**Perceptual Color Differences**

Their goal is to quantify the differences between colors as perceived by a human observer. However, the euclidean distance in the RGB or the XYZ space are not sufficient and the $\Delta E_{00}$ distance has been proposed for this task.

**State of the art**

Insufficiency of the XYZ space:

- Several transformations:
  - Complex
  - Non linear
  - 3 global approximations (one for each step)

**Local Metric Learning**

One transformation:

- Simple
- Non linear (Local transformations)
- Approximations depend of the local metrics

$\Delta E_{00}$

RGB

**Metric learning**

Learning how to compare objects: learn a new space where some constraints are fulfilled, e.g. move closer circles of the same color and keep far away circles of different colors.

- Mahalanobis distance:
  \[ \Delta(x, x') = \sqrt{\Delta(x - x')^T M(x - x') = \sqrt{(x - x')^T (x - x')}} \]

**Learning Local Metrics**

We learn $K+1$ metrics in different regions. To compute the distance between two colors, we simply have to select the matrix $M$ corresponding to their region.

**Algorithm 1: Local metric learning**

- **Input:** A training set $\mathcal{S}$ of patches, a parameter $K \geq 2$
- **Output:** $K$ local Mahalanobis distances and one global metric

1. Run $K$-means on $\mathcal{S}$ and deduce $K+1$ training subsets $\mathcal{T}_j$, $j = 0, \ldots, \bar{K}$
2. For $j = 0 \rightarrow \bar{K}$
   - Learn $M_j$ by solving the convex optimization problem:
     \[
     \arg \min_{M_j} \sum_{(x, x') \in \mathcal{T}_j} [\Delta(x, x')^2 M_j(x - x')^2]
     \]
   - Local Empirical Risk:
     \[
     \epsilon_{\mathcal{T}_j}(M_j) = \frac{1}{2n_j} \sum_{(x, x') \in \mathcal{T}_j} \Delta(x, x')^2 M_j(x - x')^2 - \Delta E_{00}(x, x')^2
     \]
   - Local True Risk:
     \[
     \epsilon_j(M_j) = \mathbb{E}_{(x, x') \sim \Delta E_{00} \sim \mathcal{T}_j} \left[ \Delta(x, x')^2 M_j(x - x')^2 - \Delta E_{00}(x, x')^2 \right]
     \]

**Theoretical Guarantees**

The bound holds with probability $1 - \delta$ with $L_B, \Delta_{\text{max}}$ and $D$ constants. It is based on the uniform stability property used in each region (Stability and Generalization, O. Bousquet and A. Elisseeff, JMLR 2002) and on the Bretagnolle-Huber-Carol inequality for multinomial distributions (Weak Convergence and Empirical Processes, A. W. van der Vaart and J. A. Wellner, Springer 2000).

**Performance of the Learned Metric in terms of Statistical Criteria**

**Dataset:** 260 images of patches with known L*a*b*. We have taken the images with 4 cameras under different acquisition conditions.

**Mean:** \[
\frac{1}{n} \sum_{(x, x') \in \mathcal{T}} |\Delta(x, x') - \Delta E_{00}(x, x')| \]

**STRESS:** quadratic criterion which penalizes more large errors than small ones.

\[
\text{STRESS} = 100 \sqrt{ \frac{\sum_{(x, x') \in \mathcal{T}} \Delta E_{00}(x, x')^2 \Delta(x, x')^2}{\sum_{(x, x') \in \mathcal{T}} \Delta E_{00}(x, x')^2 \Delta(x, x')^2} } \]

In each case, for $K = 20$, our results are significantly better with a $p$-value lower than 0.001.

**Generalization to colors**

**Generalization to cameras**

**Performance of the Learned Metric in a Segmentation Task**

- In all cases, for $K = 20$, our results are significantly better with a $p$-value lower than 0.001.

**Theoretical Guarantees**

- True Risk: $\epsilon(M) \leq \hat{\epsilon}_T(M) + L_B \sqrt{\frac{2(2 + 1) \ln 2 n + 2 \ln(2/\delta)}{n}} + (K + 1) \left( \frac{2D^4}{\lambda n} + \frac{4D^4}{\lambda} + \Delta_{\text{max}} \left( \frac{2D^2}{\lambda^2} + 2\Delta_{\text{max}} \right) \sqrt{\ln \left( \frac{4(k+1)}{2n} \right)} \right)
- Empirical Risk: $\epsilon_T(M) = \mathbb{E}_{(x, x') \sim \Delta E_{00}} \left[ \Delta(x, x')^2 M_j(x - x')^2 - \Delta E_{00}(x, x')^2 \right]$