STRUCTURAL TEXTURE SEGMENTATION USING AFFINE SYMMETRY

Heechan Park   Graham R Martin   Abhir H Bhalerao

Department of Computer Science
University of Warwick, Coventry, United Kingdom
email: {heechan, grm, abhir}@dcs.warwick.ac.uk

ABSTRACT

Many natural textures comprise structural patterns and show strong self-similarity. We use affine symmetry to segment an image into self-similar regions; that is a patch of texture (blocks from a uniformly partitioned image) can be transformed to other similar patches by warping. If the texture image contains multiple regions, we then cluster patches into a number of classes such that the overall warping error is minimized. Discovering the optimal clusters is not trivial and known methods are computationally intensive due to the affine transformation. We demonstrate efficient segmentation of structural textures without affine computation. The algorithm uses Fourier Slice Analysis to obtain a spectral contour signature. Experimental evaluation on structural textures shows encouraging results and application on natural images demonstrates identification of texture objects.

1. INTRODUCTION

Texture segmentation is a well-researched topic in image processing. It has been studied for decades and utilised in many applications including tracking, surveillance, medical imaging and robot navigation. Texture is an ambiguous term but there are two prominent definitions. One defines a primitive individual element, *texton* that comprises texture, *e.g.* blob on jaguar. The other derives primitive statistical feature vectors from a group of pixels where the visual element is difficult to observe, *e.g.* the texture of grass. In this work, we consider an approach based on the texton that exhibits a structural pattern. This follows from an observation that many natural textures comprise structural patterns and show strong self-similarity. We attempt to capture the self-similarity by affine symmetry; that is a part of a texture region can be warped to approximate other parts. Previously, Wilson and Li [1] performed texture segmentation using the same principle and demonstrated promising results. They used the warping error as one of the distance metrics in their Markov random field framework. The method required considerable computation as the affine transformation of all pairs of patches was computed. Later, Bhalerao and Wilson [2] reduced the computation by adopting the Fourier power spectrum as a single long feature vector and achieved invariance by having an affine symmetric group of blocks as centroid, the members of which were derived from a single block by scale and orientation change. However, both approaches require manual selection of the initial prototypical blocks. Park *et. al* [3] showed unsupervised block classification based on the number of directional features using independent component analysis (ICA). We present a texture segmentation algorithm that is affine invariant, requires less computation and is unsupervised.

Section 2 covers affine invariant feature extraction. Section 3 illustrates the overall process. Section 4 reports the experimental results. Lastly, conclusions are drawn.

2. AFFINE INVARIANT FEATURES

We present an efficient and affine invariant feature based on directional information of the texture pattern. We employ the Fourier transform with a $\cos^2$ window to capture directional features. The directional information is then extracted from the local spectrum using Fourier Slice Analysis. The extraction proceeds as follows:

1. Apply Fourier transform to texture block
2. Extract contour signature using Fourier slice analysis
3. Extract affine invariant features from the signature using Fourier description

An analogous approach [4] was suggested in 80s, which was developed for remote sensing of terrains but different in the contour definition.

2.1. Fourier Spectrum

A general affine transformation T in R^2 space is define as

$$T(x) = Ax + t \tag{1}$$

The Fourier transform is well utilized in [1, 2] to separate the affine transform, T into a linear part, A which affects only the magnitude spectrum linearly, and a translational part, t that is exhibited as a phase gradient. This implies that we can analyze deformation of texture in the frequency domain,

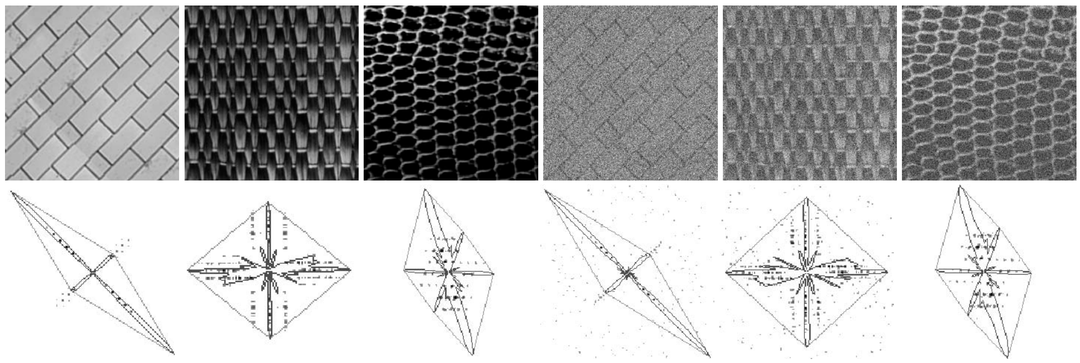

Fig. 1. Structural texture samples: `brick`, `weave`, `reptile` and the same sample with white Gaussian noise (5dB) (top) and Fourier magnitude spectrum of the texture and extracted shape (bottom)

regardless of translational difference. We extract a contour signature from the local magnitude spectrum, which is then described by an affine invariant descriptor.

2.2. Fourier Slice Analysis

The radon transform integrates a 2D function $f(\cdot)$ over a set of lines in the angle θ between the line and the axis. The transform reveals the directional strength of the texture pattern. A similar analysis with the Fourier transform is possible by the projection-slice theorem. The theorem states that an angular slice through the origin at angle, θ of the Fourier transform (*Fourier Slice*) is the Fourier transform of the Radon projection at angle θ. The slice analysis involves the computation of projection $r(\theta)$ for $0 \leq \theta < \pi$ and $\sqrt{x^2+y^2} \leq blocksize/\sqrt{2}$, which resulting an adaptively shaped closed contour. The contour signature, S, consists of a set of points as below

$$\begin{aligned} S &= \{[r(\theta)\cos\theta, r(\theta)\sin\theta]^T\} \\ r(\theta) &= c \cdot \sum\sum |F(x,y)|\delta(x\cos\theta + y\sin\theta)/e^{\sigma^2} \quad (2) \end{aligned}$$

where c and σ^2 indicate a normalization constant and the variance of the *half* Fourier slice, respectively and e^{σ^2} is the regularity term that reflects the strength of harmonic signal. Fig.1 shows the sample texture in the top row and the extracted directional shapes. The right three texture samples are identical to those on the left except for the addition of white Gaussian noise (5dB). As shown, the extracted shapes (red) are consistent regardless of the presence of noise.

2.3. Extraction of Affine Invariant Feature

Given a contour signature, its constituent boundary pixels C-$=\{[x(t), y(t)]^T\}$ are traversed to yield a parametric equation based on the affine length of a closed curve, as below. This is linear under affine transformation and also yields the same parameters, independent of the initial representation.

$$\int_C \sqrt[3]{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}dt \quad (3)$$

where the number of dots indicate the order of the derivatives. Having encoded the boundary as a function of the parameter, taking the Fourier transform of the boundary equation results in $[U, V]^T$, where U and V are Fourier coefficients referring to the x and y coordinates respectively. Since the Fourier transform is a linear operator, the equation below holds

$$[U_k, V_k]^T = A[U_k^0, V_k^0]^T \quad (4)$$

where $[U_0, V_0]^T$ denotes the same coefficients from the affine transform of the reference block. By including another coefficient and extending eq.(4) to a 2×2 matrix, obtaining their determinants reveals a linear factor. A simple division of both sides by either side produces an absolute affine invariant feature. For more details, readers are referred to [4].

3. AFFINE INVARIANT CLASSIFICATION

This section shows the overall workflow of our segmentation algorithm. Regarding the window size for texture analysis, there is a problem known as the class-boundary uncertainty *i.e.* if we confine the analysis to a small window, we get a better resolution of segmentation but we lose confidence of the texture characteristics within the window. On the other hand, a larger window allows a better analysis of texture but results in a coarse resolution of segmentation. An solution is to employ a multi-resolution approach. Texture information at the top level passes down to the lower level and is combined recursively as it proceeds. We use the Multi-resolution Fourier Transform (MFT) [5] as given below. For a given scale(σ), frequency(ω), position(ξ) and image $f(\xi)$, the 2D MFT is defined as,

$$F(\xi, \omega, \sigma) = det(\sigma I)^{\frac{1}{2}} \int_{-\infty}^{\infty} w(\sigma I(x-\xi))f(\xi)e^{-j\omega x}dx \quad (5)$$

where w denotes an appropriate window function as given below and I is the identity matrix.

$$w(x) = \cos^2[\pi p/2B]\cos^2[\pi p/2B] \quad (6)$$

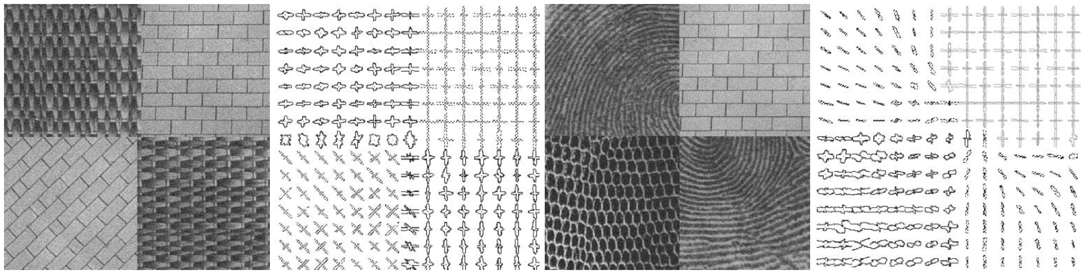

Fig. 2. Segmentation test of texture composites: texture composite and classification result in pairs, white Gaussian noise (15dB) is added to both texture samples

where $x = [p, q]^T$ and B is the block size. Having applied the MFT to a source image, the affine invariant features are extracted from the Fourier spectrum at each scale as described in section 2. The resultant features, v at the bottom scale are joined together with the quadtree parent as follows:

$$Feature(i, j) = \{p^l(V)|0 \leq l \leq k\}$$
$$p^l(V_{i,j}) = V_{\lfloor i/2^k \rfloor, \lfloor j/2^k \rfloor} \;\; V_{i,j,k} = w_k \times \{v_{i,j,k}\} \tag{7}$$

where w_k and k refer to a weight and scale respectively. The combined features are fed into a K-means clustering algorithm. Prior knowledge of the number of classes is required due to the nature of the algorithm. The K-means clustering is chosen only for simplicity as our main focus is to determine an effective affine invariant feature. A random field approach may be of interest if convergence of the number of classes is desired [1, 6].

3.1. Selection of Representative Block

The convex hull-area ratio is a new metric for selection of the most representative texture block in a class. It can reveal the strength of directional patterns in the block. This follows the similar idea in [3], where a simple isometric measure based on PCA was used to determine the most representative block in a class under two observations. One is that a block with strong directional pattern and high contrast produces less warping error. The other is that only a simple directional pattern is likely to exist as we constrain the window size (16 × 16). However, as the window size increases, the metric does not work well due to the lack of ability to deal with the mult-directional features. The new metric can indicate the strength of directional pattern well and it can also be computed efficiently by a Grahams scan. Given a set of points, s , the convex hull ? area ratio, R, is given as

$$R(s) = \frac{Area(s)}{Area(ConvexHull(s))}$$
$$Area([x, y]^T) = \frac{1}{2}\sum_{i=0}^{N-1} x_i y_{i+1} - x_{i+1} y_i \tag{8}$$

where N denotes the number of points. The convex hull of the shape is shown in (green) in Fig.1.

4. EXPERIMENTS

To illustrate the effectiveness, the algorithm was tested using simple composites of sample structural textures to which Gaussian noise (15dB) was added, as shown in Fig.2. The contour signature extraction in the presence of noise is illustrated in Fig.1.

Firstly, it is observed that the signature extracted from the Fourier transform is too jagged in the case of a severe level of noise ($0dB$) due to the scattered high-frequency coefficients, which in turn disrupts the affine invariant description and results in a poor clustering. Consequently, the scattered high-frequency coefficients are removed using minimax thresholding before the Fourier slice projection and the Gaussian smoothing filter is applied to the shape boundary. This, in fact, makes the shape-extraction robust to noise. The classification result shows the extracted shapes colored according to classes in Fig.2.

Secondly, many contour signatures from the bottom level (16× 16) of the MFT were elliptical, which makes affine invariant shape description useless. Starting with a window of 32 × 32 at the bottom level still produces an acceptable result. However, we have found that using the $area$ of the contour at the bottom level feature produces a better result. This is because the shape size increases with the strength and directionality of the feature and decreases as the directional pattern becomes less significant. It is probable that different textures with the same number of directional features fall into the same class, but it is the most information we can gather at the bottom level. The structural information from a bigger window is passed on from the parent block and combined together as discussed Section 3.

Having successfully tested the algorithm on the sample test image set, the natural images, `jaguar` and `zebra` were used as shown in Fig.3. The segmented texture was obtained by application of region growing from the block classification. The jaguar blob texture is segmented well, but there are some holes in the zebra stripes segmentation that result from the smaller window size (no texture in a window), compared with the bold stripes of the zebra. To see if the block classification result is indeed affine symmetric, the image is reconstructed

Fig. 3. Block classification test on natural images: `jaguar` (top) and `zebra` (bottom), original image, classification result, segmentation and reconstruction (by column)

by selecting a prototypical block in each class as described in Section.3.1 and replacing other blocks by the affine transformed prototype. The affine model is estimated using the two-component method [7, 3] and the prototype blocks are chosen by eq.(8). A non-affine invariant classification would have resulted in an odd image, for instance, the log featuring blobs of the jaguar. The reconstruction quality varies around 25dB in PSNR and is visually acceptable. This work not only presents an interesting approach to the segmentation task but also offers a feasible solution for efficient implementation. We intend to employ the algorithm in our affine image coding system described in [3].

5. CONCLUSION

Affine invariance has received much attention with the recent emergence of content based retrieval systems. The underlying concept has been applied to texture segmentation by many researchers. The complexity of the algorithms, however, has been a major issue in prohibiting their practical implementation. Our motivation has been to develop a computationally efficient image texture classification algorithm while maintaining the texture discriminative power of previous approaches. We have demonstrated a simple and efficient approach utilizing affine invariant shape description. Experimental evaluation indicates acceptable segmentation results for structural texture such as brick wall, jaguar and reptile skin, and also the algorithm's robustness to noise. Further study using a random field may improve results. The presented work may be of interest where efficient texture segmentation is demanded.

6. REFERENCES

[1] R. Wilson and C.T. Li, "A class of discrete multiresolution random fields and its application to image segmentation," *IEEE Trans. PAMI*, vol. 25, no. 1, pp. 42–56, 2002.

[2] A. Bhalerao and R. Wilson, "Affine invariant image segmentation," in *BMVC.*, 2004.

[3] H. Park, A. Bhalerao, G.R. Martin, and A.C. Yu, "An affine symmetric approach to natural image compression," in *Mobimedia*, 2006.

[4] J. Davernoy, "Optical digital processing of directional terrain textures invariant under translation, rotation and change of scale," *Applied Optics*, vol. 23, no. 6, pp. 828–837, 1984.

[5] R. Wilson, A.D Calway, and E.R.S Pearson, "A generalized wavelet transform for fourier analysis: themultiresolution fourier transform and its application to image and audiosignal analysis," *IEEE Trans. on Image Processing*, vol. 38, pp. 674–690, Mar. 1992.

[6] C.T. Li, "Multiresolution image segmentation integrating gibbs sampler ang region mergion algorithm," *Signal Processing*, vol. 83, pp. 67–78, 2003.

[7] T.I Hsu and R. Wilson, "A two-component model of texture for analysis and synthesis," *IEEE Trans. on Image Processing*, vol. 7, no. 10, pp. 1466–1476, October 1998.