Warplets: An Image-Dependent Wavelet Representation

Abhir Bhalerao   Roland Wilson
Signal and Image Processing Group, Department of Computer Science, University of Warwick, UK
{abhir|rgw}@dcs.warwick.ac.uk

Overview

A novel image-dependent representation, warplets, based on self-similarity of regions is introduced. The representation is well suited to the description and segmentation of images containing textures and oriented patterns, such as fingerprints. An affine model of an image as a collection of self-similar image blocks is developed and it is shown how textured regions can be represented by a single prototype block together with a set of transformation coefficients. Image regions are alligned to a set of dictionary blocks and their variability captured by PCA analysis. The block-to-block transformations are found by Gaussian mixture modelling of the block spectra and a least-squares estimation. Clustering in the Warplet domain can be used to determine a warplet dictionary. Experimental results on a variety of images demonstrate the potential of the use of warplets for segmentation and coding.

1. An Affine Image Model

An image representation that exploits self-similarity in natural images is presented. We use Gaussian Mixture Modelling (GMM) to generalize the 'two-component' affine texture model of Hsu, Calway and Wilson [1, 2]. Image regions or patches are modelled as affine transformed versions of a prototypical patch, or **texton**, plus a stochastic residual:

$$f_q(\mathbf{x}) = \alpha_{pq} f_p(A_{pq}\mathbf{x} + \mathbf{t}_{pq}) + \varepsilon_q(\mathbf{x}) \tag{1}$$

where $A_{pq}, \mathbf{t}_{pq}$ are the parameters of an affine transformation of block p onto block q and the residual is normally distributed: $\varepsilon \sim N(0, \sigma^2)$.

- A GMM of the prototype block's magnitude spectrum is fit to each target block in turn to estimate the linear transformations, A. Translations are found by correlation.
- To discover a compact set of textons to represent the image we inverse transform or **warp** all image blocks to the prototype's coordinate frame and cluster on the residual errors.
- An eigenimage analysis of the set of warped image blocks is used to model their variability. The Image Warplet is defined as the set of warping block-to-block transformations and the eigenmodes of its PCA.

2. Model-and-Match

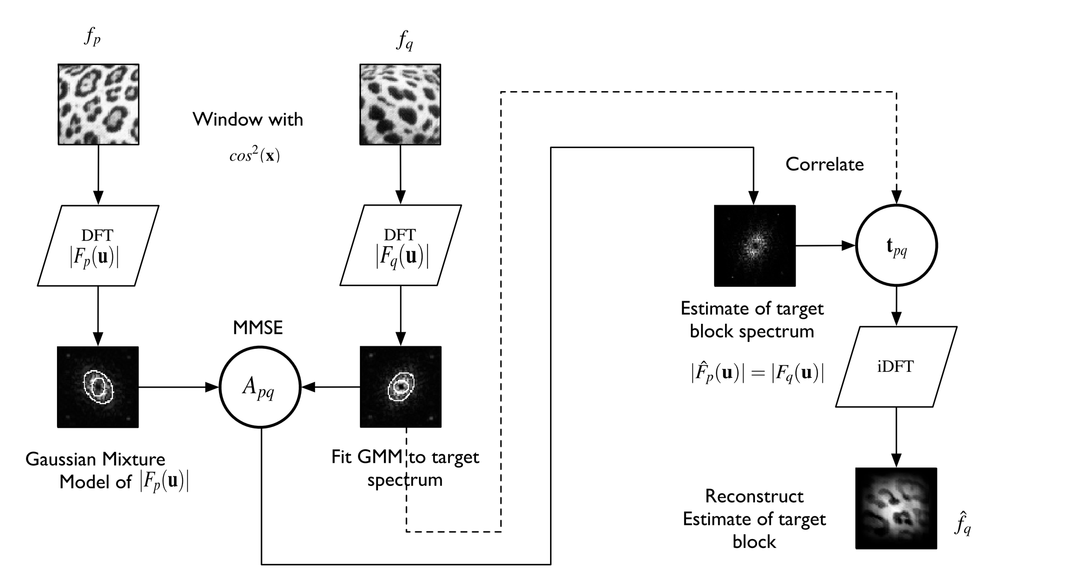


FIGURE 1. **Schematic of estimatation of block-to-block affine transformations.**

The search for the affine transformation between patches is performed by a model-and-match method (Figure 1).

- The model is a Gaussian Mixture Model of the amplitude spectrum of the patch. An M component, zero mean, mixture of the spectrum of p is:

$$G_p(\mathbf{u}) = \sum_m^M a_m \exp(-\mathbf{u}^T C_m^{-1}\mathbf{u}/2). \tag{2}$$

  Using a non-linear optimization method, its parameters $\{a_m, C_m\}$ are estimated by minimizing the residual error,

$$\sum_{\mathbf{u}} (|F_p(\mathbf{u})| - G_p(\mathbf{u}; a_m, C_m))^2. \tag{3}$$

- The model is moved by a second regression to find A_{pq} to fit the spectrum of the target patch, q minimizing $\sum_{\mathbf{u}}(G_p(A_{pq}(\mathbf{u})) - |F_q(\mathbf{u})|)^2$. This requires the gradients of $H_p(\mathbf{u}) = G_p(A(\mathbf{u}))$ w.r.t. the parameters of the linear transformation matrix:

$$\frac{dH_p(\mathbf{u})}{dA} = -H_p(\mathbf{u}) \sum_m^M C_m^{-1} A\mathbf{u}\mathbf{u}^T. \tag{4}$$

- The translation $\mathbf{t}_{pq}$ is given by correlating the synthesized spectrum of the target block, $\hat{f}_q = A_{pq} f_p$, with the actual target block:

$$\mathbf{t}_{pq} = \arg\max_{\mathbf{t}} [\hat{f}_q \star f_q]. \tag{5}$$

Model reconstructions approximate the original data up to the residual error, ε_q, (Figure 2).

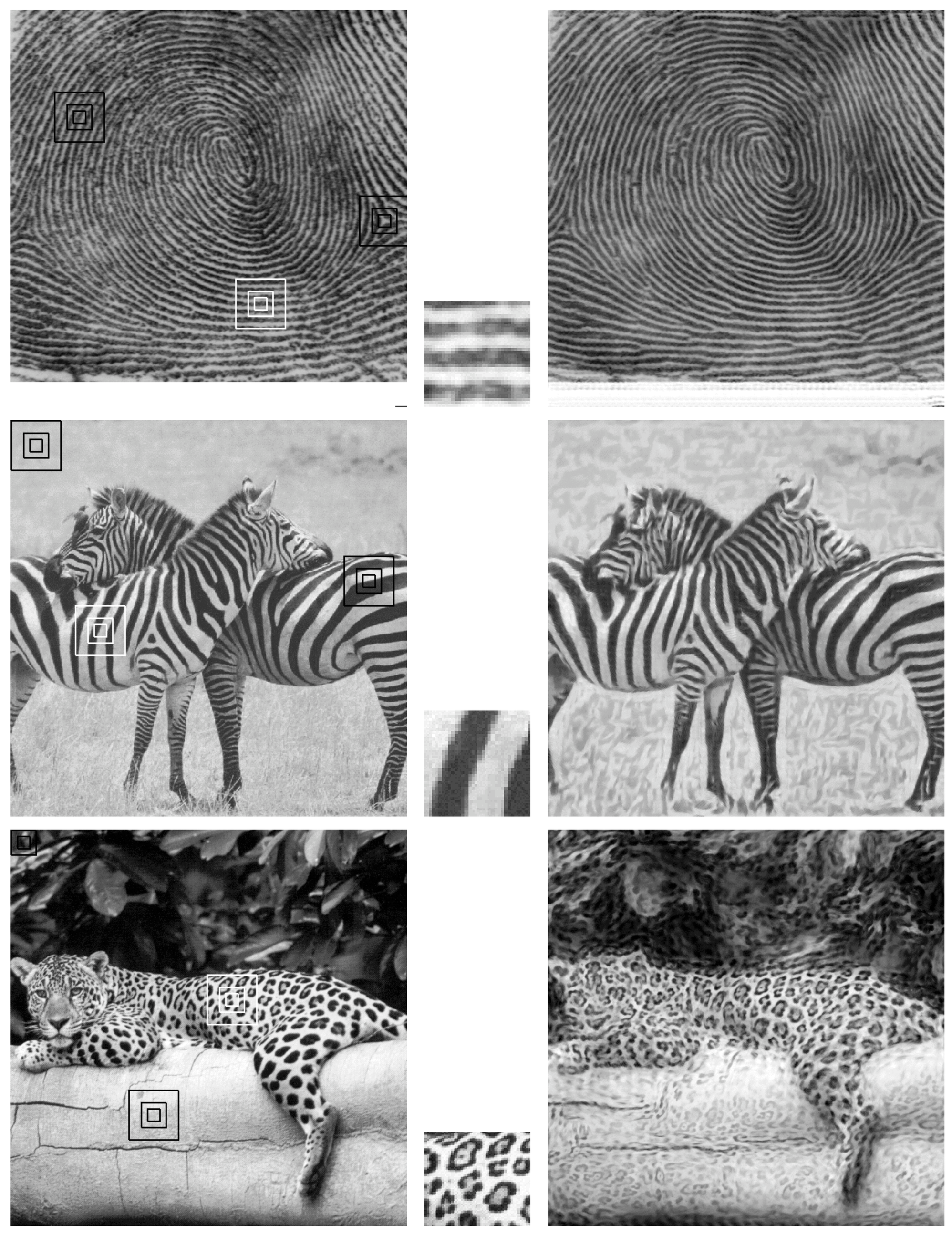


FIGURE 2. **Affine-invariant patch model reconstructions.** *Left-column shows original images and source prototype outlined in white. Middle-column shows textons used to produce synthesis results in right-hand column. Blocks sizes of 32 used for fingerprint and zebra images. Block size of 64 used for jaguar reconstruction.*

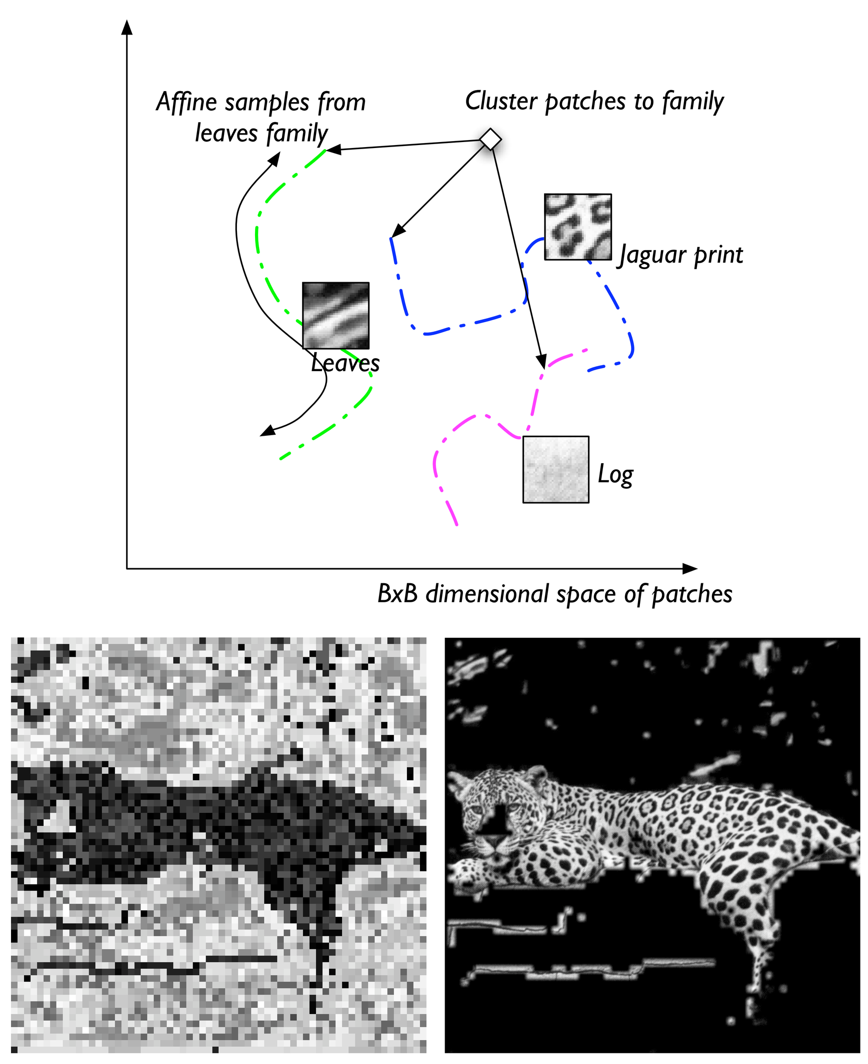


FIGURE 3. **Supervised clustering to representative textons.** *Results on jaguar image. Right-hand image shows which image blocks are labelled as the 'fur' class.*

3. Image Warplets

The **Image Warplet**, $\mathcal{W}_p$, is defined as the set of all image blocks q transformed to the coordinate frame of block p using T_{pq}^{-1},

$$\mathcal{W}_p = \{w_{pq}(\mathbf{x}) = \hat{f}_q^p(\mathbf{x}), \mathbf{x} = A_{pq}^{-1}(\mathbf{y} - \mathbf{t}_{pq}) \forall q\} \tag{6}$$

PCA (or harmonic analysis such as a DCT) of the warplet blocks, w_{pq}, is used to encode their variability (Figure 4).

Image blocks are represented by the mean warplet plus an appropriate number of modes of variation in the Warplet domain, and the corresponding block-to-block affine warps, T_{pq} (Figure 5).

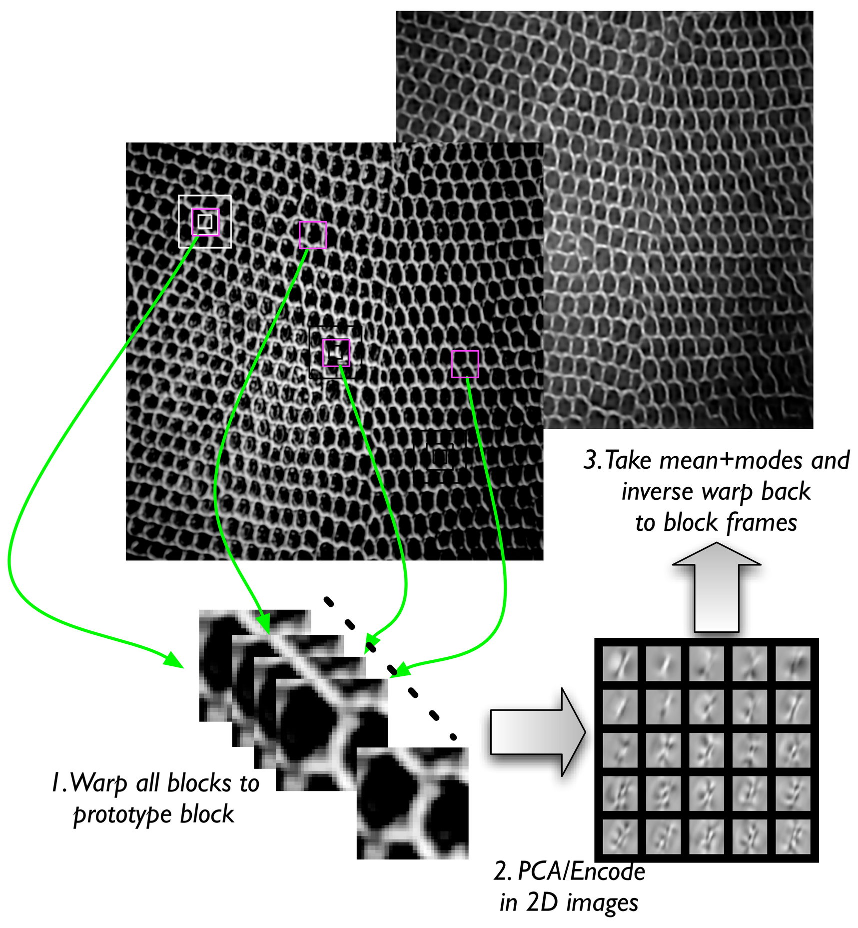
It is informative to visualise the mean and eigenmodes of the warplet decomposition. We have compared PCA with a DCT analysis of the warped blocks.

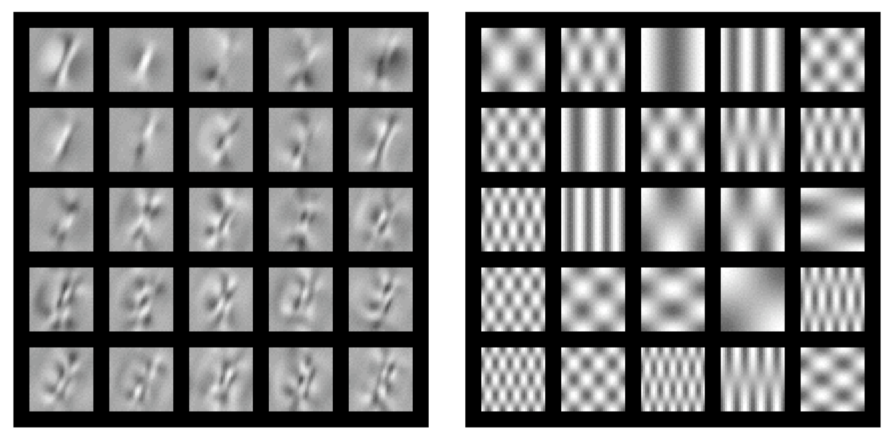
4. Segmentation and Texton Codebooks

The affine invariant image model works well if the source texton is representative. If the prototype is not representative for a region, then the residual errors are large and greater variability is captured by the eigenmodes of the warplet decomposition.

Representative textons can be discovered by:

1. Clustering in the space of affine symmetries given a likely set of textons. This can be used to classify blocks by family resemblance (Figure 3).
2. Alternatively, in the warplet domain, the residual error is an appropriate distance norm as the effects of linear transformation are eliminated by block-to-block warps to the same coordinate frame.





Basis vectors of reptile image.

FIGURE 4. **Schematic of warplets.**

PCA   DCT

PSNR 25.9 dB   PSNR 18.3dB

FIGURE 5. **Warplet reconstruction results**. *Coding and synthesis using the warplet domain using mean only, mean plus 3 modes and mean plus 15 modes on jaguar image. Comparisons with DCT coding of warplet domain using equivalent number of coefficients.*

5. Discussion

A new image-dependent wavelet image model has been introduced which can be used to synthesize image regions from representative image patches with good accuracy. The Image Warplet combines a block-based affine invariant image model with eigenimage or harmonic decomposition. Images can be reconstructed by projecting all transformed blocks on to the PCA modes and applying the inverse affine-transformations. We have shown how a clustering approach can be used to discover, i.e. segment, these prototypical image patches which can be used to construct a warplet dictionary for the image. Image warplets are good at representing regions containing natural and could have application in CBIR.

References

[1] T-I Hsu and R. Wilson. A Two-Component Model of Texture for Analysis and Synthesis. *IEEE Trans. on Image Processsing*, 7(10):1466–1476, October 1998.

[2] A. D. Calway. Image Representation Based on the Affine Symmetry Group. In *Proc. ICIP 1996*, pages 189–192, 1996.