Modeling Perceptual Color Differences by Local Metric Learning

Michaël Perrot, Amaury Habrard, Damien Muselet and Marc Sebban
Laboratoire Hubert Curien, Saint-Etienne, France
michael.perrot@univ-st-etienne.fr

Perceptual Color Differences

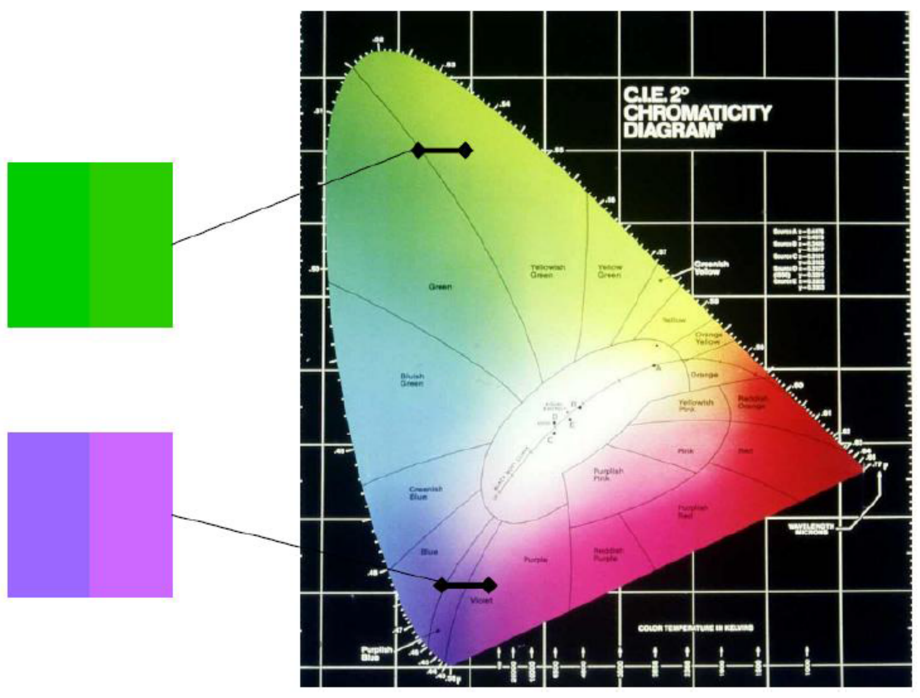
Their goal is to quantify the differences between colors as perceived by a human observer. However, the euclidean distance in the RGB or the XYZ space are not sufficient and the ΔE_{00} distance has been proposed for this task.

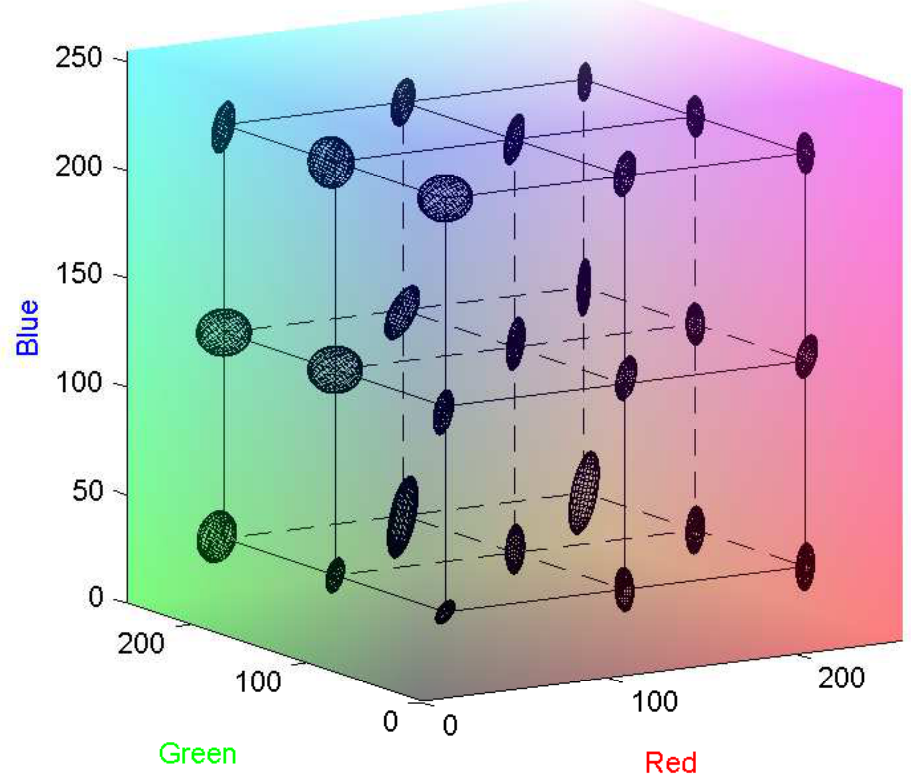
State of the art

RGB ⇓ XYZ ⇓ L*a*b* ⇓ ΔE_{00}

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

Insufficiency of the XYZ space:





Local Metric Learning

RGB ⇓ ΔE_{00}

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

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 diffrent colors.

Mahalanobis distance:

$$\Delta(\mathbf{x},\mathbf{x}') = \sqrt{(\mathbf{x}-\mathbf{x}')^T\mathbf{M}(\mathbf{x}-\mathbf{x}')} = \sqrt{(\mathbf{Lx}-\mathbf{Lx}')^T(\mathbf{Lx}-\mathbf{Lx}')}$$

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 S of patches; a parameter $K \geq 2$
output: K local Mahalanobis distances and one global metric
begin
 Run K-means on S and deduce $K+1$ training subsets T_j $(j = 0, 1 \ldots, K)$ of triplets $T_j = \{(\mathbf{x_i}, \mathbf{x_i'}, \Delta E_{00})\}_{i=1}^{n_j}$ (where $(\mathbf{x_i}, \mathbf{x_i'})$ are similar patches in region C_j with a $\Delta E_{ab} < 5$)
 for $j = 0 \to K$ **do**
  Learn $\mathbf{M_j}$ by solving the convex optimization problem:

$$\arg\min_{\mathbf{M_j}\succeq 0} \hat{\varepsilon}_{T_j}(\mathbf{M_j}) + \lambda_j \|\mathbf{M_j}\|^2_{\mathcal{F}}$$

Local Empirical Risk:

$$\hat{\varepsilon}_{T_j}(\mathbf{M_j}) = \frac{1}{n_j}\sum_{(\mathbf{x},\mathbf{x}',\Delta E_{00})\in T_j} \left|(\mathbf{x}-\mathbf{x}')^T\mathbf{M_j}(\mathbf{x}-\mathbf{x}') - \Delta E_{00}(\mathbf{x},\mathbf{x}')^2\right|$$

Local True Risk:

$$\varepsilon(\mathbf{M_j}) = \mathbb{E}_{(\mathbf{x},\mathbf{x}',\Delta E_{00})\sim P_j} \left|(\mathbf{x}-\mathbf{x}')^T\mathbf{M_j}(\mathbf{x}-\mathbf{x}') - \Delta E_{00}(\mathbf{x},\mathbf{x}')^2\right|$$

Theoretical Guarantees

True Risk

$$\varepsilon(\mathbf{M}) \leq \hat{\varepsilon}_T(\mathbf{M}) + L_B\sqrt{\frac{2(K+1)\ln 2 + 2\ln(2/\delta)}{n}} + (K+1)\left(\frac{2D^4}{\lambda n} + \left(\frac{4D^4}{\lambda} + \Delta_{\max}\left(\frac{2D^2}{\sqrt{\lambda}} + 2\Delta_{\max}\right)\right)\sqrt{\frac{\ln(\frac{4(K+1)}{\delta})}{2n}}\right)$$

Empirical Risk

The bound holds with probability 1 - δ with $L_B, \Delta_{\max}$ 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.

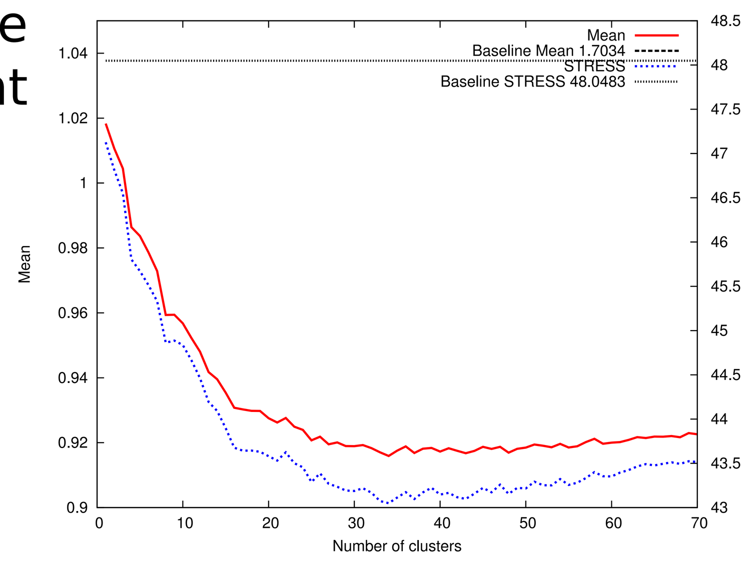
Mean: $\frac{1}{n}\sum_{(\mathbf{x},\mathbf{x}')\in TS} |\Delta_T(\mathbf{x},\mathbf{x}') - \Delta E_{00}(\mathbf{x},\mathbf{x}')|$

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

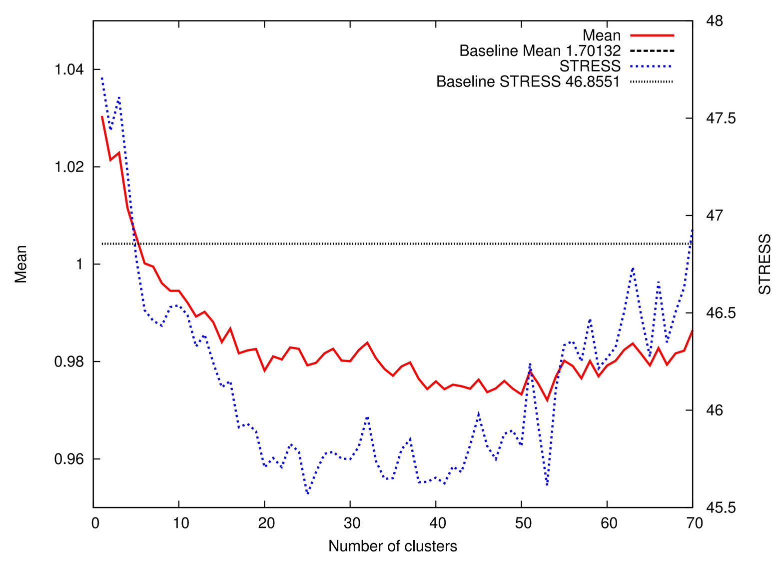
$$\mathrm{STRESS} = 100\sqrt{\left(\frac{\sum_{(\mathbf{x},\mathbf{x}')\in TS}(\Delta E_{00}(\mathbf{x},\mathbf{x}') - F\Delta_T(\mathbf{x},\mathbf{x}'))^2}{\sum_{(\mathbf{x},\mathbf{x}')\in TS} F^2\Delta_T(\mathbf{x},\mathbf{x}')^2}\right)}$$

with

$$F = \frac{\sum_{(\mathbf{x},\mathbf{x}')\in TS}\Delta E_{00}(\mathbf{x},\mathbf{x}')^2}{\sum_{(\mathbf{x},\mathbf{x}')\in TS}\Delta E_{00}(\mathbf{x},\mathbf{x}')\Delta_T(\mathbf{x},\mathbf{x}')}.$$



Generalization to colors



Generalization to cameras

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

Performance of the Learned Metric in a Segmentation Task