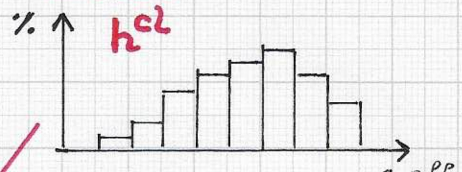
Bernd Hamann — -16- — 5/24/2022

Stratovan

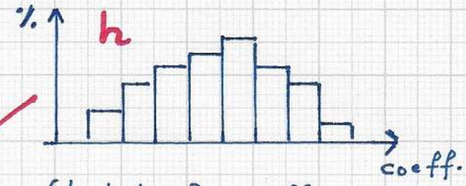
■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

Computation of p_i when scale i of a specific class is represented by only ONE class sample in the "database":



Sketch of a coeff. value histogram h^{cl} of ONE sample s, for scale i

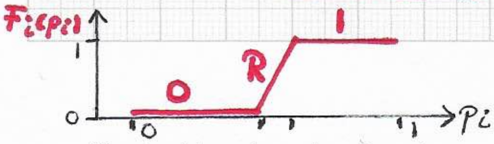


Sketch of a coeff. value histogram h of unclassified segment S, for scale i

By embedding h^{cl} and h in a vector space of finite or infinite dimension, one can compute a difference measure

$$\| h^{cl} - h \| ,$$

normalized to the interval $[0,1]$ that serves as the value for p_i, where minimal difference 0 is mapped to maximal similarity $p_i = 1$.



Possible decider function

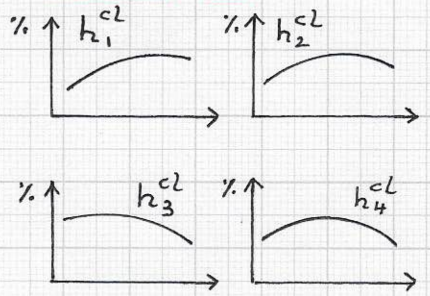
The discussion of individual univariate decider functions $F_i(p_i)$ did not address the necessary computation of the value of p_i in the desirable detailed manner. The discussion simply referred to p_i as a **measure of similarity (at scale i) of an unclassified segment S and a second, classified segment s serving as THE ONE sample image segment representing a specific material class – stored with all other sample segments in a "database."** When the "database" indeed only stores **EXACTLY ONE SAMPLE PER CLASS**, the discussion concerning the computation of a p_i-value is straightforward: At scale i, and for each class, one only needs to compute the coefficient value histogram difference for **A PAIR** of histograms – the one associated with S and the one associated with s, where s is the single, only representation of a specific class.

...

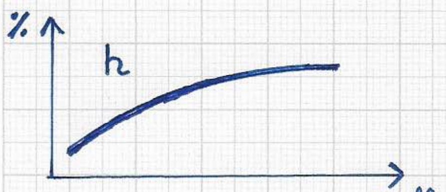
OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks:...

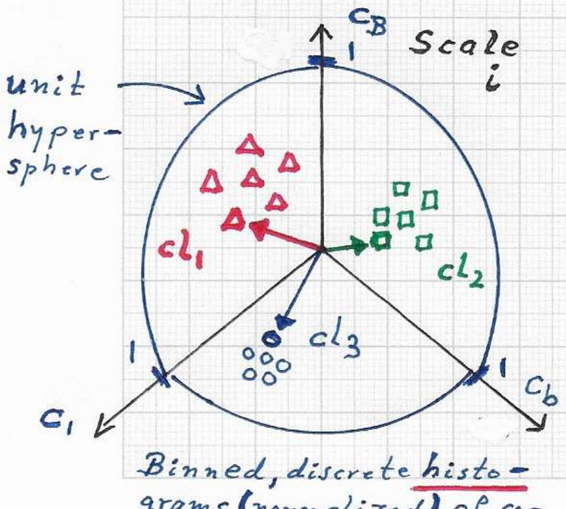
Computation of p_i when the "database" stores MORE THAN ONE class sample:



Sketch of 4 coefficient value histograms h_1^{cl}, ..., h_4^{cl} of 4 samples s_1, ..., s_4, for scale i.



Sketch of a coeff. value histogram h of unclassified segment S, for scale i.



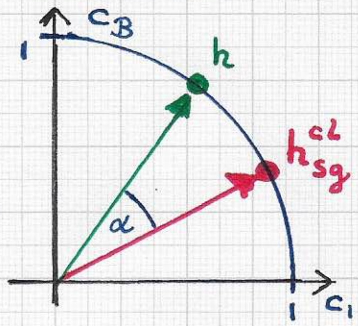
Binned, discrete histograms (normalized) of coefficient values of 6 samples of 3 classes can be viewed as points on a sphere.

Generally, MULTIPLE sample image segments, represented via their coefficient value histograms for H scales, are stored in the "database" for each class. Using the notation from the previous discussion, we have C classes ($cl_1, \ldots, cl_C$) with k_1 sample segments for class cl_1, ..., and k_C sample segments for class cl_C. IN A SETTING WHERE ONE HAS MORE THAN ONE SAMPLE SEGMENT FOR A CLASS, HOW DOES ONE / CAN ONE COMPUTE p_i-VALUES? The figures (Left) sketch the case where a class cl is represented in the sample "database" via the histogram data for four class samples. Thus, one can compute the histogram distance measures $\|h_{sg}^{cl} - h\|$, $sg = 1 \ldots 4$, for the unclassified segment S, for each scale i, $i = 1 \ldots H$. The problem to be addressed is the proper definition of a p_i-value from multiple $\|h_{sg}^{cl} - h\|$-values.

...

OBJECT AND MATERIAL EIGENFUNCTIONS — Cont'd.

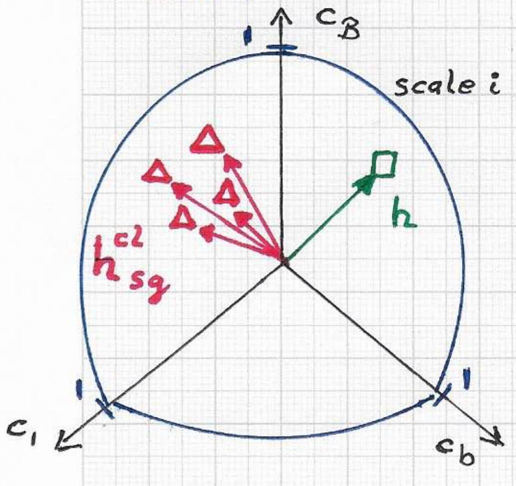
- Laplacian eigenfunctions and neural networks: ...



Using the COS angle measure to define the similarity of two histograms h and h^{cl}_{sg}:

$$\cos\alpha = \langle h, h^{cl}_{sg} \rangle.$$

This measure, as desired, yields the value **1** when the two histograms are identical and **0** when the two histograms are orthogonal to each other.



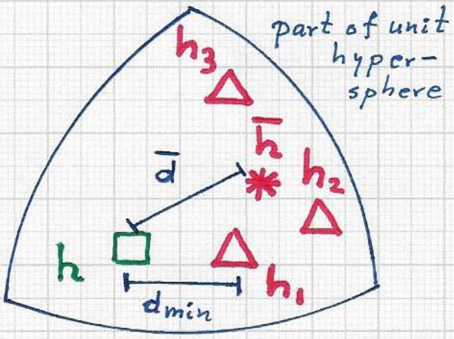
To determine the similarity between a histogram h and 4 same-class histograms h^{cl}_{sg}, $sg=1\ldots4$, one must consider all 4 cos values.

We can assume that all coefficient value histograms are binned, discrete histograms, using B bins — $bin_1, \ldots, bin_B$. Thus, all NORMALIZED histograms can be interpreted as **points** on a UNIT HYPER-SPHERE, having positional vectors with only non-negative coordinate values. The figure (bottom) on the previous page illustrates this interpretation; for a particular scale i, it shows the points representing the histograms of 6 sample segments for 3 material classes. **Based on this geometrical viewpoint, one can use a simple "angle measure" to define the difference/distance of two normalized, binned higrams in B-dimensional space.** The top figure (left) shows how one can employ the inner product of two normalized vectors — histograms — to define their similarity. The bottom figure (left) illustrates our situation: We must calculate ONE similarity value for a given histogram h and a set of histograms h^{cl}_{sg}, $sg=1\ldots4$, of the same class.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...



Two viable possibilities for computing a similarity value for h and $\{h_1, h_2, h_3\}$: use the "minimal angle distance" d_{min} (with maximal cos value) or use the angle distance to $\bar{h}$, i.e., the normalized vector of the averaged vectors h_1, h_2 and h_3.

The left figure illustrates our setting: Given a histogram h and the set of histograms $\{h_1, h_2, h_3\}$ – where h_1, h_2 and h_3 are stored sample histograms of the same class – for a specific scale i, calculate a similarity value between h and $\{h_1, h_2, h_3\}$. One could, for example, define similarity as follows:

i) $SIM_{max} = \max\{\langle h, h_i\rangle\}_{i=1}^{3}$ or

ii) $SIM_{avg} = \langle h, \bar{h}\rangle$, where

$$\bar{h} = \frac{(h_1+h_2+h_3)/3}{\|(h_1+h_2+h_3)/3\|}.$$

When one has decided which definition of similarity is "a proper and simply computable" option, OUR PROBLEM IS SOLVED: WE HAVE A MEANS FOR CALCULATING SIMILARITY BETWEEN A GIVEN, NEW HISTOGRAM h AND A SET OF HISTOGRAMS THAT CORRESPOND TO SEVERAL SEGMENTS, ALL BELONGING TO THE SAME CLASS.

Generalizing the concept of "similarity between two segments S and s" to the concept of "similarity between a segment S and a set of segments $\{s_{sg}\}$" reduces the p-values to values only depending on class and scale:

	scales: 1	...	H
$class_1$	$p_{1,1}$	...	$p_{1,H}$
⋮	⋮	⋮	⋮
$class_C$	$p_{C,1}$	...	$p_{C,H}$

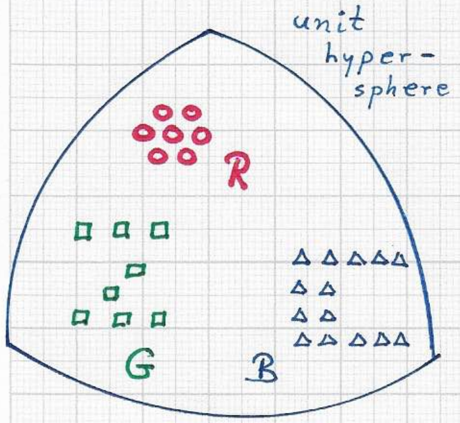
When comparing a segment S against C classes and H scales, $C \cdot H$ $p_{cl,sc}$-values result.

...

Stratovan

■ OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

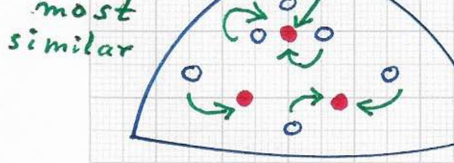
• Laplacian eigenfunctions and neural networks:...



• **Note.** Generally, one should assume that same-class, same-scale histograms of multiple sample segments form "compact clusters" - like cluster R. BUT: In principle, it might be possible that same-class, same-scale segment histograms define non-compact clusters - like clusters G and B. THUS:

The similarity measure SIM_{max} is preferable.

(See previous page.)



Definition of p_i-value distribution. ● = histograms from database, o = histograms of add'l samples.

IN SUMMARY, it is now possible to calculate p_i-values - and p_i-value distributions - for the definition of univariate decider functions $F_i(p_i)$. A p_i-value captures the similarity between an unclassified segment **S** and a classified, stored segment **s**. WE CAN NOW UNDERSTAND **s** ALSO AS A SET OF CLASSIFIED, STORED SEGMENTS $\{s_{sg}\}$. We can now reconsider the construction of a decider function $F_i(p_i)$. The left figure (bottom) illustrates the construction process. We use two datasets: (i) For a specific material class and a specific scale, the database stores **multiple** histograms associated with samples (of this class); this is the first dataset, shown as '●' in the figure. (ii) For the same material class and scale, ADDITIONAL samples (of this class) are made available and characterized via the same type of histograms, shown as 'o'. The sets {●} and {o} define a p_i-value distribution.

...