

NON-STATIONARITY DETECTION IN NATURAL IMAGES

Raghu G. Raj, Alan C. Bovik, and Wilson S. Geisler

Center for Perceptual Systems, University of Texas at Austin, Austin, Texas 78712, USA

ABSTRACT

We present a novel approach for non-stationarity detection in natural images by exploiting the prior knowledge of the independent component structure of scene statistics. Our proposed non-stationarity index is conceptually simple and is intertwined with the probabilistic structure of the image segment being analyzed. It shows consistently good results when applied to natural scenes and, we expect, will find useful applications in computer vision algorithms inasmuch as the detection of statistically non-stationary locations in images can be an important preliminary step toward the understanding of scene content and in the guiding of visual fixations.

Index Terms— Non-stationarity, Natural Scene Statistics, Textures, ICA, Vision Systems, Fixation selection

1. INTRODUCTION

A central factor that contributes to image understanding is the grouping of the image into regions that are likely to result from independent physical processes; for example, the pattern of clouds in the sky and the pattern of branches in a tree arise from independent physical processes. This impressive ability of the Human Visual System (HVS) to perform such perceptual groupings is based to a large extent on the knowledge (acquired through eons of learning and evolution) of the typical higher-level structure and semantics of natural images such as object structure, typical relationship between objects found in images etc.

However the nature of these so-called higher-level inferences (and the efficiency of resulting algorithms) are fundamentally dependent upon the characterization of the underlying spatial statistical processes that constitute natural scenes. This paper exclusively deals with the low-level statistical structure of images wherein we introduce a novel statistical measurement called *non-stationarity index* which can provide a building block that can be effectively utilized by higher-level vision algorithms.

The motivation for this statistical measurement stems from the simple observation that a region of an image corresponding to a single physical process can often be described as an approximately statistically stationary texture. More generally, for a given scale of analysis, any non-stationary region can be

decomposed as a non-linear combination of underlying spatially stationary processes. Throughout this paper we assume a simple non-linearity which is that of a partitioning of the region into spatially stationary textures. Though more complicated non-linearities (such as occlusion models) can be considered, we nevertheless find useful results in the mapping of the non-stationary structure of natural scenes as we show in the Section 3.

For the purpose of detecting non-stationarities (based on the simple image model described above) we propose a novel index that is based on measuring scene statistics within image sub-regions, using Independent Component Analysis (ICA) [1, 2]. The goal of ICA is to find directions in image-patch space which (when projected onto), produce random variables that are statistically independent. Our proposed non-stationarity detection scheme is based on the intuition that performing an ICA analysis on a local image region that contains more than one stationary texture should yield components with statistical dependences larger than the dependencies between the ICA components of the stationary textures that compose the image region. We describe this in more detail in Section 2 where we also propose our non-stationarity index. Section 3 gives simulation results demonstrating the effectiveness of our approach of non-stationarity detection in natural images. We conclude in Section 4.

Importantly, we also note that since the computation of our non-stationarity index is intertwined with ascertaining the probabilistic structure of the image segments being analyzed, detailed probabilistic information of the textural properties constituting the scene can be determined in this process which in turn can be useful in guiding image segmentation, object detection and other high-level vision algorithms.

2. NON-STATIONARITY INDEX

We define an image patch to be a set of contiguous pixels (occupying a region in the image J) whose bounding contour is a simple closed curve. Let $\{\phi_i^n\}_{i=1}^K$ be the ICA components corresponding to texture T_n (where K is assumed to be equal to the dimensionality of the data space of each texture) such that $\phi_i^n \in \mathbb{R}^{K \times 1}$ for all i , $s^n = B^n y^n$ where $y^n \in \mathbb{R}^{K \times 1}$ consists of K -dimensional samples of texture T_n , $B^n = [\phi_1^n, \dots, \phi_K^n] \in \mathbb{R}^{K \times K}$, and $s^n \in \mathbb{R}^{K \times 1}$ consists of the ICA projections of texture T_n . We say $\{T_n\}_{n=1}^N$ is a set of

unique textures if, for all k ($1 \leq k \leq N$), at least one ICA component of T_k differs from the ICA components of T_m for $m \neq k$.

We call an image region non-stationary if it can be partitioned into N non-overlapping sub-patches (of arbitrary geometry) each of which contains a unique texture T_n (for $1 \leq n \leq N$). For simplicity let us consider the case where $N = 2$. For this case, consider one such sub-patch (of the image patch T) which we call the *center patch* for which we compute the corresponding ICA components $\{\phi_i^{ctr}\}_{i=1}^K$. Let H_{ctr} be the covariance matrix of this center patch with respect to $\{\phi_i^{ctr}\}_{i=1}^K$. Now let H_{surr} be the covariance matrix (again with respect to $\{\phi_i^{ctr}\}_{i=1}^K$) of the region of the image patch that surrounds the center patch (which we call the *surround patch*). Then it is clear that $F(\eta_{ctr}, \eta_{surr}) = \frac{\eta_{ctr}}{\eta_{surr}} \geq 1$ where, $\eta_{ctr} = \frac{\text{trace}(H_{ctr})}{\|H_{ctr}\|}$ and $\eta_{surr} = \frac{\text{trace}(H_{surr})}{\|H_{surr}\|}$ are ratios that measure diagonal dominance of the center and surround covariance matrices respectively, and where $\|\cdot\|$ is the l_1 norm of the matrix. Assuming that we normalize the variances of each channel to unity, we have that $\text{trace}(H_{ctr}) = \text{trace}(H_{surr}) = K$ and thus $F(\eta_{ctr}, \eta_{surr}) = \frac{\|H_{surr}\|}{\|H_{ctr}\|}$.

Clearly $F(\eta_{ctr}, \eta_{surr})$ is unity when the image patch is a stationary texture; and the larger the deviation of F from unity the more non-stationary the image patch is. But from a theoretically point of view (even assuming perfect ICA decomposition) it cannot be guaranteed that F is different from unity for non-stationary patches due to the fact that we are only considering second order statistics. We point out, however, that when using a related Multilinear ICA (MICA) [3] model (which is a refinement of the ICA model designed to capture dependencies among the pseudo ICA filters) instead of ICA, one can demonstrate that for a related inequality of a form similar to the one above, strict inequality holds for non-stationary patches and that equality holds if and only if the center and surround patches are sub-patches of the same texture [3]. But for practical purposes we find, from the simulations below, that our proposed non-stationarity index $F(\eta_{ctr}, \eta_{surr})$ does indeed give us good results in mapping non stationary regions in natural images and textures.

Due to the fact that ICA vectors computed from natural scenes are non-ideal (i.e. does not completely factor out residual dependences between the ICA vectors), we occasionally do find that the $F(\eta_{ctr}, \eta_{surr})$ can be slightly less than unity for some 'roughly stationary' patches. In the simulations below, for the purpose of facilitating visual interpretation of the non-stationarity maps, such values of the non-stationary index are mapped to their reciprocals since it is the extent of the deviation from unity that indicates the degree of non-stationarity.

One can of course substitute the ICA vectors employed in the computation of the non-stationarity index with PCA or a fixed set of filters. But an advantage of an ICA-based approach is that with only a slight increase in the computational

load required to compute the ICA filters (i.e. for small patch sizes, as detailed in the next section), one can get an approximate characterization of probability density function of the image patches by a product of the marginals distributions—and a more accurate characterization by means of the corresponding MICA decomposition [3]. This lends to a compact probabilistic and sparse characterization of the textural regions of the image. Furthermore, these ICA filters can be utilized to derive feature vectors that can be successfully used for texture classification and segmentation purposes [4].

Another issue is the choice of the location of the center patch. Ideally we would like to place the center patch in a stationary region of the image patch. But in practice since we consider relatively small patch sizes (compared to the size of the image), the image patches can be continuously sampled in the image so that all possible locations are covered. The use of such a dense non-stationary map obviates the need for a careful choice of center patch location. Here, we obtain useful results with 32×32 center patches, centered in 64×64 image patches that densely sample 1024×1024 images.

3. SIMULATION RESULTS

As explained in the previous section, our non-stationarity index is $F(\eta_{ctr}, \eta_{surr}) = \frac{\eta_{ctr}}{\eta_{surr}}$. An image patch is deemed non-stationary if $F(\eta_{ctr}, \eta_{surr}) \geq \tau$, for some empirically determined threshold $\tau > 1$. In particular, the larger the value of $F(\eta_{ctr}, \eta_{surr})$ the greater is the degree of non-stationarity. Equivalently our non-stationary index can also be expressed as $F(\eta_{ctr}, \eta_{surr}) = \left(\frac{\eta_{ctr}}{\eta_{surr}}\right)^s$, where $s \geq 1$ (since s can be subsumed into the threshold τ).

For a given image patch under consideration, the center patch is densely sampled with $M \times M$ windows (for $M=3$ or 5). The resultant data vectors are then analyzed using Comon's algorithm [1] to obtain the ICA vectors characterizing the center patch. The running time for computing ICA vectors for a 32×32 center patch (for $M=3$) is less than 0.5 sec when simulated in Matlab using a 3GHz PC. Next the covariance matrices of the center and surround patches, with respect to these ICA vectors, are computed as explained above.

The image patch size we choose determines the scale of the analysis. Table 1 shows non-stationarity indices computed when the image patch size is 256×256 (and the center patch is 128×128) for $M=5$ and $s=4$. We observe that the non-stationary index accurately discovers the degree of non stationarity for all the cases; i.e. near unity values for visibly stationary patches and higher values for visibly non-stationary patches. However, for natural images captured under normal viewing conditions, this scale of analysis is not of much practical value since image patches will invariably contain non-stationarities. Therefore when analyzing natural images, we used (as mentioned earlier) 64×64 scenes patch sizes, with center patch size of 32×32 ($M=3$, and $s=2$), for 1024×1024 images obtained from the van Hateren database of natural im-

Name	Image	Non-st Idx
Grass		1.0972
Water		1.0803
Tree		4.2481
Chem plant		2.1692

Table 1. Non-stationarity index for sample image segments. Non-stationarity index proportional to the degree of non-stationarity. Low values of non-stationary index (close to unity) correspond to stationary patches (like grass and water textures above); higher values for non-stationary patches.

ages [5]. When analyzing specific textures, we used images from the Brodatz collection [6]. All the non-stationarity maps presented below are dense, i.e. a non-stationarity measure is assigned to every point in the image. A non-stationarity index assigned to a pixel location in the image is associated with the 64×64 patch which has this pixel location as its center. In order to reduce computation time, a sub-sampled non-stationarity map (with a sub-sample factor of two in each dimension) is computed followed by linear interpolation. The image is periodically extended along the borders so that the non-stationarity index can be computed even at the boundaries of the image. For display purposes, all the non-stationarity maps below have been non-linearly mapped by a square-root operation. This of course does not effect the inherent dynamic range of the non-stationarity maps.

Fig.1(a) shows an image containing three texture regions. Fig.1(b) shows the non-stationarity map for this image using our non-stationarity index. We observe that even although variations in the non-stationarity index are present within a given texture region due to randomness and possibly residual non-stationarities (at the scale being analyzed), higher values of the index are typically observed near the boundaries between the textures. As we see in Fig.1(b), in many cases the peak non-stationarity may not occur at the exact location of the transition between the two textures but rather may occur at locations in the vicinity of it (provided the patch window intersects the boundary between the textures as it does in Fig.1(b)).

Fig.2(a) shows a natural image and Fig.2(b) shows its corresponding dense non-stationarity map. We observe that the values of the non-stationarity indices for patches in the shrub region have similar values (close to unity) since they are a part

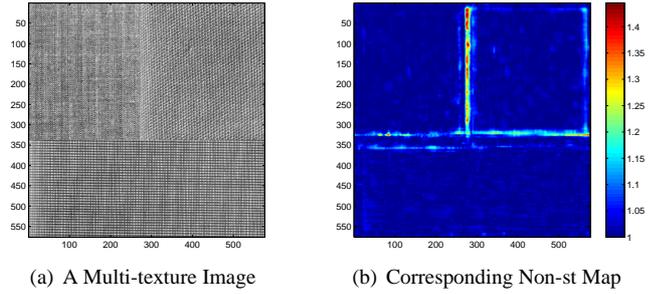


Fig. 1. A Multitexture image is shown in Fig.1(a) composed of textures D16(top-left), D77(top-right), and D21(bottom) [6]. Fig.1(b) shows the corresponding non-stationarity map.

of the same texture. Likewise, the image patches of Fig.2(b) that exhibit higher values of non-stationarity index do indeed correspond to regions that are non-stationary in the sense described in Section 2; for example, the region at the intersection of the tree and the building, at intersection of the shrub and house regions, the various space-varying structures in the house region etc. Thus the non-stationarity map in Fig.2(b) gives us a quantitative way of determining the non-stationary structure of natural images—the knowledge of which can be exploited by higher-level vision algorithms.

Fig.3(a) shows yet another image obtained from [5] for which the non-stationarity map is shown in Fig.3(b). In this case too we observe that the significant non-stationary points in the image—for example the regions at the boundary of the objects and the background texture—are detected by the non-stationarity index; and that roughly stationary regions of the image, like the shrub region, are mapped to smaller values of the non-stationary measure. However we also observe that for many 'roughly stationary' patches—for example on closer examination of parts of the texture located in the interior part of the object on the right hand side of the image—the non-stationarity index is amplified compared to the other stationary patches in the image. Our experiments have shown that this phenomenon occurs only for textural regions whose spatial interactions extend to larger spatial dimensions than the scale at which it is being analyzed (and that it *excludes* regular textures like sinusoidal patterns etc). For example, when water texture is analyzed at a large spatial scale as in Table 1, we find that it is stationary; but the same texture when analyzed at smaller scales, like 64×64 as shown in Fig.4, tends to have higher values for the non-stationarity index on an average (unlike the grass texture which is stationary at both scales). While such a phenomenon could potentially be fruitfully exploited by higher-level vision algorithms engaged in automatic scale detection, it is preferable to decouple low-level vision algorithms from this high-level scale information. One way of achieving these objectives would be to perform a multi-scale analysis of the non-stationary image i.e. to compute the non-stationarity maps of the image for different win-

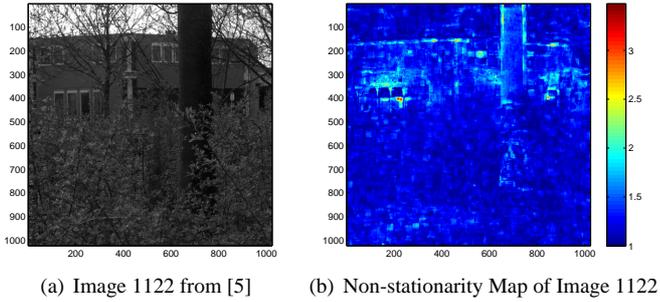


Fig. 2. Non-stationarity map of image#1122 taken from [5]

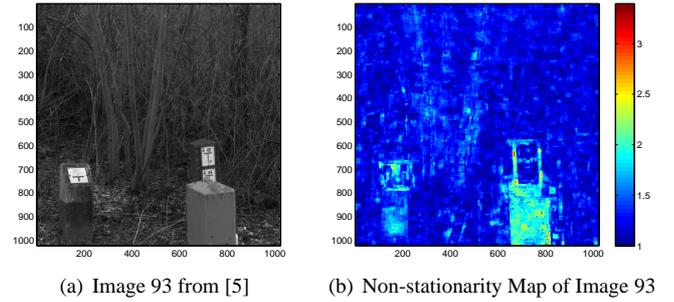


Fig. 3. Non-stationarity map of image#93 taken from [5]

dow sizes—this information when combined across the different scales can then provide a richer description of the non-stationary structure of natural images.

4. DISCUSSION

The structure of natural scenes is the result of complicated non-linear interactions of spatially stationary processes called textures (superimposed with other variations like luminosity profiles). To a first-order analysis, these non-linear interactions (that include partitioning of image into textures—which we assume throughout in this paper—or more complicated occlusion phenomena) give rise to non-stationarities in image. Thus the ability to recognize 'roughly stationary' regions in the image is central to elucidating the structure of natural images and this forms a basis for further higher level visual processing.

To this end we have presented, in this paper, a novel statistical measure called the non-stationarity index which we have shown to be effective in quantitatively determining the extent of non-stationarity across the image being analyzed. This information can be subsequently utilized for a variety of image processing and computer vision applications. One example that we single out is fixation selection in foveated visual systems based on textural information.

While previous methods of low-level fixation selection are based on cues like luminance and contrast [7], our non-stationarity index serves as a natural cue for texture based fixations for the simple reason that regions of maximum non-stationarity are precisely where maximum computational resources will have to be expended to perform detailed texture-based segmentation and pattern recognition operations, given that minimization of total fixations is the objective. Given a foveated luminance image, once defoveation is performed by means of scale-varying deblurring [8], the subsequently computed non stationarity map serves as a saliency map that can be used to determine significant fixation points within the image.

More generally, it is plausible that the HVS combines information about the non-stationarity structure of the image

together other low-level cues like contrast in order to guide visual fixations that are otherwise modulated by the needs of more specific (high-level) tasks, demands or by particularly significant high level content extracted during the fixations.

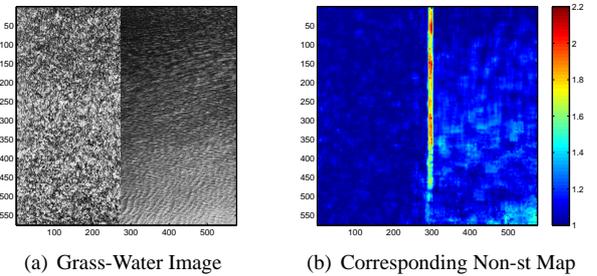


Fig. 4. A Multitexture image is shown in Fig.4(a) composed of Grass texture(Left), and Water Texture(right). Fig.4(b) shows the corresponding non-stationarity map. At this scale of analysis (i.e. 64x64) Water texture has higher non-stationarity index on an average compared to the Grass texture.

5. REFERENCES

- [1] P. Comon, "Independent Component Analysis: A New Concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [2] J-F. Cardoso, "Blind Source Separation: Statistical Principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct. 1998.
- [3] R.G. Raj and A.C. Bovik, "The Multilinear ICA Decomposition with Applications to NSS Modeling," *IEEE ICASSP 2007*, vol. 2, pp. 669–672, Apr. 2007.
- [4] T-W Lee and M.S. Lewicki, "Unsupervised Image Classification, Segmentation using ICA Mixture Models," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 270–279, Mar. 2002.
- [5] J.H. van Hateren and A. van-der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proceedings of the Royal Society of London B*, vol. 265, no. 1394, pp. 359–366, Mar. 1998.
- [6] "http://www.ux.uis.no/tranden/brodatz.html," .
- [7] R.G. Raj, W.S. Geisler, R.A. Frazor, and A.C. Bovik, "Contrast statistics for foveated visual systems: fixation selection by minimizing contrast entropy," *JOSA A*, vol. 22, no. 10, pp. 2039–2049, Oct. 2005.
- [8] A.C. Bovik and R.G. Raj, "Approximating Filtered Scale-Variant Signals," *IEEE Transactions on Image Processing*, vol. 14, no. 1, pp. 23–35, Jan. 2005.