I/O-Efficient Algorithms for Computing Contours on a Terrain

Bardia Sadri Duke University

joint work with:

Pankaj K.Agarwal Duke University Lars Arge MADALGO Thomas Mølhave MADALGO

Terrains

A terrain is the graph of a continuous bivariate function.



Terrains

A terrain is the graph of a continuous bivariate function.

In GIS the surface of earth is often represented as a terrain that interpolates collected data. Y

X

 \mathcal{Z}

Terrains

A terrain is the graph of a continuous bivariate function.

In GIS the surface of earth is often represented as a terrain that interpolates collected data.

LIDAR (Light Detection and Ranging)

- Massive (irregular) point sets (1-10m resolution)
- Becoming relatively cheap and easy to collect
- Appalachian mountains between 50GB to 5TB















The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$.



The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$.



The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$. Each connected component of a level set is called a contour.



The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$. Each connected component of a level set is called a contour. Given levels $L = \{\ell_1, \dots, \ell_k\}$, the contour map is $h^{-1}(L)$.



The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$. Each connected component of a level set is called a contour. Given levels $L = \{\ell_1, \dots, \ell_k\}$, the contour map is $h^{-1}(L)$.



The level-set \mathbb{M}_{ℓ} at height ℓ is $h^{-1}(\ell)$. Each connected component of a level set is called a contour. Given levels $L = \{\ell_1, \dots, \ell_k\}$, the contour map is $h^{-1}(L)$.



Pretty Old Stuff!

Applications at least as early as the 18th century.

Applications at least as ea

Droth Old Ctuff

XXXIII. An Account of the Calculations made from the Survey and Measures taken at Schehallien, in order to ascertain the mean Density of the Earth. By Charles Hutton, Esq. F. R. S.

This circumstance at first gave me much trouble and diffatisfaction, till I fell upon the following method by which the defect was in a great measure supplied, and by which I was enabled to proceed in the effimation of the altitudes both with much expedition and a confiderable degree of accuracy. This method was the connecting together by a faint line all the points which were of the fame relative altitude: by fo doing, I obtained a great number of irregular polygons lying within, and at some distance from, one another, and bearing a confiderable degree of refemblance to each other: thefe polygons were the figures of fo many level or horizontal fections of the hills, the relative altitudes of all the parts of them being known; and as every bafe or little fpace had

Applications at least as ea

Droth Old Ctuff

XXXIII. An Account of the Calculations made from the Survey and Measures taken at Schehallien, in order to ascertain the mean Density of the Earth. By Charles Hutton, Esq. F. R. S.



Given a set of levels $L = \{\ell_1, \ldots, \ell_k\}$, compute the contour map $h^{-1}(L)$ such that each contour is reported separately and in sorted (circular) order.



Given a set of levels $L = \{\ell_1, \ldots, \ell_k\}$, compute the contour map $h^{-1}(L)$ such that each contour is reported separately and in sorted (circular) order.



Given a set of levels $L = \{\ell_1, \ldots, \ell_k\}$, compute the contour map $h^{-1}(L)$ such that each contour is reported separately and in sorted (circular) order.



Given a set of levels $L = \{\ell_1, \ldots, \ell_k\}$, compute the contour map $h^{-1}(L)$ such that each contour is reported separately and in sorted (circular) order.



Answering Contour Queries

Preprocess the terrain to answer contour queries efficiently: Given a level $\ell \in \mathbb{R}$, return the level set $h^{-1}(\ell)$ such that each contour is reported separately and in sorted (circular) order.





Classical Complexity: Number of basic operations as a function of N.



Classical Complexity: Number of basic operations as a function of N. Disk access is about 10^6 times slower than main memory access.



Classical Complexity: Number of basic operations as a function of N. Disk access is about 10^6 times slower than main memory access. To amortize access delay, disks transfer large contiguous blocks of data.

Image: Sector Sector

Classical Complexity: Number of basic operations as a function of N. Disk access is about 10^6 times slower than main memory access. To amortize access delay, disks transfer large contiguous blocks of data.

- N = number of items in the problem instance
- B = number of items per disk block
- M = number of items that fit main memory
- T = number of items in output

Image: Sector Sector

Classical Complexity: Number of basic operations as a function of N. Disk access is about 10^6 times slower than main memory access. To amortize access delay, disks transfer large contiguous blocks of data.

- N = number of items in the problem instance
- B = number of items per disk block
- M = number of items that fit main memory
- T = number of items in output

I/O-Complexity: Number of I/Os as a function of N, B, M, and T.

 Image: Sex treme
 Image: Sex treme

 Image: Sex treme
 Image: Sex treme
</tre>

 Image:

Classical Complexity: Number of basic operations as a function of N. Disk access is about 10^6 times slower than main memory access. To amortize access delay, disks transfer large contiguous blocks of data.

- N = number of items in the problem instance
- B = number of items per disk block
- M = number of items that fit main memory
- T = number of items in output

r			
Co,		internal	external
	Scanning	N	N/B
Class	Sorting	$N \log N$	$\frac{N}{B}\log_{\frac{M}{B}}\frac{N}{B}$
Disk	Permuting	N	$\min\left\{N, \frac{N}{B}\log_{\frac{M}{B}}\frac{N}{B}\right\}$
To a	Searching	$\log_2 N$	$\log_B N$

- N = number of items in the problem instance
- B = number of items per disk block
- M = number of items that fit main memory
 - T = number of items in output

P			
Co,		internal	external
L	Scanning	N	N/B
Class	Sorting	$N \log $ Sort (Λ	$V) \qquad \frac{N}{B} \log_{\frac{M}{B}} \frac{N}{B}$
Disk	Permuting	N	$\min\left\{N, \frac{N}{B}\log_{\frac{M}{B}}\frac{N}{B}\right\}$
To a	Searching	$\log_2 N$	$\log_B N$.

- N = number of items in the problem instance
- B = number of items per disk block
- $M \ = \ \operatorname{number}$ of items that fit main memory
 - T = number of items in output

P			
Co,		internal	external
L	Scanning	N	N/B
Class	Sorting	$N \log \mathbb{Sort}(N)$	$I) \qquad \frac{N}{B} \log_{\frac{M}{B}} \frac{N}{B} \ll N$
Disk	Permuting	N	$\min\left\{N, \frac{N}{B}\log_{\frac{M}{B}}\frac{N}{B}\right\}$
To a	Searching	$\log_2 N$	$\log_B N$.

- N = number of items in the problem instance
- B = number of items per disk block
- $M \hspace{.1in} = \hspace{.1in} \operatorname{number} \operatorname{of} \operatorname{items} \operatorname{that} \operatorname{fit} \operatorname{main} \operatorname{memory}$
 - T = number of items in output

Previous Work

Answering contour queries I/O-efficiently:

	Preprocessing I/Os	Structure Size	Query I/Os
Chiang, Silva'97	O(Sort(N))	O(N)	$O(\log_B N + T/B)$
Agarwal, Arge, Murali, Varadrarajan, Vitter'98	O(N)	O(N)	$O(\log_B N + T/B)$

Previous Work

Answering contour queries I/O-efficiently:

	Preprocessing I/Os	Structure Size	Query I/Os
Chiang, Silva'97	O(Sort(N))	O(N)	$O(\log_B N + T/B)$
Agarwal, Arge, Murali, Varadrarajan, Vitter'98	O(N)	O(N)	$O(\log_B N + T/B)$
Previous Work

Answering contour queries I/O-efficiently:

	Preprocessing I/Os	Structure Size	Query I/Os
Chiang, Silva'97	O(Sort(N))	O(N)	$O(\log_B N + T/B)$
Agarwal, Arge, Murali, Varadrarajan, Vitter'98	O(N)	O(N)	$O(\log_B N + T/B)$









Find one "seed" triangle intersecting each contour and trace out the contour sequentially.

This is essentially the optimal algorithm for the RAM model.

Find one "seed" triangle intersecting each contour and trace out the contour sequentially.

This is essentially the optimal algorithm for the RAM model. I/O Complexity:

O(N/B+T).

segments in the output

Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.



Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.



Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.



Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.

For segment s at level ℓ_i store pair (ℓ_i, s) plus the segments before and after son contour containing s.

Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.

For segment s at level ℓ_i store pair (ℓ_i, s) plus the segments before and after son contour containing s.

Sort pairs on first component to separates level sets. Then use successor/predecessor-sorting to put contours in order.

Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.

For segment s at level ℓ_i store pair (ℓ_i, s) plus the segments before and after son contour containing s.

Sort pairs on first component to separates level sets. Then use successor/predecessor-sorting to put contours in order.

I/O Complexity:

 $O(N/B + \mathsf{Sort}(T)).$

Scan the triangles (in the order laid out on the disk) and generate all segments. Then sort the output.

For segment s at level ℓ_i store pair (ℓ_i, s) plus the segments before and after son contour containing s.

Sort pairs on first component to separates level sets. Then use successor/predecessor-sorting to put contours in order.

I/O Complexity:

 $O(N/B + \mathsf{Sort}(T)).$

This talk: O(Sort(N) + T/B).









If triangles were ordered on disk such that all partially generated contours in "less naïve" algorithm stayed connected, no succ/pred sorting would be needed.

 \prec : such an ordering



If triangles were ordered on disk such that all partially generated contours in "less naïve" algorithm stayed connected, no succ/pred sorting would be needed.

 \prec : such an ordering Δ_{ℓ} : triangles that intersect level ℓ



If triangles were ordered on disk such that all partially generated contours in "less naïve" algorithm stayed connected, no succ/pred sorting would be needed.

 \prec : such an ordering Δ_{ℓ} : triangles that intersect level ℓ



If triangles were ordered on disk such that all partially generated contours in "less naïve" algorithm stayed connected, no succ/pred sorting would be needed.



The restriction of \prec to Δ_{ℓ} traverses each contour of M in circular order.

1. Sweep the terrain by a horizontal plane in the +z direction.



1. Sweep the terrain by a horizontal plane in the +z direction.



1. Sweep the terrain by a horizontal plane in the +z direction.

 \prec .

2. Keep triangles that intersect the sweep plane in a search tree ordered by



- 1. Sweep the terrain by a horizontal plane in the +z direction.
- 2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec .
- 3. When passing target level $\ell_i \in L$, dump contents of tree to disk.



1. Sweep the terrain by a horizontal plane in the +z direction.



- 2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec .
- 3. When passing target level $\ell_i \in L$, dump contents of tree to disk.



- 1. Sweep the terrain by a horizontal plane in the +z direction.
- Keep triangles that intersect the sweep plane in a search tree ordered by ≺.
 Buffer Tree [Arge'95]
- 3. When passing target level $\ell_i \in L$, dump contents of tree to disk.



 $\mathsf{Sort}(N)$

- 1. Sweep the terrain by a horizontal plane in the +z direction.
- 2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec Amortized Sort(N) Buffer Tree [Arge'95]
- 3. When passing target level $\ell_i \in L$, dump contents of tree to disk.



 $\mathsf{Sort}(N)$

- 1. Sweep the terrain by a horizontal plane in the +z direction.Sort(N)2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec .Amortized Sort(N)Buffer Tree [Arge'95]
- 3. When passing target level $\ell_i \in L$, dump contents of tree to disk. $\langle T \rangle$



1. Sweep the terrain by a horizontal plane in the +z direction.Sort(N)2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec .Amortized Sort(N)Buffer Tree [Arge'95]

3. When passing target level $\ell_i \in L$, dump contents of tree to disk. $\langle T \rangle$

Using a persistent search tree, we can answer contour queries in $O(\log_B N + T/B)$ I/Os.

1. Sweep the terrain by a horizontal plane in the +z direction.Sort(N)2. Keep triangles that intersect the sweep plane in a search tree ordered by \prec . Amortized Sort(N)3. When passing target level $\ell_i \in L$, dump contents of tree to disk.

Using a persistent search tree, we can answer contour queries in $O(\log_B N + T/B)$ I/Os.

Preprocessing needs O(Sort(N)) I/Os and O(N) space.

Theorem. One can find in O(Sort(N)) I/Os a total ordering " \prec " of triangles of M, s.t. for any ℓ :

1. Triangles in Δ_{ℓ} are \prec -sorted in cw or ccw order.



Theorem. One can find in O(Sort(N)) I/Os a total ordering " \prec " of triangles of M, s.t. for any ℓ :

- 1. Triangles in Δ_{ℓ} are \prec -sorted in cw or ccw order.
- 2. Subsets of Δ_ℓ intersecting distinct contours C_1 and C_2 don't "interleave":

 $t_1 \prec t_3 \prec t_2 \quad \Longrightarrow \quad t_1 \prec t_4 \prec t_2$



Theorem. One can find in O(Sort(N)) I/Os a total ordering " \prec " of triangles of M, s.t. for any ℓ :

- 1. Triangles in Δ_{ℓ} are \prec -sorted in cw or ccw order.
- 2. Subsets of Δ_{ℓ} intersecting distinct contours C_1 and C_2 don't "interleave": $t_1 \prec t_3 \prec t_2 \implies t_1 \prec t_4 \prec t_2$

We call \prec a "level-ordering" of triangles in \mathbb{M} .



Theorem. One can find in O(Sort(N)) I/Os a total ordering " \prec " of triangles of M, s.t. for any ℓ :

- 1. Triangles in Δ_{ℓ} are \prec -sorted in cw or ccw order.
- 2. Subsets of Δ_{ℓ} intersecting distinct contours C_1 and C_2 don't "interleave": $t_1 \prec t_3 \prec t_2 \implies t_1 \prec t_4 \prec t_2$

 t_2

 t_1

We call \prec a "level-ordering" of triangles in \mathbb{M} .

 $a_1a_2b_1c_1c_2c_3b_2b_3d_1d_2b_4a_3a_4a_5$

 t_3

 t_4
Level Ordering Theorem

Theorem. One can find in O(Sort(N)) I/Os a total ordering " \prec " of triangles of M, s.t. for any ℓ :

- 1. Triangles in Δ_{ℓ} are \prec -sorted in cw or ccw order.
- 2. Subsets of Δ_{ℓ} intersecting distinct contours C_1 and C_2 don't "interleave": $t_1 \prec t_3 \prec t_2 \implies t_1 \prec t_4 \prec t_2$

 t_2

 t_1

We call \prec a "level-ordering" of triangles in \mathbb{M} .

Can separate contours using a stack in O(T/B) I/Os.

 $a_1a_2b_1c_1c_2c_3b_2b_3d_1d_2b_4a_3a_4a_5$

 $a_1 a_2 a_3 a_4 a_5 b_1 b_2 b_3 b_4 c_1 c_2 c_3 d_1 d_2$

 t_3

 t_4

































An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .





An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .





An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .





An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .



Lemma. Every cycle of \mathbb{M}^* loses precisely one edge.

An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .



Lemma. Every cycle of \mathbb{M}^* loses precisely one edge.

Lemma. The resulting dual graph is acyclic and the induced relation " \prec " a partial order. Thus we can topologically sort it into a total order.

An elementary terrain has no saddles; thus 1 max and 1 min (boundary). Take a monotone min (bd) to max path P and delete its dual from \mathbb{M}^* .



Lemma. Every cycle of \mathbb{M}^* loses precisely one edge.

Lemma. The resulting dual graph is acyclic and the induced relation " \prec " a partial order. Thus we can topologically sort it into a total order.

[Arge, Toma, Zeh'03]

A saddle is negative if it joins two disjoint connected components of its sublevel-set and positive otherwise.





A saddle is negative if it joins two disjoint connected components of its sublevel-set and positive otherwise.



A saddle is negative if it joins two disjoint connected components of its sublevel-set and positive otherwise.



If we replace h with -h, the two types switch roles.

A saddle is negative if it joins two disjoint connected components of its sublevel-set and positive otherwise.



If we replace h with -h, the two types switch roles.

[Agarwal, Arge, Yi '06] Positive and negative saddle points can be found in $O({\rm Sort}(N))$ I/Os.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

 \sim

 \sim

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.

Positive Cut-Tree: follow an ascending path in every connected component of the upper link of every positive saddle, joining paths when they collide.




















The elementary terrain \mathbb{M}' has all the triangles of \mathbb{M} plus some of "new" triangles.



The elementary terrain \mathbb{M}' has all the triangles of \mathbb{M} plus some of "new" triangles.

All contours of a level set of $\mathbb M$ are combined in a single contour in $\mathbb M'$



The elementary terrain \mathbb{M}' has all the triangles of \mathbb{M} plus some of "new" triangles.

All contours of a level set of $\mathbb M$ are combined in a single contour in $\mathbb M'$



Theorem. In a contour of \mathbb{M}' , corresponding contours of \mathbb{M} are broken (by segments from new triangles) in a nested (parenthesized) manner.

 $a_1a_2 * *b_1 * * *c_1c_2c_3 * * *b_2b_3 * d_1d_2 * *b_4 * *a_3a_4a_5$

The elementary terrain \mathbb{M}' has all the triangles of \mathbb{M} plus some of "new" triangles.

All contours of a level set of $\mathbb M$ are combined in a single contour in $\mathbb M'$



Theorem. In a contour of \mathbb{M}' , corresponding contours of \mathbb{M} are broken (by segments from new triangles) in a nested (parenthesized) manner.

 $[a_1a_2 * *b_1 * * *c_1c_2c_3 * * * *b_2b_3 * d_1d_2 * *b_4 * *a_3a_4a_5]$













A contour is red (blue) if "locally" the sublevel set is in its outside (inside).



A contour is red (blue) if "locally" the sublevel set is in its outside (inside).





A contour is red (blue) if "locally" the sublevel set is in its outside (inside).





A contour is red (blue) if "locally" the sublevel set is in its outside (inside).





Theorem. Only the red tree connects a blue contour and its red children.

A contour is red (blue) if "locally" the sublevel set is in its outside (inside).





Theorem. Only the red tree connects a blue contour and its red children.

Theorem. If we contract each contour to a point, the result is a tree.

A contour is red (blue) if "locally" the sublevel set is in its outside (inside).



Theorem. Only the red tree connects a blue contour and its red children.

Theorem. If we contract each contour to a point, the result is a tree.

The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

What about surfaces of genus ≥ 1 ? (Orientable or not)

The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

What about surfaces of genus ≥ 1 ? (Orientable or not)



The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

What about surfaces of genus ≥ 1 ? (Orientable or not)



The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

What about surfaces of genus ≥ 1 ? (Orientable or not)



The embedding of the terrain is unimportant: the set of triangles that intersect a level set only depends on function value on vertices.

What about surfaces of genus ≥ 1 ? (Orientable or not)



