A Probabilistic Definition of Intrinsic Dimensionality for Images

Michael Felsberg[1]* and Norbert Krüger[2]

[1] Computer Vision Laboratory, Dept. of Electrical Engineering, Linköping University SE-58183 Linköping, Sweden, mfe@isy.liu.se
[2] Computational Neuroscience, Department of Psychology, University of Stirling Stirling FK9 4LA Scotland, UK, norbert@cn.stir.ac.uk

Abstract. In this paper we address the problem of appropriately representing the intrinsic dimensionality of image neighborhoods. This dimensionality describes the degrees of freedom of a local image patch and it gives rise to some of the most often applied corner and edge detectors. It is common to categorize the intrinsic dimensionality (iD) to three distinct cases: i0D, i1D, and i2D. Real images however contain combinations of all three dimensionalities which has to be taken into account by a continuous representation. Based on considerations of the structure tensor, we derive a cone-shaped iD-space which leads to a probabilistic point of view to the estimation of intrinsic dimensionality.

1 Introduction

The aim of this paper is to develop a representation of the intrinsic dimensionality which is well suited for further probabilistic processing. The *intrinsic dimensionality* (iD) is a well known concept from statistics which can be defined as follows: "a data set in d dimensions is said to have an *intrinsic dimensionality* equal to d' if the data lies entirely within a d'-dimensional subspace" [1], p. 314. The term itself goes back to the late sixties [2]. The intrinsic dimensionality was introduced to image processing by Zetsche and Barth [3]. It is obtained by applying the previous definition to the spectrum of an image patch, i.e., the Fourier transform of a neighborhood. The three possible intrinsic dimensionalities in images are defined according to their local spectrum [4] (see also Fig. 1):
i0D – It is concentrated in the origin, i.e., the neighborhood is constant.
i1D – It is concentrated in a line through the origin, i.e., the neighborhood is varying in only one direction. These signals are also called *simple signals* [5].
i2D – It is neither concentrated in the origin, nor in a line.

Typical examples for i1D neighborhoods are edges, lines, sinusoids, whereas corners, junctions, line ends, spots are instances of i2D neighborhoods. As soon as we take noise into account, the discrete definition above becomes useless. Noise is i2D and every signal contains noise. Hence, every image neighborhood is i2D. But how to distinguish between noise and i2D image structures?

* This work has been supported by DFG Grant FE 583/1-2.

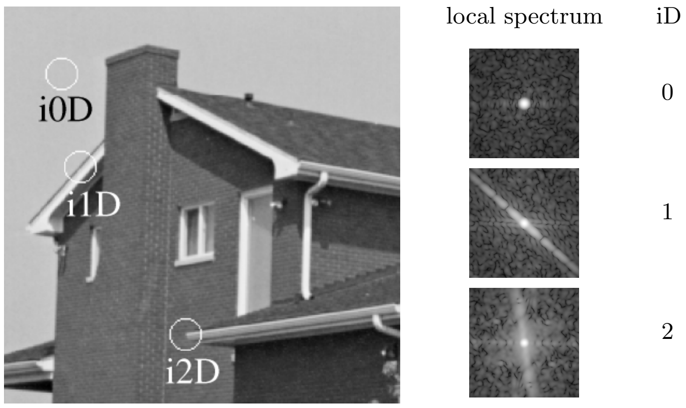

Fig. 1. Illustration intrinsic dimensionality. In the image on the left, three neighborhoods with different intrinsic dimensionalities are indicated. The other three images show the local spectra of these neighborhoods.

In recent years, there have been several attempts to define image processing operators which detect the intrinsic dimensionality in images. Note that as long as the intrinsic dimensionality is considered to be a discrete choice from the set $\{\mathrm{i0D}, \mathrm{i1D}, \mathrm{i2D}\}$, it is more appropriate to speak of *detection* rather than *estimation.* By switching from a discrete choice to a continuous model for the intrinsic dimensionality, we also switch the terminology from "detection" to "estimation". Looking at the examples for i2D patches, evidently every corner detector is a detector for i2D neighborhoods. Considering image patches at an appropriate scale, all line and edge detectors are detectors for i1D neighborhoods. Besides these two popular fields of image processing, there are other approaches to detect or measure the intrinsic dimensionality, e.g., by Volterra operators [4], tensor methods [6, 5], and generalized quadrature filters [7, 8].

Since the structure tensor is probably the most well known approach among these, our paper is based on an analysis of the latter approach. From this analysis we derive a new *continuous*, *triangular* representation of intrinsic dimensionality, making use of *barycentric coordinates.* These coordinates can be interpreted as *confidences* of the measurements or the *likelihood* that the measurement is correct. Hence, they can be used as a prior for further probabilistic processing. The new contribution of this paper compared to [5], p. 253, is the introduction of a coefficient for the i0D case, such that the coefficients add up to one and can therefore be interpreted as probabilities.

In a second step, we introduce orientation information to our model, resulting in a *cone shaped* geometry. This model allows us to average the representation while treating i1D structures with different orientations in an appropriate way, i.e., two i1D neighborhoods with different orientations result in an i2D measurement. This extension of the model makes it independent of the structure tensor: No i2D information is necessary for our model, a simple gradient estimation or quadrature filter response is sufficient. Furthermore, we switch from a deterministic preprocessing to a probabilistic estimation of the barycentric coordinates.

2 A Continuous Definition of Intrinsic Dimensionality

The structure tensor is an approach for the local analysis of images that was first proposed in 1987 [9, 10]. For our considerations, however, the derivation in [11] is most appropriate: The structure tensor can be considered as a local approximation of the auto-covariance function in the origin.

The structure tensor is typically interpreted in terms of its eigensystem. Its two eigenvalues correspond to the maximum and minimum 1D frequency spread[3] in the neighborhood of $\mathbf{x}$, i.e.,

$$\lambda_1 \sim \max_{\mathbf{e}_1} \int (\mathbf{e}_1 \cdot \mathbf{u})^2 |F_{\mathbf{x}}(\mathbf{u})|^2 \, d\mathbf{u} \quad \text{and} \quad \lambda_2 \sim \min_{\mathbf{e}_2} \int (\mathbf{e}_2 \cdot \mathbf{u})^2 |F_{\mathbf{x}}(\mathbf{u})|^2 \, d\mathbf{u} \ , \tag{1}$$

where $\mathbf{u} = (u, v)^T$ is the frequency vector and $F_{\mathbf{x}}(\mathbf{u})$ is the local spectrum. The two vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are perpendicular and $\mathbf{e}_1$ represents the main orientation of the structure. In practice, the structure tensor is typically computed by averaging the outer product of the image gradient:[4]

$$\mathbf{J}(\mathbf{x}) = \int_{\mathcal{N}(\mathbf{x})} (\nabla f(\mathbf{x}'))(\nabla f(\mathbf{x}'))^T w(\mathbf{x} - \mathbf{x}') \, d\mathbf{x}' \ , \tag{2}$$

where $w(\cdot)$ is some weighting function for the neighborhood $\mathcal{N}$ and $\nabla f(\cdot)$ is the image gradient. According to the power theorem [14] and the derivative theorem of the Fourier transform, the tensor $\mathbf{J}$ is proportional to the second moment tensor of the local Fourier spectrum, i.e.,

$$\mathbf{J} \sim \int \mathbf{u}\mathbf{u}^T |F_{\mathbf{x}}(\mathbf{u})|^2 \, d\mathbf{u} \ , \tag{3}$$

such that the eigenvalues of $\mathbf{J}$ are consistent with (1).

A classical technique for estimating the intrinsic dimensionality is to consider the rank of the structure tensor. Theoretically, the number of non-zero eigenvalues corresponds to the rank of the tensor, and therefore, to the intrinsic dimensionality of the neighborhood. In practice the eigenvalues are never zero due to noise and a commonly applied method is to *threshold* the eigenvalues [6]. This approach leads to a discrete categorization of neighborhoods according to their intrinsic dimensionality.

Indeed, it is not only noise that disturbs the evaluation of the rank, but most neighborhoods in real signals consist of combinations of i0D, i1D, and i2D signals. Hence, it is more appropriate to think of the intrinsic dimensionality as a *continuous* measure rather than a discrete set of cases. In order to define a continuous measure, it is necessary to define the *topology* of the measurement

---

[3] The frequency spread is obtained by considering the variance of the squared amplitude response [12].

[4] The are other ways to compute the structure tensor, e.g., polar separable quadrature filter [5] or polynomial expansions [13], but the gradient based approach is best suited in the context of this paper.

space. For the intrinsic dimensionality we observe that the measurement space *cannot be 1D*, since each of the intrinsic dimensionalities is adjacent to the other two. *The intrinsic dimensionality space is thus 2D.*

One approach which at least partially realizes a continuous measure is the *coherence* [6], which takes values in the interval $[0, 1]$ depending on the quotient[5]

$$c = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \quad . \tag{4}$$

The coherence is one for ideal i1D neighborhoods and tends to zero for isotropic structures, i.e., it represents the *confidence* for the presence of an i1D structure. However, the coherence does not take the i0D case into account. In the latter case, we meet a singularity for $A = \lambda_1 + \lambda_2 = 0$ (see Fig. 2, left). Therefore, the coherence is mostly combined with a threshold of A, i.e., the coherence approach is a mixture of a continuous model (i1D–i2D) and a discrete model (i0D–i1D/i2D). Notice that one of the most popular corner detectors, the Harris-Stephens detector [15], is based on the coherence measure.

Another approach to realize a continuous measure is proposed in [5], p. 253, where the tensor is decomposed into a linear combination of tensors with different (non-zero) rank. For images, the linear coefficients are given by $\lambda_1 - \lambda_2$ (i1D) and λ_2 (i2D). Considering these coefficients in an orthonormal basis[6] yields a $\pi/4$ sector (see Fig. 2, center). The same sector shaped space is obtained by multiplying the coherence by A. The center point of the sector corresponds to an i0D neighborhood, one edge corresponds to i1D neighborhoods and the other one corresponds to i2D neighborhoods. Due to the triangular structure, we avoid the problem of evaluating the coherence for $A = 0$. However, the problem with the i0D case is still present, since there is no upper bound of the coordinates. Without having an upper bound, i.e., without having a finite interval for the coordinates, we cannot normalize the coordinates. Hence, we cannot represent confidences for all three cases, and *no probabilistic interpretation* is possible.

In order to map the measured values to finite intervals, we apply a popular technique which is called *soft thresholding*, see e.g. [1, 16]. The value A is transformed by a non-linear function:

$$A' = \arctan(a \log(A) + d)/\pi + 1/2 \ \in [0, 1] \quad , \tag{5}$$

where a and d are real constants. This transformation results in a finite area for all possible measurements of the intrinsic dimensionality: a *triangle*, see Fig. 2, right, with the coordinates

$$(x, y)^T = (A', A'c)^T \ \in [0, 1] \times [0, 1] \quad \text{and } y \leq x \ . \tag{6}$$

The i0D case corresponds to the coordinates $(0, 0)$, the i1D case to $(1, 1)$, and the i2D case to $(1, 0)$. The coordinates (x, y) can easily be transformed to *barycentric*

---

[5] In [6] the coherence is defined as the square of the expression in (4).

[6] The decomposition of a 2D tensor into a rank one tensor and an isotropic tensor is not an orthogonal decomposition, since the angle between a rank one tensor and the identity is $\pi/4$.

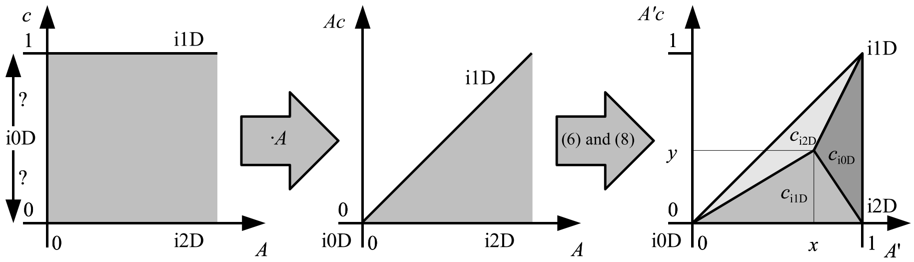

Fig. 2. About the topology of iD-space. Left: the coherence leads to an infinite stripe of width one; center: the tensor decomposition leads to a sector, right: soft thresholding yields a triangle which can be parameterized by barycentric coordinates.

coordinates [17], yielding three coordinates $(c_{\mathrm{i0D}}, c_{\mathrm{i1D}}, c_{\mathrm{i2D}})$ with $c_{\mathrm{i}k\mathrm{D}} \in [0,1]$ for $k = 0,1,2$ and $\sum_{k=0}^{2} c_{\mathrm{i}k\mathrm{D}} = 1$, i.e., the barycentric coordinates can be interpreted as *likelihoods* or *confidences*. The barycentric coordinates correspond to the areas of the opposite triangles (see Fig. 2, right) and are obtained by the formulas

$$c_{\mathrm{i0D}} = 1 - x \qquad c_{\mathrm{i1D}} = y \qquad c_{\mathrm{i2D}} = x - y \ . \tag{7}$$

Note that although the new representation of the intrinsic dimensionality has been derived from considerations of the structure tensor, it is not necessarily based on its eigenvalues. Any preprocessing method yielding two measurements, which in some way represent the isotropic and the directed part of a signal neighborhood, can be used for building up the triangle representation. Other examples besides the structure tensor are the generalized quadrature filter in [8], a combination of the Canny edge detector [18] and the Harris-Stephens corner detector [15] (both without thresholding), or a combination of local amplitude and local orientation variance [19].

3 The Intrinsic Dimensionality Cone

If we want to process the intrinsic dimensionality information further, we must assure that the representation is *consistent*, i.e., averaging of the representation for two (adjacent) neighborhoods should result in a proper representation for the joint neighborhood. Considering the different possible combinations, one problematic case pops up if two i1D neighborhoods with *different orientations* are considered. The averaged intrinsic dimensionality according to the representation defined above is again i1D, which is wrong. Two i1D neighborhoods with different orientations should give a decreased i1D likelihood and an increased i2D likelihood, depending on the orientation difference, see Fig. 3, right.

Hence, the triangle representation must be modified to be consistent. The required modification has already been mentioned implicitly in terms of linear combination of tensors. A rank one tensor does not only include information about the eigenvalue, but also about the eigenvector, i.e., the local orientation. The rank one tensor represents the orientation in double angle, which is appropriate for averaging orientation information [5]. The triangle representation

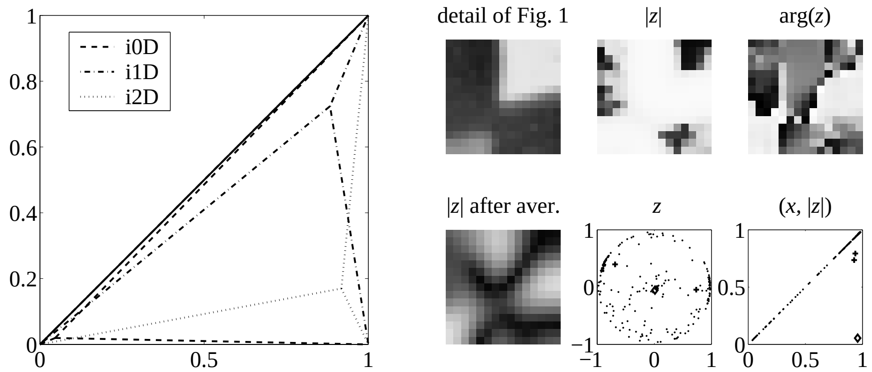

Fig. 3. Left: The triangle representation of the intrinsic dimensionality of the three selected points from Fig. 1. The estimated likelihoods are: 0.95 for the i0D case, 0.72 for the i1D case, and 0.75 for the i2D case. Right: averaging an appropriate i1D representation results in a high i2D likelihood at corners. Bottom row, center and right: histogram representations. The dots indicate the measurements before the averaging, the diamond indicates the estimate at the corner after the averaging, and the pluses indicate the estimates at the edges (five pixels from the corner) after the averaging.

is now modified by multiplying the y-coordinate by the complex double angle representation:

$$z = y \exp(i2\theta) \ , \tag{8}$$

where θ represents the local orientation, see Fig. 4. This modification leads to a *cone-shaped* space for the intrinsic dimensionality, the *iD-cone*.

Measurements for the intrinsic dimensionality are now combined with orientation information and are represented by coordinates inside the iD-cone. Each point in the iD-cone corresponds to a set of local image structures with the same orientation and the same intrinsic dimensionality. Averaging the cone coordinates over some neighborhood leads to a consistent estimate of the intrinsic dimensionality in that neighborhood. If the confidences for the different iD cases are required, we simply evaluate the barycentric coordinates in the plane given by the complex argument, i.e., y is replaced with $|z|$ in (7).

In the cone model, two i1D structures with different orientations give rise to an increased likelihood of a i2D structure. This observation implies that it is not necessary to extract i2D information in the preprocessing, i.e., we do not need coherence information. It is sufficient to estimate the orientation and the intensity, to represent this information in cone coordinates (i.e., on the cone surface), and to apply a local averaging. The i2D information then drops out as a result of the averaging process.[7] The preprocessing can be performed by a simple gradient estimation, by the Riesz transform, or by a spherical quadrature filter [20] which was used to produce the results in Fig. 3.

[7] This behavior is similar for the averaging of the outer product of gradients for computing the structure tensor [6].

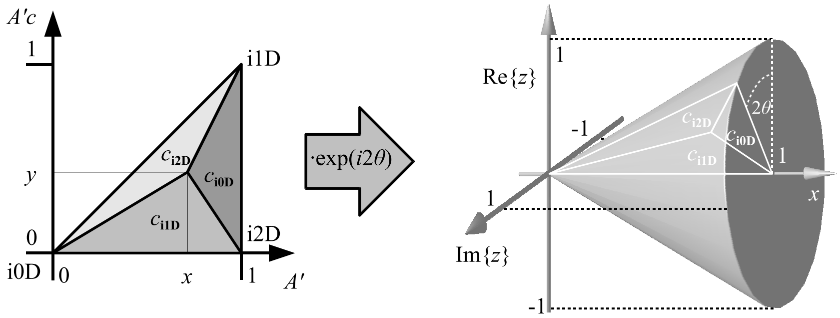

Fig. 4. Modification of the iD-triangle: multiplying y by $\exp(i2\theta)$ yields an iD-cone.

Except for the preprocessing, the described approach for estimating the intrinsic dimensionality contains only three free parameters: the constants a and d in (5) and the size of the neighborhood for the local averaging. In our experiments we set $a = 5$, $d = 0$, and the local averaging is obtained by Gaussian smoothing with variance 2, but the method is very stable with respect to changes in the parameters. Due to the simple analytic description of the approach, it is straightforward to optimize the free parameters for certain applications, e.g., in a similar way as it is done in the back-propagation algorithm [1]. From this point of view and since the barycentric coordinates are obtained from an averaging process, the iD-cone model can be considered as a probabilistic approach. In a larger system the estimated likelihoods for the different intrinsic dimensionalities can be used as priors for the subsequent processing steps. The new model can be applied in a wide range of applications, e.g., corner detection [8][8], edge detection, curve fitting, and segmentation. The introduced definition of intrinsic dimensionality has also been applied as a descriptor for the 'edgeness' or 'junctioness' of local image patches in a new kind of multi-modal image representations [21].

4 Conclusion

In this paper we have derived a continuous representation of the intrinsic dimensionality of images. Although our considerations are based on the structure tensor approach, our model is independent of a specific preprocessing method, which need not even contain i2D information. The derived finite cone-shaped iD-space is easily interpreted in terms of a probabilistic representation by using barycentric coordinates. The model contains only few parameters, which might be obtained by a learning method, depending on a specific application.

Acknowledgment

We like to thank Klas Nordberg for all the discussions on tensor representations.

[8] To be precise, we did not make explicit use of the here proposed model in [8], but the applied algorithm is compatible the iD triangle.

References

1. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, New York (1995)
2. Trunk, G.V.: Representation and analysis of signals: statistical estimation of intrinsic dimensionality and parameter identification. General System **13** (1968) 49–76
3. Zetzsche, C., Barth, E.: Fundamental limits of linear filters in the visual processing of two dimensional signals. Vision Research **30** (1990)
4. Krieger, G., Zetzsche, C.: Nonlinear image operators for the evaluation of local intrinsic dimensionality. IEEE Transactions on Image Processing **5** (1996) 1026–1041
5. Granlund, G.H., Knutsson, H.: Signal Processing for Computer Vision. Kluwer Academic Publishers, Dordrecht (1995)
6. Jähne, B.: Digitale Bildverarbeitung. Springer, Berlin (1997)
7. Bülow, T.: Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images. PhD thesis, Christian-Albrechts-University of Kiel (1999)
8. Felsberg, M., Sommer, G.: Image features based on a new approach to 2D rotation invariant quadrature filters. In Heyden, A., Sparr, G., Nielsen, M., Johansen, P., eds.: Computer Vision - ECCV 2002. Volume 2350 of Lecture Notes in Computer Science., Springer (2002) 369–383
9. Förstner, W., Gülch, E.: A fast operator for detection and precise location of distinct points, corners and centres of circular features. In: ISPRS Intercommission Workshop, Interlaken. (1987) 149–155
10. Bigün, J., Granlund, G.H.: Optimal orientation detection of linear symmetry. In: Proceedings of the IEEE First International Conference on Computer Vision, London, Great Britain (1987) 433–438.
11. Förstner, W.: Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung. Number 370 in C. Verlag der Bayerischen Akademie der Wissenschaften (1991)
12. Ramanathan, J.: Methods of Applied Fourier Analysis. Birkhäuser (1998)
13. Farnebäck, G.: Fast and accurate motion estimation using orientation tensors and parametric motion models. In: Proceedings of 15th International Conference on Pattern Recognition. Volume 1., Barcelona, Spain, IAPR (2000) 135–139
14. Bracewell, R.N.: The Fourier transform and its applications. McGraw Hill (1986)
15. Harris, C.G., Stephens, M.: A combined corner and edge detector. In: 4th Alvey Vision Conference. (1988) 147–151
16. Krüger, N.: Learning object representations using a priori constraints within ORASSYLL. Neural Computation **13** (2001) 389–410
17. Coexeter, H.S.M.: Introduction to Geometry. 2nd edn. Wiley & Sons (1969)
18. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence **8** (1986) 679–698
19. Krüger, N., Felsberg, M.: A continuous formulation of intrinsic dimension. In: British Machine Vision Conference. (2003) submitted.
20. Felsberg, M.: Low-Level Image Processing with the Structure Multivector. PhD thesis, Institute of Computer Science and Applied Mathematics, Christian-Albrechts-University of Kiel (2002) TR no. 0203, available at `http://www.informatik.uni-kiel.de/reports/2002/0203.html`.
21. Krüger, N., Lappe, M., Wörgötter, F.: Biologically motivated multi-modal processing of visual primitives. 'AISB 2003 Convention: Cognition in Machines and Animals', Wales (2003)