Fast, Adaptive Mahalanobis Distance-Based Search and Retrieval in Image Databases

Sharadh Ramaswamy & Kenneth Rose
Signal Compression Lab, UCSB

Multimedia Databases

- Handle video/image/audio/text data

  e.g. in Picsearch, Youtube, Picasa,Facebook
- Often “metadata” lacking – needs to be extracted

  - typical of scientific data e.g. in genomics, bio-molecular imaging
- Organized based on object *content*
- Today, multimedia data management critical

  - with availability of cheap storage

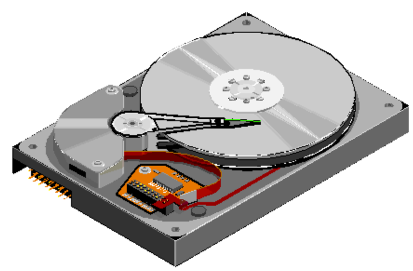
  - widespread use of multimedia devices

  e.g. dig. still and video cameras, camcorders, MP3 players

Challenges in Image Database Indexing

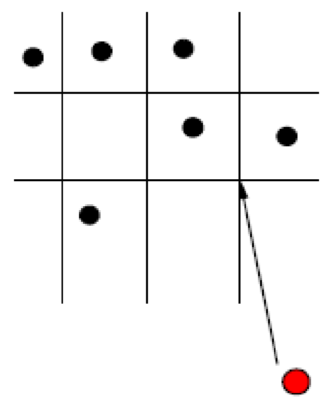
- "Interesting regions" in images apriori unknown
- High dimensional descriptors needed
  - color, color layout, shape, texture, SIFT etc.
  - image similarity $\propto$ feature vector distance
- Large volume of data
  - search engines index billions of webpages,
  - millions of photos uploaded ***each day*** to Facebook, Flickr, Picasa etc.
- Feature vectors stored "offline"
  - Secondary storage (hard drives) slower
  - I/O time dominates search

$\Rightarrow$ need efficient indexing



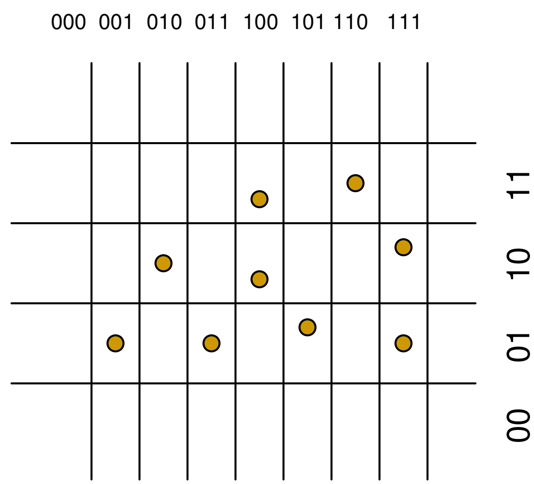
Indexing High-dimensional Spaces

- State-of-the-art indexes based on compression
  - Search compressed version of database
- Scalar quantization methods
  - VA-file (VLDB 1998)
- Clustering/VQ methods
  - VQ is optimal in compression
    - Compact representation of data-set
  - VQ exploits correlations across dimensions
  - Used extensively in approx kNN search
- Focus is on exact kNN search



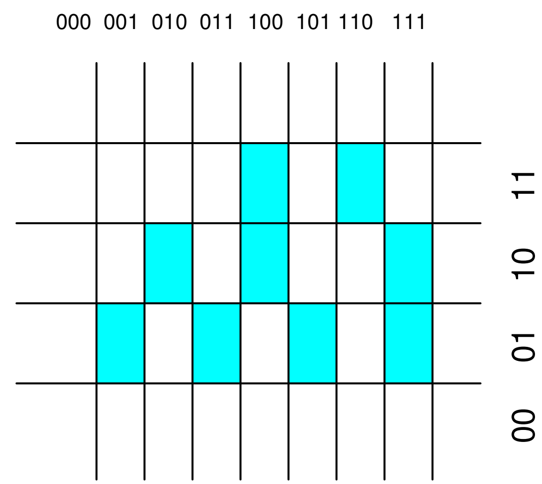
VA-File based Indexing

- Quantize each dimension uniformly
- Quantize each element of data-set



VA-File based Indexing

- Quantize each dimension uniformly
- Quantize each element of data-set
- Create approximation file
  - store quantization bit-strings for each element

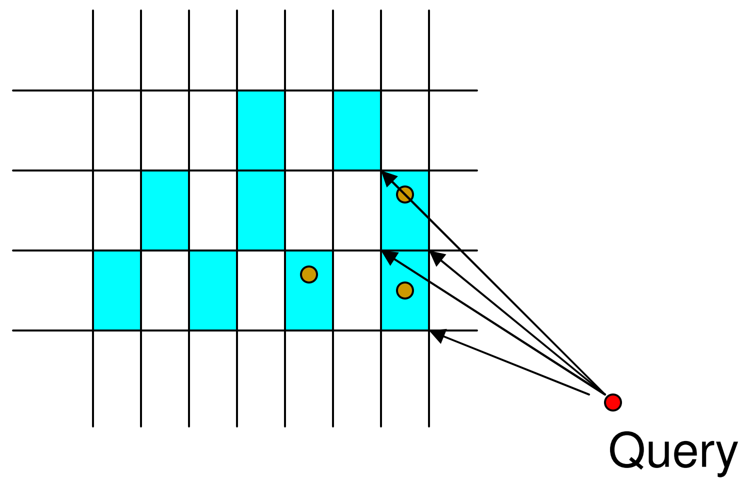


VA-File –Query processing

1. Read approximation file
2. Establish lower and upper distance bounds to occupied cells
3. Eliminate irrelevant cells
4. Access all survivors in order of lower bounds
5. If k^{th} lowest distance found so far, less than next lower bound,

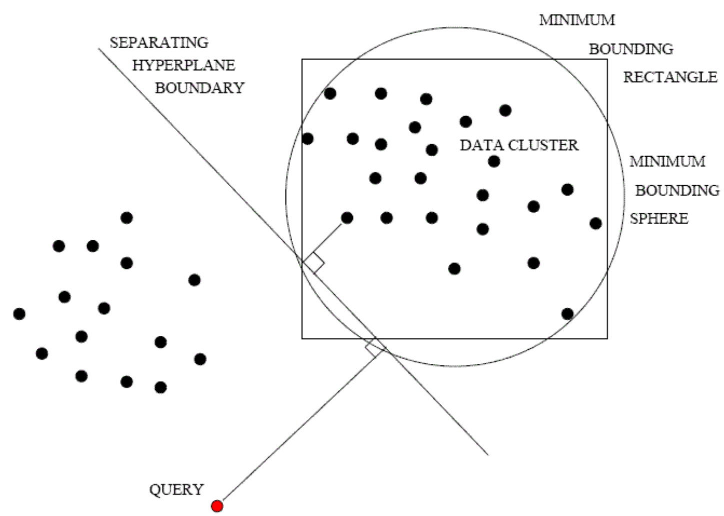
   STOP (kNNs found)

   Else read next survivor.



Clustering for Exact NN Search

- Bound distance to cluster
- Retrieve nearest clusters (till kNNs found)
- Bounds with rectangles and spheres loose
  - “curse of dimensionality”
  - Cluster-distance bounding (Ramaswamy & Rose ICIP 2007)



Perceptual Accuracy in Image Retrieval

- Quality of retrieved images important
- Euclidean distance/ l_2 norm
  - typically, perceptually poor
- Mahalanobis distance

  $$d_w(x,y)=[(x-y)^T W(x-y)]^{1/2}, W>0$$

  - More degrees of freedom
  - Perceptually better similarity measure
  - Also a metric (useful in indexing)
  - $W = ?$

Learning Optimal W

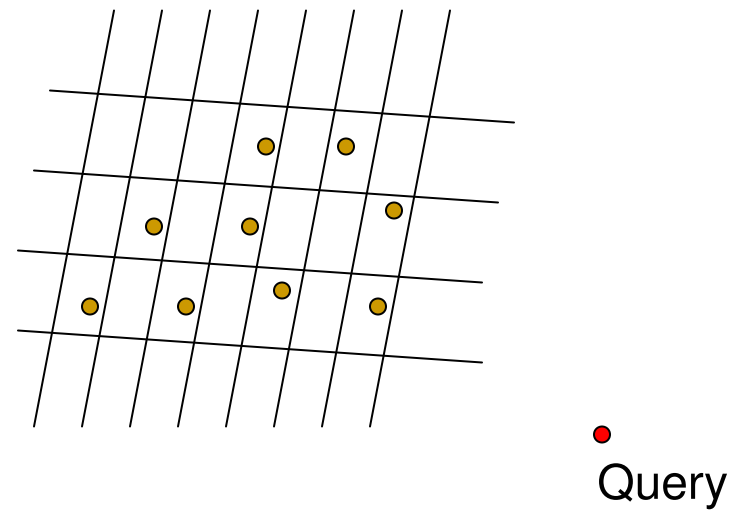
- Useful W possibly independent of database
- Users mark relevant / irrelevant results
- System learns from feedback
  - Update W for each user (Rui et. al. CVPR 2000)
  - Update W in batch mode (Davis et. al. ICML 2007)

Indexing for Relevance Feedback

- Normally, W is known prior to indexing
  - $W = U^T \Lambda U$
    - Rotate & skew data-set prior to indexing
    - Index new feature set using Euclidean distance
- W changes $\Rightarrow$ need to re-create index?
  - Most indexes fail to adapt

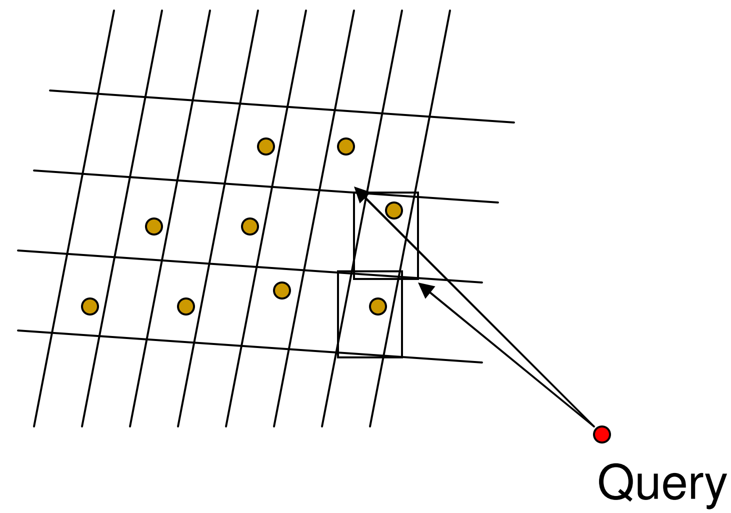
VA-File with Relevance Feedback

- Mahalanobis Distance $d_W(x,y)$
- W apriori unknown
  - estimated from user feedback
- Cells now skewed and rotated
  - Distance bounding complicated
  - $O(Nd^3)$ calculations



VA-File with Relevance Feedback

- Mahalanobis Distance $d_W(x,y)$
- W apriori unknown
  - estimated from user feedback
- Cells now skewed and rotated
  - Distance bounding complicated
  - $O(Nd^3)$ calculations
- Fit bounding rectangles cells on skewed cells (Sakurai et. al. VLDB 2001)
  - Bound distance to bounding rectangles
  - $O(Nd)$ calculations
  - Bounding rectangles overlap
    - ⇒ Distance bounds loosened
    - ⇒ (possibly) more disk accesses



Can clustering support relevance feedback?

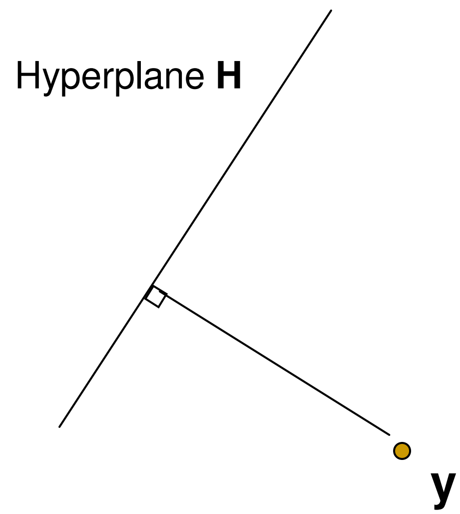
Can clustering support relevance feedback?

- Bound cluster-distance with changing W?

Point-to-hyperplane Distance

$$H(\mathbf{a}, b) = \{\mathbf{x} : \mathbf{a}^T\mathbf{x} + b = 0\}$$

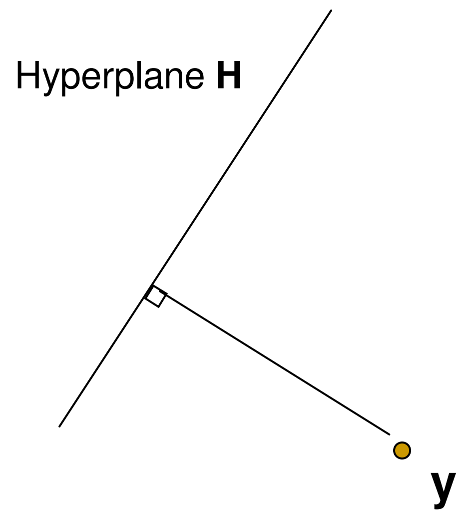
$$d_W(\mathbf{y}, H) = \min_{\mathbf{x} \in H} d_W(\mathbf{y}, \mathbf{x}) = \frac{|\mathbf{a}^T\mathbf{y} + b|}{\sqrt{\mathbf{a}^T W^{-1}\mathbf{a}}}$$



Point-to-hyperplane Distance

$$H(\mathbf{a}, b) = \{\mathbf{x} : \mathbf{a}^T\mathbf{x} + b = 0\}$$

$$\begin{aligned} d_W(\mathbf{y}, H) &= \min_{\mathbf{x}\in H} d_W(\mathbf{y}, \mathbf{x}) \\ &= \frac{|\mathbf{a}^T\mathbf{y} + b|}{\sqrt{\mathbf{a}^T W^{-1}\mathbf{a}}} \end{aligned}$$



Useful Invariance Property:

$$\frac{d_{W_1}(\mathbf{y}, H)}{d_{W_2}(\mathbf{y}, H)} = \sqrt{\frac{\mathbf{a}^T W_2^{-1}\mathbf{a}}{\mathbf{a}^T W_1^{-1}\mathbf{a}}}$$

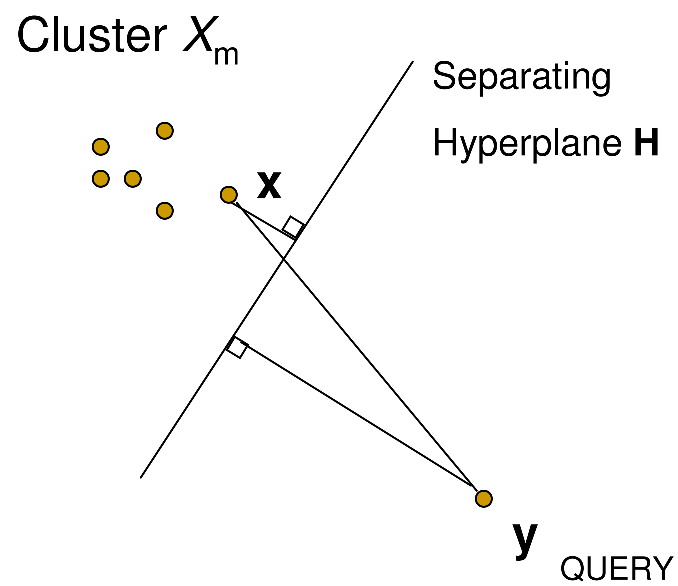
Bounding Query-Cluster Distance

Euclidean distance:

$$d(\mathbf{y},\mathbf{x}) \quad \geq \quad d(\mathbf{y},H) + d(\mathbf{x},H)$$

Mahalanobis distance:

$$d_W(\mathbf{y},\mathbf{x}) \quad \geq \quad d_W(\mathbf{y},H) + d_W(\mathbf{x},H)$$

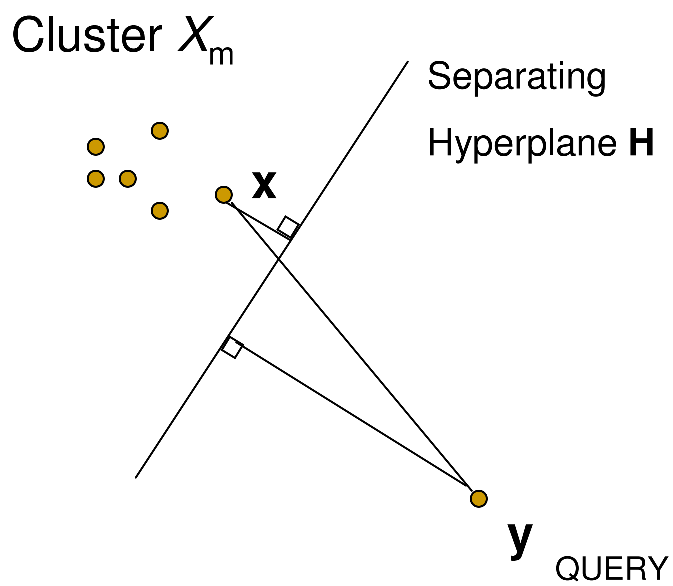


Bounding Query-Cluster Distance

$$d(\mathbf{y}, \mathbf{x}) \quad \geq \quad d(\mathbf{y}, H) + d(\mathbf{x}, H)$$

$$d_W(\mathbf{y}, \mathbf{x}) \quad \geq \quad d_W(\mathbf{y}, H) + d_W(\mathbf{x}, H)$$

$$\begin{aligned} \Rightarrow d_W(\mathbf{y}, X_m) &= \min_{X_m} d_W(\mathbf{y}, \mathbf{x}) \\ &\geq d_W(\mathbf{y}, H) + \min_{X_m} d_W(\mathbf{x}, H) \end{aligned}$$

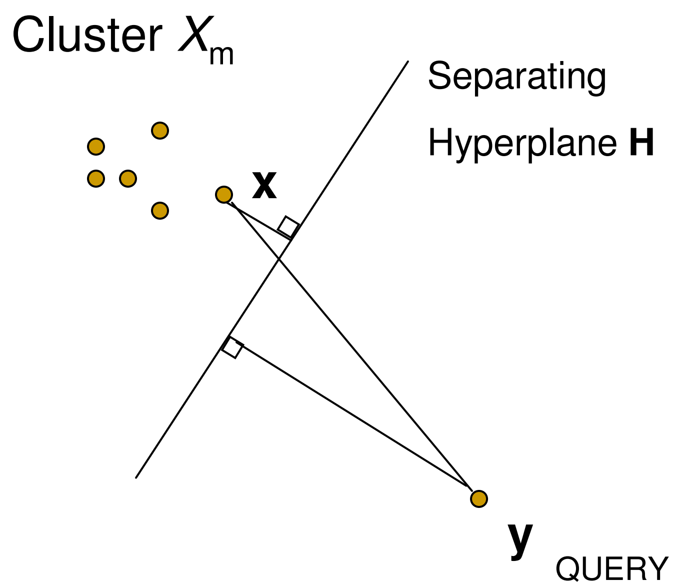


Bounding Query-Cluster Distance

$$d(\mathbf{y}, \mathbf{x}) \quad \geq \quad d(\mathbf{y}, H) + d(\mathbf{x}, H)$$

$$d_W(\mathbf{y}, \mathbf{x}) \quad \geq \quad d_W(\mathbf{y}, H) + d_W(\mathbf{x}, H)$$

$$\begin{aligned} \Rightarrow d_W(\mathbf{y}, X_m) \quad &= \quad \min_{X_m} d_W(\mathbf{y}, \mathbf{x}) \\ &\geq \quad d_W(\mathbf{y}, H) + \min_{X_m} d_W(\mathbf{x}, H) \end{aligned}$$



Distance-to-cluster ≥ Query-Hyperplane distance + Cluster-Hyperplane Distance

("Support")

Support evaluated offline & stored

- W changes ⇒ re-evaluate support?

Adaptive Support Estimation

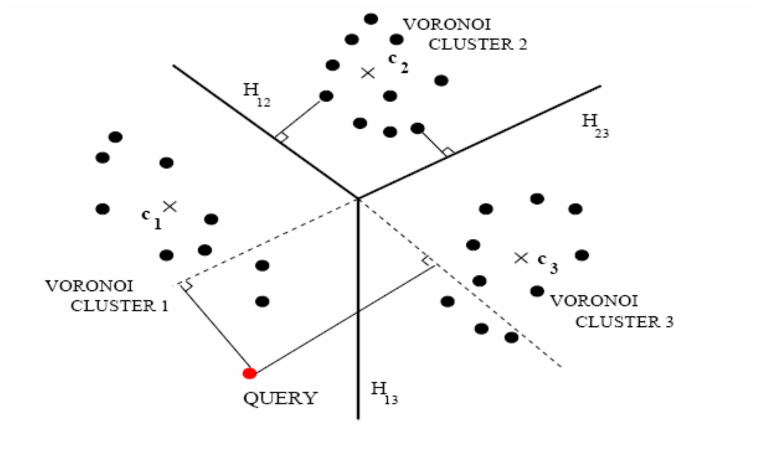
Invariance property $\Rightarrow d_W(x,H) = \sqrt{\frac{\mathbf{a}^T W_0^{-1}\mathbf{a}}{\mathbf{a}^T W^{-1}\mathbf{a}}} d_{W_0}(x,H)$

$$\Rightarrow \min_{X_m} d_W(x,H) = \sqrt{\frac{\mathbf{a}^T W_0^{-1}\mathbf{a}}{\mathbf{a}^T W^{-1}\mathbf{a}}} \min_{X_m} d_{W_0}(x,H)$$

$$d_W(X_m,H) = \sqrt{\frac{\mathbf{a}^T W_0^{-1}\mathbf{a}}{\mathbf{a}^T W^{-1}\mathbf{a}}} d_{W_0}(X_m,H)$$

New support can be found without accessing X_m!

Adaptive Cluster Distance Bounding



- Bound distance with multiple hyperplanes
  - use tightest distance bound
- Cluster boundaries are linear
  - use them as (separating) hyperplanes
  - no need to store hyperplanes

Proposed Indexing Scheme

1. Cluster data-set through VQ/K-means
   - "nearest neighbor" partitioning for linear boundaries
   - evaluate cluster "supports" for current W
2. Bound query-cluster distance with hyperplane bound
   - change in W $\Rightarrow$ scale cluster support
3. Retrieve clusters in order of distance
   - IF kNN distance so far < distance to next cluster

     STOP (kNNs found)

     ELSE read next cluster (till all clusters read)

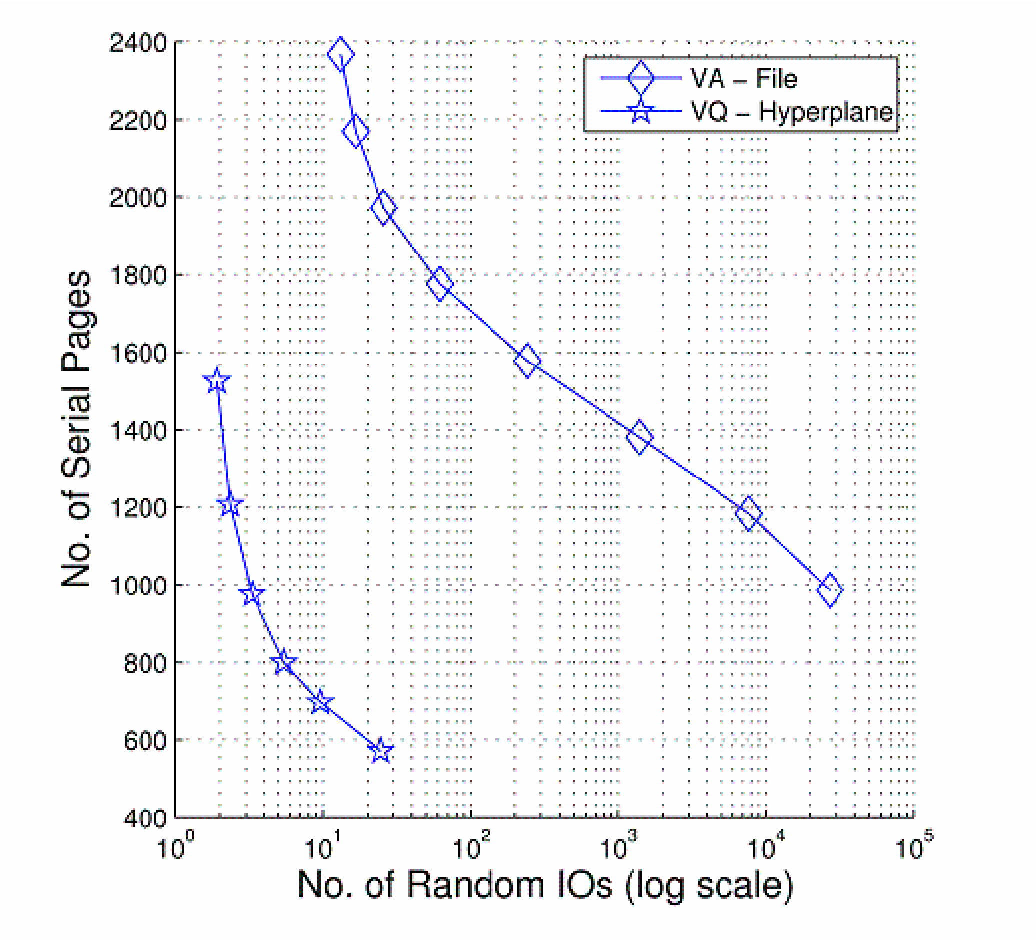
Set-up of Experiments

- Data-set BIO-RETINA
  - Gabor texture features extracted from feline retina
    - 208,506 elements
    - 62 dimensional
  - Clustered with squared Euclidean distance
- 10NN queries mined
- Search with new $W=U^T \Lambda U$
  - U – orthonormal matrix randomly generated
  - Λ – eigenvalues uniformly distributed [0,10]
- Compared VQ/Clustering vs. VA-File

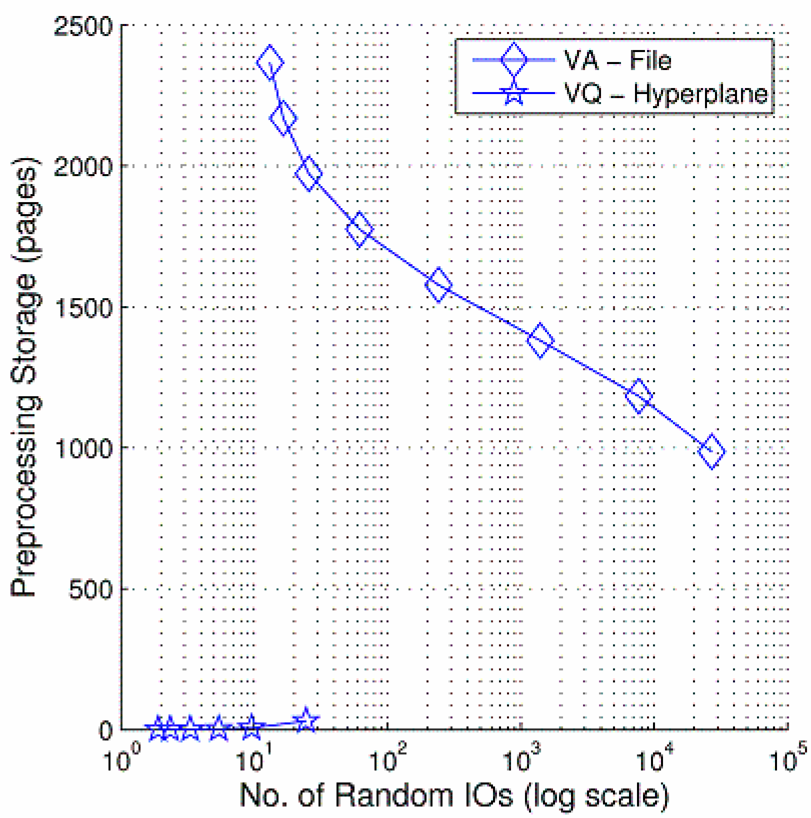
Experiments

- Data retrieved in blocks/pages
  - Page size – 8 kB $\Rightarrow$ 34 feature vectors per page
- Disk access: Sequential vs. Random disk IOs
  - Every access = 1 random IO + rest serial IOs
- Parameter varied
  - VA-File quantization bits per dimension (3-12 bits)
  - Number of clusters in cluster-distance bounding (10-600 clusters)

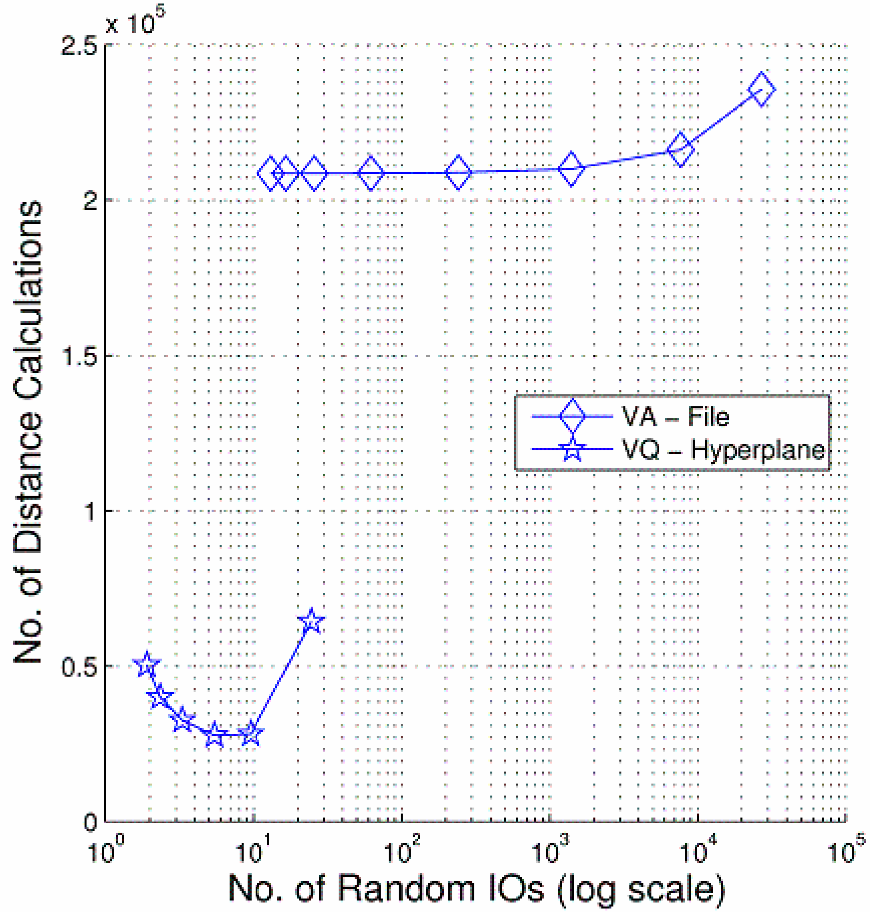
Results: IO Performance



Results: Preprocessing Storage



Results: Computational Costs



Summary

- Indexing image databases is a challenge
- Mahalanobis distance $d_W(x,y)$
  - perceptually better
  - trained/tuned with user/relevance feedback
- Relevance feedback complicates indexing
- Derived distance ratio invariance property
  - combined with cluster-distance bounding
- Proposed index outperforms VA-File
  - lower IO, storage and computation costs