1. Similar result is achieved with web cameras and phone cameras as well. The color of the I_{20} once again indicates the spatial orientation of the symmetric pattern detected.
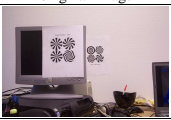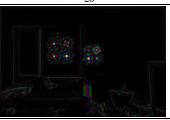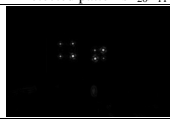

	Original Image	I_{20}	Detected patterns I_{20}/I_{11}
d=1.5m			
d=2 m			
d=2 m $\alpha=\pi/4$			

Fig. 5. Detection of symmetric patterns in real images using symmetry filters

Rotation (in-depth)	Distance from image and accuracy			
	2 meters		3 meters	
	d	α	d	α
0°	±1 pixel	±2°	±1 pixel	±5°
30°	±1 pixel	±3°	±1 pixel	±8°
45°	±2 pixel	±6°	±2 pixel	±12°
60°	±3 pixel	±15°	±4 pixel	±20°

Table 1. Average results of recognition of symmetric patterns from multiple views. d=localization error and α=orientation error. The test is performed on 12 different images, e.g. Figure 5 captured by a 2.1 megapixel CMOS camera. Each of the images are subjected to zooming and in-depth rotation as in Figure 4 but naturally

4.2 Recognition of symmetric patterns using Scale Invariant Feature Transform-SIFT

Lowe [4] proposed features known as SIFT to match images representing different views of the same scene by using histograms of gradient directions. The features extracted are often used for matching between different views severed by scale and in-depth local rotation as well as illumination changes. SIFT feature matching is one of the most popular object detection methods.

The SIFT approach uses the following four steps to extract the location of a singularity and its corresponding feature vector from an image and store them for subsequent matching.

1. Scale-space extrema detection: this is the first step where all candidate points that are presumably scale invariant are extracted using arguments from scale-space theory. The implementation is done using Difference of Gaussian (DoG) function by successively subtracting images from its Gaussian smoothed version within an octave;
2. Keypoint localization: the candidate points from step 1 that are poorly localized and sensitive to noise, especially those around edges, are removed;
3. Orientation assignment: in this step, orientation is assigned to all key points that have passed the first two steps. The orientation of the local image around the key point in the neighborhood is computed using image gradients;
4. Extracting keypoint descriptors: the histograms of image gradient directions are created for non-overlapping subsets of the local image around the key point. The histograms are concentrated to a feature vector representing the structure in the neighborhood of the key points to which the global orientation computed in step 3 is attached.

The SIFT demo software[1] can be used to extract the necessary features to automatically recognize patterns in an image such as those shown in Figure 5. To this end, we used real images (containing symmetric patterns), e.g. the 2nd

---

[1] SIFT Demo http://www.cs.ubc.ca/ lowe/keypoints/

and 3rd rows of Figure 4, so that a set of SIFT features could be collected for each image. However, keypoint extraction failed often presumably. The method returned a few key points or in some cases failed to return any key point at all in the extraction of the SIFT features.

SIFT features are often successful in extraction of discriminative features in images and are widely used in computer vision. The key points at which these features are extracted are essentially based on their lack of linear symmetry (orientation of lines) in the respective neighborhood, e.g. to detect corner like structures. These keypoints as well as the corresponding features are organized in such a way that they could be matched against keypoints in other images with similar local structure. However, the lack of linear symmetry does not describe the presence of a specific model of curves in the neighborhood such as parabolic, circular, spiral, hyperbolic etc. In our case, lack of linear symmetry in addition to existence of known types of curve families as well as their orientation can be precisely determined, as demonstrated in Figure 4. Although these patterns are structurally different, SIFT keypoints consider them as the same often with only one key point - the center of the pattern leaving the description of the neighborhood type to histograms of gradient angles(SIFT features). The center of the pattern is chosen as a key point by SIFT since that is where there is a lack of linear symmetry. However, SIFT features apparently cannot be used to identify what patterns are represented around the key point because all orientations equally exist in the local neighborhood for all curve families despite their obvious differences in shape.

Two of the images from Figure 1 are used to test the capability of SIFT features in detecting the patterns in real images. Additive noise is applied to the images to study the change in extraction of keypoints as well as the corresponding SIFT features. The clean images returned 1 and 6 key points and the noisy images returned 89 and 101 key points, see Figure 6. Although, 89 and 101 key points are extracted from the two noisy images, none of these points actually match to the patterns in the real scene which contains these patterns, last row of Figure 6.

5   Conclusion and further work

In conclusion, the strength of the responses of symmetry filters in detecting symmetric patterns that are rotated (planar and in-depth) is investigated. It is shown via experiments that images of symmetric patterns (see Figure 5) used as artificial landmarks in a realistic environment can be localized and their spatial orientation simultaneously detected by symmetry filters from as far as 3 meters and in-depth rotation of 45 degrees. The images are captures by a low resolution commercial 2.1 megapixel Kodak CMOS camera. The results of this experiment illustrated that symmetry filters are resilient to in-depth rotation and scale changes in symmetric patterns. On the other hand, it is shown that SIFT features lack the ability to extract keypoints from these patterns as they look for lack of linear symmetry (existence of corners) and not the presence of
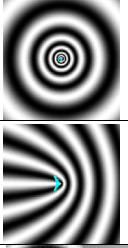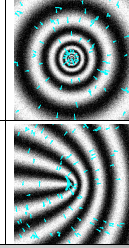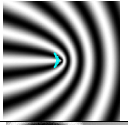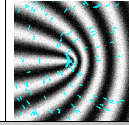
	Original Image with extracted keypoints	No of Keypoints extracted	Noisy Image with extracted keypoints	No of keypoints extracted
$g(z) = log(z)$		1		89
$g(z) = z^{1/2}$		6		101
Result of SIFT based matching using the Demo software				89, 921

Fig. 6. Extraction and matching of keypoints on Symmetric patterns and their noisy counterparts using SIFT

certain types of known symmetries. SIFT feature extraction fails because all orientations equally exist around the center of the image which makes it difficult for SIFT feature to find differences in the gradients in the local neighborhood.

The findings of this work can be applied for automatic camera calibration where symmetric patterns are used as artificial markers in a non-planar arrangement in a world coordinate system to automatically determine the intrinsic and extrinsic parameter matrices of a camera by point correspondence. Other possible applications include generic object detection and encoding and decoding of numbers using local orientation and shape of symmetric images.

References

1. J. Bigun. Vision with direction. Springer, 2006.
2. J. Bigun, T. Bigun, and K. Nilsson. Recognition of symmetry derivatives by the generalized structure tensor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(12):1590–1605, December 2004.
3. C. Harris and M. Stephens. A combined corner and edge detector. In *Fourth Alvey Vision Conference, Manchester, UK*, pages 147–151, 1988.
4. D. G. Lowe. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 60(2):91–110, 2004.
5. S. Michel, B. Karoubi, J. Bigun, and S. Corsini. Orientation radiograms for indexing and identification in image databases. In *European conference on signal processing (Eupsico), Trieste*, pages 693–696, September 1996.
6. K. Nilsson and J. Bigun. Localization of corresponding points in fingerprints by complex filtering. *Pattern Recognition Letters*, 24:2135–2144, 2003.
7. C. Schmid and R. Mohr. Local gray value invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530–534, 1997.