Representing and Matching Molecular Surfaces

Thomas Funkhouser
Princeton University
CS597A, Fall 2005

Outline

Molecular surfaces
- Definitions
- Representations

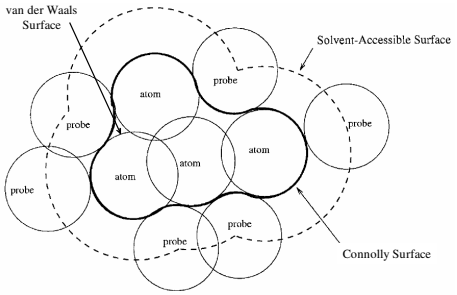
Surface matching
- Continuous surface mappings
- Discrete point correspondences

Surface retrieval
- Shape descriptors

Results

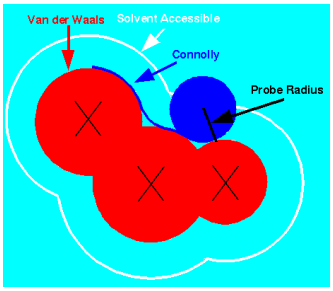
Discussion

Molecular Surfaces



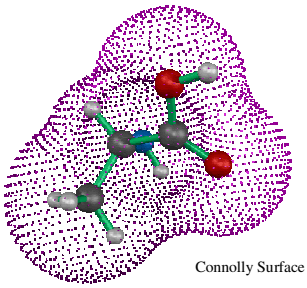
[Cai98]

Molecular Surfaces



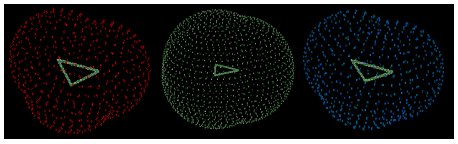
http://www.simbiosys.ca/sprout/eccc/cangaroo.html

Molecular Surfaces



http://www.netsci.org/

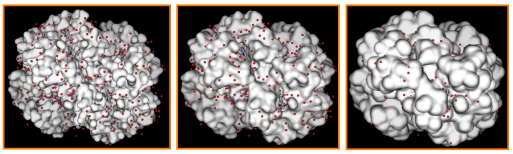
Molecular Surfaces



van der Waals Solvent accessible Connolly

http://www.simbiosys.ca/sprout/eccc/cangaroo.html

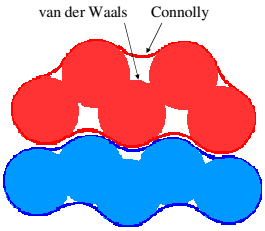
Molecular Surfaces



van der Waals | Connolly | Solvent accessible

http://www.chemaxon.com/shared/MarvinSpace/gallery.html

Molecular Surfaces



Connolly surfaces of bound molecules are complementary

http://www.netsci.org/

Outline

Molecular surfaces
- Definitions

ØRepresentations

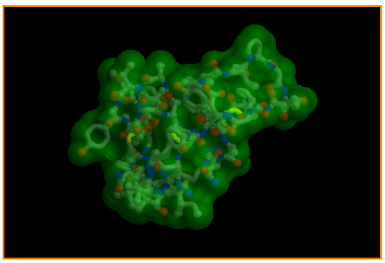
Surface matching
- Continuous surface mappings
- Discrete point correspondences

Surface retrieval
- Shape descriptors

Results

Discussion

Molecular Surface Representation

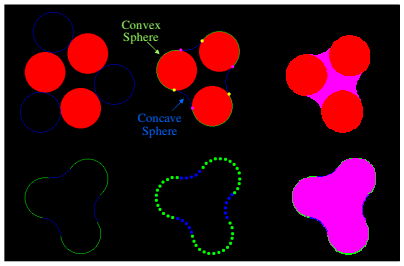


How should we store the surface description in a computer?

http://www.biohedron.com

Molecular Surface Representations

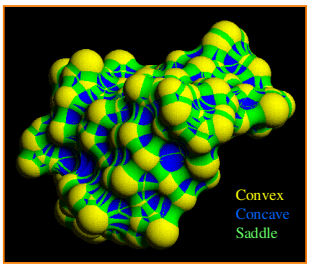
Union of partial spheres and tori



http://www.netsci.org/

Molecular Surface Representations

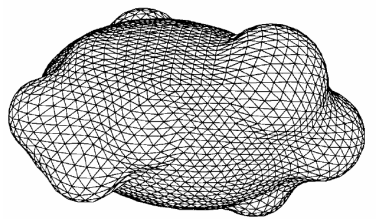
Union of partial spheres and tori



[Connolly83]

Molecular Surface Representations

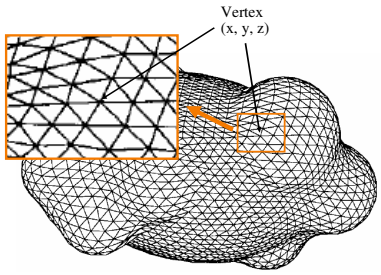
Triangle mesh



[Cai98]

Molecular Surface Representations

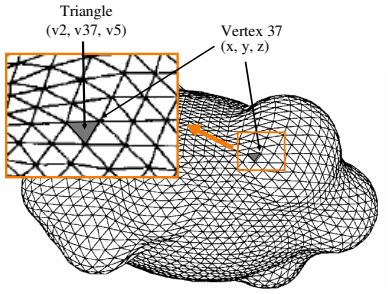
Triangle mesh



[Cai98]

Molecular Surface Representations

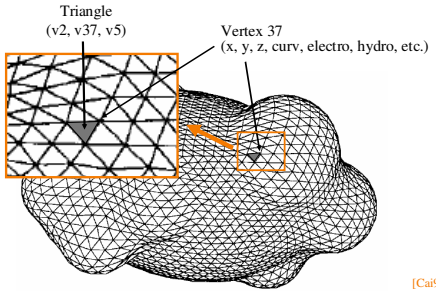
Triangle mesh



[Cai98]

Molecular Surface Representations

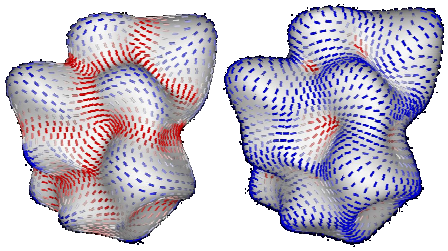
Triangle mesh (with properties at every vertex)



[Cai98]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)

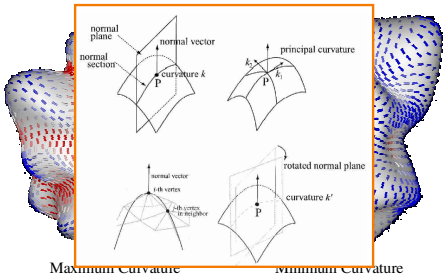


Maximum Curvature Minimum Curvature

[Duncan93]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)

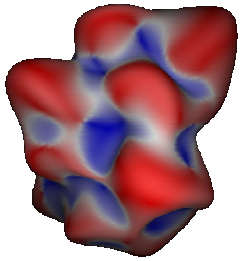


Maximum Curvature Minimum Curvature

[Kinoshita03]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)

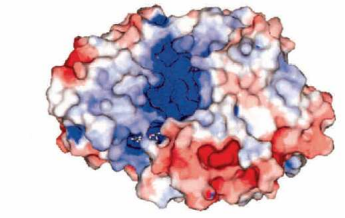


Mean Curvature

[Duncan93]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)

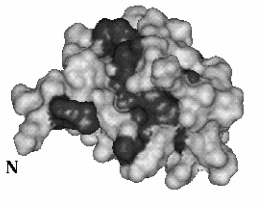


Electrostatic Potential

[Kinoshita03]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)

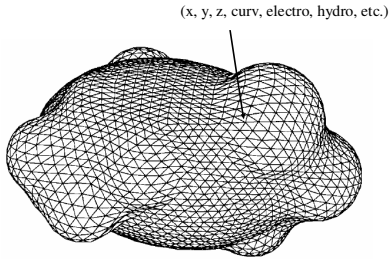


Hydrophobicity

Exposure of hydrophobic surface before binding of calcium to calmodulin. [CaBP Data Library]

Molecular Surface Representations

Triangle mesh (with properties at every vertex)



[Cai98]

Outline

Molecular surfaces
- Definitions
- Representations

Surface matching ←
- Continuous surface mappings
- Discrete point correspondences

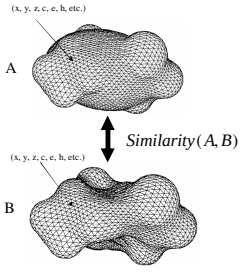
Surface retrieval
- Shape descriptors

Results

Discussion

Surface Matching

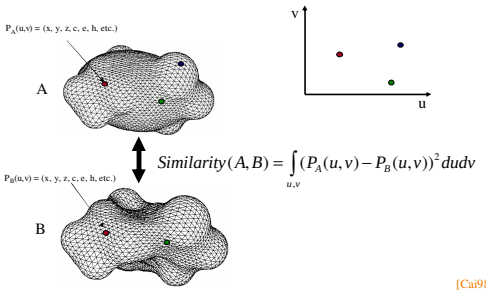
Goal: compute the similarity between two molecular surfaces with attributes at vertices



[Cai98]

Surface Matching

Challenge: find a consistent parameterization

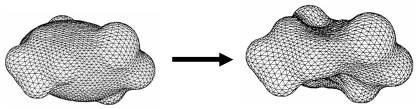


$$Similarity(A,B) = \int_{u,v} (P_A(u,v) - P_B(u,v))^2 dudv$$

[Cai98]

Continuous Surface Mappings

Map one surface onto the other, and measure the cost of the mapping

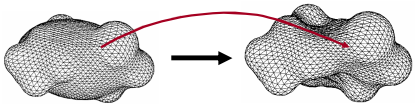


[Cai98]

Continuous Surface Mappings

Earth mover's distance

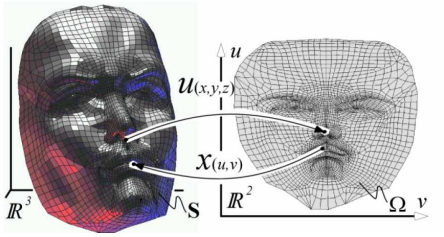
- Amount of "work" required to deform A onto B



[Rubner00] [Cai98]

Continuous Surface Mappings

Map all surfaces to planar parameterization

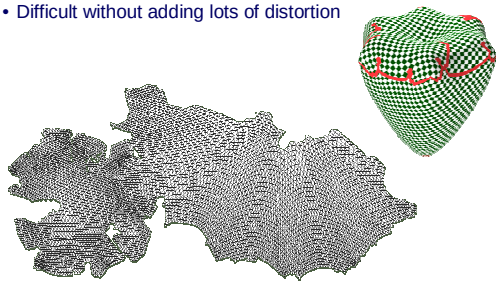


[Sheffer]

Continuous Surface Mappings

Map all surfaces to planar parameterization

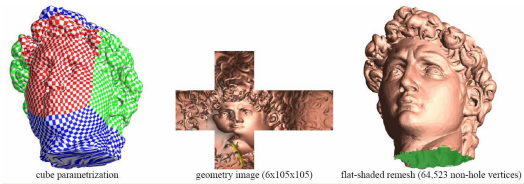
- Difficult without adding lots of distortion



[Sorkine02]

Continuous Surface Mappings

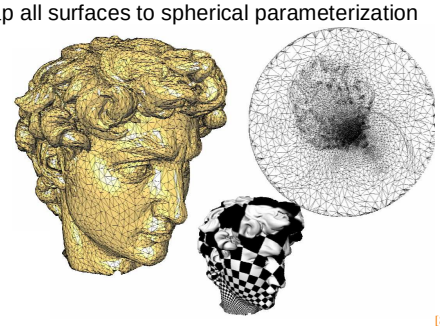
Map all surfaces to cubic parameterization



[Praun03]

Continuous Surface Mappings

Map all surfaces to spherical parameterization



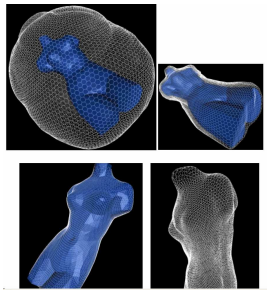
[Sheffer]

Continuous Surface Mappings

Harmonic map to spherical parameterization

- Minimizes

$$\frac{d^2P_i}{dt^2} = -\gamma\frac{dP}{dt} + \vec{F}_{ext} + \vec{F}_{int}$$

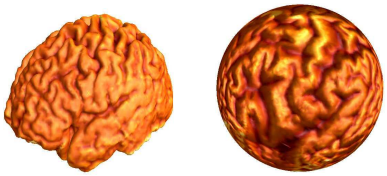


[Shum][Kalnins00]

Continuous Surface Mappings

Conformal map to spherical parameterization

- Minimizes distortion to angles

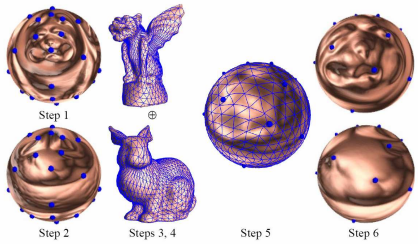


http://www.cise.ufl.edu/~gu/yau_talk/

Continuous Surface Mappings

Consistent spherical parameterization

- Guarantees feature point correspondences

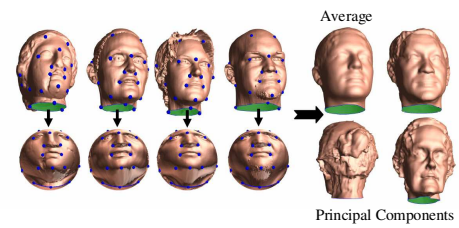


[Asirvatham05]

Continuous Surface Mappings

Consistent spherical parameterization

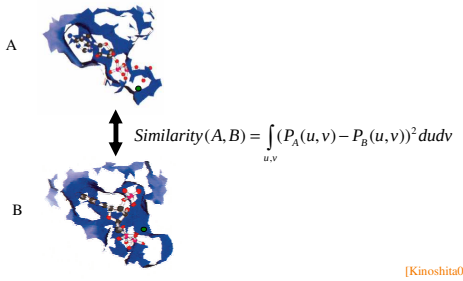
- Guarantees feature point correspondences



[Asirvatham05]

Continuous Surface Mappings

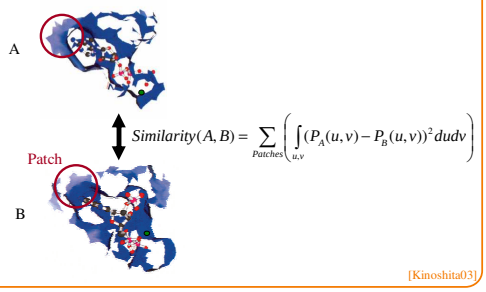
Partial matches and flexible surfaces don't fit well into this framework



[Kinoshita03]

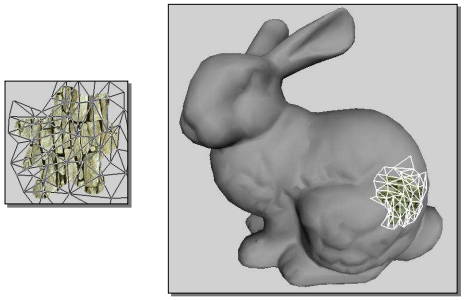
Piecewise Continuous Mappings

Use continuous surface matches for local patches



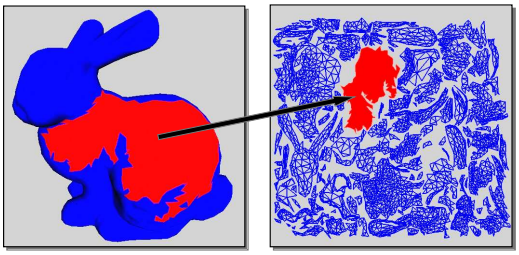
[Kinoshita03]

Piecewise Continuous Mappings



[Praun00]

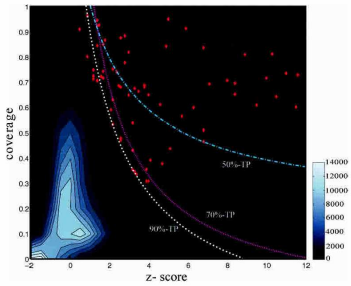
Piecewise Continuous Mappings



[Praun00]

Piecewise Continuous Mappings

Need to balance coverage vs. quality of match



[Kinoshita05]

Outline

Molecular surfaces
- Definitions
- Representations

Surface matching
- Continuous surface mappings

ØDiscrete point correspondences

Surface retrieval
- Shape descriptors

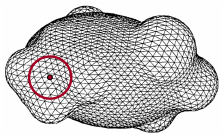
Results

Discussion

Surface Matching at Discrete Points

Sample the surfaces at discrete sets of points

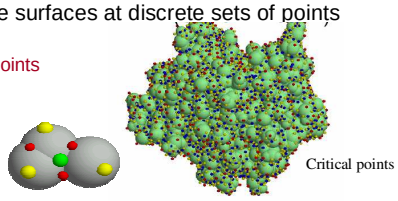
ØVertices
- Critical points
- Features
- etc.



Match surface patches near sampled points
- Association graphs
- Geometric hashing
- Iterative closest points

Surface Matching at Discrete Points

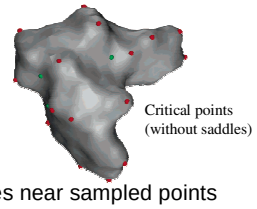
Sample the surfaces at discrete sets of points
- Vertices
- Critical points
- Features
- etc.



Match surface patches near sampled points
- Association graphs
- Geometric hashing
- Iterative closest points

Surface Matching at Discrete Points

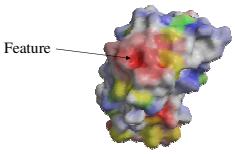
Sample the surfaces at discrete sets of points
- Vertices
- Critical points
- Features
- etc.



Match surface patches near sampled points
- Association graphs
- Geometric hashing
- Iterative closest points

Surface Matching at Discrete Points

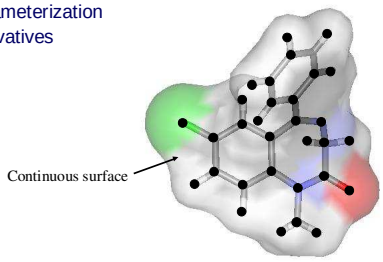
Sample the surfaces at discrete sets of points
- Vertices
- Critical points
- Features
- etc.



Match surface patches near sampled points
- Association graphs
- Geometric hashing
- Iterative closest points

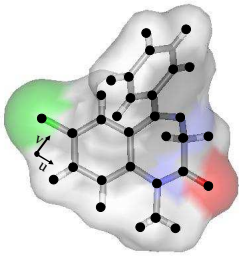
Surface Matching at Discrete Points

Differences from matching atoms/pseudo-centers?
- Point samples represent continuous set of points
- Parameterization
- Derivatives



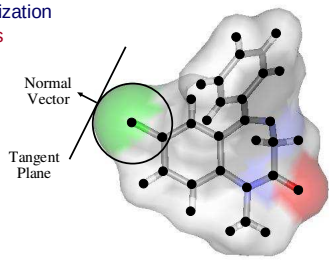
Surface Matching at Discrete Points

Differences from matching atoms/pseudo-centers?
- Point samples represent continuous set of points
- Parameterization
- Derivatives



Surface Matching at Discrete Points

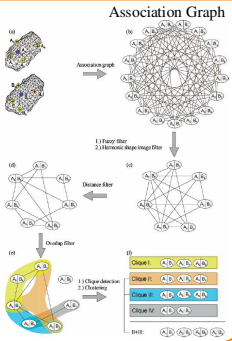
How is it different than matching a set of points?
- Point samples represent continuous set of points
- Parameterization
- Derivatives



Surface Matching at Discrete Points

Associate pairs of points only if they have the same:

- Intra-molecular distances
- Electrostatic potentials
- Lipophilic potentials
- Principal curvatures
- Harmonic shape images (geometric representation of local neighborhood shape)
- Relative orientations of harmonic shape images



[Hofbauer04]

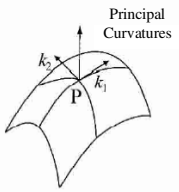
Surface Matching at Discrete Points

Associate pairs of points only if they have the same:

- Intra-molecular distances
- Electrostatic potentials
- Lipophilic potentials

ØPrincipal curvatures

- Harmonic shape images (geometric representation o local neighborhood shape)
- Relative orientations of harmonic shape images



[Kinoshita03]

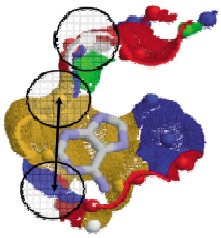
Surface Matching at Discrete Points

Associate pairs of points only if they have the same:

- Intra-molecular distances
- Electrostatic potentials
- Lipophilic potentials
- Principal curvatures

ØHarmonic shape images (geometric representation of local neighborhood shape)

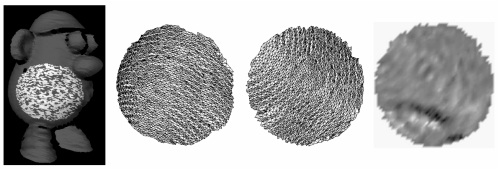
- Relative orientations of harmonic shape images



[Hofbauer04] [Shulman-Peleg04]

Surface Matching at Discrete Points

Harmonic shape image



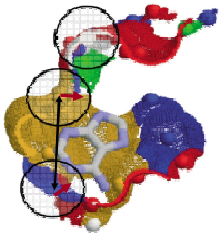
Patch | Mesh | Harmonic Image | Harmonic Shape Image

Surface Matching at Discrete Points

Associate pairs of points only if they have the same:

- Intra-molecular distances
- Electrostatic potentials
- Lipophilic potentials
- Principal curvatures
- Harmonic shape images (geometric representation of local neighborhood shape)

ØRelative orientations of harmonic shape images

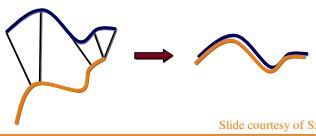


[Hofbauer04] [Shulman-Peleg04]

Surface Matching with ICP

Iterate until convergence:

1. Select source points (from one or both surfaces)
2. Correspond to closest points on other surface
3. Weight the correspondences
4. Reject outlier point pairs
5. Compute an error metric for the current transform
6. Minimize the error metric w.r.t. transformation

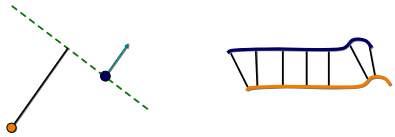


Slide courtesy of Szymon Rusinkiewicz

Surface Matching with ICP

Using point-to-plane distance for surfaces instead of point-to-point distance

- Lets flat regions slide along each other



Slide courtesy of Szymon Rusinkiewicz

Outline

Molecular surfaces
- Definitions
- Representations

Surface matching
- Continuous surface mappings
- Discrete point correspondences

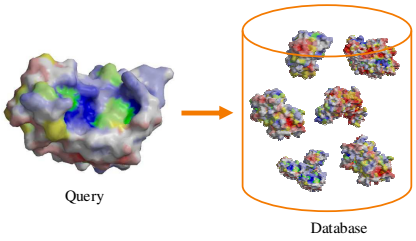
Surface retrieval
- Shape descriptors

Results

Discussion

Surface Retrieval

Goal: search a database of surfaces for the ones most similar to a query

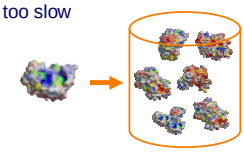


[Kinoshita03]

Surface Retrieval

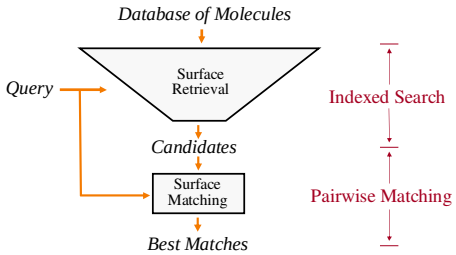
Requirements:
- Should be fast (indexed)
- Should be conservative (don't miss any good matches)
- Can be approximate (check best matches in more detail)

Observation:
- Finding explicit correspondences/mappings for every pair of surfaces is too slow



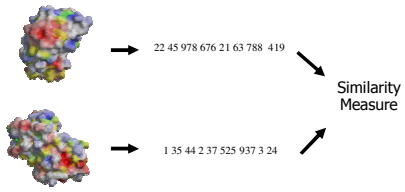
[Kinoshita03]

Surface Retrieval Pipeline



Surface Retrieval

General strategy:
- Compute shape descriptor for each surface
- Search for most similar shape descriptors



Surface Retrieval

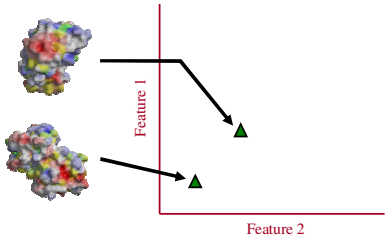
Shape descriptor:

- Quick to compute
- Indexible
- Concise
- Invariant to translation
- Invariant to rotation
- Insensitive to small features
- Discriminating

Shape Descriptors

Avoid finding point correspondences

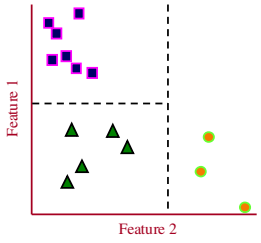
- Map surfaces to a feature space where correspondences are easy to find



Shape Descriptors

Avoid finding point correspondences

- Map surfaces to a feature space where correspondences are easy to find and classes are distinguishable

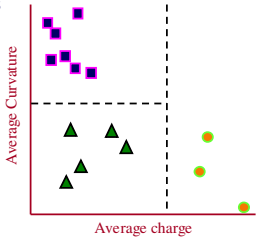


Shape Descriptors

Examples:

ØSet of geometric properties

- High-order moments
- Shape histograms
- etc.



Shape Descriptors

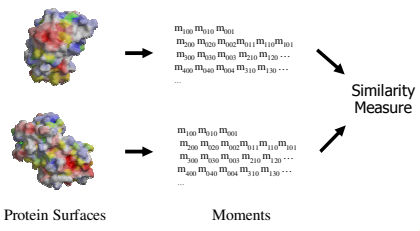
Examples:

- Set of geometric properties

ØHigh-order moments

- Shape histograms
- etc.

$$m_{pqr} = \int_{surface} x^p y^q z^r \, dxdydz$$



Shape Descriptors

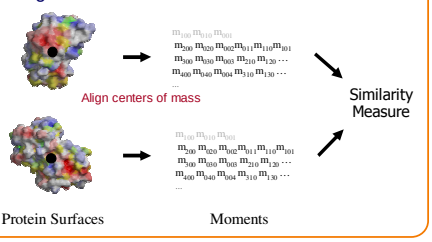
Examples:

- Set of geometric properties

ØHigh-order moments

- Shape histograms
- etc.

$$m_{pqr} = \int_{surface} x^p y^q z^r \, dxdydz$$



Shape Descriptors

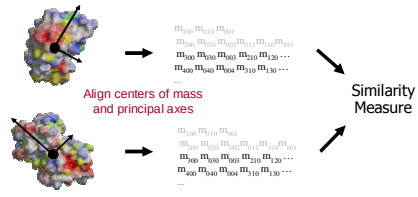
Examples:

- Set of geometric properties

ØHigh-order moments

- Shape histograms
- etc.

$$m_{pqr} = \int_{surface} x^p y^q z^r \, dxdydz$$



Protein Surfaces Moments

Shape Descriptors

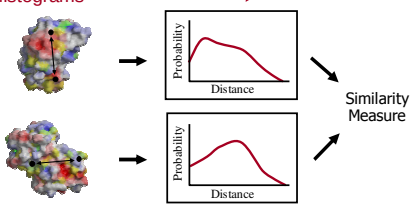
Examples:

- Set of geometric properties
- High-order moments

ØShape histograms

- etc.

These histograms are invariant under translations and rotations



Protein Surfaces Shape Histogram

[Osada01]

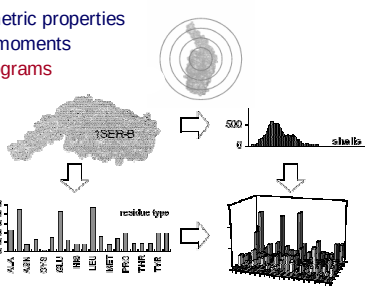
Shape Descriptors

Examples:

- Set of geometric properties
- High-order moments

ØShape histograms

- etc.



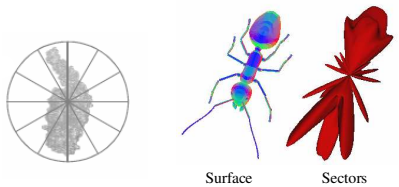
[Ankerst99]

Spherical Shape Descriptors

Examples:

ØShape histograms (sectors)

- Spherical extent function
- Extended Gaussian image
- Spherical attribute image
- etc.



Surface Sectors

[Ankerst99]

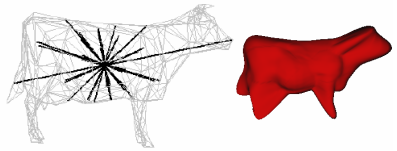
Spherical Shape Descriptors

Examples:

- Shape histograms (sectors)

ØSpherical extent function

- Extended Gaussian image
- Spherical attribute image
- etc.



Surface Spherical Extent Function

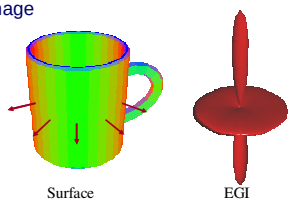
Spherical Shape Descriptors

Examples:

- Shape histograms (sectors)
- Spherical extent function

ØExtended Gaussian Image

- Spherical attribute image
- etc.



Surface EGI

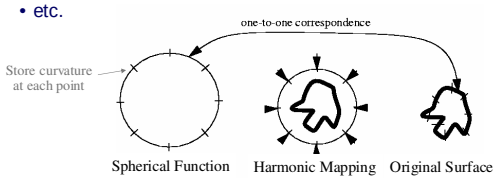
[Horn86]

Spherical Shape Descriptors

Examples:
- Shape histograms (sectors)
- Spherical extent function
- Extended Gaussian image

ØSpherical attribute image

- etc.



[Ikeuchi95]

Outline

Molecular surfaces
- Definitions
- Representations

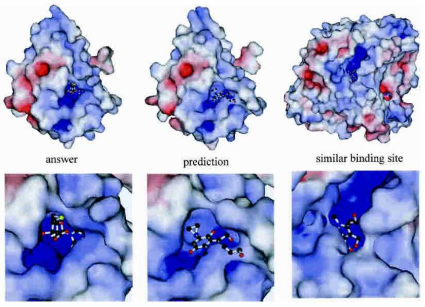
Surface matching
- Continuous surface mappings
- Discrete point correspondences

Surface retrieval
- Shape descriptors

Results ←

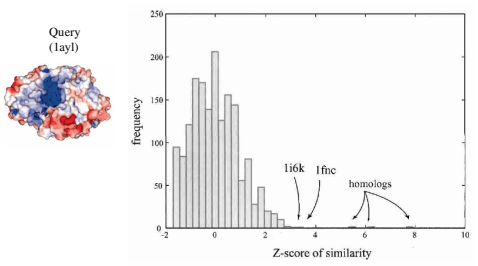
Discussion

Results



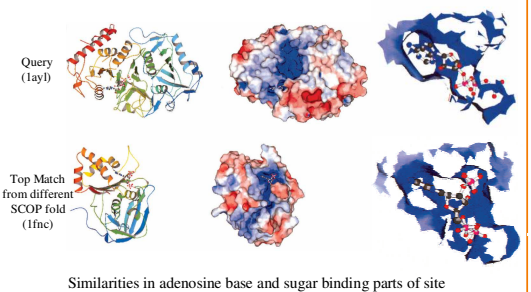
[Kinoshita05]

Results



[Kinoshita03]

Results

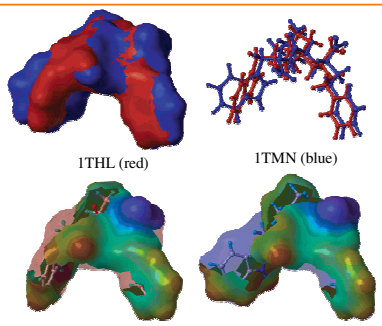


Similarities in adenosine base and sugar binding parts of site

"The relationship between their biochemical functions and the observed similarities is not clear."

[Kinoshita03]

Results



Surfaces colored by electrostatic potential

[Hofbauer04]

Discussion

?

References

[Ankerst99] Mihael Ankerst, Gabi Kastenmüller, Hans-Peter Kriegel, Thomas Seidl, "3D Shape Histograms for Similarity Search and Classification in Spatial Databases," SSD, 1999, pp. 207-226.

[Asirvatham05] A. Asirvatham, E. Praun, and H. Hoppe, "Consistent spherical parameterization", Computer Graphics and Geometric Modeling (CGGM) Workshop, 2005.

[Cai98] W. Cai, M. Zhang, B. Maigret, "New approach for representation of molecular surface," J. Comput. Chem, 19, 1998, pp. 1805-1815.

[Connolly83] M.L. Connolly, "Solvent-accessible surfaces of proteins and nucleic acids," Science, 221, 1983, pp. 709-713.

[Duncan93] B.S. Duncan, A.J. Olson, "Approximation and characterization of molecular surfaces," Biopolymers, 33, 1993, pp. 219-229.

[Hofbauer04] C. Hofbauer, H. Lohninger, A. Aszodi, "SURFCOMP: A Novel Graph-Based Approach to Molecular Surface Comparison," J. Chem. Inf. Comput. Sci., 44, 2004, pp. 837-847.

[Horn86] B.K.P. Horn, ``Extended Gaussian Images," Proceedings of the IEEE, Vol. 72, No. 12, December 1984, pp. 1671--1686.

[Ikeuchi95] K. Ikeuchi and M. Hebert. "Spherical Representations: from EGI to SAI," Tech. report CMU-CS-95-197, Computer Science Department, Carnegie Mellon University, October, 1995.

[Kinoshita05] K. Kinoshita,, H. Nakamura, "Identification of the ligand binding sites on the molecular surface of proteins," Protein Science, 14, 2005, pp. .

[Kinoshita03] K. Kinoshita,, H. Nakamura, "Identification of protein biochemical functions by similarity search using the molecular surface database eF-site," Protein Science, 12, 2003, pp. 1589-1595.

[Kinoshita02] K. Kinoshita, J. Furui, H. Nakamura, "Identification of Protein Functions from a Molecular Surface Database, eF-site," J. Struct. Func. Genomics, 2, 1, 2002, pp. 9-22.

[Lin94] S.L. Lin, R. Nussinov, D. Fischer, H.J. Wolfson, "Molecular-Surface Representations By Sparse Critical-Points," Proteins-Structure Function and Genetics, 18, 1994, pp. 94-101.

[Praun03] E. Praun and H. Hoppe, "Spherical Parameterization and Remeshing", SIGGRAPH, 2003.

[Rubner00] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance," International Journal of Computer Vision, 2000.

[Shulman-Peleg04] A. Shulman-Peleg, R. Nussinov, H.J. Wolfson, "Recognition of functional sites in protein structures," J Mol Biol, 339, 2004, pp. 607-633.

[Sorkine02] Olga Sorkine, Daniel Cohen-Or, Rony Goldenthal and Dani Lischinski, "Bounded-distortion Piecewise Mesh Parameterization", IEEE Visualization, 2002.