## Hyperconvexity and Metric Embedding

**David Eppstein** Univ. of California, Irvine Computer Science Department

### Hyperconvex metric spaces

Defined by Helly property on metric balls Include trees,  $L_{\infty}$  metrics, other interesting geometric spaces

### Tight span

Embed any metric space into a hyperconvex space "Convex hull" for metric spaces Bounded faces of a Euclidean convex polytope

### Algorithms

Planar tight spans and Manhattan embedding Minimum dilation stars



## Outline

Hyperconvexity

Tight spans

### Planar tight span construction

### Minimum dilation stars

CC-BY-NC photo by Makz on Flickr, http://www.flickr.com/photos/makz/119762730/



# Outline

#### Hyperconvexity

Tight spans

Planar tight span construction

### Minimum dilation stars

CC-BY-NC photo by Makz on Flickr, http://www.flickr.com/photos/makz/119762730/

### Helly's theorem [Helly 1923]

Given a family of convex objects in d-dimensional Euclidean space:

If each (d+1)-tuple has a common intersection, so does the whole family



## k-Helly family

Any family of sets such that, for any subfamily, if all k-tuples in the subfamily intersect, then the whole subfamily has a common intersection

(Like convex sets in (k-1)-dimensional Euclidean space)

## Helly family

Special case of a 2-Helly family

(the simplest nontrivial case of a k-Helly family, like intervals of the real line)

## Hyperconvex (aka injective) metric spaces:

[Aronszajn and Panitchpakdi 1956; Isbell 1964]

Balls form a Helly family, and

If two balls have radii adding to at least their center separation, they intersect (equivalently, any two points have a geodesic connecting them)



Euclidean plane: not hyperconvex



L<sub>1</sub> plane: hyperconvex

## **Example: Chebyshev distance**

Points = d-dimensional vectors of real numbers Distance = maximum coordinatewise difference ( $L_{\infty}$  norm)



#### Metric balls = axis-aligned cubes Family of balls has common intersection iff they intersect in each coordinate

GFDL image by Nevit Dilmen on Wikimedia commons, http://commons.wikimedia.org/wiki/File:1000\_cubes.jpg

# Example: Paris metric (hedgehog space)

Points = plane with polar coordinates

All geodesics follow rays through origin

Distance( $(\rho_1, \theta_1)$ ,  $(\rho_2, \theta_2)$ ) =  $|\rho_1 - \rho_2|$  if  $\theta_1 = \theta_2$ , =  $|\rho_1| + |\rho_2|$  otherwise





CC-BY-SA images by Mortadelo2005 and Benh LIEU SONG on Wikimedia commons, http://commons.wikimedia.org/wiki/File:Area\_metropolitana\_paris.png and http://commons.wikimedia.org/wiki/File:Paris\_Night.jpg

Hyperconvexity and metric embedding

D. Eppstein, 2009

# **Example: Real trees**



Metric spaces in which

- Any two points are endpoints of a unique simple curve
- Curve length = distance

Include:

- Paris metric
- Acyclic connected undirected graphs, with edges replaced by line segments
- Diffusion-limited aggregation

"3D Diffusion Limited Aggregation", CC-BY-SA-NC image by Simon Chorley on Flickr, http://www.flickr.com/photos/mylaboratory/363532702/

Hyperconvexity and metric embedding

D. Eppstein, 2009

### Example: Manhattan orbifolds [E., arXiv:math/0612109]



Space locally modeled on the Manhattan ( $L_1$  or  $L_\infty$ ) plane

Topologically, must be a simply-connected 2-manifold

Geometrically, most points must have neighborhood isometric to L<sub>1</sub> neighborhood of the origin

Discrete "cone points" allowed (e.g. cone of order 3 on boundary of  $L_1$  orthant, depicted)

All cone points must have order > 4

## Manhattan orbifolds are hyperconvex

Main idea: study two different types of polygon

Orthogonal polygon (shown): boundary locally looks like slope-1 line segment Must have  $\ge 4$  interior angles of  $\pi/2$  (Gauss-Bonnet)

Orthoconvex: all interior angles π/2, can include boundary of manifold All metric balls are orthoconvex Orthoconvex regions have Helly property (analyze possible configurations of crossings, use orthogonal polygon angle count to rule out bad cases)



D. Eppstein, 2009

# Manhattan orbifold approximation to hyperbolic plane

(analogous to Manhattan plane as an approximation to Euclidean)

Form tesselation by quadrilaterals meeting five to a vertex Fill each quadrilateral by an  $L_1$  metric square





# Outline

#### Hyperconvexity

### Tight spans

### Planar tight span construction

### Minimum dilation stars

CC-BY-NC photo by Makz on Flickr, http://www.flickr.com/photos/makz/119762730/

## The big picture:

Any metric space (X,d) can be embedded without distortion into a minimal hyperconvex space, its tight span

Each point of the tight span represents embedding of (X,d) into a star metric (allowing distortion)



The 200-inch telescope at Mt. Palomar, DE, 08/01

D. Eppstein, 2009

## **Star metrics**

Defining property: there exists a hub h that is between every other pair of points



For all s and t, d(s,t) = d(s,h) + d(h,t)Once distances from the hub are known, all other distances are determined

### Star embedding

A map from a metric space (X,d) to a star may be represented as a function D from X to  $R^+$ (D(x) = distance from image of x to hub)

Embedding: a map that may distort but does not decrease distances Equivalently: for all p and q,  $D(p) + D(q) \ge d(p,q)$ 

> Technical condition for infinite X: |D(x) - d(x,0)| must be bounded for an arbitrary base point 0

The embedding is minimal if no hub distance D(x) may be decreased i.e., for all p,  $inf_q D(p) + D(q) - d(p,q) = 0$ 

Photo by underwaterguy on Flickr, http://www.flickr.com/photos/underwaterguy/2136586253/, CC-BY-NC

### Interstellar distances

Let D(x) and E(x) be two star embeddings (distances from x to hub centers)

Define Distance(D,E) =  $\sup_{x} |D(x) - E(x)|$ 

The original metric space (X,d) embeds without distortion into the space of minimal stars:

Map point x to function  $D_x(y) = d(x,y)$ 

Distance( $D_x, D_y$ ) = sup<sub>z</sub> | d(x,z) - d(y,z) | = d(x,y) by triangle inequality for d

http://xkcd.com/482/

Hyperconvexity and metric embedding



# Hyperconvexity of minimal stars

The distance on star embeddings is just the  $L_{\infty}$  metric on the functions from X to R

Helly: intersection of pairwise intersecting balls may be found coordinatewise

Constraint  $D(x) + D(y) \ge d(x,y)$ defines an (unbounded) convex set, within which Helly property still holds

Minimal stars lie on bounded faces of this polytope and still have Helly property: If some non-minimal embedding D belongs to a set of balls centered at minimal stars, decrease distances in D to find a minimal embedding in the intersection

Ghost quartz, CC-BY-SA photo by Didier Descouens, http://commons.wikimedia.org/wiki/File:QuartzFantome.jpg Hyperconvexity and metric embedding D. Eppstein, 2009 The tight span of a metric space (X,d) [Isbell 1964, Dress 1984, Chrobak & Larmore 1994]

Equivalently (up to isometry), may be defined as:

1. The hyperconvex metric space of minimal star embeddings of (X,d)

- 2. The points on the bounded faces of the polytope  $D(x) + D(y) \ge d(x,y)$ (with the  $L_{\infty}$  metric)
- 3. The tropical convex hull of the distance vectors [Develin & Sturmfels 2004]

4. The smallest hyperconvex superspace

(if (X,d) is a subspace of any hyperconvex space, so is the tight span)

5. The injective hull in the category of metric spaces and distance-decreasing functions

### Tight spans in the Manhattan plane

Orthogonal convex set: intersects any axis-parallel line in an interval

Orthogonal convex hull: Intersection of all orthogonally convex connected supersets [Montuno & Fournier 1982; Nicholl et al 1983; Ottman et al 1984; Karlsson & Overmars 1988]

> For any subset of the L<sub>1</sub> plane with a connected orthogonal hull, tight span = orthogonal convex hull

If orthogonal hull is not connected, tight span = components of hull connected together by monotone curves


# Outline

Hyperconvexity

Tight spans

### Planar tight span construction

### Minimum dilation stars

CC-BY-NC photo by Makz on Flickr, http://www.flickr.com/photos/makz/119762730/



Easy warm-up problem:

Given a finite metric space can it be represented as distances in Euclidean plane?

Find 3 non-collinear points Place them in a triangle Use distances from these points to uniquely place each other point

Then verify all distances correct

Total time O(n<sup>2</sup>) linear in size of input distance matrix

Much additional research exists when distance matrix is partial or corrupted...

Same problem for Manhattan plane  $(L_1 \text{ or } L_\infty \text{ distance})$ :

Given a finite metric space can it be represented as a set of points in the L<sub>1</sub> plane?

More difficult because of ambiguities: three points do not uniquely determine the location of the rest

Edmonds [Disc. Comput. Geom. 2008]: algorithm with O(n<sup>2</sup> log<sup>2</sup> n) time

New result: O(n<sup>2</sup>), linear in input size based on finding planar tight span then testing whether span has Manhattan embedding



## Main ideas of algorithm:

Build up metric space by adding points one by one

Maintain tight span of added points as a complex of rectangles with L<sub>1</sub> geometry (not necessarily edge-to-edge)

Characterize the rectangle complexes that have Manhattan embeddings



Piet Mondrian, 1921

# Important special case: tight span of a geodesic path pq and one added point s

Always has an L<sub>1</sub> plane embedding:

x = (d(p,r) + d(r,s) - d(p,s))/2y = (d(p,s) - d(p,r) + d(r,s))/2

If distances from s form a sawtooth (function with derivative +/- 1) as function of path position, then can be represented as a rectangular complex

#rectangles = #sawtooth breakpoints - 1

S

## Using landmarks to measure distances within rectangles

Each rectangle corner  $c_i$  stores a site  $s_i$  having c as closest point in the rectangle

If r is in rectangle and p is anywhere,  $dist(p,r) = max_i d(p,s_i) - d(r,c_i) - d(c,s_i)$ 

Same idea works with two landmarks for edges of complex



Hyperconvexity and metric embedding

D. Eppstein, 2009

## Planar tight span algorithm

maintain a rectangular complex with landmarks (initially empty)

for each point p of the metric space (in an arbitrary order):

compute distances to p within each cell of the complex

if p is at zero distance to some point of complex: add it there, verify distances to other sites

else if local minimum of distance is interior to a rectangle: tight span is not planar

else:

use distance computations to determine attachment path compute tight span of path+point and add to complex

### Which planar tight spans have a Manhattan embedding?

Several easily-checked necessary and sufficient conditions, amounting to:

- Each point of the rectangle complex must be locally Manhattan

- Each biconnected component of the complex must have exactly four extreme boundary edges (both endpoints convex)

- Articulation points must lie on boundary edges
- Any biconnected component must have at most one neighboring component in each of its four directions (biconnected neighbors take up two directions each)

### Planar tight span summary

The number of new rectangles created by each point insertion is balanced by the number of vertices removed from the boundary

Therefore, total size of the rectangular complex = O(n)

Inserting each point takes time proportional to size of updated complex Therefore, total time to create tight span =  $O(n^2)$ , linear in input size

Testing whether the span has a Manhattan embedding also takes O(n)



# Outline

Hyperconvexity

Tight spans

Planar tight span construction

### Minimum dilation stars

CC-BY-NC photo by Makz on Flickr, http://www.flickr.com/photos/makz/119762730/



# How to measure the quality of a metric embedding?

Dilation: how much farther you would have to travel in the new metric

Scale the new metric so that all distances are at least as large as they were in the old metric

Dilation = maximum ratio new distance / old distance among all pairs of input points

(factor by which some distances grow)

Photo by mbk on Flickr, http://www.flickr.com/photos/mbk/2782424469/, CC-BY-NC-SA

### **Minimum dilation star problem** [E., Wortman] Given an n by n matrix D[s,t] of distances in a metric space

Produce a vector H[s] of distances from each input point to a new hub

Satisfying, for all s and t,  $D[s,t] \le H[s] + H[t]$ (scale output distances to be at least as big as input)

> Minimize max<sub>s,t</sub> (H[s] + H[t])/D[s,t] (find the star with optimal dilation)

[s] + H[t])/D[s,t] optimal dilation)

We solve this in **strongly polynomial time** O(n<sup>3</sup> log<sup>2</sup> n)

Photo by underwaterguy on Flickr, http://www.flickr.com/photos/underwaterguy/2136586253/, CC-BY-NC

D. Eppstein, 2009

## Related work: Euclidean min dilation star [E & W, SoCG'05]



Given n points in Euclidean plane or higher dimensional space

Find a hub within that space minimizing dilation of star network with Euclidean distances as lengths

Equivalently, maximize min eccentricity of ellipses passing through the hub having pairs of input points as foci

O(n log n) in any fixed dimension if hub can be any point of the space

O(n  $2^{\alpha(n)} \log^2 n$ ) if hub must be one of the input points (2d only)

### The difference between geometric and metric stars

Example: input is an equilateral triangle (three points with equal distances)





Euclidean min dilation star dilation = 1.155 Metric min dilation star dilation = 1

# Solution ideas, I: Express MDS as a linear program

```
Find H[x] and \Delta
```

```
Satisfying

H[x] \ge 0

H[x] + H[y] \ge distance(x,y)

H[x] + H[y] \le \Delta distance(x,y)

for all x and y
```



Minimizing  $\Delta$ 

### Not in form for known strongly-polynomial LP algorithms more than O(1) variables total more than two variables in some inequalities

Image by Sdo on Wikimedia, http://commons.wikimedia.org/wiki/File:Simplex-method-3-dimensions.png, CC-BY-SA

## Solution ideas, II: Characterize LP basis

Any even cycle in original metric space lower-bounds the dilation: dilation  $\geq$  (sum of even edge lengths) / (sum of odd edge lengths)

(because in star, both sums are forced to equal each other)

Optimal dilation turns out to equal worst cycle of this type



# Solution ideas, III: transform to a graph problem

Parametric negative cycle detection:

digraph where edge weights are linear functions of a parameter  $\lambda$ , find smallest value of  $\lambda$  such that all cycle lengths are non-negative

Optimal  $\lambda$  = optimal dilation  $\Delta$  of original metric embedding problem



# Solution ideas, IV: strongly polynomial solution for parametric negative cycle detection problem

Megiddo's parametric search [Megiddo, JACM 1983]:

Simulate parallel algorithm for optimization problem as if it were given the optimal parameter value as input using a decision algorithm to help simulate each branch step

Simulated algorithm [Savage, Ph.D. thesis 1977]:

Compute all pairs shortest paths by repeated matrix squaring

Decision algorithm for comparing parameter value to optimum:

Bellman-Ford, detect negative cycles in non-parametric graph

### Solution ideas, V: cycle detection details

Store a matrix of piecewise linear parametric functions

Represents lengths of paths with at most 2<sup>i</sup> hops

Initially: i = 0, matrix stores the edge length functions

Repeat log n times, until  $2^i > n$ :

Square in (min,+) matrix arithmetic to increment i

 $O(n^3 \log n)$  time to combine piecewise linear functions

Binary search for the optimal parameter value among the breakpoints of the path-length functions simplifying matrix entries back to non-piecewise functions

O(log n) calls to Bellman-Ford

### Solution ideas VI: finding the actual embedding

Now that we know the correct dilation...

Plug it into the same parametric graph

Add an extra "source" vertex to the graph

Compute distances to all other vertices (Bellman-Ford)

H[x] = (half of) difference between two distances from source to two vertices representing x

Some algebra + triangle inequality shows this is a valid embedding

Lower bound shows this is the optimal embedding

## Conclusions

Hyperconvexity is...

- ...as central to metric spaces as convexity is to Euclidean spaces
   ...important for the development of approximation algorithms
   already used as such in k-server online algorithms
   ...a unifying point of view for finding metric embeddings
   exact embeddings into Manhattan plane
   optimal-dilation embeddings into stars
  - ...an interesting subject for more algorithmic research