Research Report RR0309-01, Computer Science, September 2003
Available online from http://www.dcs.qmul.ac.uk/~fabri

Ranklets: a Complete Family of Multiscale, Orientation Selective Rank Features

Fabrizio Smeraldi
Department of Computer Science
Queen Mary, University of London
Mile End Road, London E1 4NS, UK

Abstract

Ranklets are a family of non-parametric rank features designed in close analogy with Haar wavelets. With these they share the same pattern of orientation selectivity, the multiscale nature and a suitable notion of completeness. Ranklets are defined based on the Wilcoxon statistics, are computationally efficient and admit an interpretation in terms of pairwise pixel comparisons.

In this report we review the definition of ranklets and their geometric properties. We then introduce a notion of completeness applicable to rank features and give proofs for one and two dimensional ranklets. Finally, we present a few considerations on the algorithmic complexity of the representation and its memory efficiency.

1 Introduction

Ranklets are a family of multiscale rank features characterized by a strong analogy to Haar wavelets. This extends to the orientation selectivity of the features, their multiscale nature and, according to a result we introduce in this report, to completeness.

In Section 2 we review the definition of ranklets, which is based on the Wilcoxon statistics, along with their geometric properties and their interpretation in terms of pairwise comparisons of pixel values.

Since rank features operate on the relative order of pixel intensities rather than on the intensity values proper, any "completeness" result must be preceded by the introduction of a suitable definition of completeness itself. It would seem reasonable to say that a set of rank features is complete if, by knowing the feature values over a given image, the ranking of all the pixels in the image can be unambiguously deduced. In other words, the computation of a complete set of features should be equivalent to a sorting of the image. Note how this constitutes a "maximal" requirement, since by definition only order information is accessible to rank features. This notion of completeness is formalized in Section 3, where proofs are also given for one and two-dimensional ranklets.

Finally, efficiency considerations along with a discussion of the compactness of the representation are presented in Section 4.

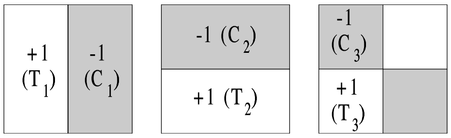

Figure 1: The three two-dimensional Haar wavelets $h_1(\vec{x})$, $h_2(\vec{x})$ and $h_3(\vec{x})$ (from left to right). Letters in parentheses refer to "treatment" and "control" pixel sets (see Sect. 2.2).

2 Ranklets: a Family of Wavelet-style Rank Features

As all other rank features, ranklets are computed starting from a ranking of the N samples representing (a suitable part of) a digitized signal, i.e. from a permutation π of the integers from 1 to N expressing the relative order of the sample values.

In the next section we recall the definition of a classic non-parametric hypothesis test from which ranklets are derived, namely the Wilcoxon rank-sum test.

2.1 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" observations $\{x_i\}$ and m "control" observations $\{y_j\}$ (according to the standard terminology). We ask whether the treatment observations are significantly higher than the controls. To this end, we rank the observations in increasing order and define the Wilcoxon statistics $\mathcal{W}_s$ as the sum of the treatment ranks: $\mathcal{W}_s = \sum_{i=1}^{n} \pi(x_i)$, where π indicates the ranking. The logic behind this definition is that high values of the treatment observations relative to the controls will result in a large value of $\mathcal{W}_s$.

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

In the next Section, we compute the Wilcoxon statistics over the samples from one or two-dimensional signals, with the aim of describing intensity variations between groups of samples. We will use the term "pixels" instead of the more generic "samples" wherever the considerations presented refer more directly or exclusively to two or higher dimensional signals.

2.2 Definition of Ranklets

Given a (one or two-dimensional) signal I, we indicate with $\pi(I^{\mathsf{W}}(\vec{x}))$ the rank of the sample $I(\vec{x})$ among those in a suitably sized window W, with $\vec{x} \in \mathsf{W}$. To simplify matters, we assume that no two samples have the same value (for practical purposes, ties can often be broken at random when they occur; for a more rigorous discussion see Section 3.5).

The Wilcoxon test can be used to determine intensity variations among conveniently chosen subsets of the samples in W. Ranklets are obtained by splitting the N samples (pixels) in W in two subsets of $N/2$ elements each, thus assigning half of the samples to the "treatment" group and half to the "control" group. This introduces a new degree of freedom, namely the geometric arrangement of the two subsets in W. In the case of two-dimensional (or indeed higher dimensional) signals, this can be exploited to obtain orientation selective features.

Let us consider, for definiteness, the case of two-dimensional ranklets. We construct the "treatment" and "control" sets starting from the Haar wavelets $h_j(\vec{x}), j = 1,2,3$ displayed in Figure 1 (see [1]). To start with, we identify the local neighbourhood W with the support of the h_j. We then define the set of "treatment" pixels T_j as the counter-image of $\{+1\}$ under h_j: $\mathsf{T}_j = h_j^{-1}(\{+1\})$, and the set of "control" pixels C_j as the counter-image of $\{-1\}$: $\mathsf{C}_j = h_j^{-1}(\{-1\})$. For each of the three resulting partitions of W, $\mathsf{W} = \mathsf{T}_j \cup \mathsf{C}_j$, we compute the Wilcoxon statistics as

$$\mathcal{W}_s^j = \sum_{\vec{x} \in \mathsf{T}_j} \pi(I^{\mathsf{W}}(\vec{x})) \tag{1}$$

(the simpler construction for the one-dimensional case involves only one partition of W induced by the one-dimensional Haar "mother wavelet").

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

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

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

$$\mathcal{R}^j = 2\frac{\mathcal{W}_{YX}^j}{N^2/4} - 1. \tag{3}$$

Thus the value of $\mathcal{R}^j$ 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 pixel is brighter than the second. If the opposite is true, $\mathcal{R}^j$ will equal -1.

2.3 Multiscale Ranklets

Due to the close correspondence between Haar wavelets and ranklets, the multiscale nature of the former directly extends to the latter. To each translation and scaling of the h_j specified by the parameters $(\vec{x}_0, s)$ we associate the sets of treatment and control pixels defined by

$$\begin{aligned} \mathsf{T}_{j;(\vec{x}_0,s)} &= \{\vec{x} \mid h_j((\vec{x} - \vec{x}_0)/s) = +1\}, \\ \mathsf{C}_{j;(\vec{x}_0,s)} &= \{\vec{x} \mid h_j((\vec{x} - \vec{x}_0)/s) = -1\}. \end{aligned} \tag{4}$$

We can then compute the value of $\mathcal{R}_{\mathsf{W}}^j$ over the local neighbourhood $\mathsf{W}_{(\vec{x}_0,s)} = \mathsf{T}_j \cup \mathsf{C}_j$, with $N = \#\mathsf{W}_{(\vec{x}_0,s)}$.

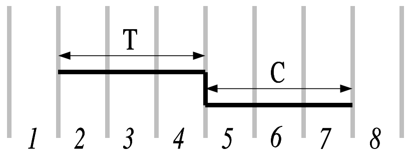

Figure 2: Graphic notation for one-dimensional ranklets. The upper and lower line identify the T and C sets respectively. Sample numbers are indicated in italics.

2.4 Geometric Interpretation

The geometric interpretation of the ranklets $\mathcal{R}^j$ is straightforward given the properties of $\mathcal{W}^j_{YX}$ and of the Haar wavelets h_j. Consider for instance $\mathcal{R}^1$ and suppose that the local neighbourhood W straddles a vertical edge, with the darker side on the left (where C_1 is located) and the brighter side on the right (corresponding to T_1). Then $\mathcal{R}^1$ will be close to $+1$, as many pixels in T_1 have a higher intensity than those in C_1. Conversely, $\mathcal{R}^1$ will be close to -1 if the dark and bright sides of the edge are reversed. Horizontal edges and other patterns with no definite left-right variation of intensity will give a value close to zero. Therefore, $\mathcal{R}^1$ responds to vertical edges in the images. By a similar argument $\mathcal{R}^2$ is tuned to horizontal edges, while $\mathcal{R}^3$ is sensitive to corners formed by horizontal and vertical lines (as well as to 45° and 135° edges). These response patterns closely match those of the corresponding Haar wavelets h_j.

Note that there is a large arbitrariness involved in our way of partitioning W into a "treatment" and a "control" set. A different choice would lead, for instance, to the rank transform of Zabih and Woodfill [8] (the authors, however, fail to clarify the relation between their transform and the Wilcoxon test; for a discussion of this topic, see [5]).

3 Completeness

In this Section we introduce a notion of completeness suitable for rank features. These features are defined in terms of the relative order of the sample values (or of a subset thereof); since the values themselves are disregarded, signal reconstruction is evidently beyond reach (however, see [4] for the description of an approximate image reconstruction based on statistical assumptions on the sample values). The most that can be achieved is a reconstruction of the global ranking of all the samples, which clearly exhausts all the information available. The notion of completeness we propose is in this spirit: a set of rank features is complete if their value is sufficient to unambiguously specify the order of the sample values.

Formally, let $(1, 2, \ldots, N)$ be the set of the sampling points. For our purposes, we can conveniently identify a signal with the order of its sample values, i.e. with an element of the symmetric group S_N on the sampling points. Let now $\mathcal{C} = \{\mathsf{W}_1, \mathsf{W}_2, \ldots, \mathsf{W}_m\}$ be a covering of $(1, 2, \ldots, N)$; we define the ranklet decomposition $\mathcal{R}$ as the map

$$\mathcal{R} : S_N \to \mathbb{R}^m \tag{5}$$

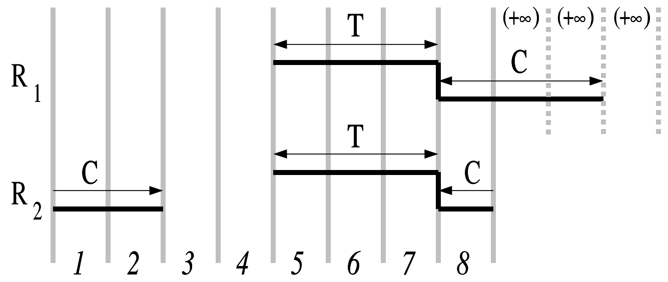

Figure 3: Two ways to deal with border problems for finite signals. In R_1, the missing samples in C are assumed to have an infinite value. We call this continuation "infinity padding". Ranklet R_2 is computed assuming a wraparound topology.

that sends a permutation $\pi \in S_N$ into the vector $\mathcal{R}(\pi) = (\mathcal{R}^i_{\mathsf{W}_j})_{\forall i, \forall j}$ obtained by computing each admissible ranklet $\mathcal{R}^i$ (depending on the dimensionality of the signal) on every window $\mathsf{W}_j \in \mathcal{C}$. We can now state the following:

Definition 1 *The decomposition $\mathcal{R}$ is complete if it is invertible over its range, i.e. if it is one-to-one.*

In other words, we will say that the decomposition $\mathcal{R}$ is complete if there exists a map $\tilde{\mathcal{R}} : \mathcal{R}(S_N) \to S_N$ such that for every signal (image) I

$$\tilde{\mathcal{R}}(\mathcal{R}(\pi_I)) = \pi_I, \tag{6}$$

where π_I is the ranking of I. The map $\tilde{\mathcal{R}}$ is the closest analogue for rank features of the reconstruction operator.

In the following we will prove completeness for ranklets, working our way from the one to the two-dimensional case. Before doing so, however, we will introduce some shorthand graphic notation in the next Section.

3.1 Graphic Notation and Data Padding

For the purpose of demonstrating the completeness of ranklets, it is convenient to introduce a graphic notation.

In the one-dimensional case, we identify a ranklet $\mathcal{R}$, its support window W and the sets of treatment and control samples T and C using a stylized Haar wavelet, as outlined in Figure 2.

Since we only deal with finite sets of samples, we need a convenient mechanism for handling border conditions, i.e. the case that the support window W of a ranklet partly falls outside the sampling window S of the signal or image at hand. In the case of linear filters, the standard devices are zero padding and wraparound.

We can easily introduce two corresponding mechanisms for ranklets. Zero padding is conveniently replaced by "infinity padding": we assume that the samples in $\mathsf{W} \setminus \mathsf{S}$ have an infinite value, i.e. we systematically assign to them the highest ranks (or, alternatively, the lowest ranks). This situation is illustrated in Figure 3 ($\mathcal{R}_1$), which also introduces the other option, namely wraparound ($\mathcal{R}_2$). In the latter case, W is split in two connected components, to simulate the periodic continuation of the data. Calculations are performed in the usual way, based on the ranks of the sample values. In the rest of this work we will chiefly make use of infinity padding.

3.2 The One-dimensional Case

Consider a one dimensional signal described by the N samples $(1, 2, \ldots, N)$. A set of $N-1$ ranklets forming a complete decomposition is shown in Figure 4, for the special case that $N = 8$ (generalization to an arbitrary finite number of samples is straightforward).

To prove this, we proceed "ab absurdo" by assuming that the decomposition $\mathcal{R}$ represented in Figure 4 is not invertible, so that for at least two permutations π_1 and π_2 (corresponding to two signals with different ranking)

$$\mathcal{R}(\pi_1) = \mathcal{R}(\pi_2) \quad \text{with} \quad \pi_1 \neq \pi_2. \tag{7}$$

Let σ be the permutation that maps π_1 onto π_2, i.e. $\pi_2 = \sigma\pi_1$: we will show that $\sigma = \mathrm{Id}$, which contradicts Equation 7.

With reference to Figure 4, consider $\mathcal{R}_{N-1}$: since by Equation 7 we have $\mathcal{R}_{N-1}(\pi_1) = \mathcal{R}_{N-1}(\pi_2)$, σ cannot move $\pi_1(N)$ (otherwise the value of $\mathcal{W}_s^j$ in Equation 1 would be changed). Next, consider $\mathcal{R}_{N-2}(\pi_1)$ that, by Equation 7, must be equal to $\mathcal{R}_{N-2}(\pi_2)$. Since $\pi_1(N)$ is fixed, this implies that σ does not move $\pi_1(N-1)$. By iterating over $\mathcal{R}_{N-3}, \ldots, \mathcal{R}_2, \mathcal{R}_1$ we see that σ fixes all the elements of $\{\pi_1(1), \pi_1(2), \ldots, \pi_1(N)\}$, or, which is the same, of $(1, 2, \ldots, N)$. We conclude that $\sigma = \mathrm{Id}$ and that $\pi_1 = \pi_2$, which contradicts our assumption (Equation 7) and thus proves our thesis.

3.3 Uniqueness

The decomposition shown in Figure 4 is not the only complete ranklet decomposition possible over the sampling points $(1, 2, \ldots, N)$. A first alternative is straightforwardly obtained by symmetrizing the arrangement of ranklets in Figure 4 around the midpoint of the sampling window. It is also possible to swap the "treatment" and the "control" sets T and C of one or more ranklets independently; all the decompositions thus obtained are complete.

Interestingly, we can also find complete decompositions with a "wraparound" topology, as exemplified in Figures 5, 6 and 7.

3.4 The Two-dimensional Case

In this section we outline how a complete decomposition can be constructed for a two-dimensional image. We will closely follow the argument given in Section 3.2 for one-dimensional signals. For simplicity, we only discuss the case of square images; generalization to rectangular frames is straightforward. The proof given below requires the use of ranklets defined on rectangular neighbourhoods W; these are easily obtained using the Haar wavelets with corresponding rectangular support to specify the respective T and C sets, in analogy with Equation 4.

Given a square image I of size $M \times M = N$, let us start by considering the three ranklets $\mathcal{R}^1_{(M-1,M-1)}$, $\mathcal{R}^2_{(M-1,M-1)}$ and $\mathcal{R}^3_{(M-1,M-1)}$ as shown in Figure 8

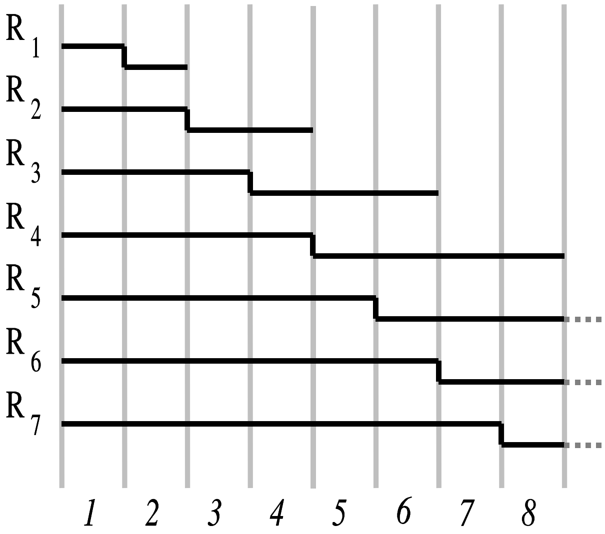

Figure 4: A complete set of ranklets for the one-dimensional case. The dashed lines indicate that infinity padding should be performed outside the sampling window.

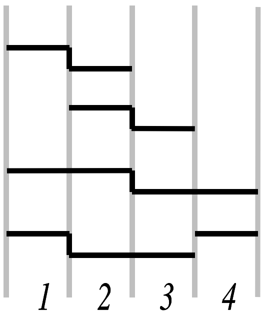

Figure 5: A complete set of ranklets for a one-dimensional 4-sample signal, wraparound topology.

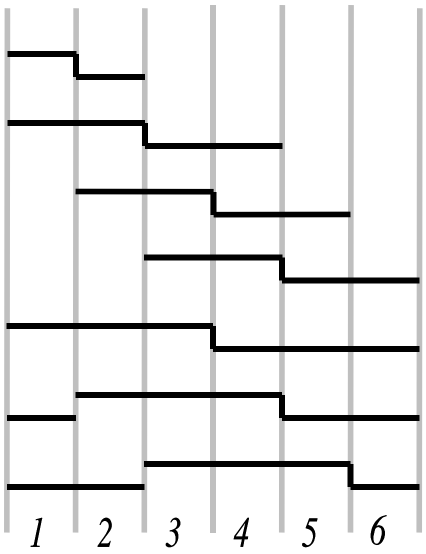

Figure 6: A complete set of ranklets for a one-dimensional 6-sample signal, wraparound topology.

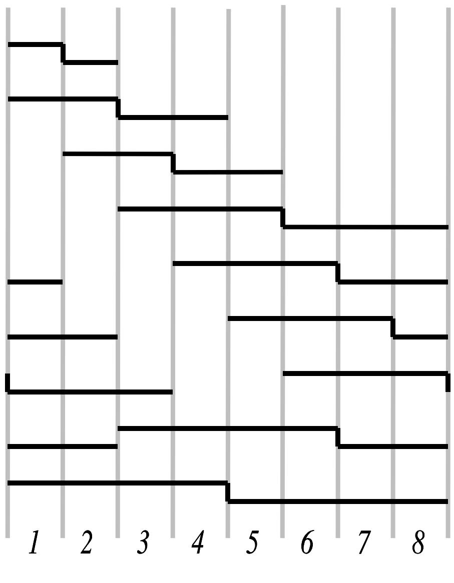

Figure 7: A complete set of ranklets for a one-dimensional 8-sample signal, wraparound topology.

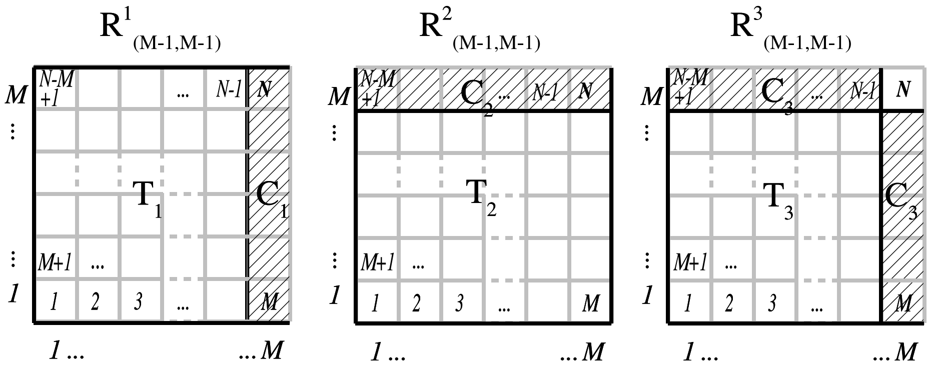

Figure 8: Ranklets $\mathcal{R}^1_{(M-1,M-1)}$ (left), $\mathcal{R}^2_{(M-1,M-1)}$ (centre) and $\mathcal{R}^3_{(M-1,M-1)}$ (right). Infinity padding is required on the "open" (gray) boundaries. Pixels in the dashed area belong to the C sets, while the T sets are left white.

(in the following, the lower indices will be omitted for brevity). Let π_1 and π_2 denote two distinct permutations on which these three ranklets take on the same value, that is

$$\mathcal{R}^i(\pi_1) = \mathcal{R}^i(\pi_2) \quad \text{with} \quad \pi_1 \neq \pi_2. \tag{8}$$

Again, let us write $\pi_2 = \sigma\pi_1$. We will now concentrate on pixel N in Figure 8 and show that the assumption that σ moves $\pi_1(N)$ is incompatible with Equation 8. Let $\delta = \sigma\pi_1(N) - \pi_1(N)$ be the rank shift of pixel N as the result of σ. Since $N \in \mathsf{C}_1$, the sum of the ranks in $\mathsf{C}_1 \setminus \{N\}$ must change by a corresponding $-\delta$ in order for σ to leave $\mathcal{R}^1$ unaltered (see Equation 1). Similarly, we have $N \in \mathsf{C}_2$, meaning that the sum of the ranks in $\mathsf{C}_2 \setminus \{N\}$ must also change by $-\delta$ for the value of $\mathcal{R}^2$ to remain constant. Therefore, the rank sum for $\mathsf{C}_1 \cup \mathsf{C}_2$ is changed by $\delta - 2\delta = -\delta$. Since the sum of the ranks over the entire image is constant, it follows that the sum over the pixels in $\mathsf{T}_3 \setminus \{N\}$ is increased by δ. Thus the rank sum for T_3 is changed by 2δ, and Equation 8 will hold for $\mathcal{R}^3$ only if $\delta = 0$. We conclude that σ fixes $\pi_1(N)$.

Next, we extend Equation 8 to include the two ranklets $\mathcal{R}^1_{(M-2,M-1)}$ and $\mathcal{R}^3_{(M-2,M-1)}$ and consider them together with $\mathcal{R}^2_{(M-1,M-1)}$, as shown in Figure 9. Remembering that $\pi_1(N)$ is fixed by σ, it is easy to see that σ should also fix $\pi_1(N-1)$ in order for Equation 8 to hold for these three ranklets. The key consideration is again that pixel $N-1$ is the only element in the intersection of C_1, C_2 and T_3 (as defined in Figure 9) the rank of which can be moved by σ.

Following the procedure outlined in these paragraphs, we can easily construct a complete decomposition of the image I. As in the one-dimensional case, a total of $N-1$ ranklets is required, where $N = M \times M$ is the number of pixels.

3.5 The Treatment of Ties

So far, we have assumed that no two samples (intensity values) in I are equal. This may in fact be the case in pattern recognition applications, as long as the size of W is small compared to the number of gray levels. In practice, the

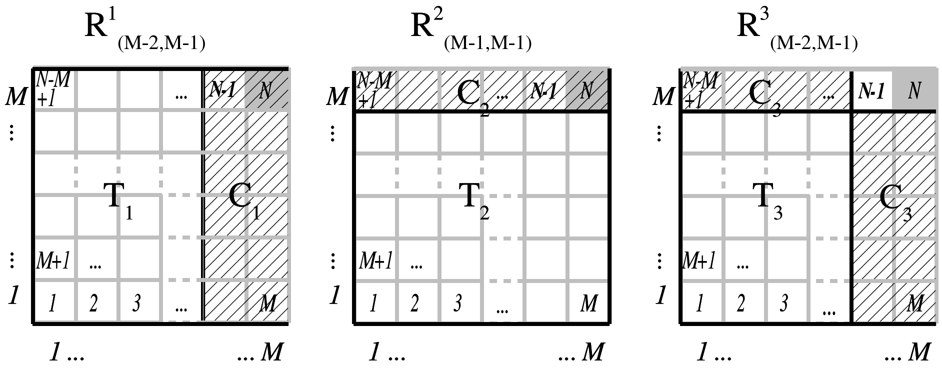

Figure 9: Ranklets $\mathcal{R}^1_{(M-2,M-1)}$ (left), $\mathcal{R}^2_{(M-1,M-1)}$ (centre, same as in Figure 8) and $\mathcal{R}^3_{(M-2,M-1)}$ (right). Infinity padding is required on the "open" (gray) boundaries. Pixels in the dashed area belong to the C sets, while the T sets are left white. The rank of pixel N (in gray) is fixed by σ (see text).

number of ties will be small and an effective strategy will be to break them at random. However, more and more ties are bound to occur as the size of the window W increases, so that especially in proving completeness we should be concerned with ties.

The standard approach to tied values in rank statistics involves the use of midranks for the computation of $\mathcal{W}_s$. Midranks are obtained by assigning to each group of tied values the the average of the ranks it occupies. Consider for instance a signal consisting of the eight sample values

$$I = (10, 128, 128, 132, 164, 164, 164, 255): \tag{9}$$

the corresponding midranks are

$$r^*(I) = (1, 2.5, 2.5, 4, 6, 6, 6, 8). \tag{10}$$

Here the midrank of the samples tied at 128, which is 2.5, is obtained by averaging ranks 2 and 3.

In the case of midranks, it is no longer convenient to model an N-sample signal as an element of S_N, since any internal permutation of a set of tied pixels leaves the ranking of the image unaltered. Indeed, once we have established which samples belong to each group of tied values, each of these groups behaves, from the point of view of the ranking, as an element in its own right. We can therefore conveniently replace S_N with the group S_{N^*}, where N^* is the number of distinct midranks (5 in the example given above).

A ranklet decomposition can then be described as a mapping

$$\mathcal{R}^* : S_{N^*} \to \mathbb{R}^m. \tag{11}$$

With this modification, the procedures outlined in Sections 3.2 and 3.4 lead to the construction of maps $\mathcal{R}^*$ that are injective, and therefore represent complete ranklet decompositions in the sense specified in Section 3.

A possible criticism to this approach is that the domain of the decomposition is defined "a posteriori" based on the number of tied samples in the observed signal. However, this reflects the standard procedure for the treatment of ties in rank statistics: the distribution of the Wilcoxon statistics $\mathcal{W}_s^*$ is computed conditionally given the number of observations tied at each value [3].

4 Efficiency Considerations

4.1 Computing Ranklets

As seen Sect. 2.2, $\mathcal{W}_{YX}^j$ equals the number of pairs $(\vec{x}_p, \vec{y}_q) \in \mathsf{T}_j \times \mathsf{C}_j$ such that the intensity at $\vec{x}_p$ is larger than the intensity at $\vec{y}_q$. Since $N^2/4$ such pairs can be formed out of the N samples in W, it would seem that the number of comparisons required should grow with the square of the cardinality of W.

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

4.2 An Algorithmic Viewpoint

Although a complete ranklet decomposition is not meant to provide an efficient way of sorting an image, it is interesting to investigate the relation between the representation and the classical sorting algorithms.

Consider the one-dimensional covering displayed in Figure 4. It should be evident that the only quantity required for the computation of $\mathcal{R}_7$ is the rank of sample 8, that can be obtained by comparing it to all the other elements and counting how many of them have a lower value.

Once the rank of sample 8 is established, $\mathcal{R}_6$ can be computed in the same way by counting how many sample values are lower than sample 7 (note that there is no need to compare elements 7 and 8, since the rank of the latter is already known). Thus, the computation of the seven ranklets in Figure 4 essentially amounts to performing an *enumeration sort* of the pixels. The only marginal difference occurs when the support of a ranklet is smaller than the sampling window. In the computation of $\mathcal{R}_3$, for instance, the ranks of samples 5 and 6 (obtained during the computation of $\mathcal{R}_4$ and $\mathcal{R}_5$) must be modified to account for the fact that samples 7 and 8 should no longer be considered in the ranking. However, this can be done without additional comparisons. As shown in [2], the complexity of the enumeration sort algorithm is $O(N^2)$.

4.3 Compactness of the Representation

We will now analyze the ranklet representation of a signal in terms of its compactness. The most straightforward way of describing a permutation of N elements consists in listing the elements $(1, 2, \ldots, N)$ in the appropriate order. Since each integer n requires $\ln_2 n$ bits for its representation, the total length of the description is $O(\ln_2 N!)$. Alternatively, one might list the $N!$ possible permutations in any given order, say lexicographic, and then address them in an optimal way through a binary tree. Interestingly, even in this case $O(\ln_2 N!)$ bits are required.

In the case of ranklets, we must allocate enough memory to represent the maximum possible value of each ranklet, which is obtained when the top ranks are all occupied by the samples in T (for simplicity, we omit the normalization and concentrate on the value of $\mathcal{W}_s^j$ as defined in Equation 1). With reference to the one-dimensional covering in Figure 4, we see that the maximum possible value of $\mathcal{W}_{s(n)}^j$, corresponding to ranklet $\mathcal{R}_n$, is given by

$$\mathcal{W}_{s(n)}^j = \sum_{i=n+1}^{2n} i = \frac{1}{2}n(3n+1). \tag{12}$$

Therefore, the number of bits required by the $N-1$ ranklets in Figure 4 is bounded by

$$\sum_{n=1}^{N-1} \ln_2 \left[\frac{1}{2}n(3n+1)\right] = \ln_2 \prod_{n=1}^{N-1} \frac{1}{2}n(3n+1) \tag{13}$$

which again is $O(\ln N!)$ (note that we could spare a few bits by discarding the padded samples in the computation of $\mathcal{W}_s^j$. The reason why this is possible is that the rank of the padded pixels is exclusively linked to their position and thus carries no information on the signal).

5 Conclusions

We have reviewed above the main properties of ranklets, a family of rank features defined in terms of the Wilcoxon (or, equivalently, the Mann-Whitney) statistics. Ranklets admit an intuitive interpretation in terms of pairwise pixel comparisons, are naturally multiscale and feature an orientation selectivity pattern similar to the orientation tuning of Haar wavelets.

The analogy with Haar wavelets is further developed by the proofs of completeness introduced in this report for one and two-dimensional ranklets. Completeness is intended in the sense that the computation of a ranklet decomposition should be equivalent to a ranking of the entire image.

Contrary to wavelets, exact image reconstruction is of course not an issue in the case of rank features. Rather, we believe the significance of our completeness result to lie in the following points: firstly, it shows that the entire information available to rank features is effectively captured by ranklets. Secondly, it further clarifies the analogy between ranklets and Haar wavelets. Thirdly, it is of relevance for computational models of biological vision such as [4], which suggest that temporal order coding might form a rank-based image representation in the visual cortex.

The discussion of specific applications is beyond the scope of this report. The use of ranklets for recognizing deformable patterns is demonstrated in [5] and [6], where face detection applications are discussed. A generalization of ranklets to hexagonal pixel lattices is described in [7].

References

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

[2] D. E. Knuth. *The art of computer programming*, volume III. Addison–Wesley, 1973.

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

[4] R. Van Rullen and S. J. Thorpe. Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Computation*, 13:1255–1283, 2001.

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

[6] F. Smeraldi. A nonparametric approach to face detection using ranklets. In *Proceedings of the 4th International Conference on Audio and Video-based Biometric Person Authentication, Guildford, UK*, pages 351–359, June 2003.

[7] F. Smeraldi and M. A. Rob. Ranklets on hexagonal pixel lattices. In *Proceedings of the British Machine Vision Conference, Norwich, UK*, volume 1, pages 163–170, September 2003.

[8] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proceedings of the 3rd ECCV*, pages 151–158, 1994.