Ordinal Coding of Image Microstructure

A. Koloydenko - Division of Statistics, the University of Nottingham, NG7 2RD, UK
D. Geman - Department of Applied Mathematics and Statistics, Johns Hopkins University, USA

Abstract -*Applications of rank-order-based methods to image and signal analyses have primarily focused on filtering. Classical median, min, and max filters have long been part of standard image processing toolboxes. More recent work has focused on more elaborate versions of such filters and associated computational issues. However, the application of these nonlinear methods to problems such as image interpretation has been scarce. We attempt to show that simple rank-order-based methods for coding image patches provide informative and computationally efficient local image descriptors.*
Keywords: Coding, informative descriptors, microstructure, rank-ordering.

1. Introduction

This paper is about ordinal, rank-based coding of microimages. Since this concept is not generally familiar in the computer vision community, we begin with an informal definition and an example. A *coder* or *quantizer* is a mapping F from $n \times n$ subimages to $n \times n$ integer-valued matrices with entries in $\{0, 1, \ldots, n^2 - 1\}$. Consider the 3×3 subimage s on the left-hand side of (1) below and suppose a threshold or granularity t has been fixed in advance, say $t = 16$. Then

$$s = \begin{smallmatrix} 33 & 8 & 32 \\ 11 & 15 & 3 \\ 14 & 65 & 67 \end{smallmatrix} \quad \overset{(F,t)}{\mapsto} \quad c = \begin{smallmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 2 \end{smallmatrix}. \tag{1}$$

The procedure can be easily explained (*efficient implementations can differ significantly*) as follows: The intensities of the subimage are first ranked: $3 \le 8 \le 11 \le 14 \le 15 \ll 32 \le 33 \ll 65 \le 67$, where $\ll$ indicates a jump of magnitude above $t = 16$. We code this subimage by assigning 0 to pixel(s) with the lowest intensity, hence $3(= s_{2,3}) \to 0(= c_{2,3})$, and proceed in the ascending order, counting only the significant transitions, i.e., those $> t$. Thus, $8, 11, 14, 15$ are all coded by 1; $32, 33$ by 2; and $65, 67$ by 3. We will refer to this subimage coding scheme (F, t) as an *ordinal quantizer*. Note that for any t, (F, t) is photometrically invariant: the photometric translates of a patch are coded identically. Another property of F that holds independently of t is that F preserves relative brightness: $s_u \le s_v \Rightarrow c_u \le c_v$.

We will also refer to these matrix-valued codes as *patterns* and the totality of possible codes as the *codebook* $\mathcal{C}$. Note that not every $n \times n$ matrix with entries from $\{0, 1, \ldots, n^2 - 1\}$ can be an (F, t) pattern, hence $\mathcal{C}$ is a proper subset of the above matrices. Finally, when n is very small, we refer to a subimage s and its pattern $F(s)$ as a "microimage" and a "micropattern", respectively. The "all zeros" pattern is special and represents small contrast noise or clutter with respect to (F, t).

1.1. Previous Related Work

In image and signal analyses, applications of order statistics [10] have primarily been to filtering "impulsive" noise that would be too difficult to eliminate with linear filters without over-smoothing useful structures. More recent work has focused on associated computational issues [12], more elaborate versions of such filters [11], and has also extended these filters to detectors of simple objects [13] for SAR imagery. Overall, however, the application of these nonlinear methods to problems involving image interpretation has been scarce.

The general ordinal quantizers (F, t) and "microimage codes" described earlier appear to have been originally introduced in [3] in the context of *natural image statistics*. At about the same time, [15] apply "local binary patterns" (LBP) obtained by an adaptive version of the uniform, pixel independent quantization of patches, to texture segmentation. It is also argued in [3] that (F, t) can produce better descriptors than patterns obtained via the uniform quantization.

Subimage quantizers F of [3] are apllied in [2] for registration of ultrasound breast images. Whereas acknowledging an improved performance in image registration relative to using the image intensities directly, the authors in [2] point out an undesirable sensitivity (at least in their application) to local contrast variations. Here (in §5.2) we consider general variable thresholds t, enabling F to *adapt* to local contrast variations, explicitly "factoring out" the contrast from the descriptor (implicitly, some contrast information can remain via statistical correlations). This approach aims to allow the geometric (2D) information content of the coder to be measured more accurately than, for example, if the coder of [15] is used. (The latter coder is adapted to the patch by setting the threshold to the intensity of a single (central) pixel.)

In [3], the probability distribution of F-codes is reported to vary insignificantly with the imaging domain, spatial and intensity scales, in effect suggesting that this distribution can be used as a *universal prior* for higher-level visual tasks. We shall return to the crucial issue of invariance in §3.

In [4], large samples of 3×3 subimages randomly sampled from images of natural scenes are considered. The natural subimage signal S is observed to concentrate in very "small" subspaces of its nine-dimensional ambient space. Specifically, according to [4], the high-contrast component of S washes out low-dimensional surfaces corresponding to ideal primitive microstructure (e.g. edgelets, blobs, bars, etc.). In the search for a compact description of the natural microimage data, their analysis relies on advanced techniques for sampling on nonlinear surfaces. Interestingly, the set of microstructure descriptors obtained in [4] as surface representations appear to correspond well with the codebook $\mathcal{C}$ (with the zero pattern removed). Thus, for the population of natural images, the simple (F, t) map can be viewed as a discrete, coarse representation of a continuous mapping of the microimage data to its principal surfaces. Clearly, the optimality of such a discretization depends on t; using a single t for coding many different locations of a complex image is unlikely to perform best for any specific purpose (see below). This provides another motivation to investigate adaptations of t to the local and global image contrasts.

Much has been learned about how biological visual processing adapts to contrast variations in natural images, presumably in order to maintain efficiency ([6],[7]). In [8] in particular, the authors show how the basic vision task of contour integration might be adapted to local contour statistics of natural images, and how this adaptation would then allow integration to begin earlier in the process than previously believed. We are thus also motivated by the above hypothesis that the functional architecture of an efficient vision system might allow for almost simultaneous detection and integration of microscopic structure from natural stimuli. Specifically, this leads us to assess (§3) suitability of (F, t) codes for their efficient integration.

1.2. Contributions

In view of the current status of ordinal methods, work on natural image statistics, and the somewhat isolated observations mentioned above, we attempt to integrate these ideas and findings into a coherent assessment of the suitability of ordinal microimage coders for investigating image content.

The information content, and perceptual distortions, of coding microimage populations by (F, t) can be studied in the context of empirical probability distributions $\{p_{F|I}(c|i)\}_{c\in\mathcal{C}}$ of F-codes within individual images i, or even over whole imaging domains (with $\{p_F(c)\}_{c\in\mathcal{C}}$ being the domain average of $\{p_{F|I}(c|i)\}_{c\in\mathcal{C}}$). Using 2×2 as the smallest non-trivial microimage configuration, we estimate $\{p_F(c)\}_{c\in\mathcal{C}}$ and its entropy $\mathcal{H}(p_F)$ from thousands of high resolution natural images. The estimates are reported with *simultaneous confidence intervals*. Obtained at standard significance levels, these intervals appear rather tight, which is not surprizing given the findings in [3] on stability of the microimage distribution. (In fact, *simultaneous confidence intervals* are used to estimate the p_F probability vector.) For F with adaptive thresholds, we additionally estimate the mutual information $\mathcal{M}(S; F)$ between S, the original microimage signal, and its F coding. These and other experimental results are presented in §5 and include comparisons of different adaptive versions of F by mutual information $\mathcal{M}(S; F)$. In §6, we also suggest to measure (average) perceptual distortions $d(S, F)$ for a more complete comparison of such coders.

2. Ordinal Quantization

Putting aside computational efficiency, the F transformation of $n \times n$ subimages $s \in \mathbb{R}^{n^2}$ into codes is formally described by the following steps:

1. Compute the ranks $s^{(1)} \leq s^{(2)} \leq \ldots \leq s^{(n^2)}$.
2. Derive the (discrete) derivatives $0, s^{(2)} - s^{(1)}, s^{(3)} - s^{(2)}, \ldots, s^{(n^2)} - s^{(n^2-1)}$.
3. Binarize by thresholding with t: $0, \mathbb{I}_{\{s^{(2)}-s^{(1)}>t\}}, \mathbb{I}_{\{s^{(3)}-s^{(2)}>t\}}, \ldots, \mathbb{I}_{\{s^{(n^2)}-s^{(n^2-1)}>t\}}$, where $\mathbb{I}_A$ stands for the indicator of set A.
4. Integrate the resulting derivative chain, producing (2) below.
5. Compose the code matrix by placing these in their proper original pixel locations.

$$0, \mathbb{I}_{\{s^{(2)}-s^{(1)}>t\}}, \mathbb{I}_{\{s^{(2)}-s^{(1)}>t\}} + \mathbb{I}_{\{s^{(3)}-s^{(2)}>t\}}, \ldots, \sum_{j=1}^{n^2-1} \mathbb{I}_{\{s^{(j+1)}-s^{(j)}>t\}}. \quad (2)$$

Summarizing:

Definition 1 *Let L represent the number of pixels in patch $s \in \mathbb{R}^L$ and let $t > 0$. The* ordinal quantizer $(F, t) : \mathbb{R}^L \to \mathbb{R}^L$ *is defined component-wise by* (3) *below:*

$$F_l(s) \quad = \quad \sum_{j=1}^{r_l-1} \mathbb{I}_{\{s^{(j+1)}-s^{(j)}>t\}}, \quad (3)$$

where $s^{(1)}, s^{(2)}, \ldots, s^{(n^2)}$ are the (ascending) order statistic of the intensities in patch s, and r_l is the ascending order rank of the l^{th} pixel.

Note that the assignment (3) is in fact independent of possible ties in ranking and hence well-defined. Note also that, for (unbounded) continuous intensities, all the maps F have the same range $\mathcal{C}$ independently of t. Namely, if c is a pattern under (F,t) for some t, than it is also a pattern for all $t>0$. However, in practice s is almost always discrete, hence, depending on the initial intensity quantization, the range of F for one value of t might be properly contained in the range of F for $t'<t$. Then, $\mathcal{C}$ corresponds to the initial intensity quantization granularity (e.g. 1 for integer gray levels), and its size can easily be calculated as follows:

$$|\mathcal{C}| = \sum_{m=1}^{L} \sum_{\substack{k_l>0\ 1\le l\le m \\ k_1+\ldots+k_m=L}} \binom{L}{k_1\ k_2 \cdots k_m}, \qquad (4)$$

where m represents the number of distinct subregions comprising the patch. The index m is also the number of subchains in the order statistic separated by significant jumps, and can be thought of as a local estimate of the depth of the image surface (§3). In the case of 2×2 patches ($n=2, L=4$), for example, there are $|\mathcal{C}|=75$ micropatterns.

3. Relevance for Visual Processing

We show elsewhere that (F,t) coding emerges naturally as satisfying a set of "perceptual axioms" based on the brightness partial order relations on the subimage and pattern sets. A central axiom is that F needs to commute with pixel permutations σ: $\sigma^{-1}(F(\sigma(s)))=F(s)$. We find this axiom sensible for "early vision", thinking of the primitive visual processor capable only of sorting out the received luminance, and not equipped with a basis for direct estimation of directional derivatives.

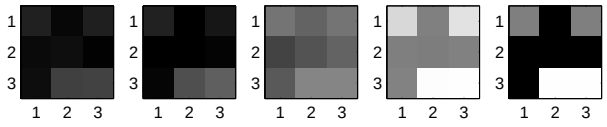
3.1. Invariance

An important issue here is that of invariance: Many subimages may share a code c and there are both perceptual and statistical relationships among subimages within one such "fiber".

When the semantic explanation of image structure is of direct importance, a high level of photometric invariance is desired and illumination is a nuisance parameter. Often, this type of invariance is considered at the global, i.e. image scale, requiring an operation to yield the same results for image I as for its (constant) translates $I+const$. For every t, F is clearly invariant to photometric translation: $F(s)=F(s+const)$. Additionally, if the same coding (i.e. keeping t fixed) is carried out for every disjoint $n\times n$ subimage(s) s of I (or a subset thereof), the collective result clearly remains unchanged even if the individual subimages were "brightened" unequally (by $const(s)$).

Photometric translation invariance simply means that, as a coder of an isolated subimage, (F,t) conveys no information about the absolute brightness of the coded patch. However, when coding a *population* of subimages, some of the original brightness information may be statistically correlated with structure, and hence may still be preserved indirectly.

In fact, the photometric shifts are only a subset of all patch transformations that do not affect the F-codes. This can, for example, be seen from Figure 1. It can also be seen



33	8	32	34	0	23	117	100	117	217	129	227	128	0	128
11	15	3	0	0	5	68	84	100	128	126	130	0	0	0
14	65	67	6	80	96	90	134	134	131	254	255	0	255	255

Figure 1: Subimages that are mapped to the same right-hand pattern of (1) by $(F,16)$ are rendered with 256 levels of gray (top) and numerically transcribed (bottom).

(as mentioned earlier) that under (F,t), any two patches receive a common code if they answer the same way to all the queries: "Is pixel u brighter than pixel v by more than t?" At the same time, answering all such queries identically is not necessary for two patches to be coded identically. Thus, there is still more invariance than explained by aggregating patches based on the binary queries above. This remaining invariance can intuitively be described based on the structure of the patch order statistic (see, for example, the commentary to (1) above): In two patches with the same code, a pair of pixels from a subchain enclosed by significant jumps ($\ll$) may respond differently to the same query.

Unlike the photometric shift invariance, the additional modes of invariance above do depend on t, which might be exploited in applications by applying F with variable t. This latter possibility has largely motivated this work.

3.2. Information Content

We reiterate that, despite being defined purely based on intensity ordering, the "non-flat" (F,t)-micropatterns do carry primitive directional information as well as primitive surface depth information. This can be seen from Figure 2.

Most binary patterns are trivially associated with one of the eight directions (on the $\pi/4$ angular scale). (The "ridgelets" $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$ can be associated with pairs of opposite directions, but occur at the microscopic scale less frequently; see §5.) Any non-zero non-binary pattern c can be "pulled back" up to the level of its coarser, binary ancestors (although non-uniquely) in order to "estimate" the

captured direction.

The uncertainty in simultaneous estimation of position and direction is well-known to be generally inevitable. This scheme provides a simple mechanism to exercise the trade-off: Suppose, for instance, the pattern $\begin{smallmatrix}01\\00\end{smallmatrix}$ is signaling the same diagonal direction as $\begin{smallmatrix}11\\01\end{smallmatrix}$. In that case the two would be each other's translates along the normal to the signaled direction. Lowering the granularity t, would then refine $\begin{smallmatrix}01\\00\end{smallmatrix}$ to $\begin{smallmatrix}12\\01\end{smallmatrix}$.

A competing interpretation might link $\begin{smallmatrix}01\\00\end{smallmatrix}$ to $\begin{smallmatrix}11\\00\end{smallmatrix}$, in which case $\begin{smallmatrix}01\\00\end{smallmatrix}$ would resolve into $\begin{smallmatrix}12\\00\end{smallmatrix}$. In general, under this scheme, refining the interpretation of a code corresponding to an internal node, requires lowering the granularity t. Thus, a contour integrator might exercise this granularity control, for example, in a feedback loop iteratively attempting to connect neighbors if their codes have compatible directions. The directional ambiguity of the coarser codes would then lead to more aggressive explorations of continuity, perhaps resulting in illusory contours. When, and if, the number of contouring hypotheses reaches a predefined limit, the codes of the more ambiguous junctions would be refined. Any discovered misalignment would then result in discarding the affected contours.

The depth of the code in the $\mathcal{C}$ hierarchy appears to provide rudimentary information about the depth of the coded structure. We refer to this depth as $m(c)$ ($m = 1, \ldots, 4$ as determined by $n = 2$). There is certainly even more ambiguity in interpreting the depth information than the directional one. However, short of real estimation of curvature, the (F, t) scheme does appear to be sufficient in the sense of providing the *2.5D primary sketch* [9]. Parallel implementations of this coding (with larger patches) might then also allow for fast scene exploration and local contour integration to develop simultaneously.

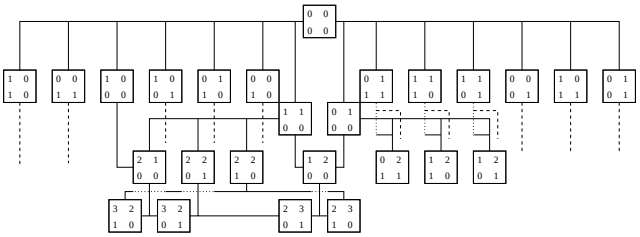


Figure 2: A fragment of $\mathcal{C}$ for $n = 2$ illustrating, among other features, how the direction and depth of the microstructure can be induced from the "brightness" order relation on $\mathcal{C}$ preserved from the original such order on the patch space.

Let us apply F to every distinct 2×2 microimage of image in Figure 3. Figure 4 marks as white all the "flat" (i.e. background) sites of image i, setting $t = contr(i)$, the spatial contrast of the image according to the following

Figure 3: A natural image (log intensities are used due to the very large image size).

definition:

$$contr^2(i) = \frac{1}{M(N-1) + N(M-1)} \sum_{k \sim l} (i_k - i_l)^2 \quad (5)$$

where the summation is over all $M(N-1) + N(M-1)$ neighboring (vertical and horizontal but not diagonal) pairs (k, l) of image sites. This definition of contrast is advocated in [4], in particular for being (a discrete version of) the only scale invariant norm on image spaces. Note that $contr^2(i)$ is also the average of the (unit lag) vertical and horizontal (sample) variograms. We also find this definition helpful in general for comparing intensity images presented on different intensity scales, and also for comparing subimages within an image.

As shown in Figure 5, the unit contrast coincides (for this image on the logarithmic intensity scale) with the 80-th percentile of the empirical distribution of the (absolute) differences between the horizontally and vertically neighboring pixels. Let us denote by G the corresponding distribution function, and by G^{-1} - its inverse, the quantile function.

4. Microimage and Micropattern Distributions and Information

The statistical context of our experiments requires some notation. The reader familiar with information theory might skip the latter parts of this section.

Let Q represent the original quantized intensity scale, in our case $Q = \{0, \ldots, 2^{16} - 1\}$. Let $\mathcal{I} = Q^{MN}$ be our image space for some positive image dimensions M and N ($M = 1024$, $N = 1536$ in our experiments), and let i refer to an individual image from $\mathcal{I}$. The space of

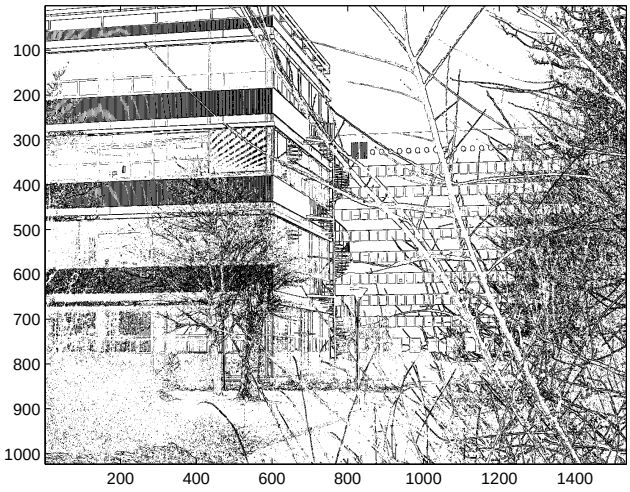


Figure 4: The background class (white) occupies 74% of image in Figure 3, according to F with $t = contr(I)$.

microimages s is denoted by $\Omega = Q^{n^2}$. We are going to think of the microimage signal S as a random variable with values in Ω distributed according to $p_{S|I}$, the distribution of the population of *all the microimages s observed in image i*:

$$p_{S|I}(s|i) = \frac{\sum_{s'} \mathbb{I}_{\{s'=s\}}}{(M-n+1)(N-n+1)}, \qquad (6)$$

where the summation is over all $n \times n$ subimages of i.

Let $\mathcal{C}$ be a finite set, image i be fixed and let $F : \Omega \to \mathcal{C}$.

Definition 2 *The F-microcode distribution of image i (over codebook $\mathcal{C}$) is the distribution of the population of all the micropatterns $F(s)$ observed in i:*

$$p_{F|I}(c|i) = \frac{\sum_{s} \mathbb{I}_{\{F(s)=c\}}}{(M-n+1)(N-n+1)}. \qquad (7)$$

Given a probability distribution P on $\mathcal{I}$, the expectations

$$p_S(s) = \int_{\mathcal{I}} p_{S|I}(s|i)dP(i), \qquad p_F(c) = \int_{\mathcal{I}} p_{F|I}(c|i)dP(i)$$

are *the mean microimage and microcode distributions*, respectively.

Similarly defined are $p_{S,F|I}(s,c|i)$ and $p_{S,F}(s,c)$, the *joint distributions* of the microimage signal S and its F-coding within a single image i, and averaged with respect to the image distribution P, respectively. The entropy of the random variables S and F (within image i) are written as $\mathcal{H}(S|I=i)$ and $\mathcal{H}(F|I=i)$, respectively. (For a comprehensive reference on information theory, see [1].) The corresponding conditional entropies are also given by

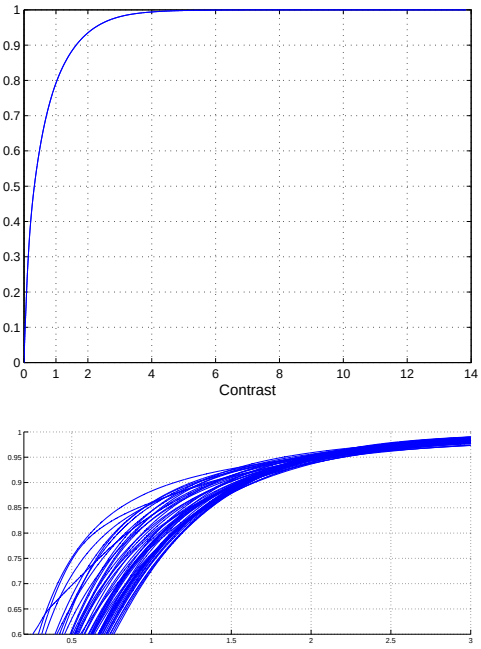


Figure 5: Cumulative distribution functions G for the population of the (absolute) differences between the horizontal and vertical neighbor pixels of image in Figure 3 (top), and other 50 random images (below) from the same data set (§5.3). The x-values are scaled to unit contrast (5)

$\mathcal{H}(S|I) = \int_{\mathcal{I}} \mathcal{H}(S|I=i)dP(i)$ $\mathcal{H}(F|I) = \int_{\mathcal{I}} \mathcal{H}(F|I=i)dP(i)$, respectively, where the random image I is assumed to follow P.

Finally, given $i \in \mathcal{I}$, the mutual information between the microimage signal S and its F-coding is defined in (8) below:

$$\mathcal{M}(S;F|I=i) = \qquad (8)$$
$$\sum_{s\in\Omega, c\in\mathcal{C}} p_{S,F|I}(s,c|i) \log \frac{p_{S,F|I}(s,c|i)}{p_{S|I}(s|i)p_{F|I}(c|i)}$$

and the conditional information (relative to P) is then

$$\mathcal{M}(S;F|I) = \int_{\mathcal{I}} \mathcal{M}(S;F|I=i)dP(i). \qquad (9)$$

5. Experiments

We are interested in estimating probability distributions of (F,t) codes for various methods of defining t. In particular, we are interested in assessing variations of these estimates from image to image. We are also interested in estimating the corresponding information measures ($\mathcal{H}(F|I)$

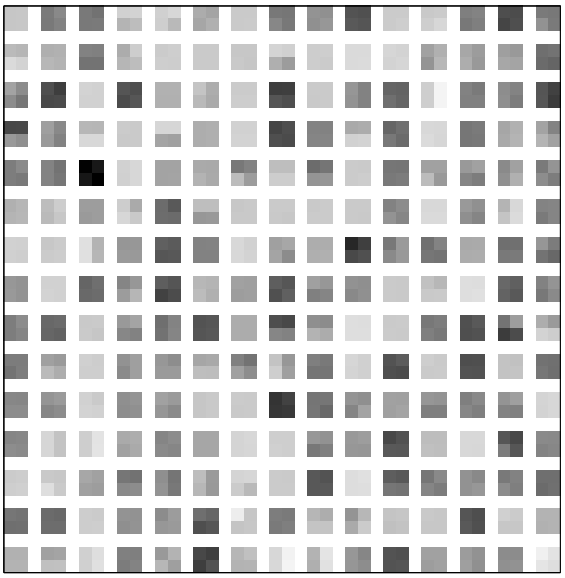


Figure 6: A random sample of 2×2 subimages from image in Figure 3.

and $\mathcal{M}(S;F|I))$ in view of their dependence on the t selection. Finally, we would like to assess these coders perceptually, similarly to how Figure 4 explicates the notion of microstructure according to a particular t-selection method.

5.1. Data

Our test images come from the popular van Hateren's collection of 4167 still natural stimuli: 1024×1536, two bytes/pixel, raw images of natural and urban landscapes obtained with a Kodak DCS420 camera, "linearized with the lookup table generated by the camera for each image" [5]. We have used both the truly linear version as well as the PSF-corrected ("deblurred") version, but present here only the results for the latter, "deblurred" images. We have also excluded 49 irregular images (42 of which appear extremely blurred, with the other seven being incorrectly oriented), arriving at the image sample size of $N_{im} = 4118$. Thus, we assume that $\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_{N_{im}}$ are i.i.d. random images distributed according to a hypothetical natural image distribution P on $\mathcal{I}$.

5.2. Adaptive Ordinal Quantization

We have considered two categories of methods of adaptation of (F,t) coders to contrast variations within image i. *First*, the coder can be adapted globally, i.e. t is computed based on some global image function. E.g. $t = \gamma \cdot contr(i)$, a fraction of the spatial contrast, or $t = G^{-1}(\gamma)$, the $\gamma \cdot 100\%$-th percentile of the distribution of the pairwise absolute differences. Note that in any of these cases every instance of configuration s in the image receives the same code $F(s)$ ($\mathcal{M}(S;F|I=i) = \mathcal{H}(F|I=i)$).

Second, (F,t) coders can be adapted *locally*. For instance, t can be set to $\gamma \cdot contr(\mathcal{N}_s)$ or $G^{-1}_{\mathcal{N}_s}(\gamma)$, where $\mathcal{N}_s$ is some neighborhood of the coded patch s. Note that in this case, it is possible for two identical patches $s = s'$ extracted from different locations in the image to be coded differently $F(s) \neq F(s')$ ($\mathcal{M}(S;F|I=i) = \mathcal{H}(F|I=i) - \mathcal{H}(F|S,I=i) < \mathcal{H}(F|I=i)$).

Graphs in Figure 8 show dependence of the coded information on the level γ with $t = G^{-1}(\gamma)$, using global (F_∞) and local (F_8 - 8×8 neighborhoods) adaptations. The estimates are obtained for $K = 16$ levels of γ and are presented along with their 90% *simultaneous* confidence intervals computed according to (10) below, using the multivariate normal asymptotic ([14]):

$$\mathcal{M}_\gamma(S;F) = \overline{\mathcal{M}_\gamma(S;F)} \quad \pm \quad s_\gamma \sqrt{\frac{K\mathbf{F}_{K,N_{im}-K}(\alpha)}{N_{im}-K}},$$

where $\overline{\mathcal{M}_\gamma(S;F)} = \frac{1}{N_{im}} \sum_i \mathcal{M}(S;F|I=i)$ and $s^2_\gamma = \frac{1}{N_{im}} \sum_i (\mathcal{M}_\gamma(S;F|I=i) - \overline{\mathcal{M}_\gamma(S;F)})^2$ are the sample mean and variance of the $\mathcal{M}_\gamma(S;F)$, respectively, and $\mathbf{F}_{K,N_{im}-K}(\alpha)$ is the $(1-\alpha) \times 100$-percentile of the F-distribution with K and $N_{im} - K$ degrees of freedom; $\alpha = 0.1$.

5.3. Estimation of p_F

Recall (4) that with $n = 2$, we have $|\mathcal{C}| = 75$ patterns. We compute the 75-dimensional probability vectors $p_{F|I}(\cdot)$ for every image in the data set, and for several (F,t) coders distinguished by the threshold selection method, and we estimate the corresponding means $p_F(\cdot)$ (over the natural image ensemble) with simultaneous confidence intervals similar to the ones in (10) (replacing K by $|\mathcal{C}|-1$). In Figure 9, p_F is estimated when F is adapted to 8×8 contexts $\mathcal{N}$ via $t = contr(\mathcal{N})$.

6. Summary and Conclusions

We have discussed a class of ordinal, nonlinear methods for coding microscopic structure in intensity images with a view toward image interpretation. Our goal has been to collect theoretical and empirical evidence in order to assess practical suitability of these methods. Our extensive statistical analysis allows us to conclude that the considered coders capture microscopic structure in various natural domains, and with different image preprocessing, consistently: High frequency patterns in any one such domain, or with any one particular intensity scale, remain comparably frequent in other domains and with other intensity scales

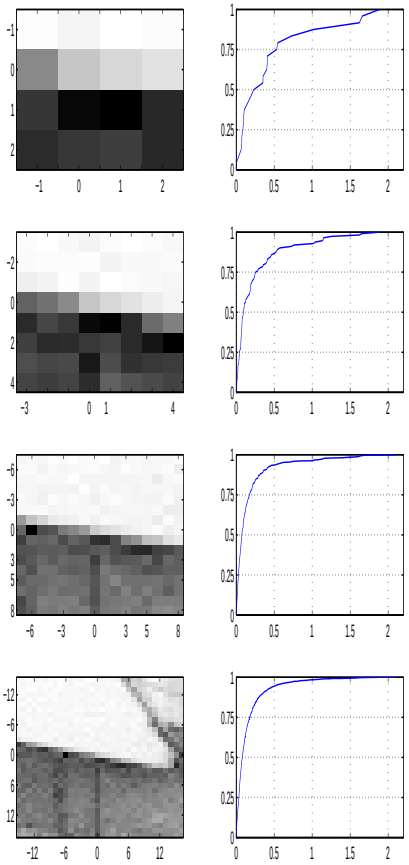


Figure 7: Left: A 2×2 central block s is presented in its variable size natural contexts $\mathcal{N}_s$ extracted from image in Figure 3. The block is coded either as $\begin{smallmatrix}12\\00\end{smallmatrix}$ or $\begin{smallmatrix}11\\00\end{smallmatrix}$, depending on the context $\mathcal{N}_s$ and the significance level γ for threshold selection. Right: Cumulative distribution functions $G_{\mathcal{N}_s}$ for the corresponding populations of the (absolute) differences between the intensities of the horizontal and vertical neighbors.

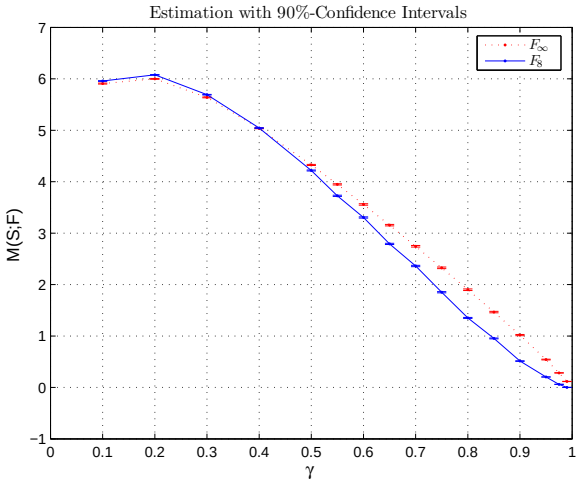


Figure 8: Comparison of the information contents of global (F_∞) and local (F_8) adaptations across several γ levels. Simple perceptual assessment (e.g. Figure 4) suggests that it is only the coders with γ levels of about 0.8 and above that filter out microscopic noise. Thus, there appears to be only 1-2 bits of "useful" information at the microscopic level.

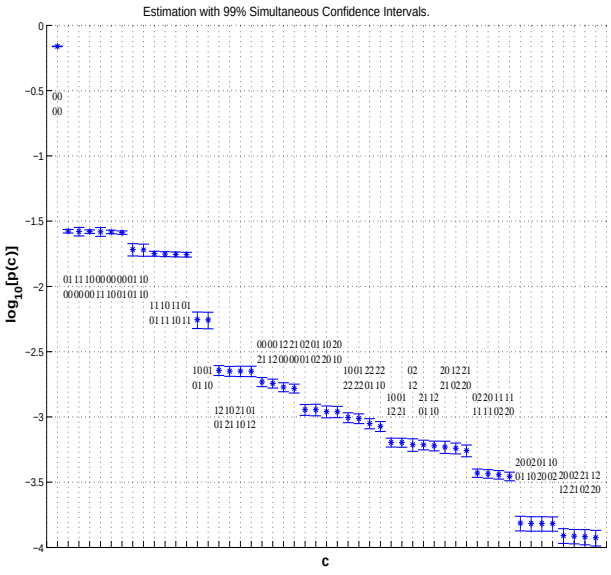


Figure 9: Local adaptation, $t = contr(\mathcal{N})$ with 8×8 contexts $\mathcal{N}$. Principal masses (i.e. 50 most frequent out of 75 total patterns) total to $\approx 99\%$.

and under variable preprocessing regimes. The information content of these coders is consequently stable with respect to the above conditions.

Additional adaptations have been introduced in order to emphasize flexibility of the quantization methods as well as to estimate $\mathcal{H}(F|S)$, the coding uncertainty due to the local contrast variations.

The presented comparisons of information contents $\mathcal{M}(S;F)$ and $\mathcal{M}(S;F')$ of coders F and F' can certainly be expanded. For example, in addition to a global perceptual analysis, a *distortion* measure $d(s,c)$ cane be introduced to relate micro- and macroscopic perceptual losses, similarly to the rate distortion analysis ([1]). For instance, the considered (F,t) coders preserve relative brightness. However, this might become an undesirable constraint in the face of a plausible competition with spatial continuity (in larger patches). Thus, coders with $\mathcal{M}(S;F) = \mathcal{M}(S;F')$ but of different (suitably defined) distortions $d(S,F(S)) \neq d(S,F'(S))$ might need to be compared. There are several relevant choices for such measures ([1]), with (10) being but one example:

$$d(s,c) = \frac{2}{n(n-1)} \sum_{k \neq l} G((s_k - s_l)^+) \mathbb{I}_{\{c_k \leq c_l\}}, \quad (10)$$

where G is as before the cdf of the population of the (absolute) difference between horizontal or vertical neighbors in a given image;$(s_k - s_l)^+ = \max(0, s_k - s_l)$. This scales the distortion of the relative brightness between two pixels to $[0,1]$.

Computing these codes efficiently is no less important for applications, and the necessity of executing many sortings, however optimally [12], might still be unattractive for applications. The fact that the set of answers to all binary queries $s_u - s_v > t$ within patch s uniquely identifies the code of s appears crucial: For a given subimage, a small subset of these answers might be sufficient. Depending on the subimage population (i.e. on the joint distribution of the query bitstrings and the F-codes), different subimages would generally require different subsets of queries for computational efficiency in the sense of minimizing the average number of the queries evaluated in determining the code. For this purpose, a suboptimal tree-based Vector Quantizer based on the above queries at internal nodes and the F-codes at terminal nodes might be learned from the training subimage data; for instance, one may use the same strategy as in building decision trees, namely greedy entropy reduction. Regarding the internal nodes as approximate, yet more invariant, F-codes might also be a viable option for generating a coarse-to-fine hierarchy of local image features for object recognition and classification. Such VQ trees are but one example of defining local image features based on ordinal methods.

References

[1] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991

[2] H. F. Neemuchwala, A. O. Hero, and P. L. Carson "Image matching using alpha-entropy measures and entropic graphs" *Signal Processing. Special Issue on Content-based Visual Information Retrieval*, Vol. 85, 2005.

[3] D. Geman and A. Koloydenko. "Invariant statistics and coding of natural microimages" In S. C. Zhu, editor, *IEEE Workshop on Statistical and Computational Theories of Vision*, web published, 1999

[4] A. Lee, K. Pedersen, and D. Mumford. "The complex statistics of high contrast patches in natural images" In S. C. Zhu, *IEEE Workshop on Statistical and Computational Theories of Vision*, web published, 2001

[5] J. H. van Hateren and A. van der Schaaf. "Independent component filters of natural images compared with simple cells in primary visual cortex" *Proc. R. Soc. Lond.*, Vol. B 265, 1998.

[6] M. A. Webster and E. Miyahara. "Contrast adaptation and the spatial structure of natural images" *J. Opt. Soc. Am. A*, Vol. 14 9, pp. 2355-2366, 1997.

[7] N. Brady and D. Field. "Local contrast in natural images: normalisation and coding efficiency" *Perception*, Vol. 29 9, 2000.

[8] R. F. Hess and A. Hayes, and D. Field. "Contour integration and cortical processing" *Journal of Physiology - Paris*, Vol. 97 2-3, 2003.

[9] D. Marr, *Vision*, W. H. Freeman and Co., N. Y, 1982.

[10] G. Heygster "Rank filters in digital image processing" *Computer Graphics and Image Processing*, Vol. 19 2,1982.

[11] Y. Nie and K. E. Barner "Fuzzy transformation and its applications" *Proc. Int. Conf. on Image Processing*, Vol. 1,2003

[12] C. C. Lin and C. J. Kuo "Two-dimensional rank-order filter by using max-min sorting network" *IEEE trans. on circuits and systems for video technology*, Vol. 8 8, 1998

[13] A. K. Mitra and T. L. Lewis, and A. Shaw "Rank-order based filters for FOPEN target detection" *IEEE Signal Processing letters*, Vol. 11 2, 2004

[14] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley, series in probability and mathematical statistics, 2d edition, 1984.

[15] T. Ojala and M. Pietikäinen "Unsupervised texture segmentation using feature distributions" *Pattern Recognition*, 32, 1999