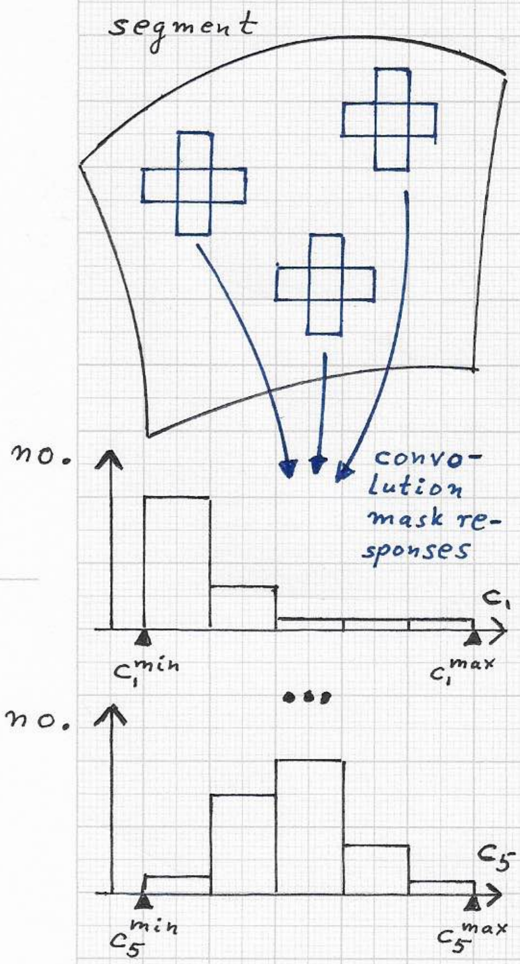
■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

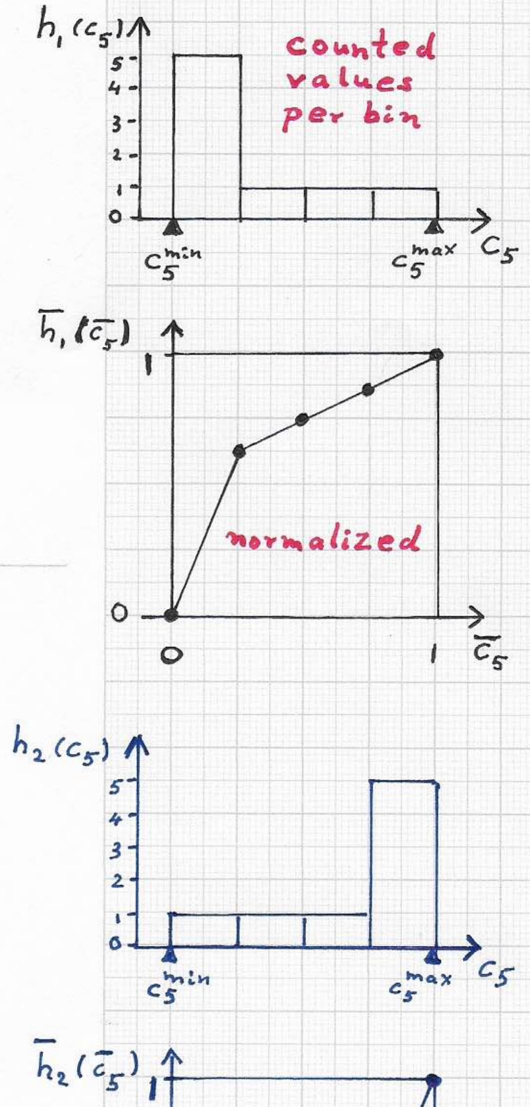
•Laplacian eigenfunctions: ... The definition of this difference tuple dI_I and its length $\|dI_I\|$ makes possible the pairwise comparison of band-specific behavior of functions. The application of a 5-pixel mask to a segment consisting of many pixels yields many (convolution response) values for each c_i-coefficient ($i=1,\ldots,5$). Thus, for c_1, c_2, c_3, c_4 and c_5 one obtains an entire distribution/histogram of values. In other words, THE MULTI-SCALE EIGENFREQUENCY-BASED SIGNATURE OF A SEGMENT IS A SET OF DISTRIBUTIONS/HISTOGRAMS DESCRIBING HOW OFTEN CERTAIN VALUES OF **COEFFICIENTS c_i** OF **LOCAL SEGMENT** EXPANSIONS $\sum_i c_i e_i$ OCCUR. One must keep in mind that two segments, usually consisting of different numbers of pixels (voxels), generate different numbers of values for each coefficient c_i. Therefore, to compare quantitatively the value distributions of a coefficient c_i, one must define and calculate a proper difference value for two discrete distributions. ...



Statistical approach to recording responses of a 5-pixel convolution mask when applied to a material segment. Response values generated by the mask for coefficient c_i lie between c_i^{min} and c_i^{max}. Using a specific number of "bins" to subdivide this interval produces histograms that capture the distributions of c_i-values.

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

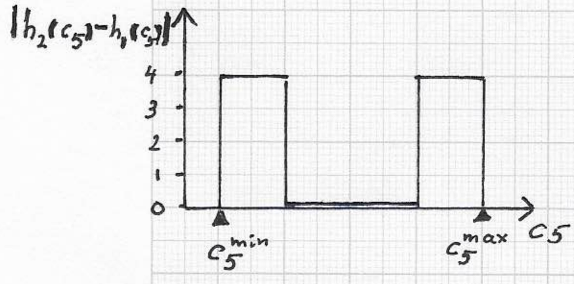
• Laplacian eigenfunctions: … In order to characterize, compare and eventually classify multiple segments of arbitrary geometry and consisting of different numbers of pixels (voxels) in a multi-scale fashion, one should employ a statistical approach. For example, when applying the **5**-pixel convolution mask to all possible **5**-pixel subsets of all the pixels of a segment, the mask returns values for the coefficients $c_1, c_2, \ldots, c_5$, where the values lie in certain intervals $[c_i^{min}, c_i^{max}]$. These intervals could possibly defined a prior for a specific type of image and material, or one could determine the intervals experimentally. Recorded c_i-values can be used to establish discrete histograms, based on a defined number of bins for c_i-values. On a high level, the histogram for values of coefficient c_i determines whether **its** associated eigenfunction is "present" in a segment's expansion and, **if so**, "at what level of strength" the eigenfunction is "expressed" in the segment. …



Sketch of two distributions of c_5-values, $h_1(c_5)$ and $h_2(c_5)$, and associated cumulative distributions.

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions: ... Characterizing a segment via histograms of calculated c_i-values requires us to employ appropriate measures for differences/distances between the histograms of c_i-values resulting for two segments. Part 34 of these notes (pp. 3-7, Nov. 18-21, 2019) discusses distribution functions, cumulative distribution functions and the Wasserstein distance; they are important for our discussion of segment comparison when segments are characterized by histograms of coefficient values c_i of set of local multi-scale eigenfunction expansions. On the previous page distribution functions $h(c_5)$ and cumulative distribution functions $\overline{h}(\overline{c_5})$, normalized, are shown for two segments with significantly different c_5-value distributions. For example, assume that B bins are used for two c_5-value histograms for two segments, with all values lying between c_5^{min} and c_5^{max}. Further, the first segment produces N_1 c_5-values and segment two produces N_2 c_5-values. For comparison one must properly normalize the binned data. The computations are summarized (left). ...



Absolute value of difference function of two "histogram functions" h_1 and h_2 – for values of coefficient c_5. **Note:** Generally, two segments generate different total numbers of c_5-values, and histogram data also must be normalized.

• **Histograms = vector data:**

$\mathbb{h}_1(c_5) = (b_1', \ldots, b_B')^T$

B = no. of bins

$N_1 = b_1' + \ldots + b_B'$

b_i' = no. of values in bin i

$\mathbb{h}_1 = \mathbb{h}_1(c_5) / N_1$

... $\mathbb{h}_2 = \mathbb{h}_2(c_5) / N_2$

⇛ **Difference vector:**

$$\mathbb{d} = \mathbb{h}_2 - \mathbb{h}_1$$

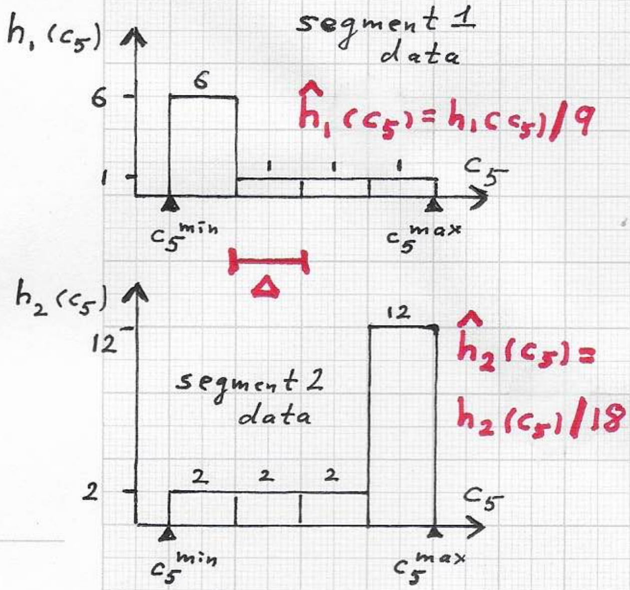
$$\| \mathbb{d} \| = (\langle \mathbb{d}, \mathbb{d} \rangle)^{1/2}$$

⇛ **Histogram difference:**

$$d = \| \mathbb{d} \|$$

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions: ...



Example with 4 bins:

$$\mathbb{h}_1(c_5) = (6, 1, 1, 1)^T$$

$$N_1 = 9$$

$$\mathbb{h}_1 = (2/3, 1/9, 1/9, 1/9)^T$$

$$\mathbb{h}_2(c_5) = (2, 2, 2, 12)^T$$

$$N_2 = 18$$

$$\mathbb{h}_2 = (1/9, 1/9, 1/9, 2/3)^T$$

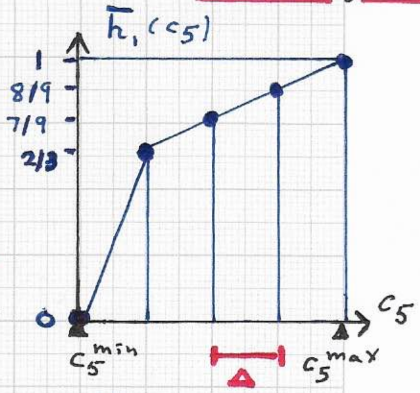
$$\Rightarrow \mathbb{d} = (-5/9, 0, 0, 5/9)^T$$

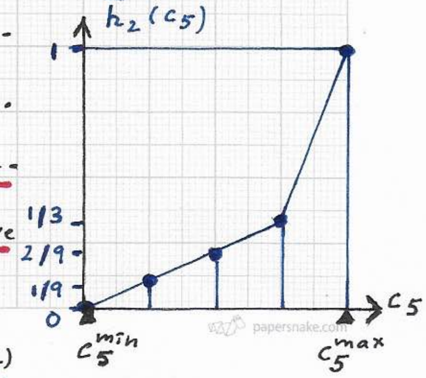
$$\|\mathbb{d}\| = (50/81)^{1/2} = 5/9\sqrt{2}$$

$$\Rightarrow \underline{d = 5/9\sqrt{2}}$$

or: $$d = \left(\int_{c_5^{min}}^{c_5^{max}} \left(\hat{h}_2(c_5) - \hat{h}_1(c_5) \right)^2 dc_5 \right)^{1/2} = \underline{5/9\sqrt{2\Delta}}$$

We discuss a simple numerical example for histogram data from two segments. Values recorded for coefficient c_5 lie between c_5^{min} and c_5^{max}, and this value range is subdivided into four bins of equal width Δ. Since the two segments produce different numbers of c_5-values — nine and 18, respectively — one must perform a proper normalization to make possible histogram distance computation that does not depend on the numbers nine and 18. The two histograms $h_1(c_5)$ and $h_2(c_5)$ are sketched here (left, top), and possible distance values (d) are provided, where the second value considers bin width Δ. Corresponding normilized cumulative distributions $\bar{h}_1(c_5)$ and $\bar{h}_2(c_5)$ are also shown (right).





...

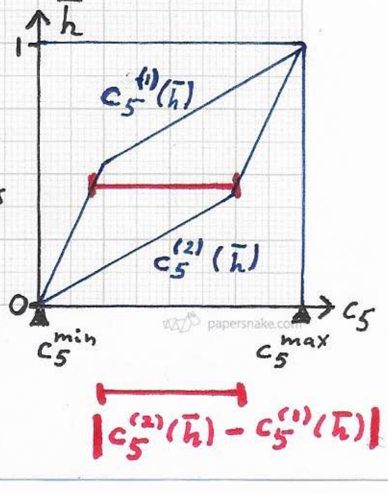
■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

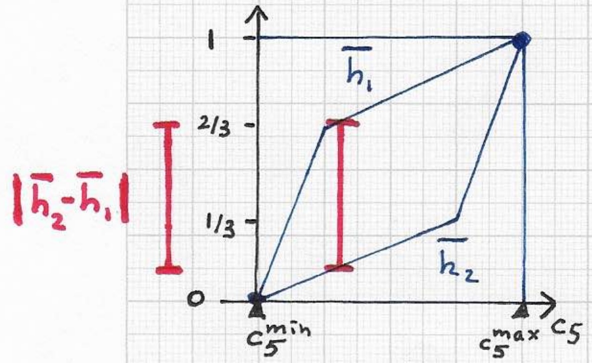
• Laplacian eigenfunctions: ... One can also define a distance measure based on the normalized cumulative distribution functions $\bar{h}_1(c_5)$ and $\bar{h}_2(c_5)$ over their shared domain interval $[c_5^{min}, c_5^{max}]$. A general form of such a distance is

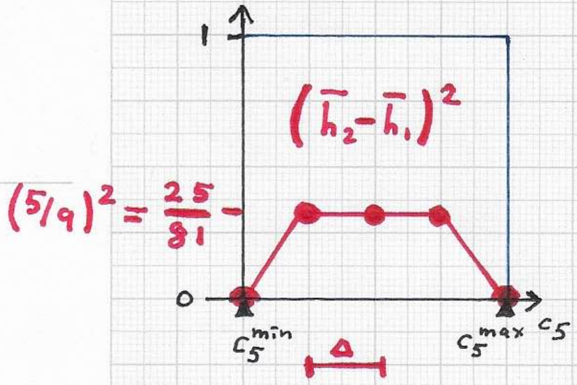
$$\bar{d} = \left(\int_\Omega (\bar{h}_1 - \bar{h}_2)^2 \, dc_5 \right)^{1/2} .$$

Depending on the needs of the specific classification problem, one should "experiment with" the meaningful options and eventually choose the distance measure leading to the best classification results. The Wasserstein distance is another measure one can consider. Here, distance is measured in c_5-direction, and one needs to use the inverse functions of $\bar{h}_1(c_5)$ and $\bar{h}_2(c_5)$, called $c_5^{(1)}(\bar{h})$ and $c_5^{(2)}(\bar{h})$, see the figure (right). The Wasserstein distance is

$$d_W = \int_0^1 \left| c_5^{(2)}(\bar{h}) - c_5^{(1)}(\bar{h}) \right| \, d\bar{h} .$$







Distance measure based on squared difference of cumulative distributions $\bar{h}_1(c_5)$ and $\bar{h}_2(c_5)$.

$$\bar{d} = \left(\int_{c_5^{min}}^{c_5^{max}} \left(\bar{h}_2(c_5) - \bar{h}_1(c_5) \right)^2 dc_5 \right)^{1/2}$$

$$= \left(3\Delta \, (5/9)^2 \right)^{1/2}$$

$$= 5/9 \sqrt{3\Delta}$$