Proceedings of the British Machine Vision Conference,
Norwich, UK, vol. 1, pages 163–170, September 2003

Ranklets on Hexagonal Pixel Lattices

Fabrizio Smeraldi and Mohammed A. Rob
Department of Computer Science
Queen Mary, University of London
Mile End Road
London, E1 4NS, UK

Abstract

Ranklets are a family of multiscale, orientation-selective rank features suitable for characterising complex patterns. On square pixel lattices, ranklets bear a strong similarity to Haar wavelets. This extends to the sensitivity to horizontal and vertical edges. We propose a generalisation of ranklets to hexagonal pixels. The sixfold rotational symmetry of the lattice translates into features that are tuned to three preferential directions in the image. We present experimental results on a pattern recognition task (face detection) over a large image database, using both square and hexagonal pixels. Comparison of the results in the two cases confirms the consistent performance of the generalised ranklets.

1 Introduction

Hexagonal sampling of two-dimensional bandlimited signals is desirable for mathematical reasons. From the signal theoretic point of view, a hexagonal lattice requires the least number of samples to represent a circularly bandlimited function, a class that models all images formed by circularly symmetric lenses [5]. Also, a few well known connectivity problems arising in the digital geometry of square lattices [7] do not occur in the hexagonal case.

Until recently, the use of hexagonal pixels has been hindered by the scarce availability of sensors. However, recent technical progress has brought hexagonal sampling into the domain of consumer electronics, with still camera manufacturers implementing the technology in their products. This is likely to foster a renewed interest in vision algorithms specifically designed for hexagonal lattices.

Ranklets are a family of multiscale, orientation-selective, nonparametric rank features based on the Wilcoxon rank-sum test [8]. In this paper, we present a generalisation of ranklets to hexagonal pixels. In the case of traditional square-pixel images, the orientation selectivity pattern of ranklets resembles that of Haar wavelets. The sixfold rotational symmetry of the hexagonal lattice allows the construction of features tuned to three different orientations.

We evaluate hexagonal-lattice ranklets by comparing them with the original square-lattice features in a face detection experiment over the $24'045$ test images of the MIT-CBCL face database. Hexagonal ranklets are applied to a resampled version of the images. Our results confirm the consistently good performance of ranklets independent of the sampling geometry.

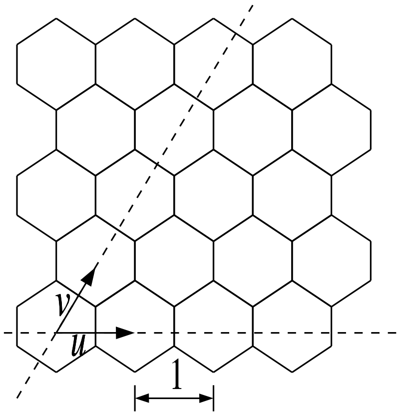

Figure 1: A hexagonal pixel lattice displayed with the skew basis $(\vec{u}, \vec{v})$.

2 Hexagonal Pixel Lattices

Hexagonal pixels represent the sampling solution with lowest sampling density for circularly bandlimited images. Not surprisingly, the arrangement of cones in the fovea and of the ommatidia in the compound eyes of insects follow a hexagonal pattern.

In this work, we make use of a lattice of regular hexagonal pixels of unit width. The points of the lattice can be obtained as $\vec{p} = u\vec{u} + v\vec{v}$ where u and v are their (integer) coordinates referred to the skew basis $(\vec{u}, \vec{v})$ shown in Figure 1, with $\vec{u} = (1,0)$ and $\vec{v} = (1/2, \sqrt{3}/2)$ (see [4, 9]).

Resampling is performed by separable interpolation of the pixels of the original rectangular lattice. We employ a separable Sinc function $\psi(x,y)$ multiplied by a Hanning window, that is $\psi(x,y) = \phi(x) \cdot \phi(y)$ with

$$\phi(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \left(\frac{1}{2} + \frac{1}{2}\cos\left(\frac{2\pi x}{W}\right)\right) & \text{if } |x| < W \\ 0 & \text{if } |x| \geq W. \end{cases} \tag{1}$$

The window size is fixed at $W = 8$ (see [10]).

Note that, although the hexagons are regular, the sampling distances along the vertical and horizontal axes differ. For this reason, the originally 19×19 images of the $MIT - CBCL$ database employed in our tests occupy an area of 19×22 hexagonal pixels. Our choice of a lattice spacing amounts to a slight oversampling of the image. However, it seemed convenient to use hexagonal pixels of unit width because the arrangement of the nodes of $\phi(x)$ can be exploited to reduce the number of computations.

3 Ranklets: a Family of Wavelet-style Rank Features

The first stage in computing ranklets involves ranking the intensity values in (part of) the image, i.e. finding the permutation π of the integers from 1 to N that expresses the relative order of the intensity values of the N pixels in a suitable image window.

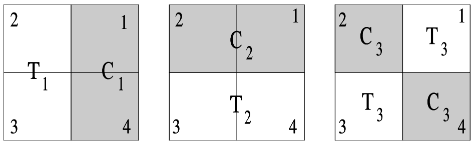

Figure 2: The three partitions for a 2×2 square-pixel window. Note the similarity with the three Haar wavelets that can be obtained by assigning the value $+1$ to the light areas and -1 to the shaded areas).

In the next Section we recall the definition of a classic non-parametric hypothesis test, the Wilcoxon rank-sum test, from which ranklets derive. We will then apply the underlying ideas to the characterisation of intensity variations in image neighbourhoods in Section 3.2.

3.1 The Wilcoxon Rank-sum Test

The Wilcoxon rank-sum test (also known as the Mann-Whitney U-test) is a hypothesis test designed for the comparison of two treatments [3]. Suppose that N quantities are split in two groups of n "treatment" and m "control" observations (according to the standard terminology). We ask whether the treatment observations are significantly larger than the controls. To this end, we define the Wilcoxon rank sum statistic $\mathscr{W}_s$ as the sum of treatment ranks: $\mathscr{W}_s = \sum_{i=1}^{n} \pi(i)$. The logic behind this definition is that large values of the treatment observables relative to the control observables will result in a large value of $\mathscr{W}_s$.

After an experiment is performed, the treatment values are judged to be significantly larger than the controls if $\mathscr{W}_s$ is above a critical value τ. The choice of τ defines the confidence level of the test.

In the next Section, we apply this test to pixel values to determine intensity variations between image regions.

3.2 Ranklets on Square Pixels

Given an image I, we indicate with $\pi(\vec{x})$ the rank of the intensity $I(\vec{x})$ of pixel $\vec{x}$ among the intensities of the pixels in a suitably sized window W. To simplify matters, we assume that no two pixels have the same intensity; ties can be broken at random when they occur.

The Wilcoxon test can be used to determine brightness variations among any two subsets of the pixels in W. On square lattices, such subsets can be conveniently chosen in close analogy with Haar wavelets [2, 8]. The resulting three partitions $W = T_j \cup C_j,\ j = 1,2,3$ for the case of a 2×2 window are displayed in Figure 2. For each of them we can compute the Wilcoxon statistics as $\mathscr{W}_s^j = \sum_{\vec{x} \in T_j} \pi(\vec{x})$, which yields $\mathscr{W}_s^1 = \pi(2) + \pi(3)$, $\mathscr{W}_s^2 = \pi(3) + \pi(4)$ and $\mathscr{W}_s^3 = \pi(1) + \pi(3)$ (pixels are labelled as shown in Figure 2).

The geometric interpretation of these quantities is straightforward. Consider for instance $\mathscr{W}_s^1$ and suppose that the local neighbourhood W straddles a vertical edge, with the darker side on the right (where C_1 is located) and the brighter side on the left (corresponding to T_1). Then $\mathscr{W}_s^1$ will be large, as many pixels in T_1 have a higher intensity and therefore rank higher than those in C_1. Conversely, $\mathscr{W}_s^1$ will be low if the dark and bright

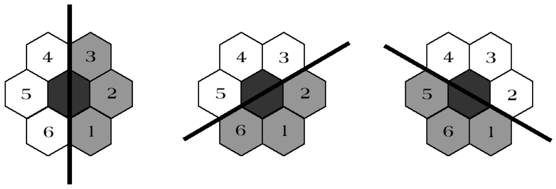

Figure 3: The three partitions for the smallest hexagonal window W_1. Shaded pixels indicate the C sets; the central pixel is discarded. The black line indicates the orientation to which each ranklet is maximally responsive.

sides of the edge are reversed. Horizontal edges or other patterns with no global left-right variation of intensity will give an intermediate value. Therefore, $\mathscr{W}_s^1$ responds to vertical edges in the images. By a similar argument $\mathscr{W}_s^2$ detects horizontal edges, while $\mathscr{W}_s^3$ is sensitive to corners formed by horizontal and vertical lines. These response patterns closely match those of the corresponding Haar wavelets obtained as indicated in Figure 2.

3.3 Ranklets on Hexagonal Pixels

We can easily generalise the approach outlined in the preceding section to hexagonal pixel lattices. In this case, we can exploit the sixfold rotational symmetry of the lattice to obtain ranklets tuned to three preferential directions (a set of linear kernels over hexagonal pixels with analogous directional selectivity is discussed in [11]). For the smallest possible window W_1 shown in Figure 3 we obtain three partitions $T_j \cup C_j \cup \{\vec{x}_0\}$, where the central pixel $\vec{x}_0$ of W_1 is not considered in the ranking. The corresponding values of $\mathscr{W}_s$ are $\mathscr{W}_s^1 = \pi(4) + \pi(5) + \pi(6)$, $\mathscr{W}_s^2 = \pi(3) + \pi(4) + \pi(5)$ and $\mathscr{W}_s^3 = \pi(2) + \pi(3) + \pi(4)$.

We can conveniently replace $\mathscr{W}_s^j$ with the equivalent Mann-Whitney statistics

$$\mathscr{W}_{YX}^j = \mathscr{W}_s^j - (N/2+1)N/4, \tag{2}$$

which has an immediate interpretation in terms of pixel comparisons. As can be easily shown [3], $\mathscr{W}_{YX}^j$ is equal to the number of pixel pairs $(\vec{x}_p, \vec{y}_q)$ with $\vec{x}_p \in T_j$ and $\vec{y}_q \in C_j$ such that $I(\vec{x}_p) > I(\vec{y}_q)$. Its possible values therefore range from 0 to the number of pairs $(\vec{x}_p, \vec{y}_q) \in T_j \times C_j$, which is $N^2/4$. For this reason, ranklets are conveniently defined as

$$\mathscr{R}_j = \frac{\mathscr{W}_{XY}^j}{N^2/4} - 1 \tag{3}$$

(this also applies to square-pixel ranklets [8]). Thus the value of $\mathscr{R}_j$ will be 1 if and only if, for all the possible pairs of one pixel in T_j and one pixel in C_j, the first is brighter than the second. If the opposite is true, $\mathscr{R}_j$ will be -1.

3.4 Multiscale Ranklets

In view of the similarity between Haar wavelets and ranklets it is straightforward, in the case of square pixel lattices, to consider multiscale versions of ranklets corresponding to the different scalings of a Haar wavelet.

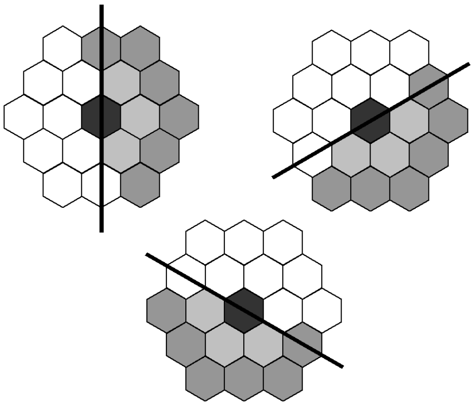

Figure 4: The three partitions for window W_2, including up to the second order neighbours. Shaded pixels indicate the C sets; light areas correspond to the T sets, and the central pixel is discarded. The black line indicates the orientation to which each ranklet is maximally responsive. Note how the directional selectivity pattern shown in Figure 3 is preserved.

This can be easily generalised to hexagonal pixel lattices by constructing around pixel $\vec{x}_0$ the window W_i comprising its neighbours up to the i-th order, and subsequently partitioning it into the sets $T_j \cup C_j \cup \{\vec{x}_0\}$ as suggested in Figure 4 for the case of neighbours up to the second order (note that care has been taken to preserve the same directional selectivity pattern as in Fig. 3). Generalisation to higher values of i is straightforward.

3.5 Efficiency Considerations

As we have seen in Section 3.3, $\mathscr{W}_{YX}^j$ equals the number of pairs $(\vec{x}_p, \vec{y}_q) \in T_j \times C_j$ such that the intensity at $\vec{x}_p$ is larger than the intensity at $\vec{y}_q$. Since $N^2/4$ pairs can be formed out of the N pixels in $W \setminus \{\vec{x}_0\}$, it would seem that the number of comparisons required should grow with the square of the area of W.

Notice however that these pairwise comparisons are never explicitly carried out; the value of $\mathscr{W}_{YX}^j$ is obtained by subtracting a constant from the Wilcoxon statistics $\mathscr{W}_s^j$ (see Equation 2). In turn, $\mathscr{W}_s^j$ is computed by ranking the pixels in W, which only requires order of $N \log N$ operations. The intuitively appealing interpretation in terms of pairwise pixel comparisons turns out to be a “free” by-product of the sorting operation.

4 Experimental Results

In order to validate our extension of ranklets to hexagonal pixel lattices, we present comparative experimental results in a pattern recognition task. More specifically, we discuss a face detection experiment over the images of the MIT-CBCL face database [6], that consists of low-resolution grey level images ($19 \times 19 = 361$ square pixels) of faces and non-faces. Square-pixel ranklets have been shown to outperform several other algorithms in tests over this dataset [8].

Figure 5: Training faces (left), test faces (centre), and test non-faces (right) from the MIT-CBCL set. Notice how a few of the test faces in the database are incorrectly framed (centre, lower right).

The database consists of a training set of $2'429$ faces and a test set of 472 faces and $23'573$ non-faces. All facial images nearly occupy the entire frame; considerable pose changes are represented (Fig. 5). The database also contains a set of $4'548$ non-faces intended for training. In our experiments we have discarded this training set of non-faces, since the notion of "non-face prototypes" appears to be problematic.

We employ two copies of the database, namely the original dataset and a hexagonal-pixel version obtained by resampling all the images as described in Section 2. Images are encoded as a vector of ranklets according to the lattice type. In both cases, the minimal window W displayed in Figure 2 (or respectively Figure 3) is centred at all possible locations in the image. Each location contributes to the feature vector the three ranklets $\mathscr{R}_j$ corresponding to the three partitions of the respective window, according to Equation 3.

For the sake of simplicity, as well as to evidence the descriptiveness of the features, we adopted a straightforward distance-based classification algorithm. The feature vector extracted from a test image is compared with those obtained from the training images (obviously sampled in the same way). The image is classified as a face if the distance from the closest example is lower than a threshold τ. By varying τ, we obtain the Receiver Operating Characteristic (ROC) curves displayed in Figure 6. As can be seen, the performance of the two types of features is comparable. The Equal Error Rates (EERs) for hexagonal and square pixel ranklets are 19% and 17% respectively. Considering the noise introduced by the resampling process, these two figures are essentially equivalent.

In Figure 6 we also display the ROC curve for multiscale hexagonal ranklets. In this case, the feature vector was constructed using 5 windows W_i of increasing size including all neighbours of the central pixel up to the 5th order, as outlined in Section 3.4. For each scale, the window was centred at all possible locations in the image. The EER in this case is 17%. The corresponding EER for multiscale square-pixel ranklets is 16% [8]. The behaviour of multiscale hexagonal-pixel ranklets is thus essentially equivalent to that of the analogous square-pixel features.

The error rates reported above represent an improvement over published results obtained with other algorithms, including linear Support Vector Machine classifiers and the Sparse Network of Winnows (SNoW) [1].

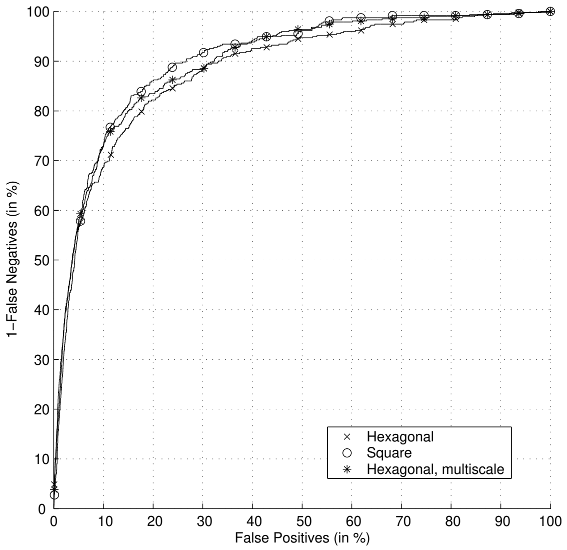

Figure 6: ROC curves for ranklets over hexagonal pixels ($\times$), ranklets over square pixels ($\circ$) and multiscale ranklets over hexagonal pixels ($*$).

5 Conclusions

We have presented a generalisation of ranklets to hexagonal pixel lattices that maintains a very close analogy with the original features. Hexagonal and square-pixel ranklets are both defined in terms of the Mann-Whitney statistics, and admit the same interpretation in terms of pairwise comparisons of pixel intensity values. Also, both features easily lend themselves to multiscale implementations.

Orientation selectivity is one of the main advantages of both square and hexagonal-pixel ranklets. However, the precise response pattern changes with the lattice. Traditional ranklets, similarly to Haar wavelets, are sensitive to horizontal and vertical lines. Hexagonal ranklets, due to the sixfold rotational symmetry of the lattice, are tuned to three different orientations in the image.

We have experimentally validated the soundness of our generalisation by comparing hexagonal and square ranklets in a challenging pattern recognition task, namely face detection. Our experiments, involving the large collection of test images (over $24'000$) from the publicly available MIT-CBCL database, show that hexagonal ranklets perform consistently with square-pixel ranklets. The very similar results obtained on the two pixel lattices also represent an improvement on other published approaches.

Recent advances in CCD technology are introducing hexagonal sampling into consumer electronics applications. For this reason, we believe the existence of a generalisation to hexagonal pixel lattices to represent a positive contribution to the usability and practical value of the ranklet algorithm.

References

[1] M. Alvira and R. Rifkin. An empirical comparison of SNoW and SVMs for face detection. Technical Report AI Memo 2001-004 – CBCL Memo 193, MIT, January 2001. http://www.ai.mit.edu/projects/cbcl.

[2] I. Daubechies. *Ten lectures in wavelets*. Society for industrial and applied mathematics, Philadelphia, USA, 1992.

[3] E. L. Lehmann. *Nonparametrics: Statistical methods based on ranks*. Holden-Day, 1975.

[4] E. Luczak and A. Rosenfeld. Distance on a hexagonal grid. *IEEE Transactions on Computers*, C-25(5):532–533, 1976.

[5] R. M. Mersereau. The processing of hexagonally-sampled two dimensional signals. *Proceedings of the IEEE*, 67(6):930–949, 1979.

[6] MIT Center for Biological and Computational Learning. CBCL face database no. 1. http://www.ai.mit.edu/projects/cbcl, 2000.

[7] A. Rosenfeld. Connectivity in digital pictures. *Journal of the ACM*, 17(1):146–160, 1970.

[8] F. Smeraldi. Ranklets: orientation selective non-parametric features applied to face detection. In *Proceedings of the 16th International Conference on Pattern Recognition, Quebec, CA*, volume 3, pages 379–382, August 2002.

[9] W. E. Snyder, H. Qi, and W. Sander. A coordinate system for hexagonal pixels. In *Proceedings of SPIE Medical Imaging: Image Processing, Pts. 1 and 2*, pages 716–727, february 1999.

[10] P. Thévenaz, T. Blu, and M. Unser. Image interpolation and resampling. In I.N. Bankman, editor, *Handbook of Medical Imaging, Processing and Analysis*, chapter 25, pages 393–420. Academic Press, San Diego CA, USA, 2000.

[11] A. B. Watson and Jr. A. J. Ahumada. A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex. *IEEE Transactions on Biomedical Engineering*, 36(1):97–106, 1989.