■ OBJECT AND MATERIALS EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

• Thus, the to-be-classified 3D image segment is STATISTICALLY characterized by H binned histograms/distributions (properly normalized in the domain value range, [0,1], and with respect to its integral norm, ||·||). These histograms describe the segment's low-, medium- and high-frequency behaviors; they are the segment's H "fingerprints."

• Subsequent segment comparison and classification steps are based on a metric that makes possible a meaningful and efficient computation of similarity of / distance between two eigenfunction-based coefficient value histograms, for the same scale of two material segments. "Degree of similarity" (concerning a specific scale) becomes the foundation of match-probabilities for the to-be-classified segment and classified segments.

• A database of materials to be recognized contains data for C classes cl, cl=1...C. (Class 0 is understood as the class that is NOT class 1, ..., NOT class C.) Class cl is represented in the database by k_{cl} material samples.

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... •**Note.** The k_{cl} material samples stored in the database for class **cl** must be "REPRESENTATIVE" for class **cl** in the following sense: As far as "features," "texture," "multi-scale characteristics" etc. of material class **cl** is concerned, the class-**cl** segment data stored in the database must capture the allowed and practically possible "variability spectrum" of class **cl** in all its appearances (in 3D image data). Further, the relative numbers of allowed samples for class-**cl** should reflect the "segment types" of class-**cl** AS OBSERVED IN A REAL-WORLD SETTING." For example, if material class **cl** exhibits a certain 3D checkerboard texture in 50% of all observed real-world cases, then it will be imperative to also populate the database accordingly, i.e., 50% of the class **cl** samples in the database will have to exhibit such a checkerboard texture. In other words, the practically possible "segment types" of class **cl** and their relative occurrences must be reflected, preserved by the chosen class-**cl** samples in the database. This database design goal, to be achieved via proper training/learning, is CRUCIAL FOR THE COMPUTATION OF MATCH-PROBABILITY VALUES FOR A TO-BE-CLASSIFIED SEGMENT throughout the classification process.

...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks:...

- Each material segment — every segment in the database and the one to-be-classified segment — is represented by H normalized (binned) histograms (over a domain interval [0,1]). Thus, every segment has H "fingerprints" generated by moving a convolution mask over the entire segment; determining the mask-associated coefficient values of many local eigenfunction-based density function expansions; and establishing the H coefficient value histograms.
- It is possible to compare each of the H histograms of the to-be-classified segment with the same-scale histograms of all the segments in the database (all classes cl, all samples per class). The result of this comparison is a set of "similarity values" defining the scale-specific values of the chosen distance metric for histograms.
- One can now calculate match-probabilities P_{cl}^{sc} that consider similarity at only one scale sc and define a probability of belonging to class cl for the to-be-classified segment. (This probability considers and therefore depends on the "representative" k_{cl} samples in the database.)

...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:...

• **Note.** Considering merely similarity for one scale vs. considering all (or several selected) scales leading to multiple scale-specific similarities impacts the definition and calculation of match-probabilities. The to-be-classified image segment is statistically characterized for H scales. Therefore, some of these scales are more and some are less similar to the scales of a specific material class' sample representatives in the database. (It is important to keep in mind that each material class in the database is captured via a representative set of class samples; and each sample is identified with its corresponding H scale histograms. This aspect is relevand for the computation of a similarity between a scale of the to-be-classified segment and the multitude of same-scale histograms of the set of samples of a certain class.)

The crucially important definition(s) in this context concerns **SIMILARITY** and **CLOSEST SAMPLE** in the database of class segment samples. We understand a coefficient value histogram as a discrete, binned normalized histogram $\mathbb{h}$, having, for example, 64 coordinates (all non-negative). Since $\|\mathbb{h}\|_2 = 1$, one can interpret $\mathbb{h}$ as a positional vector from the origin to a point on a unit hyper-sphere.

...

■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... The to-be-classified image segment is represented via H scales, i.e., via H histograms. One can write the set of histograms as a "vector of histograms $\mathbb{h}_1, \dots, \mathbb{h}_H$":

$$\mathbb{H} = (\mathbb{h}_1, \mathbb{h}_2, \dots, \mathbb{h}_H) .$$

SIMILARITY of two image segments for a specific scale is defined by histogram similarity. In our setting, we can use **COS-SIMILARITY**: The similarity of two histograms $\mathbb{h}_a$ and $\mathbb{h}_b$ is

$$s(\mathbb{h}_a, \mathbb{h}_b) = \cos(\mathbb{h}_a, \mathbb{h}_b) = \langle \mathbb{h}_a, \mathbb{h}_b \rangle ,$$

where $\langle \cdot, \cdot \rangle$ is the inner product and $0 \le s \le 1$. The value $s = 0$ indicates orthogonality and the value $s = 1$ equality of the two histograms. Among the stored k_{cl} histograms **given** for the k_{cl} material samples of class **cl**, one must determine that histogram vector ${}^{sg}\mathbb{H}_{cl}$ that is closest/most similar to $\mathbb{H}$. (We use the notation ${}^{sg}\mathbb{H}_{cl}$ to denote the histogram vector of a specific sample segment **sg** belonging to class **cl**.) Here,

$${}^{sg}\mathbb{H}_{cl} = ({}^{sg}\mathbb{h}^1_{cl}, {}^{sg}\mathbb{h}^2_{cl}, \dots, {}^{sg}\mathbb{h}^H_{cl}) .$$

We can now use the **cos**-measure to compute similarities for each component of vectors of histograms:

$$\begin{aligned} {}^{sg}\mathbb{S}_{cl} &= (s(\mathbb{h}_1, {}^{sg}\mathbb{h}^1_{cl}), \dots, s(\mathbb{h}_H, {}^{sg}\mathbb{h}^H_{cl})) \\ &= ({}^{sg}s^1_{cl}, \dots, {}^{sg}s^H_{cl}) . \end{aligned}$$

For the vector-valued quantity ${}^{sg}\mathbb{S}_{cl}$, one can compute the norms

$$i)\ \|{}^{sg}\mathbb{S}_{cl}\|_{max} = \max_{sc=1\dots H} \{ {}^{sg}s^{sc}_{cl} \} ;\quad ii)\ \|{}^{sg}\mathbb{S}_{cl}\|_2^2 = \sum_{sc=1}^{H} ({}^{sg}s^{sc}_{cl})^2 .$$

...