Computational Geometry Neighborhoods for Local Learning

Maya Gupta
Eric Garcia
James (Yihua) Chen
Erika Chin

Dept. of Electrical Eng.
Univ. of Washington

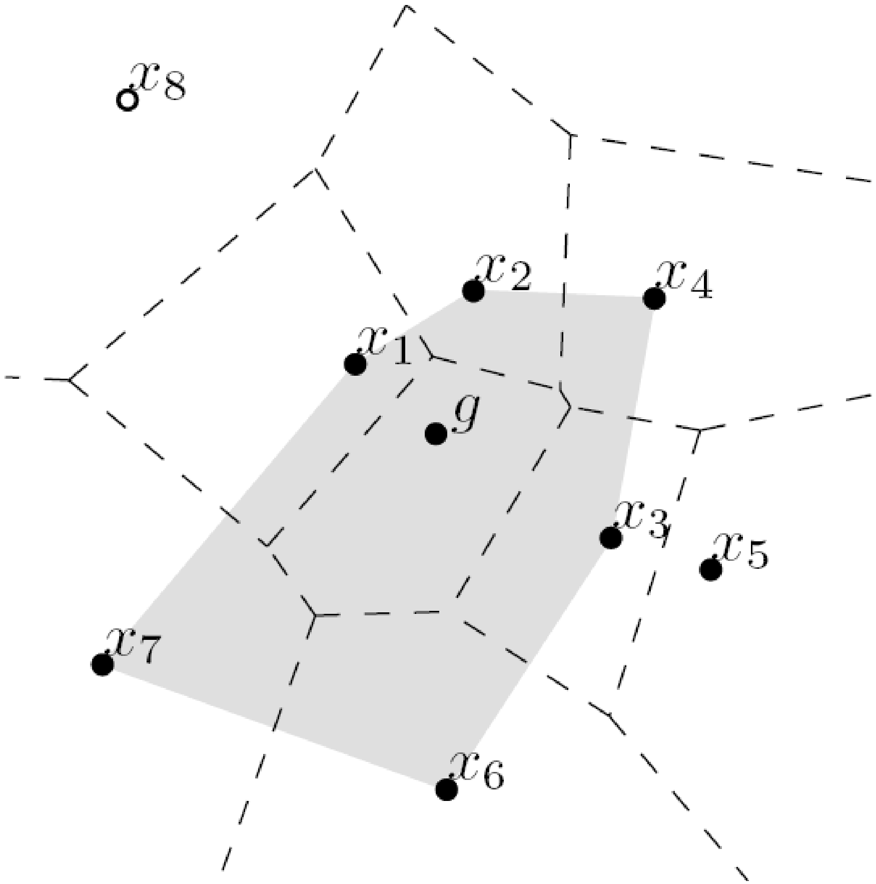
Problem: How to choose a neighborhood for local learning?

Goal 1) neighbors are near

Goal 2) neighbors are diverse in feature space (i.e. “surround the test point”)

Prior Work: Natural Neighbors

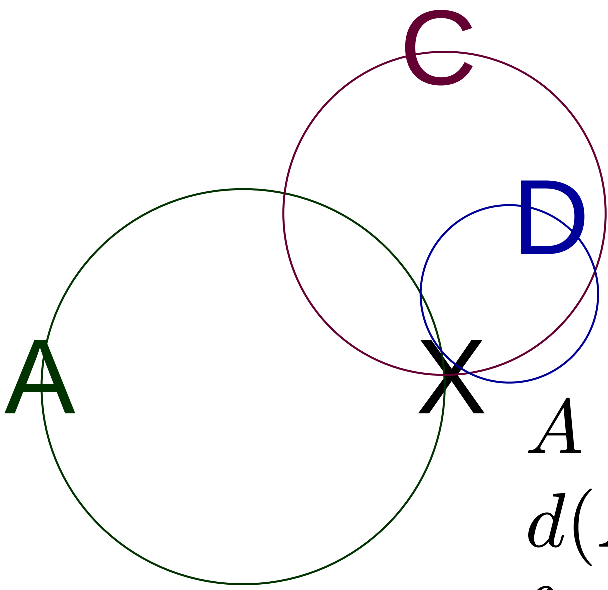
Natural Neighbors of g
is the set of all points
whose Voronoi cells
are adjacent to
the cell of g.



Sibson proposed
1981 as neighborhood
for linear interpolation.

Natural Neighbors Inclusive (Gupta '07):
Add all nearer neighbors to the natural neighbors.

Prior Work: Gabriel Neighbors



X's Gabriel neighbors are A and D

A and B are Gabriel neighbors if $d(A,B)^2 \leq d(A,x_j)^2 + d(B,x_j)^2$ for all x_j.

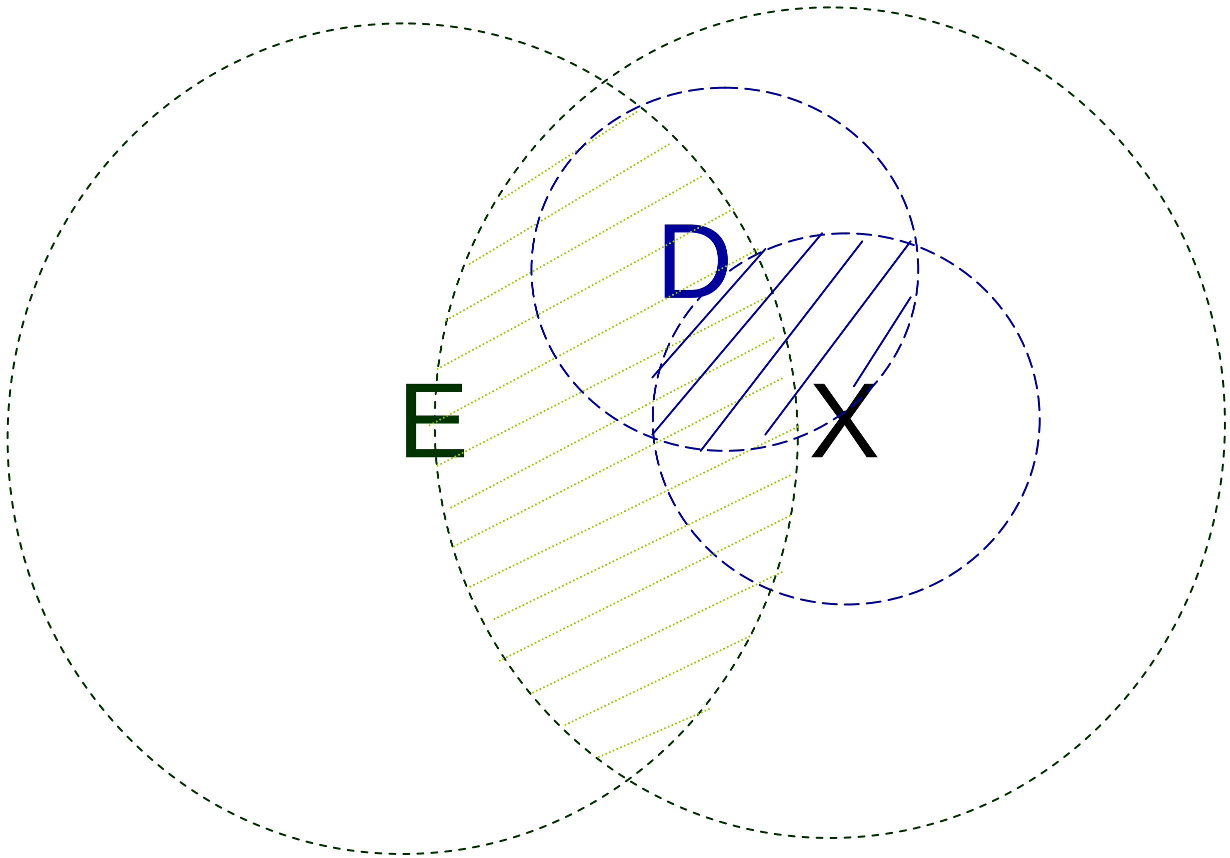
Shown to sometimes work better than k-NN for classification by Sanchez et al. 1997

Prior Work: Relative Neighbor Graphs

A and B are RNG neighbors if
$d(A, B) \leq \max\left(d(A, x_j), d(B, x_j)\right)$
for all x_j.

Shown to work consistently better than k-NN for classification by Sanchez et al. 1997

D is a RNG neighbor of X but E is not.

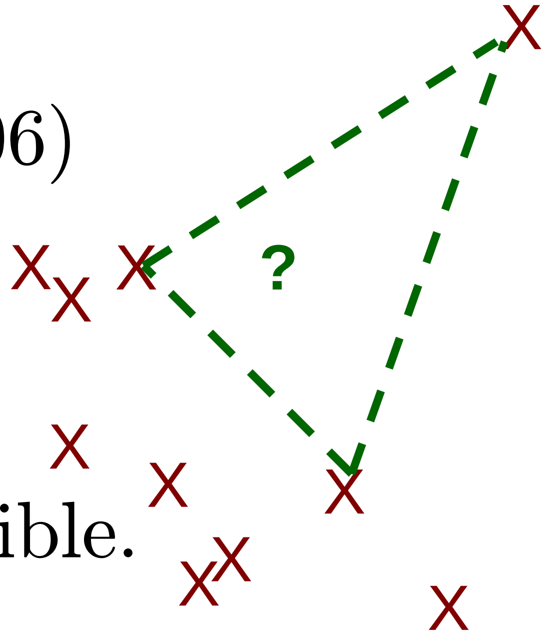


Enclosing Neighborhoods *(Gupta et al. '07, '08)*

Neighborhood definition that produces neighbors that enclose a test point in their convex hull if such a neighborhood exists.

Ex: Natural neighbors

Ex: Enclosing k-NN (Gupta 06)



Ensures interpolation rather than extrapolation when possible.

Enclosing Neighborhoods Have Bounded Estimation Variance

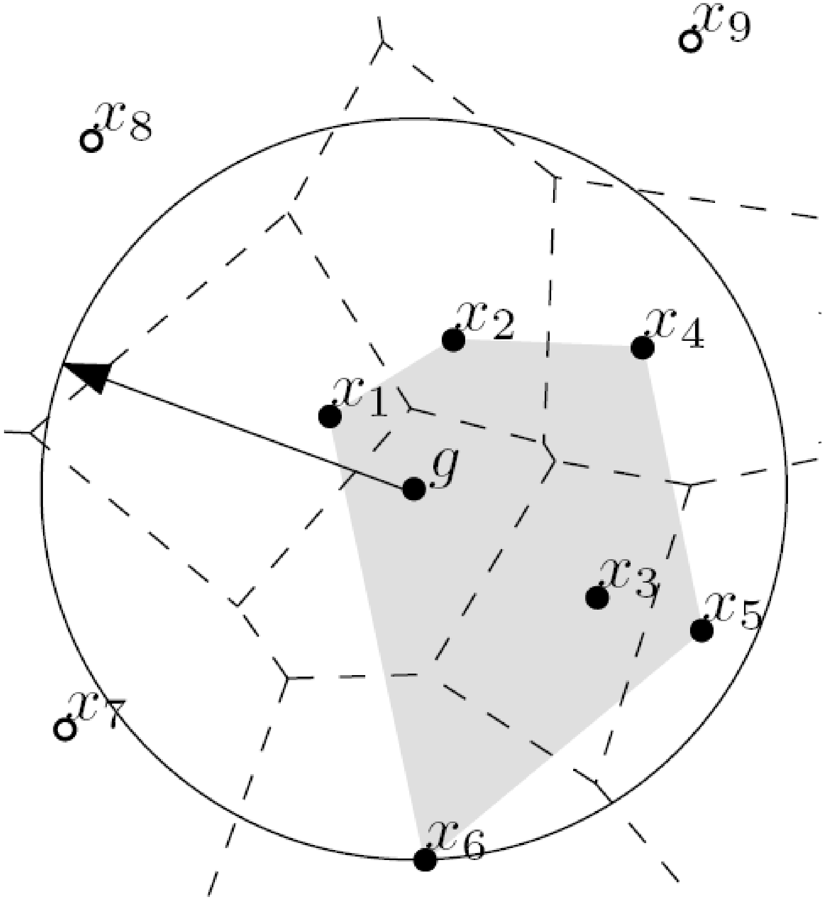
Theorem *(Gupta Garcia Chin 08)*:
Local least-squares fitted hyperplane to an enclosing neighborhood has estimation variance bounded by σ^2 if the true output values are linear with additive white noise $N \sim \mathcal{N}(\mu, \sigma^2)$:

$$E_N\left[\left(\hat{f}(x) - E_N[\hat{f}(x)]\right)^2\right] \leq \sigma^2$$

Enclosing k-NN *(Chin Garcia Gupta '07)*

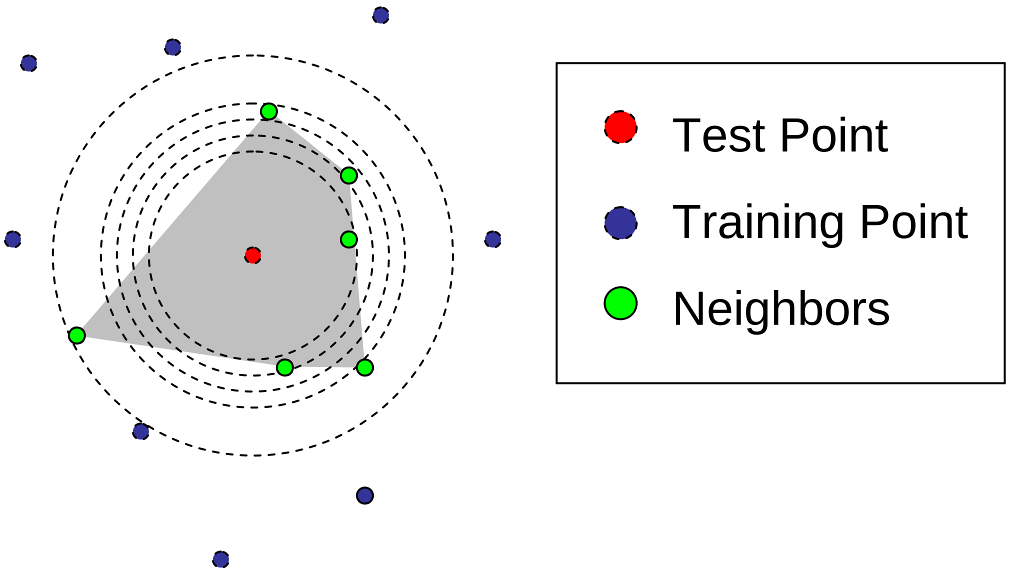
The neighborhood about *g* with the fewest number of neighbors k that achieves the minimum distance to their convex hull.

(The neighborhood about g with the fewest number of neighbors k that enclose g in their convex hull.)



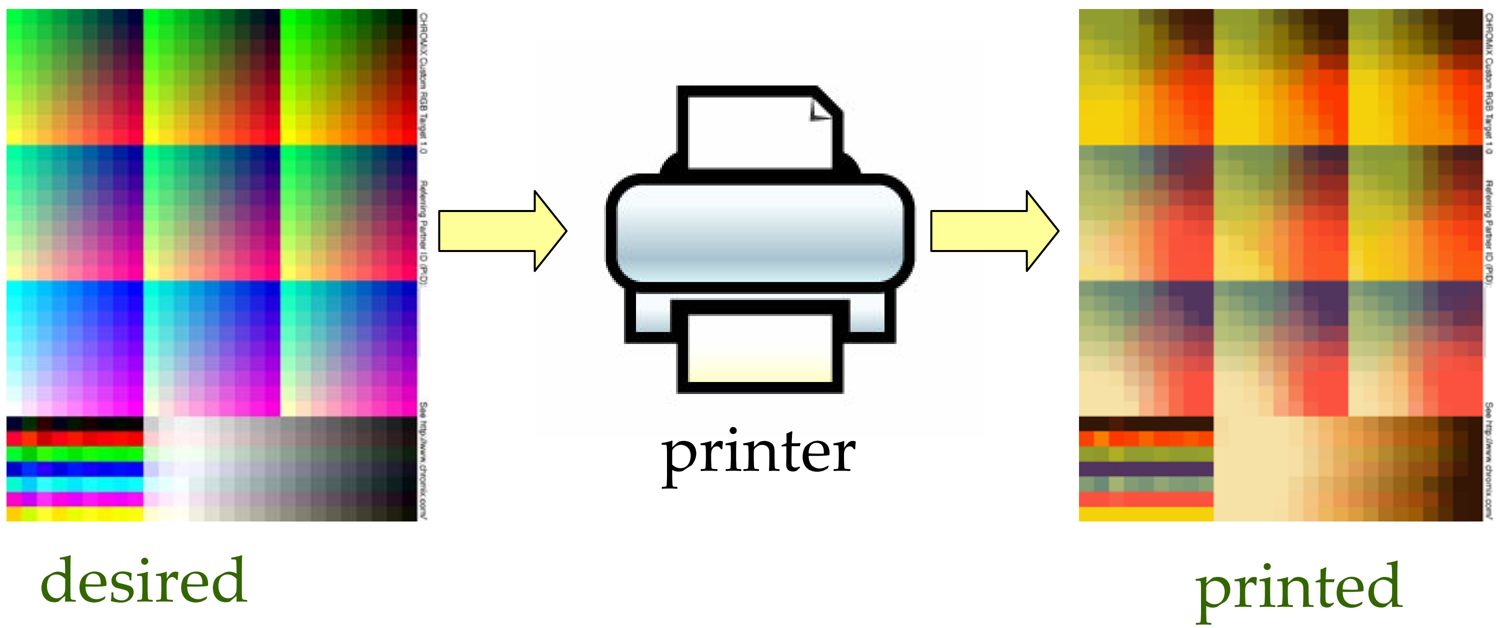
Expected Size of Enclosing k-NN

Theorem (Gupta Garcia Chin '08): If d-dimensional training points drawn uniformly randomly around test point, the enclosing k-NN will have an average of $2d + 1$ neighbors.



Color management

Problem: How do you get a printer to print colors “correctly”?



Motivation: art, catalogs, product quality,
skin tones, preserve contrast and image quality
scientific and information visualization

Printer color management

Step 1: Input RGB patches and measure CIEL*a*b* values

Step 2: *Estimate RGB inputs corresponding to each color in 3D CIEL*a*b* grid*

Step 3: Given a desired CIEL*a*b* color, interpolate the 3D LUT for best RGB input

Printer color management summary

Step 1: Input RGB patches and measure CIEL*a*b* values

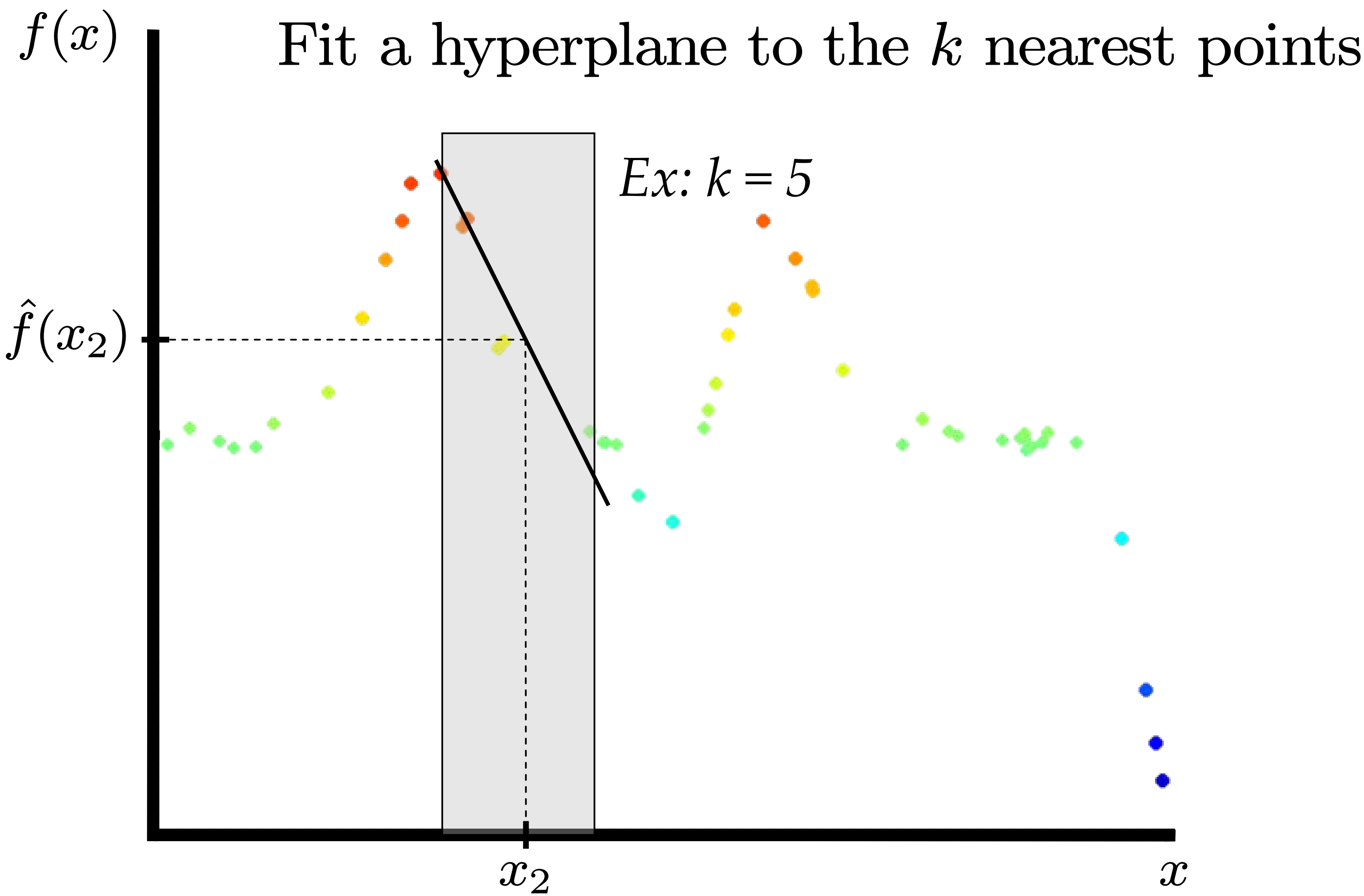
Step 2: *Estimate RGB inputs corresponding to each color in 3D CIEL*a*b* grid*

Bala 2003: best results with local linear regression compared to neural nets, polynomial regression, or splines.

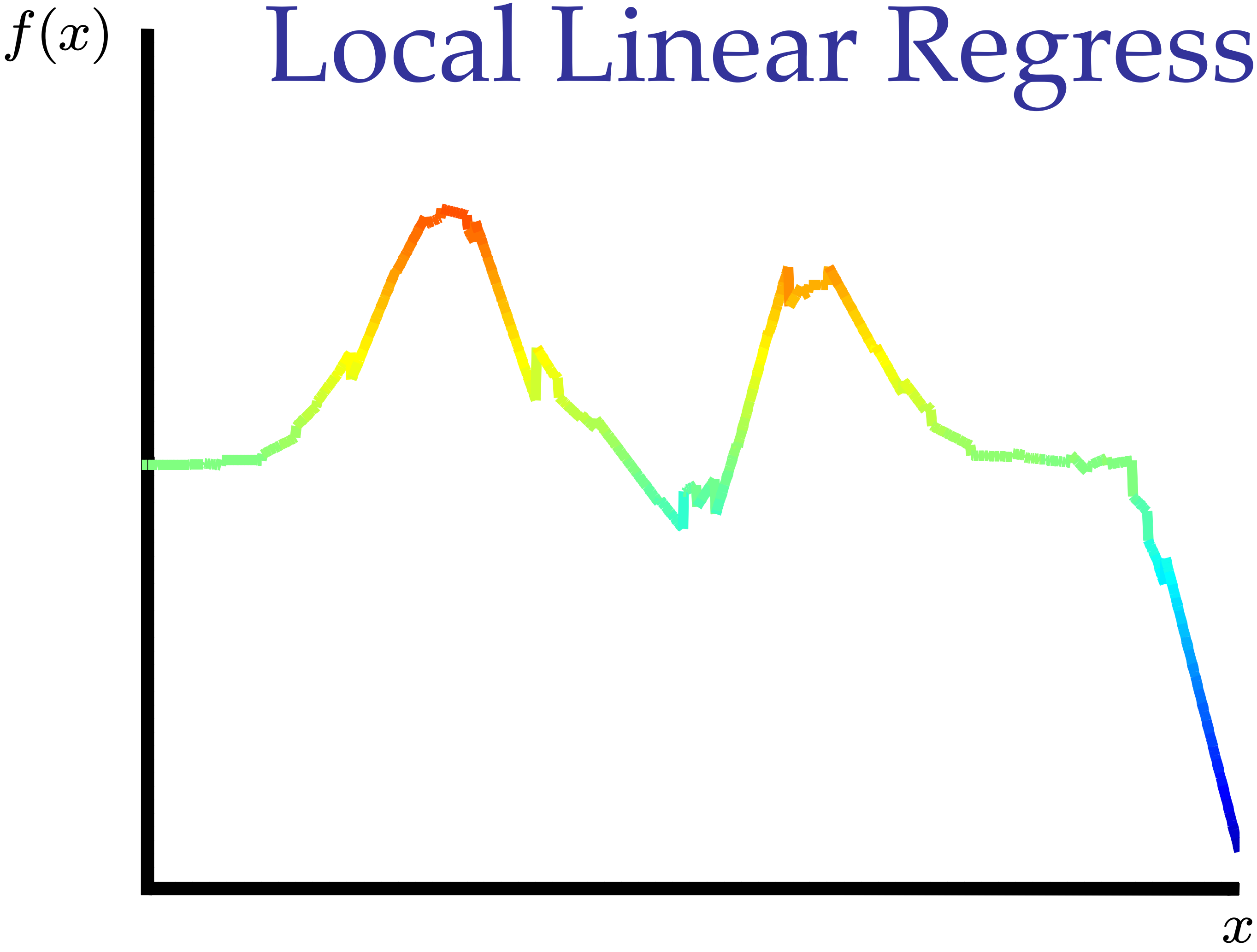
Our work: better results by regularized local linear regression with **enclosing neighborhoods** *(Gupta et al. IEEE Trans Image Proc 2008)*

Step 3: Given a desired CIEL*a*b* color, interpolate the 3D LUT for best RGB input

Local Linear Regression



Example Result of Local Linear Regression



Local Ridge Regression

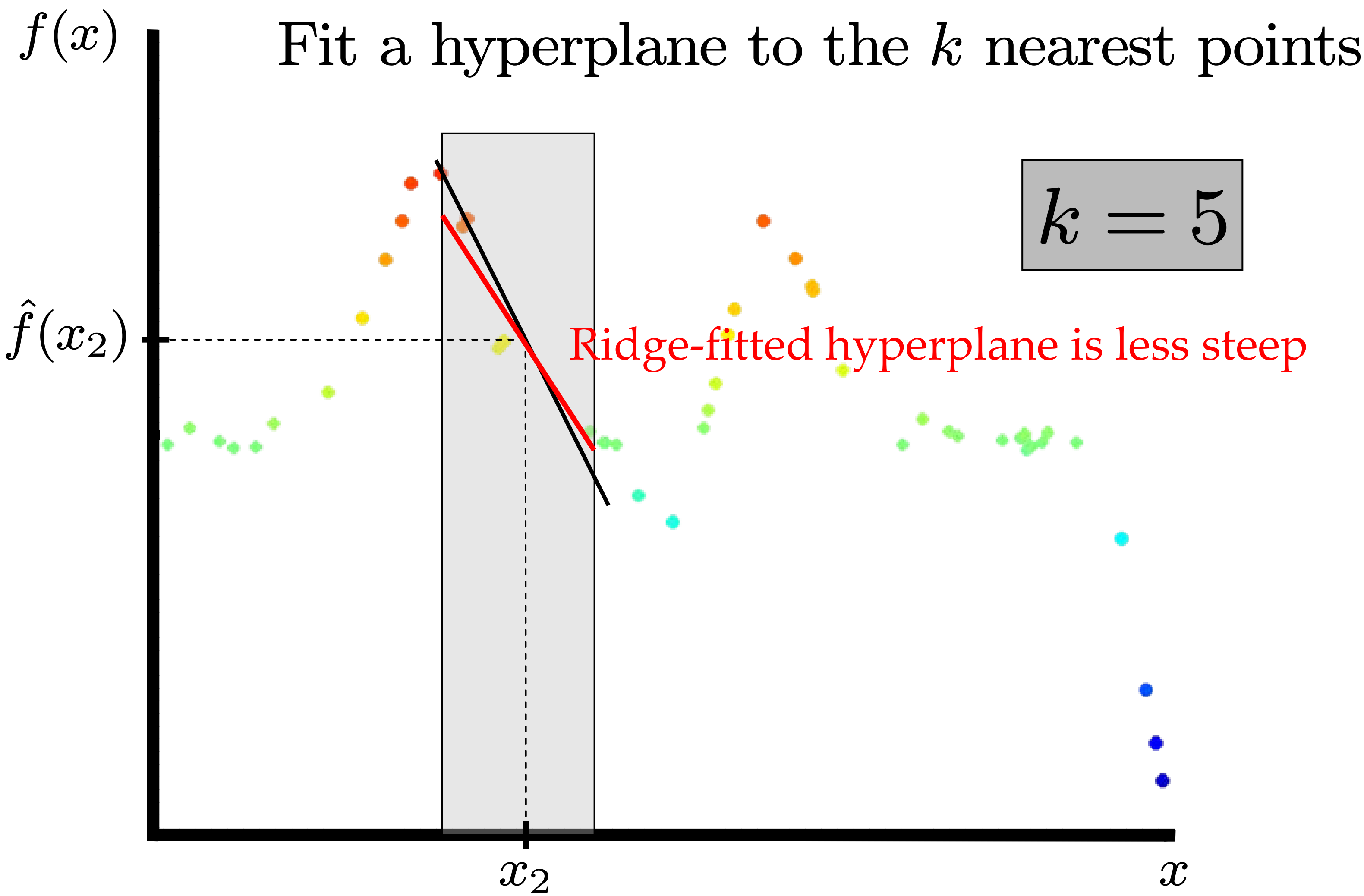
Local Linear Regression: $\hat{f}(x) = {\beta^*}^T x$

$$\beta^* = \arg\min_{\beta} \sum_{\text{neighbors } x_i} (\beta^T x_i - f(x_i))^2$$

Local Ridge Regression: $\hat{f}(x) = {\beta^*}^T x$

$$\beta^* = \arg\min_{\beta} \sum_{\text{neighbors } x_i} (\beta^T x_i - f(x_i))^2 + \lambda \beta^T \beta$$

Local Linear Regression

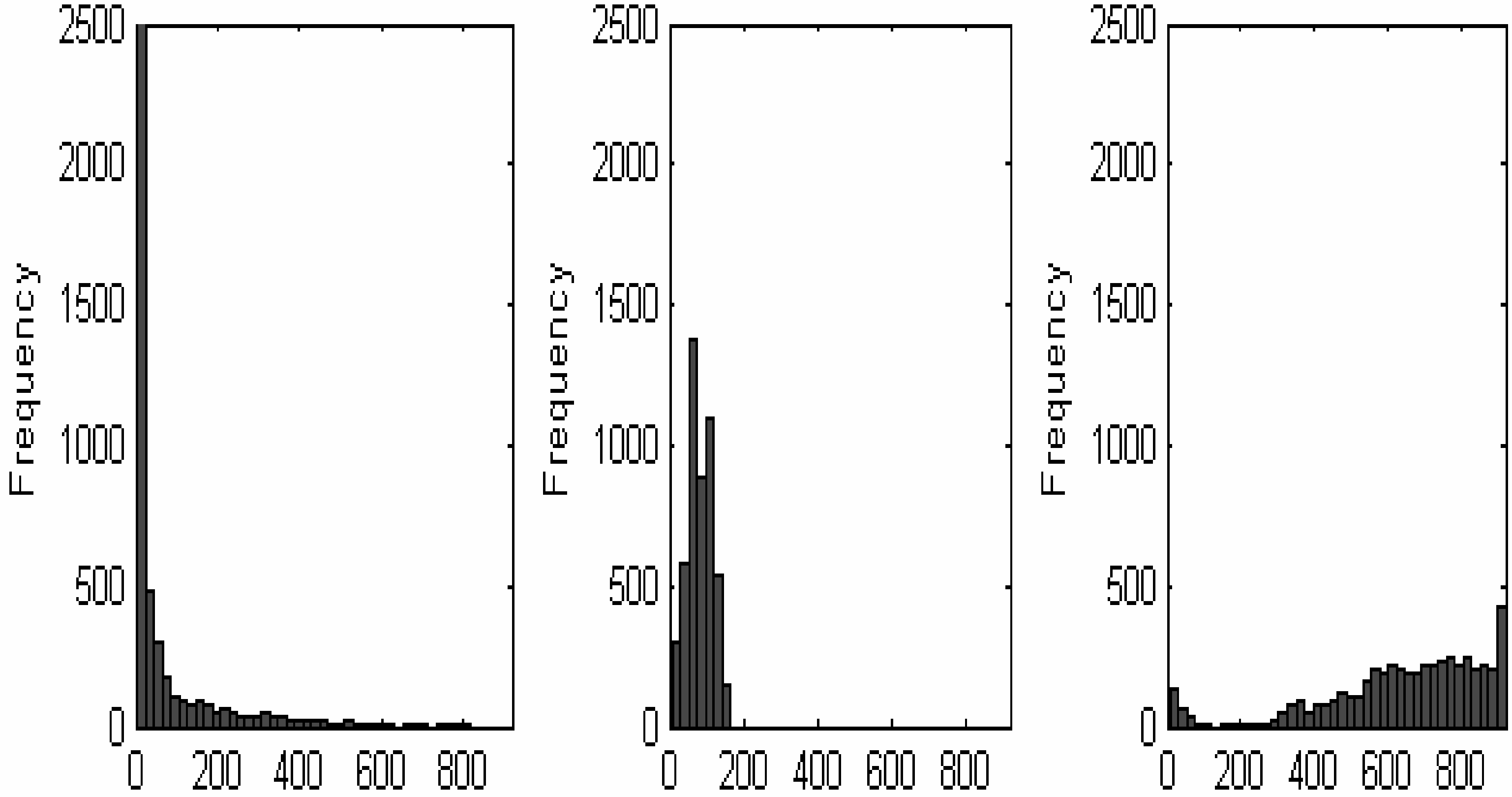


Example Color Management Results

Ricoh Laser Printer, 918 training patches,
729 in-gamut test patches, regularization parameter fixed at .1

Neighborhood	Mean ΔE Error	95th %-ile ΔE Error
Enclosing k-NN	3.7	7.4
Enclosing k-NN Minimum 15	**3.5**	**6.8**
Natural Neighbors	3.7	7.1
Baseline: 15 Neighbors	4.2	8.6

Example Neighborhood Sizes for 3-dimensional color samples



Enclosing k-NN

Natural Neighbors (Voronoi)

Natural Neighbors Inclusive

Enclosing Neighborhood Problems

Computationally Expensive

Natural Neighbors

Voronoi tessellation of entire training set and test point.

Worst Case: $O(n^{\lfloor \frac{d}{2} \rfloor})$

Enclosing k-NN

QP at each step to find distance to neighbors.

Worst Case: $O(n^4)$

Test point outside convex hull of data

Sometimes in 3D, common in higher dimensions

Papers (available at idl.ee.washington.edu)

M. R. Gupta, Custom Color Enhancements, Proc. Intl. Conf. on Image Processing 2005.

E. Chin, E. K. Garcia, M. R. Gupta, Color Management of Printers by Regression over Enclosing Neighborhoods, Proc. Intl. Conf. on Image Processing 2007.

M. R. Gupta, E. K. Garcia, E. Chin, Adaptive Local Linear Regression with Application to Printer Color Management, To appear, IEEE Trans. On Image Processing (2008).