Minimal curves and surfaces for segmentation

Ben Appleton & Hugues Talbot

`talboth@esiee.fr`

Google Inc. – ESIEE

Problem and motivation

- Segmentation – one of the fundamental problems in image analysis.
  - Tessellate the image into consistent regions, by some measure
  - Delineate objects of interest
- We want an accurate, versatile and efficient method for segmentation.

Illustration

Original

Final

Methods

- Contour-based
  - Edge map linking, e.g. Canny (1986).
  - Traditionnal active contours Kass et al. (1988)
- Region-based
  - Watershed (Beucher and Lantuéjoul, 1979), region growing (Adams and Bischof, 1994)
  - Geodesic active contour (PDE) (Caselles et al., 1997a)
- Many, many others (over 1,000 different methods published in the litterature), sorry if not cited.

Quick Literature review

1. Snakes/active contours
2. Level sets methods
3. Geodesic active contours

Geodesic active contours

- Proposed by Caselles et al. (1997b) ;
- Sensible model: the functional to minimize is simply:

$$E(C) = \int g(C(s))ds$$

  where s is the Euclidean arc length.
- The LS formulation is:

$$\phi_t = \text{div}\left(g\frac{\nabla\phi}{|\nabla\phi|}\right)|\nabla\phi| \tag{1}$$

- This flow deforms an initial curve towards the path of minimal weighted length, where the arc-length is measured by $g^2ds^2 = g^2(x,y)(dx^2+dy^2)$.
- Relatively fast implementation (Goldenberg et al., 2001).

Metric

The metric g is derived from local features, it should be low on the border of objects and high elsewhere.
Caselles et al. (1997a) proposed the following metric:

$$g = \frac{1}{1 + |\nabla G_\sigma \star I|^p} + \epsilon. \tag{2}$$

Here $|\nabla G_\sigma \star I|$ represents the magnitude of the image derivatives at scale σ, $p = 1$ or 2 is the power to which the gradient is raised, and ϵ controls the regularity of the surface.

Globally optimal geodesic active contour

- Really the above is a minimal path problem, already pointed out in (Cohen and Kimmel, 1997).
- Idea: solve with a dynamic programming approach.
- Difficulties:
  - number of contour candidates grows exponentially
  - relatively easy to find shortest path between two fixed points, but hard to find shortest closed contours.
  - discretization problems.

GAC about a point

- Find a closed contour of minimal energy containing an interior point p_i
- That contour must pass through a parallel to positive x-axis passing through p_i at some point
- We *cut* the plane $I\!R^2$ along this line from $-\infty$ to p_i so that the two sides of the cut are disconnected (new manifold).
- We compute the *surface of minimal action* from an arbitrary starting point p_s on one side of the cut to anywhere on the new manifold:

$$U(p) = \inf\{E(C) | C(0) = p_s, C(L) = p\}$$

GAC about a point (cont.)

- We use the fast marching method (FMM) to solve the corresponding Eikonal equation for U:

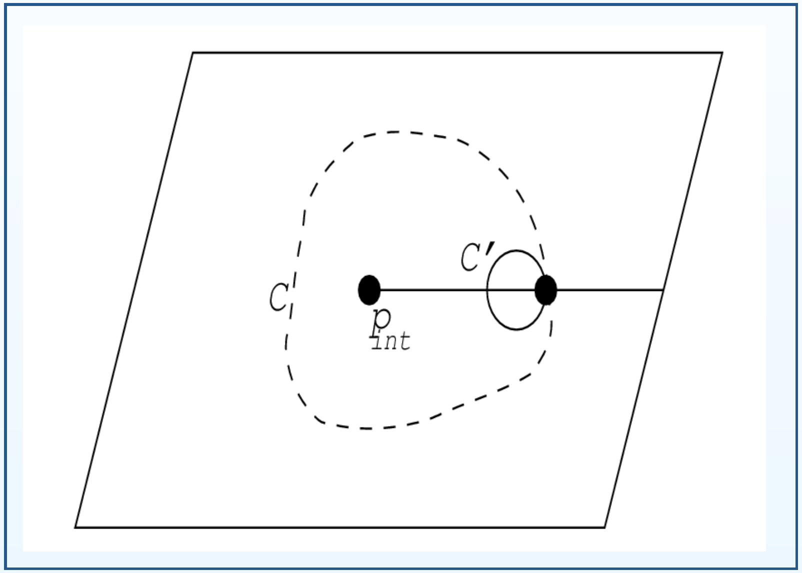
$$|\nabla U| = g, U(p_s) = 0$$

- Once U has been computed, we determine the minimal geodesic C as follows:

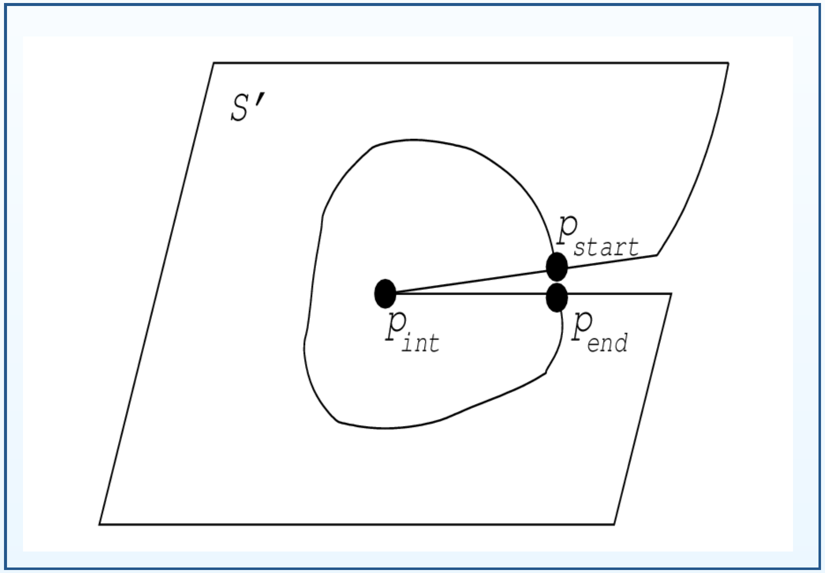
$$\frac{\partial C}{\partial s} = -\frac{\nabla U}{|\nabla U|}, C(L) = p_e$$

- We constrain p_e to be opposite p_s (as in CSP)
- To obtain the minimal geodesic, we try for all possible p_s.

GAC about a point (cont.)



(a)

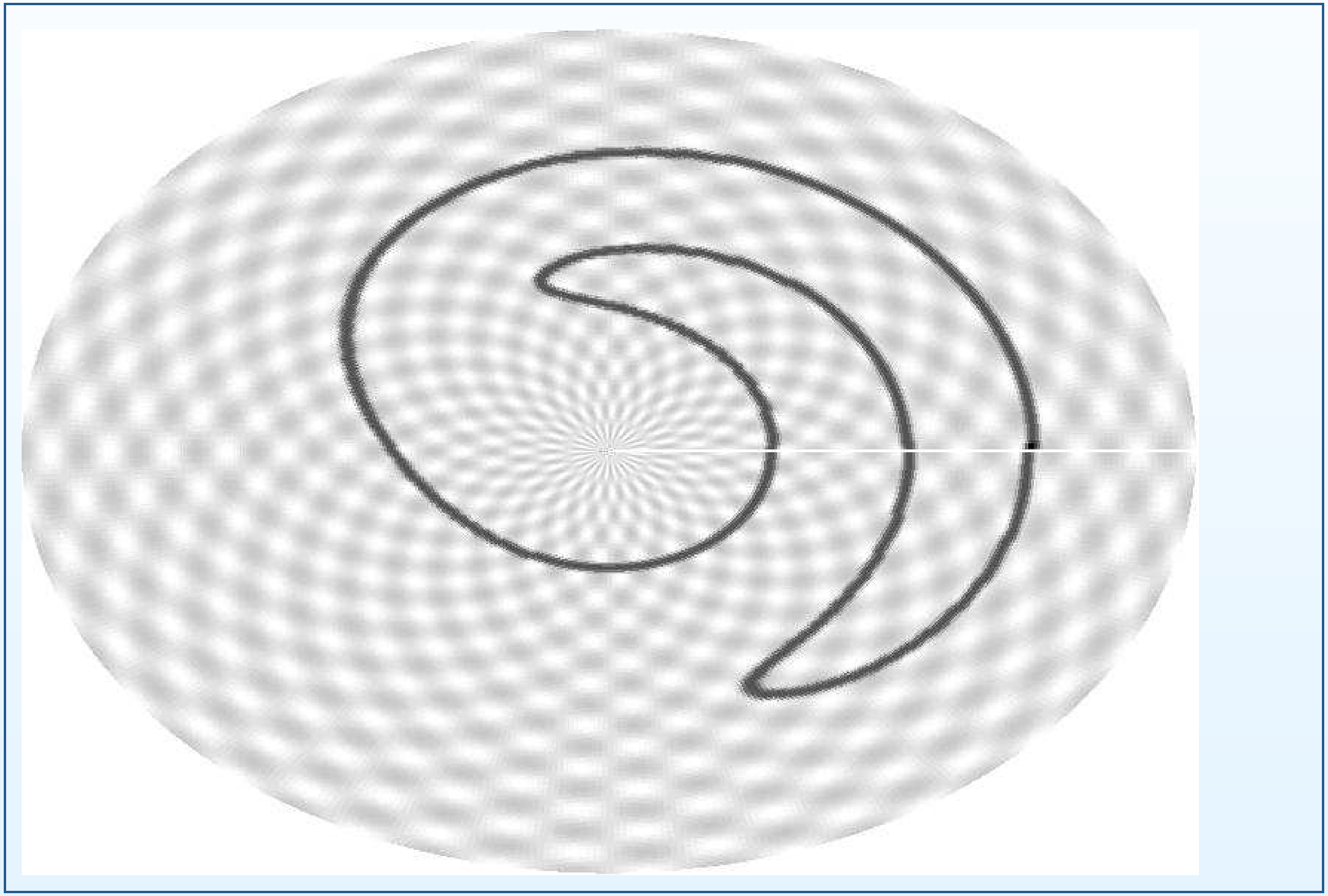


(b)

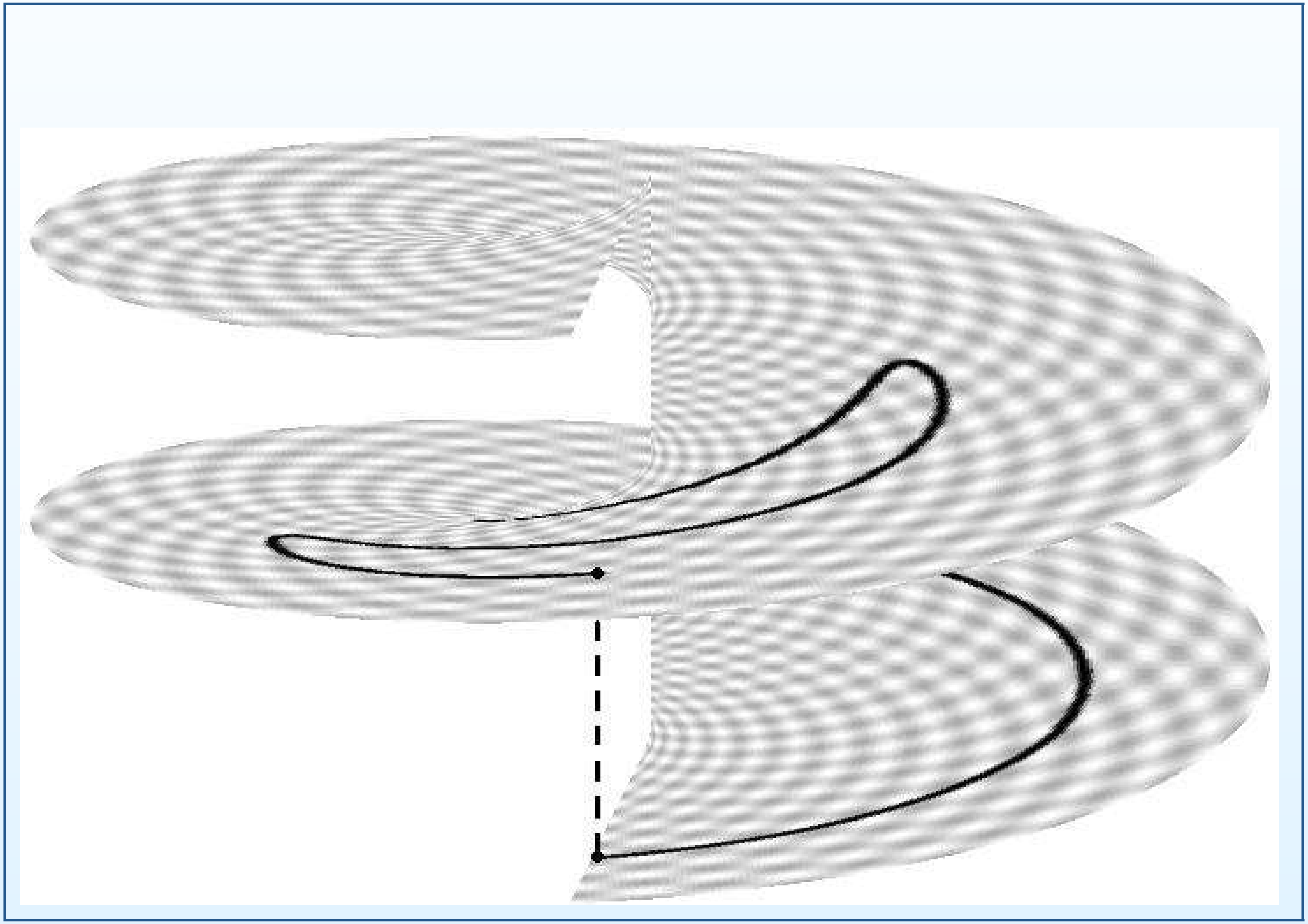
Cut-concave paths

- Path can only go through the x axis once
- To solve this we use an helicoidal manifold to compute U
- We halt the FMM when it reaches p_s the first time.

Cut-concave paths (cont.)



Cut-concave paths on helicoidal manifold



Choice of metric

- The standard g GAC metric won't work because the optimal GAC is the null contour.
- Instead we use $g' = g/r$, i.e:

$$g' = \frac{1}{r}\left(\frac{1}{1 + |\nabla G_\sigma * I|^n}\right)$$

- With such a metric, a small closed contour around the central point has positive energy:

$$E(C) = \lim_{r \to 0} \oint_C g'(C(s))ds = 2\pi g(p_i)$$

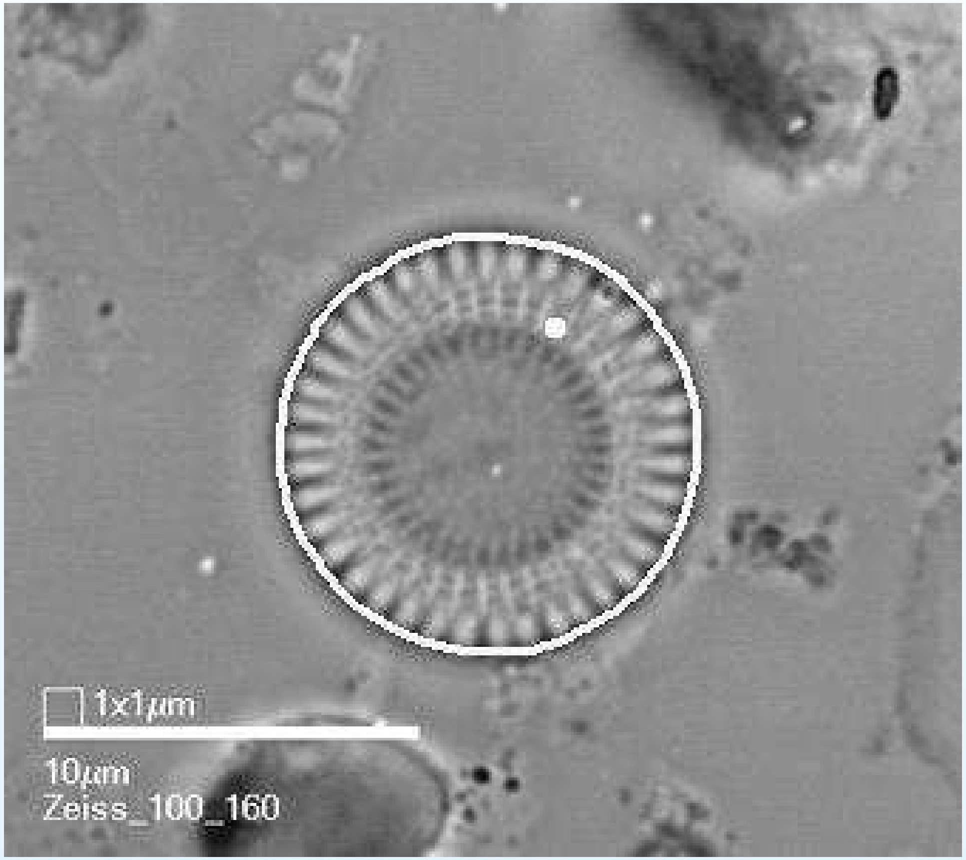
Choice of metric (cont.)

- this $1/r$ term is related to the polar transform of the contour

$$E(C) = \int_C g(C(r,\theta))d\theta = \int_C \frac{g(C(s))}{r} \langle ds, \vec{\theta} \rangle$$

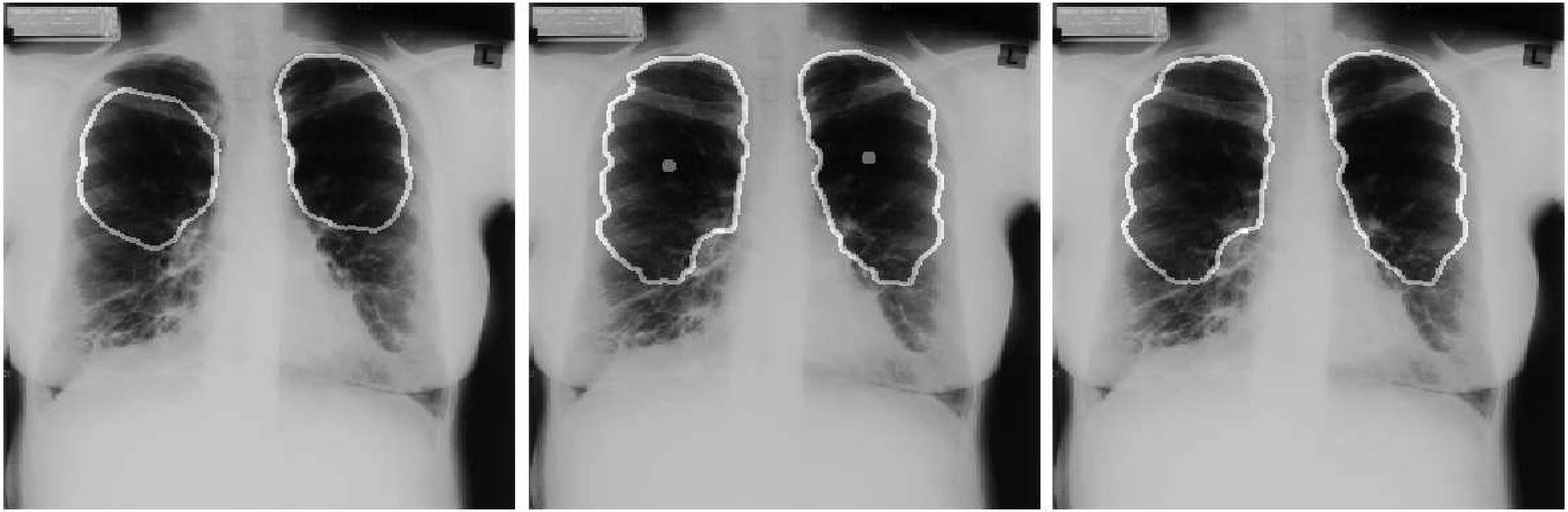
- It acts as a natural "balloon force", that had to be added to the standard GAC anyway.

Example



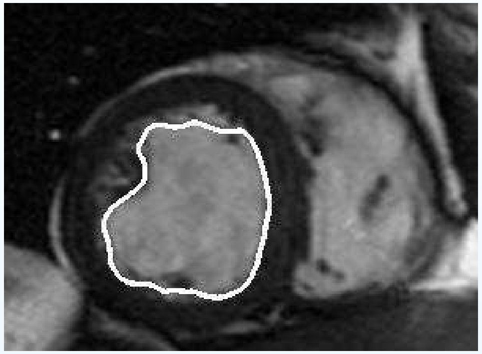
diatom

GOGAC vs. GAC

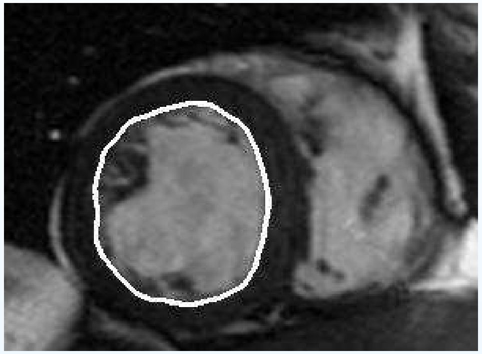


(a) (b) (c)

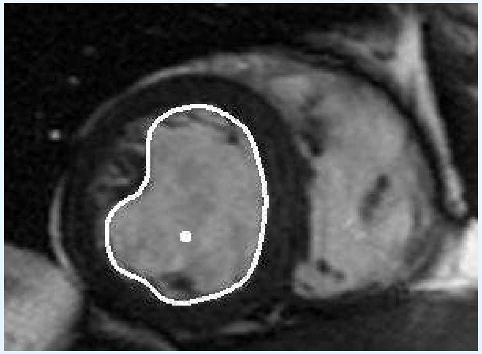
GOGAC vs. GAC (cont.)



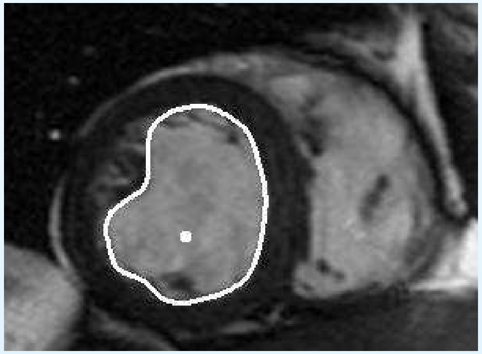
gac



aos-gac



gogac



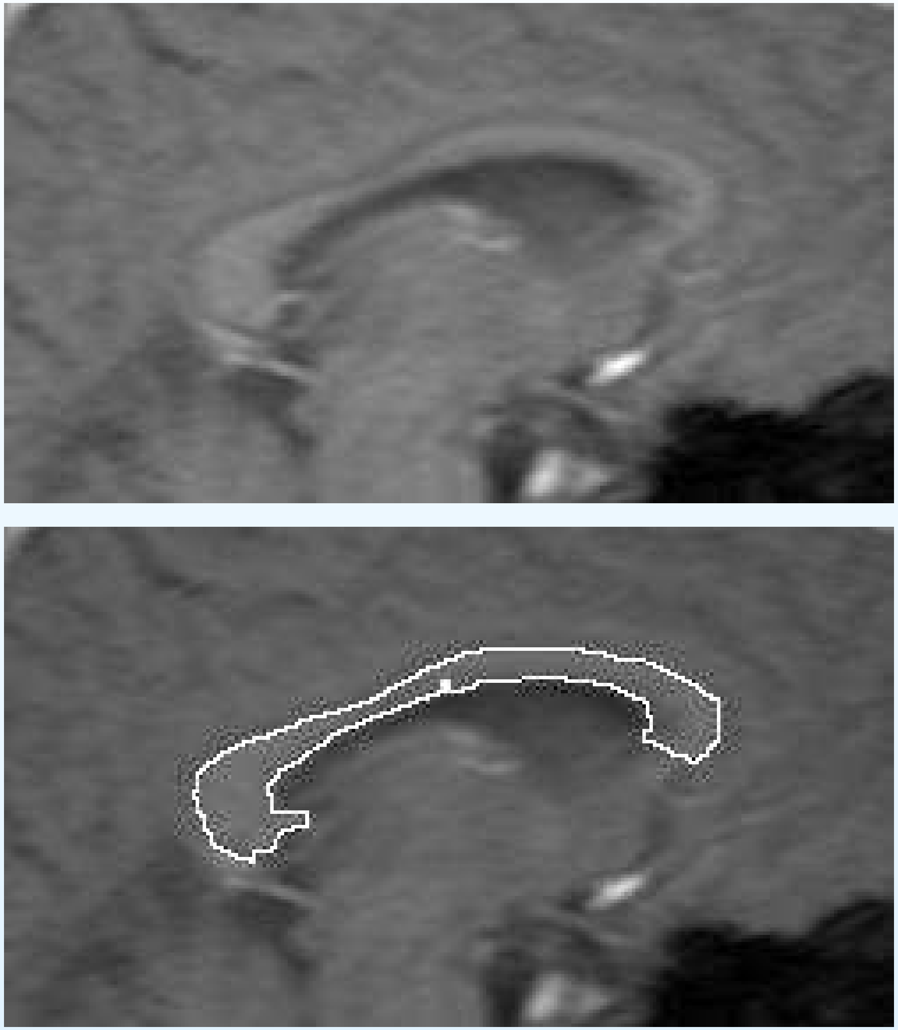
gogac cut-convex

Speed

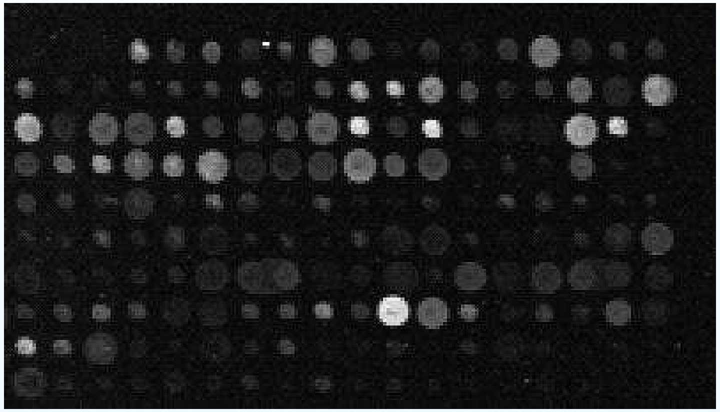
	exact	approx.
GAC	4.5s	2.0s
GOGAC	3.4s	0.48s

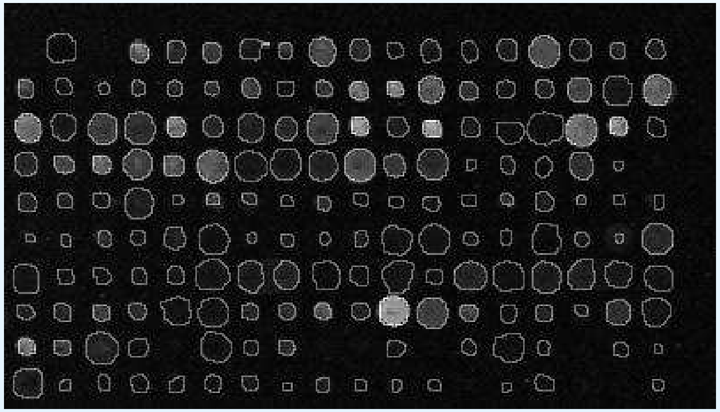
This is on a P-III 700MHz. Image subset is 256×256

Point-concavity: corpus callosum

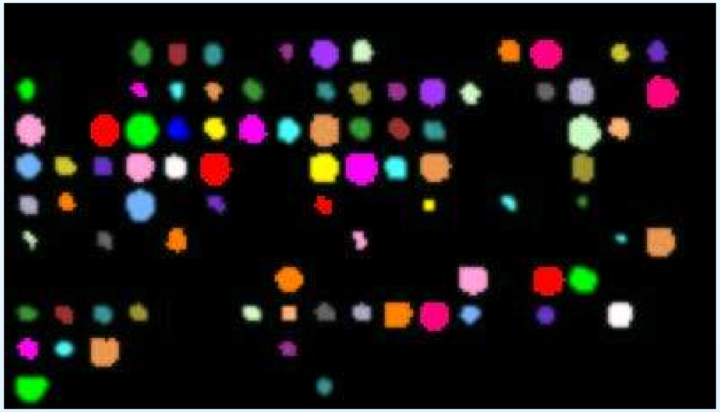


Current application: microarrays





gogac



srg

Properties of GOGAC

- GOGAC compares favourably to classic variational approaches ;
- GOGAC is constrained to single simple closed curves ;
- Constrained to 2D (classical GAC works very well in 3D) ;
- Optimal contours not necessarily meaningful.

Optimal minimal surfaces

- We want to extend that approach to 3D
- The dual approach to minimal path in the graph framework is the maximum flow algorithm (Ford and Fulkerson, 1962),

Graph maximum flow algorithm

A partitioning of a graph G decomposes its vertex set into a collection $\Gamma_G = \{V_1, V_2, \ldots\}$ of disjoint subsets:

$$\bigcup_{V_i \in \Gamma_G} V_i = V, \quad V_i \cap V_j = \emptyset \quad \text{for} \quad i \neq j.$$

To each partition Γ_G we associate a cost $C(\Gamma_G)$ which is the total cost of the edges whose endpoints lie in different partitions,

$$C\left(\Gamma_G\right) = \sum_{e \in E^*} C_E\left(e\right).$$

Here the *cut* $E^* \subseteq E$ is the set of edges crossing the partition.

Minimal cut

The s-t minimal cut problem seeks the partitioning of minimal cost such that the disjoint vertices $s,\ t \subseteq V$ lie in different partitions. Although we do not give the construction here, it is simple to extend these algorithms to the case of multiple sources and sinks (Sedgewick, 2002).

Maximum flow

Let G be a graph with edge costs C_E now reinterpreted as *capacities*. A flow $F : E \to I\!R$ from a *source* $s \in V$ to a *sink* $t \in V$ has the following properties:

- Conservation of flow: The total (signed) flow in and out of any vertex is zero.
- Capacity constraint: The flow along any edge is less than or equal to its capacity:

$$\forall e \in E, \quad F(e) \leq C_E(e).$$

An edge along which the flow is equal to the capacity is described as *saturated*. Following Sedgewick (2002) we implicitly add a directed edge connecting $t \to s$ of infinite capacity to conserve flow uniformly throughout G.

Max flow = min cut

- A maximum flow in a weighted graph G maximises the flow through the $t \rightarrow s$ edge, equivalently maximising the flow from the source to the sink through the graph. Ford and Fulkerson (1962) demonstrated that the maximum s–t flow equals the minimal s–t cut, with the flow saturated uniformly on the cut.
- Note: this works for a binary partition, i.e. decomposing the problem into 2 components. The multi-component mincut problem is known to be NP-hard.

Illustration maxflow/mincut

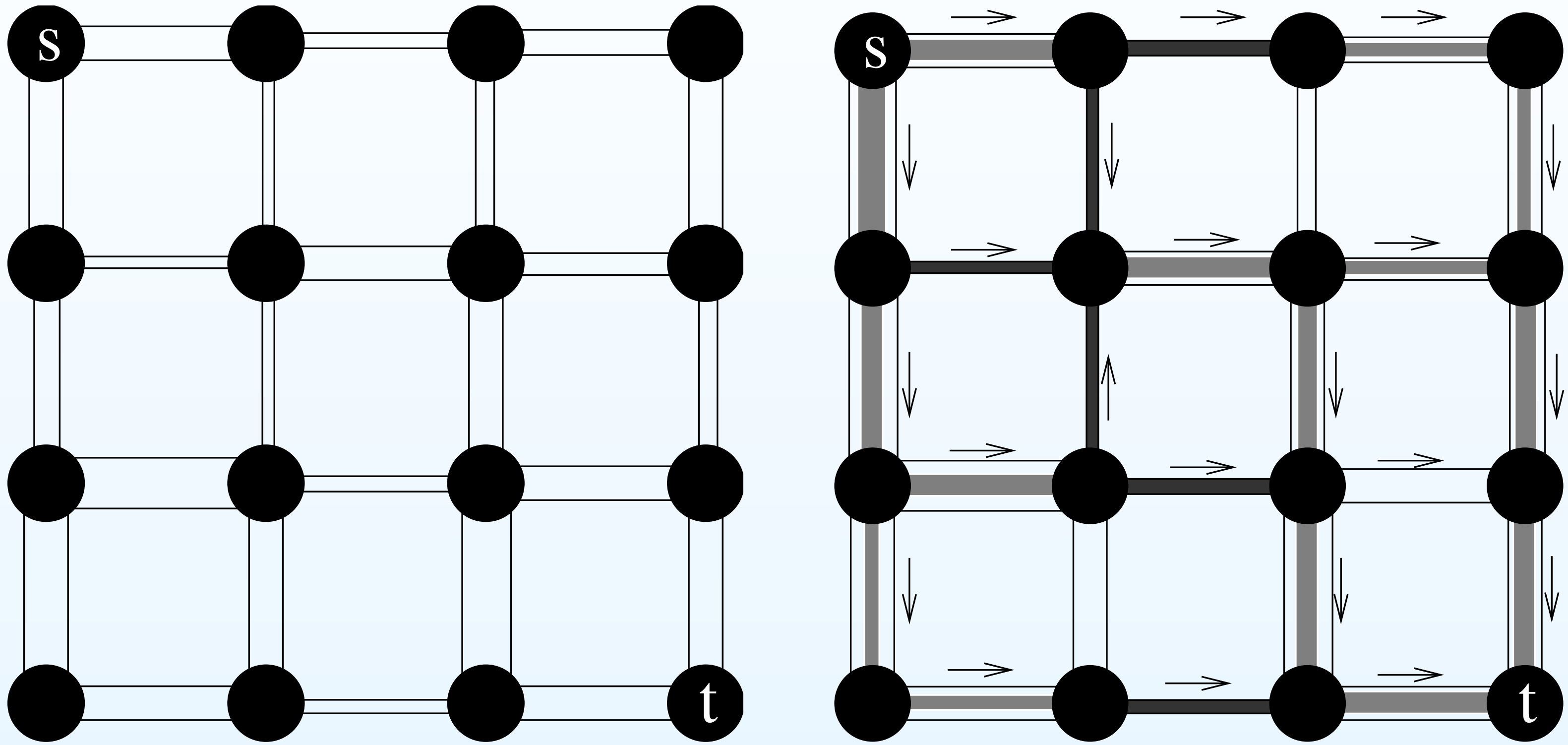


Figure 1: (a) A graph with source s and sink t. Edge capacities are depicted by their thickness. (b) An s-t maximum flow. Mincut edges are saturated.

Classical maxflow algorithm

Ford and Fulkerson (1962) also proposed a general maxflow computation algorithm. It builds a maximum flow from s to t by repeatedly locating paths along which more flow may be pushed, under the constraint that the flow be feasible (satisfying the conservation constraint).
The flow is maximal once there are no more unsaturated paths between the source and the sink.

Initialisation:

- *Set* $F = 0$ *on each edge*

Loop:

- *Search for an* s–t *path along which more flow may be pushed*
- *If no such path exists, halt*
- *Otherwise, increase the flow uniformly along this path until at least one edge becomes saturated*

Notes on maxflow algorithms

- The max flow is not necessarily unique, and there exist other alternative algorithms (e.g. pre-flow push by Goldberg and Tarjan, etc).
- There exists (restricted) dualities between max-flow, distance functions and minimal paths algorithms in 2D (if the graph is planar, including the s-t edge).

Minimal surfaces in continuous space

- In the augmenting path methods for graphs, at each step we augment the flow along an unsaturated path between the source and the sink, subject to capacity constraints. The flow is deemed incompressible.
- In the preflow-push algorithm (Goldberg and Tarjan, 1988), this constraint is relaxed, a new variable is introduced to ensure that the flow converges to an incompressible solution.
- The preflow-push update method is local, it only depends on its local neighbours.
- Using the notions of local pre-push flows, relaxing the conservation constraint and adding an extra variable to the flow system at each point, we can extend this approach to the continuous case using a system of PDEs.

Flow-surface duality

Strang (1983) and Iri (1979) have explored the extension of maximal flow to continuous domains ; continuous flows have the following properties:

- Conservation of flow: $\nabla \cdot \vec{F} = 0$
- Capacity constraint: $\left|\vec{F}\right| \leq g$

Let $\vec{F}$ be any flow and S be any simple, closed and smooth surface containing the source s. Let $\vec{N}$ denote the normal to the surface S and $\nabla \cdot \vec{F}_s$ the net flow out of the source s. Then, combining the two properties stated above, we obtain

$$\nabla \cdot \vec{F}_s = \oint_S \vec{F} \cdot \vec{N} \mathrm{d}S \leq \oint_S g \mathrm{d}S. \tag{3}$$

Continuous maximal flow properties

From this we derive two properties:

- All flows are bounded from above by all smooth, simple and closed surfaces separating the source and sink.
- All simple closed surfaces have weighted area bounded from below by all flows from source to sink.

In fact, Iri showed that under very general continuity assumptions the maximal flow F_{max} is strictly equal to the minimal surface S_{min}. For such a flow and surface, the flow must uniformly saturate the minimal surface:

$$\forall \mathbf{x} \in S_{min}, \quad F_{max}\left(\mathbf{x}\right) = g\left(\mathbf{x}\right) \vec{N}\left(\mathbf{x}\right). \tag{4}$$

Interpretation

The duality between continuous maximal flows and minimal surfaces has a simple interpretation:

- Any surface forms a bottleneck for a flow, limiting the flow to be less than the capacity or weighted area of that surface.
- The maximal flow is limited by all possible surfaces, and therefore must be less than or equal to the cost of the minimal surface.
- This duality states that the maximal flow is indeed equal to the minimal surface, and therefore that a maximal flow saturates the minimal surface.

Neither Strang or Iri propose a solution to the general minimal surface problem.

A system of PDEs for maximal flow

$$\frac{\partial P}{\partial t} = -\mathrm{div}\vec{F}, \tag{5}$$

$$\frac{\partial \vec{F}}{\partial t} = -\nabla P, \tag{6}$$

subject to

$$\left\|\vec{F}\right\|_2 \leq g. \tag{7}$$

For boundary conditions we fix the scalar field P at the source s and sink t: $P_s = 1$ and $P_t = -1$. These particular values are chosen without loss of generality to maintain symmetry between the source and sink sets.

Discussion on this system of PDEs

- We have relaxed the requirement that $\vec{F}$ have zero divergence.
- The field P stores the excess flow at every point in the domain
- This may be interpreted as a linear model of the dynamics of an idealised fluid with pressure P and velocity $\vec{F}$.
- Convection terms have been ignored in this interpretation.
- Because the magnitude constraint is harsh we can produce non-differentiable fields, and so to regulate the equations in the continuous domain we may add a dissipative term:

$$\frac{\partial \vec{F}}{\partial t} = -\nabla P + \epsilon \Delta \vec{F},$$

where $\Delta \vec{F}$ denotes the vector Laplacian of $\vec{F}$.
However in practical discrete implementations this does not appear to be necessary.

Properties: conservation of potential

Let $P_A = \int_A P \mathrm{d}A$ denote the total integral of P in a given region A not including the source and sink sets. Then, for differentiable P and $\vec{F}$,

$$\begin{aligned}\frac{\partial P_A}{\partial t} &= \int_A \frac{\partial P}{\partial t} \mathrm{d}A \\ &= -\int_A \mathrm{div} \vec{F} \mathrm{d}A \\ &= -\oint_{\partial A} \vec{F} \cdot \vec{N}_{\partial A} \mathrm{d}\left(\partial A\right).\end{aligned}$$

So P is conserved in the interior of any sourceless region A, that is, any region not including the source s or sink t.

Properties: monotonic reduction of energy

Consider the time derivative of the total quantity of energy $\frac{1}{2}(P^2 + \|\vec{F}\|_2^2)$ in a sourceless region A. For differentiable P and $\vec{F}$,

$$\begin{aligned}
\frac{\partial}{\partial t}\int_A \frac{1}{2}\left(P^2 + \|\vec{F}\|_2^2\right)\mathrm{d}A &= \int_A \frac{\partial}{\partial t}\frac{1}{2}\left(P^2 + \|\vec{F}\|_2^2\right)\mathrm{d}A \\
&= \int_A \left(\frac{\partial P}{\partial t}P + \frac{\partial \vec{F}}{\partial t}\cdot\vec{F}\right)\mathrm{d}A \\
&= -\int_A \left(P\mathrm{div}\vec{F} + (\nabla P)\cdot\vec{F}\right)\mathrm{d}A \\
&= -\oint_{\partial A} P\vec{F}\cdot\vec{N}_{\partial A}\mathrm{d}\left(\partial A\right).
\end{aligned}$$

The time derivative of the total energy is negative. This suggests that the system converges.

Properties: correctness at convergence (I)

At convergence all temporal derivatives are zero. We may then restate the system of PDEs as:

$$\mathrm{div}\vec{F} = 0,$$

$$\nabla P = \begin{cases} 0 & & \text{if} \|\vec{F}\|_2 < g \\ -\lambda\vec{F} & \text{where } \lambda \geq 0 & \text{if} \|\vec{F}\|_2 = g \end{cases}$$

- The first equation states that we obtain an incompressible flow, as required.
- The second states that where $\vec{F}$ is not saturated, P must have zero derivative and therefore be constant.
- The second equation also states that where $\vec{F}$ is not saturated, the potential gradient ∇P must be aligned so that $\vec{F}$ can neither change direction nor decrease in magnitude.

Properties: correctness at convergence (II)

- From this we can derive the property

$$\nabla P \cdot \vec{F} \leq 0.$$

  This indicates that P is a monotonic function along the flow lines of $\vec{F}$.
- In an incompressible fluid flow lines may only initiate at the source or terminate at the sink, so P has no local extrema in the interior of the domain Ω.
- Consider the closed region A_p obtained from P by the application of a threshold $-1 < p < 1$, $A_p = \{\mathbf{x} \mid P(\mathbf{x}) \geq p\}$. Due to the monotonicity of P this must be a connected region containing the source s. The boundary $S = \partial A_p$ of this region is the isosurface of value p of the potential field. On this surface we have $\nabla P \neq \vec{0}$ by construction.

Properties: correctness at convergence (III)

- Therefore the flow is uniformly saturated in the outward direction on this surface. We may make use of the incompressibility of the flow to evaluate the net outward flow through this surface,

$$\mathrm{div}\vec{F}_s = \oint_S \vec{F} \cdot \vec{N} \mathrm{d}S = \oint_S g \mathrm{d}S.$$

  Hence $\vec{F}$ and S satisfy with equality the optimality condition.
- At convergence any isosurface of P is a globally minimal surface. In the usual case of a unique minimal surface, S_{min} will be the only isosurface at convergence and hence P will approach an indicator function for the interior and exterior of S_{min}.
- Without loss of generality we may select the zero level set of P to obtain S_{min}.

Metric weighting function

- Similar to the 2D case we want to avoid bias towards small volumes.
- We have a method for deriving the metric function depending on the geometry of s and t. There isn't a systematic $w = 1/r^2$ weighting, but it is similar in principle.
- We wish to derive an unbiased flow $\vec{F}$ from which to obtain $w = \|\vec{F}\|$.
- This flow is produced by the source set s and absorbed by the sink set t, defined by

$$\nabla \cdot \vec{F} = \rho, \tag{8}$$

  where ρ is a distribution that is zero in the interior of the domain, positive on the source set s and negative on the sink set t, with total source weight $\int_s \rho \mathrm{d}V = +1$ and sink weight $\int_t \rho \mathrm{d}V = -1$.

Weighting function (II)

- We select a flow that minimises a measure of the weighting function

$$\begin{aligned} E[w] &= \int_V \frac{1}{2} w^2 \mathrm{d}V, \\ &= \int_V \frac{1}{2} \|\vec{F}\|_2^2 \mathrm{d}V. \end{aligned}$$

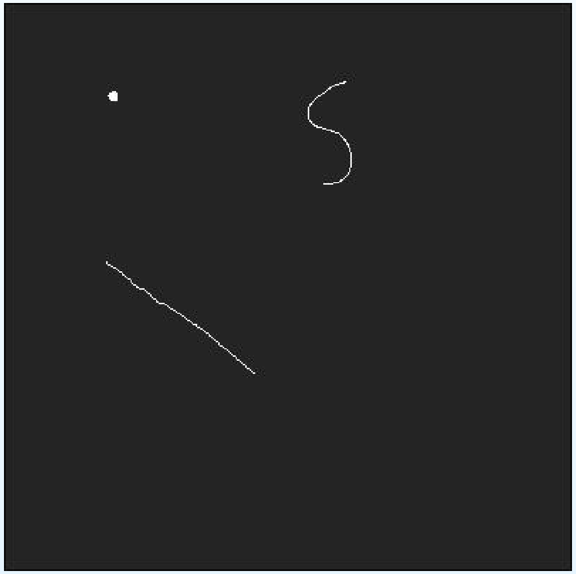
  In this way we will ensure that the weighting function is not arbitrarily large.

- We find that $\frac{\partial \vec{F}}{\partial t}$ is a conservative field, equivalently a potential flow. So we set $\frac{\partial \vec{F}}{\partial t} = \nabla \phi_t$ then, and replace the divergence of the flow $\vec{F}$ in equation (9) by the Laplacian of ϕ to obtain

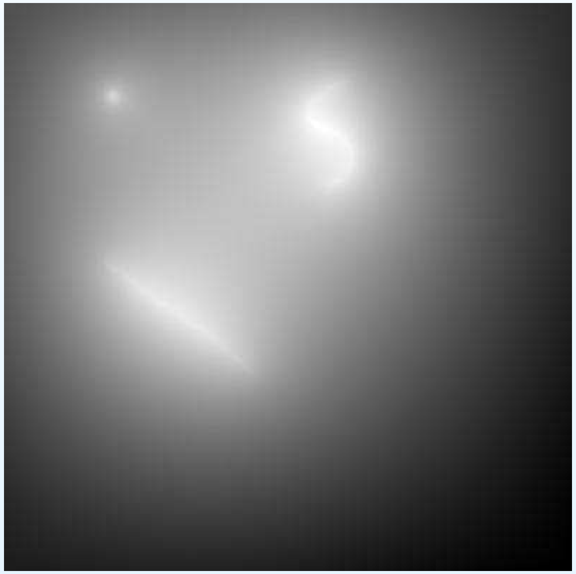
$$\Delta \phi = \rho. \tag{9}$$

Metric weighting (III)

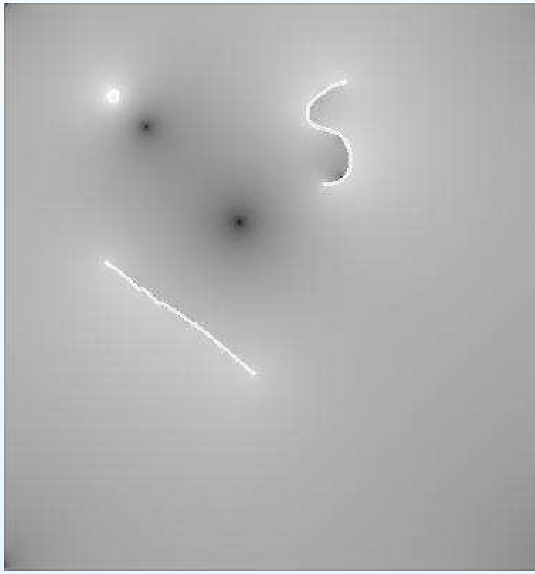
- This latter equation is classical in electrostatics, can be solved by convolving ϕ with a known impulse response $\phi_\odot$
- In 2D, $\phi_\odot = \frac{1}{2\pi}\ln(r)$ and in 3D, $\phi_\odot = -\frac{1}{4\pi r}$.



seeds



convolution

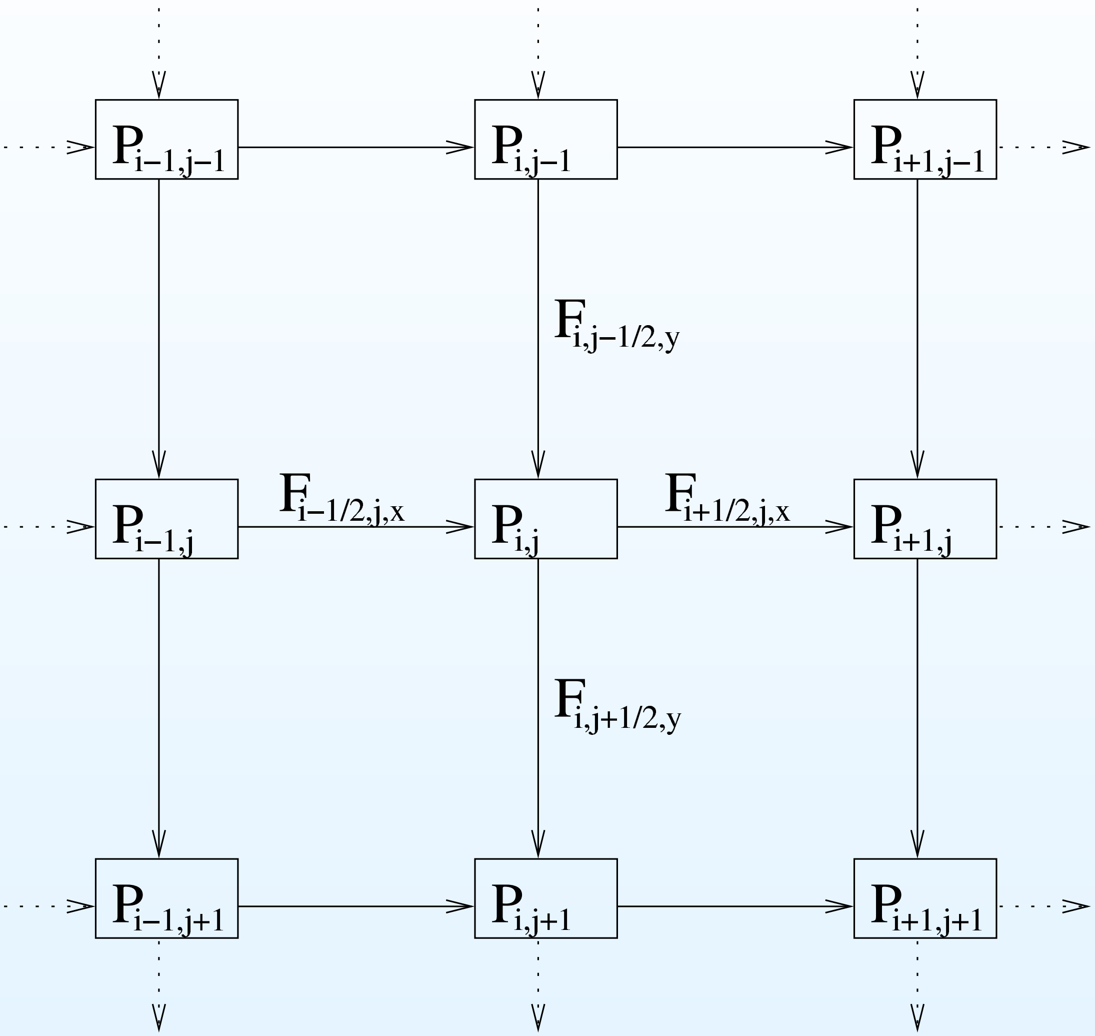


weight

Implementation

- Equations (5) and (6) are discretised on a staggered grid using an explicit first-order scheme in time and space. The scalar field P is stored on grid points while the vector field $\vec{F}$ is stored by component on grid edges.
- This allows for the computation of derivatives without interpolation
- The system of equations is iterated sequentially with the flow magnitude constraint enforced after each timestep.

Implementation (II)



Implementation (III)

Conservation PDE:

$$P_{i,j}^{n+1} = P_{i,j}^{n} - \Delta t \left((F^{n}_{i+\frac{1}{2},j,x} - F^{n}_{i-\frac{1}{2},j,x}) + (F^{n}_{i,j+\frac{1}{2},y} - F^{n}_{i,j-\frac{1}{2},y}) \right), \tag{10}$$

Driving PDE:

$$\begin{aligned} F'^{n+1}_{i+\frac{1}{2},j,x} &= F^{n}_{i+\frac{1}{2},j,x} - \Delta t (P^{n+1}_{i+1,j} - P^{n+1}_{i,j}) \\ F'^{n+1}_{i,j+\frac{1}{2},y} &= F^{n}_{i,j+\frac{1}{2},y} - \Delta t (P^{n+1}_{i,j+1} - P^{n+1}_{i,j}). \end{aligned} \tag{11}$$

The magnitude constraint is applied immediately following the update of the flow field by the preceeding equation.

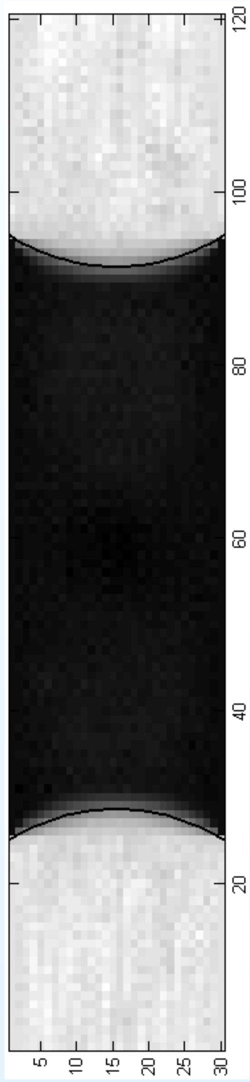
Grids and convergence test

- The previous description works for rectangular n-D grids (finite differences)
- A similar approach would work for arbitrary meshes (finite elements) but is of limited interest in usual segmentation problems.
- At convergence, P is an indicator function for the interior of the surface. In practice, we deem the convergence is reached if the sum of the relative areas of potential $|A_{P\geq(1-\gamma)}|$ and $|A_{P\leq(\gamma-1)}|$ is greater than $\mu\%$. For instance we may set $\gamma = 0.03$ and $\mu = 99$.

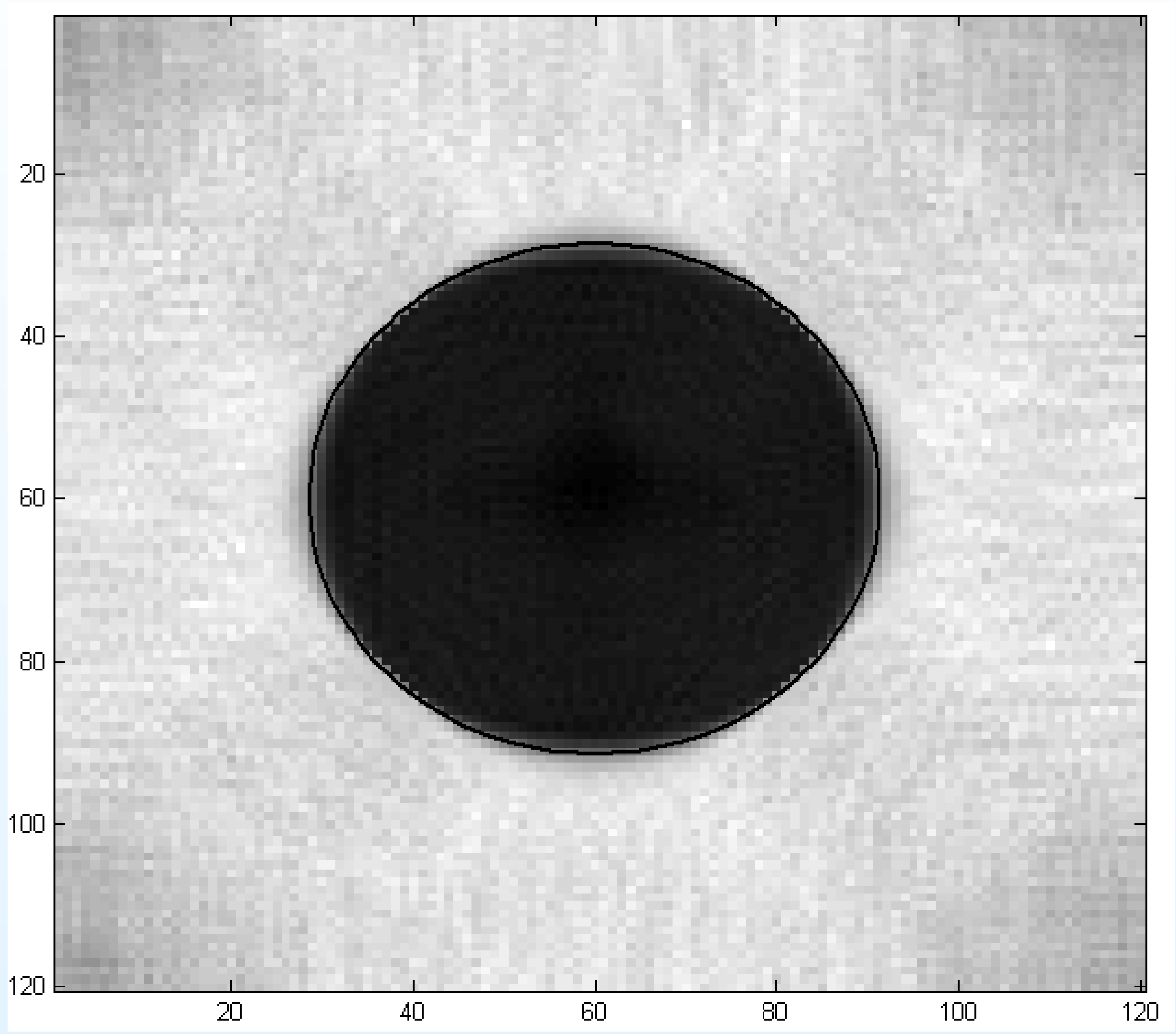
Non-trivial minimal surface

- We take the source to be two identical parallel disks separated by some distance λ in the axis perpendicular to the planes of the disk only.
- In a parallelipedic 3-D box, this source is situated in the first and last planes (only).
- The drain is formed by the 4 sides of the box the sources are not in.
- The metric is uniformly flat ($g = 1$)
- The theoretical continuous solution of this problem is a catenoid, one of the few minimal surface that can be found analytically.
- Here we discretize the box and let our algorithm run until close enough to convergence (e.g: $\#$ of voxels > 0.90 + $\#$ of voxels $< 0.1 > 0.9\times$ total number of voxels)
- The solution is the 0.5–isosurface.

What the image result looks like



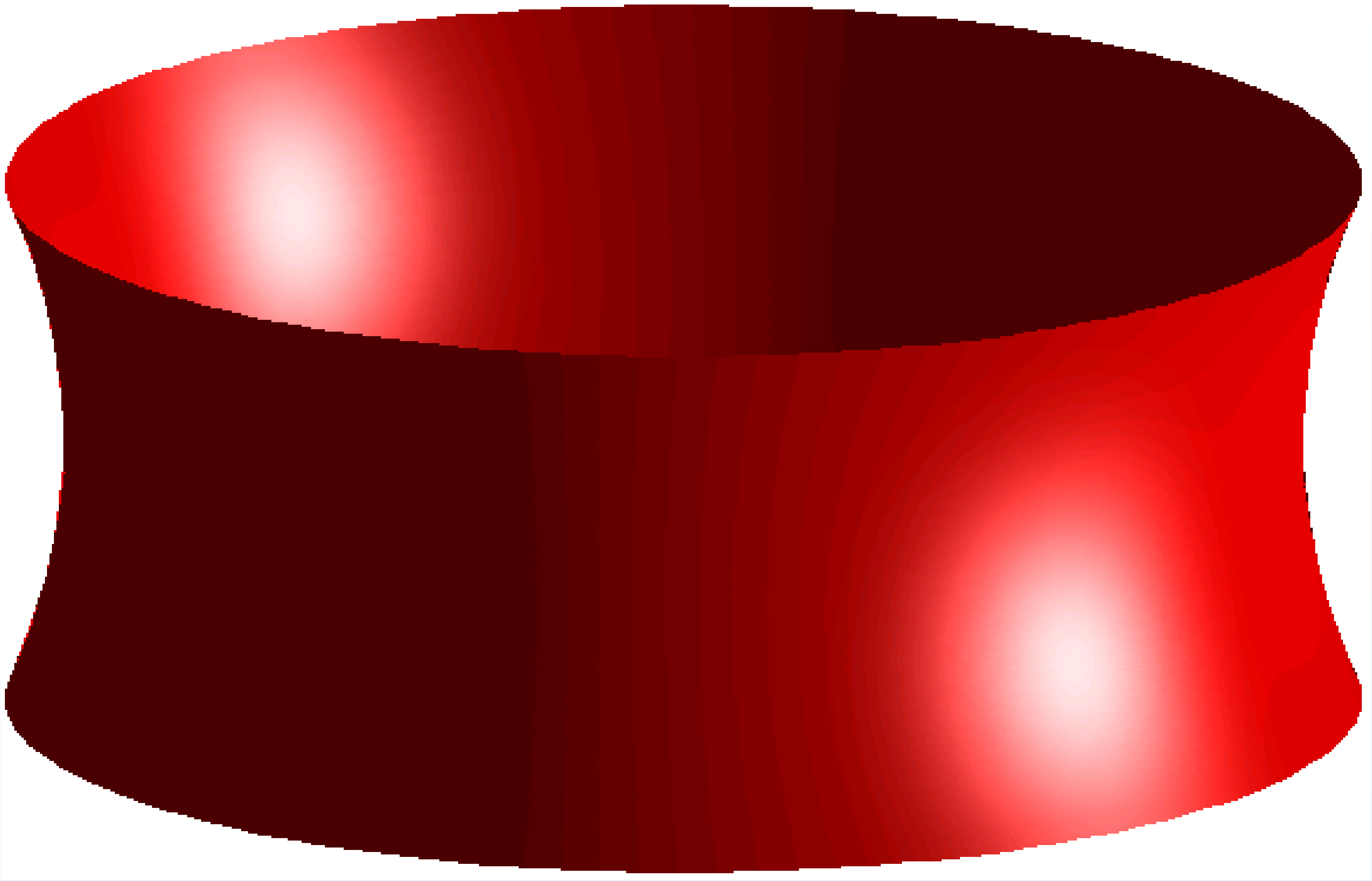
v



h

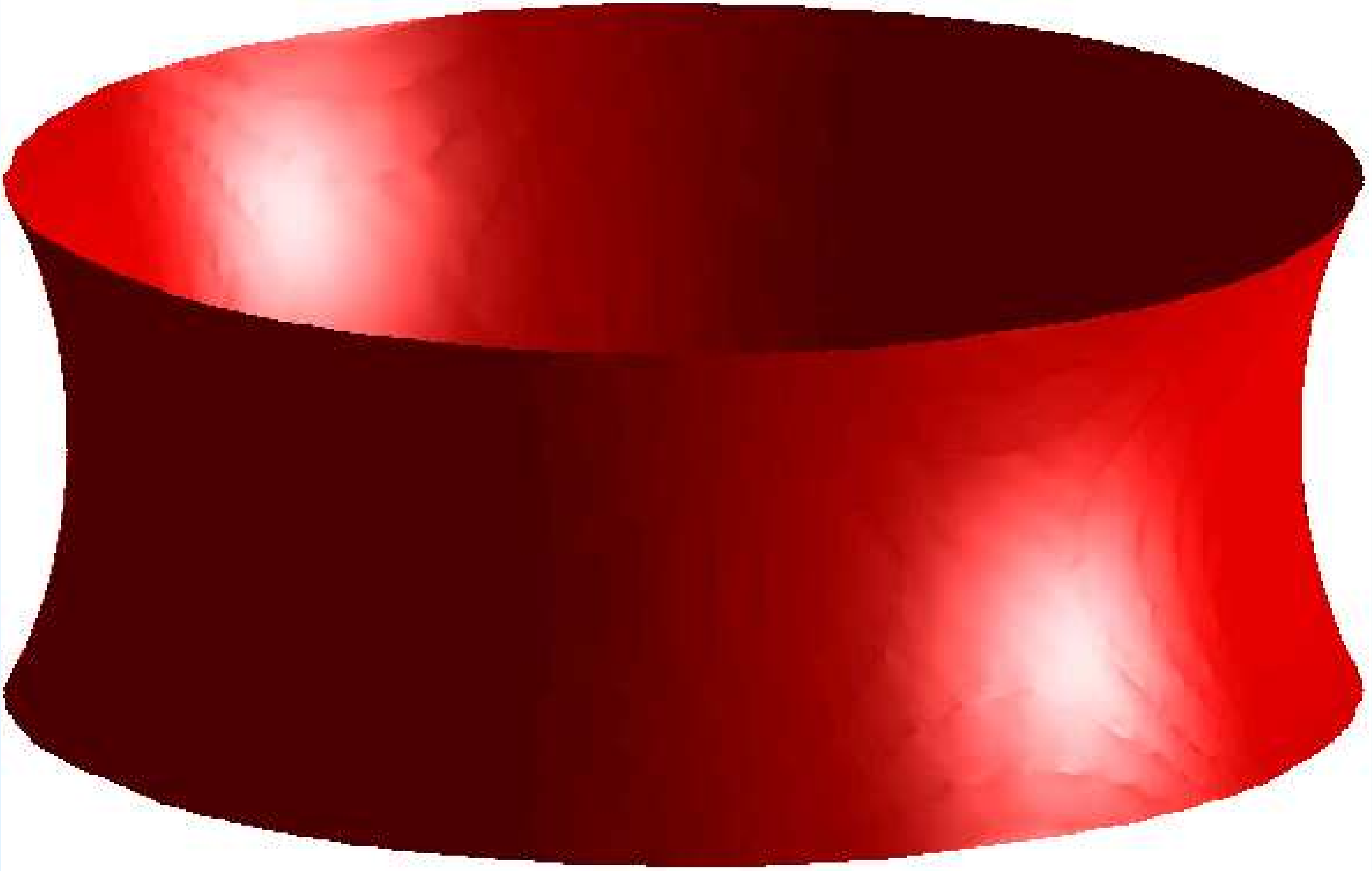
In the parallepiped of data, the correct solution is highlighted in dark.

Isosurface



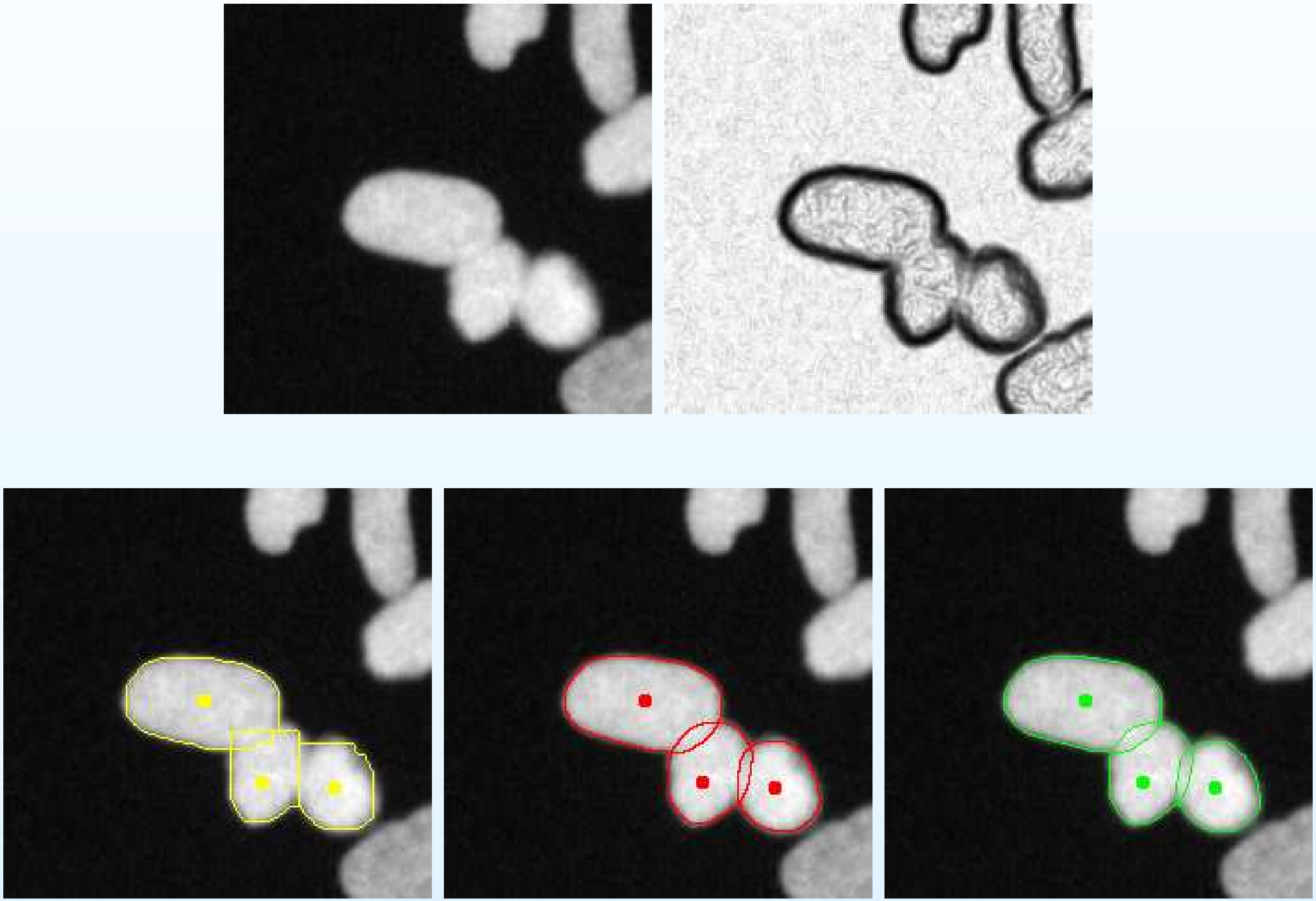
True catenoid

Isosurface



Computed catenoid

2-D results



cells, metric ; discrete GC, GOGAC and CMF solutions.

3-D results (I)

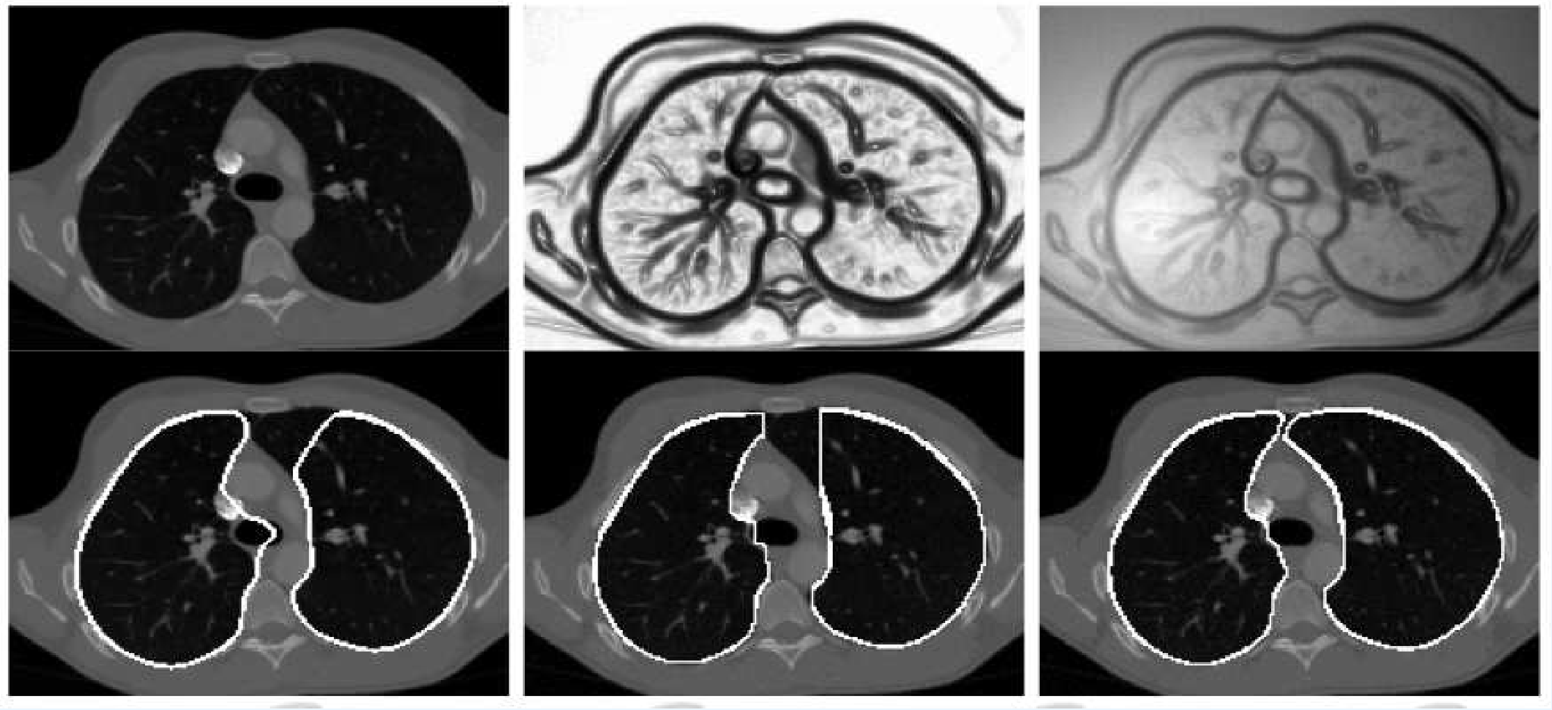
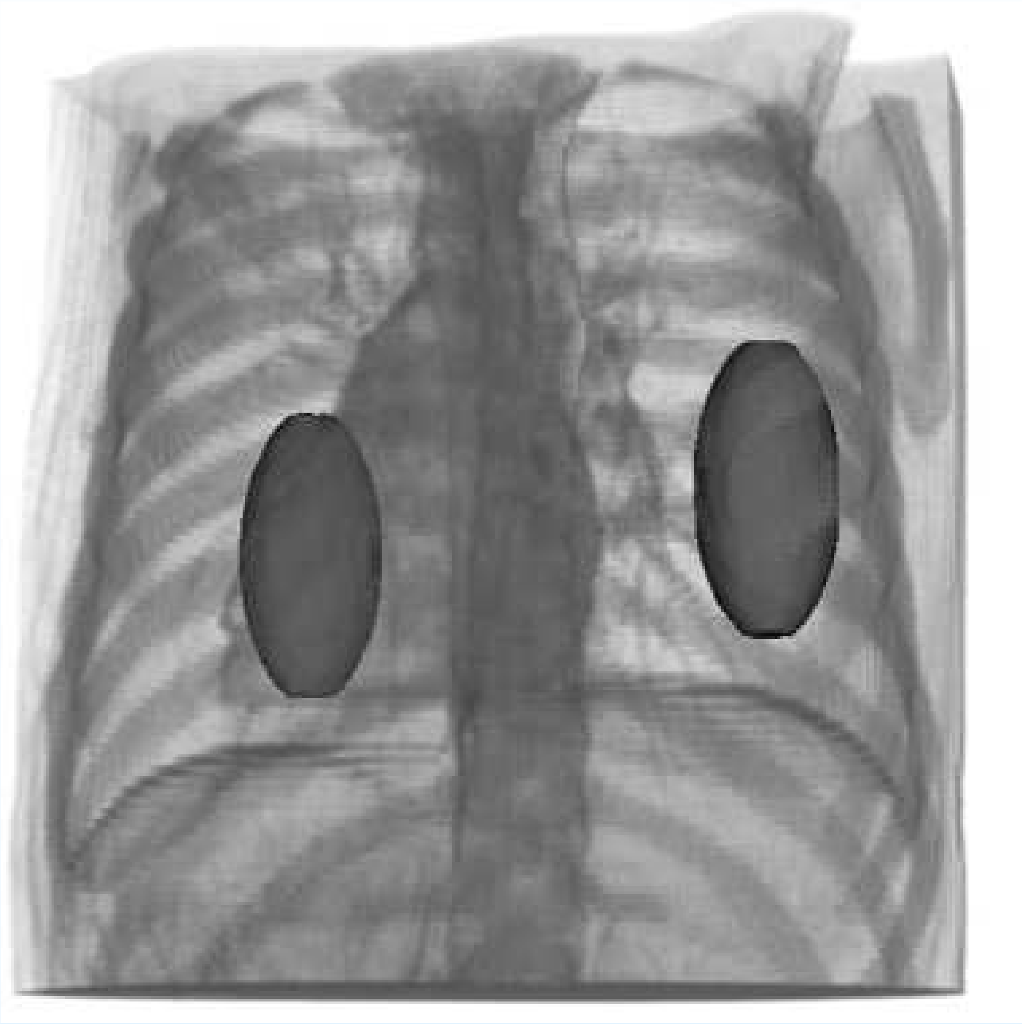


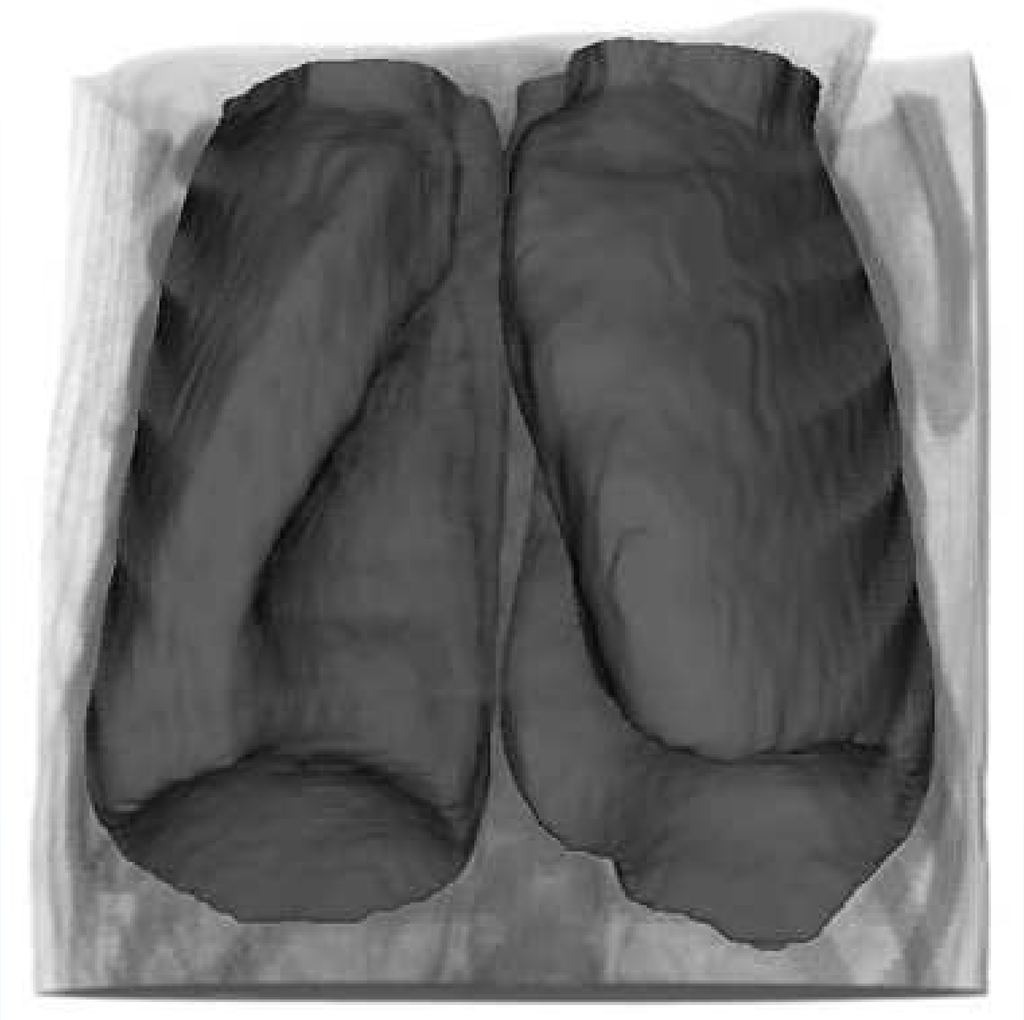
Figure 1: 3D slice of lung segmentation : slice, metric, discrete GC, LS and CMF solutions

3-D results (II)



Seeds

3-D results (II)

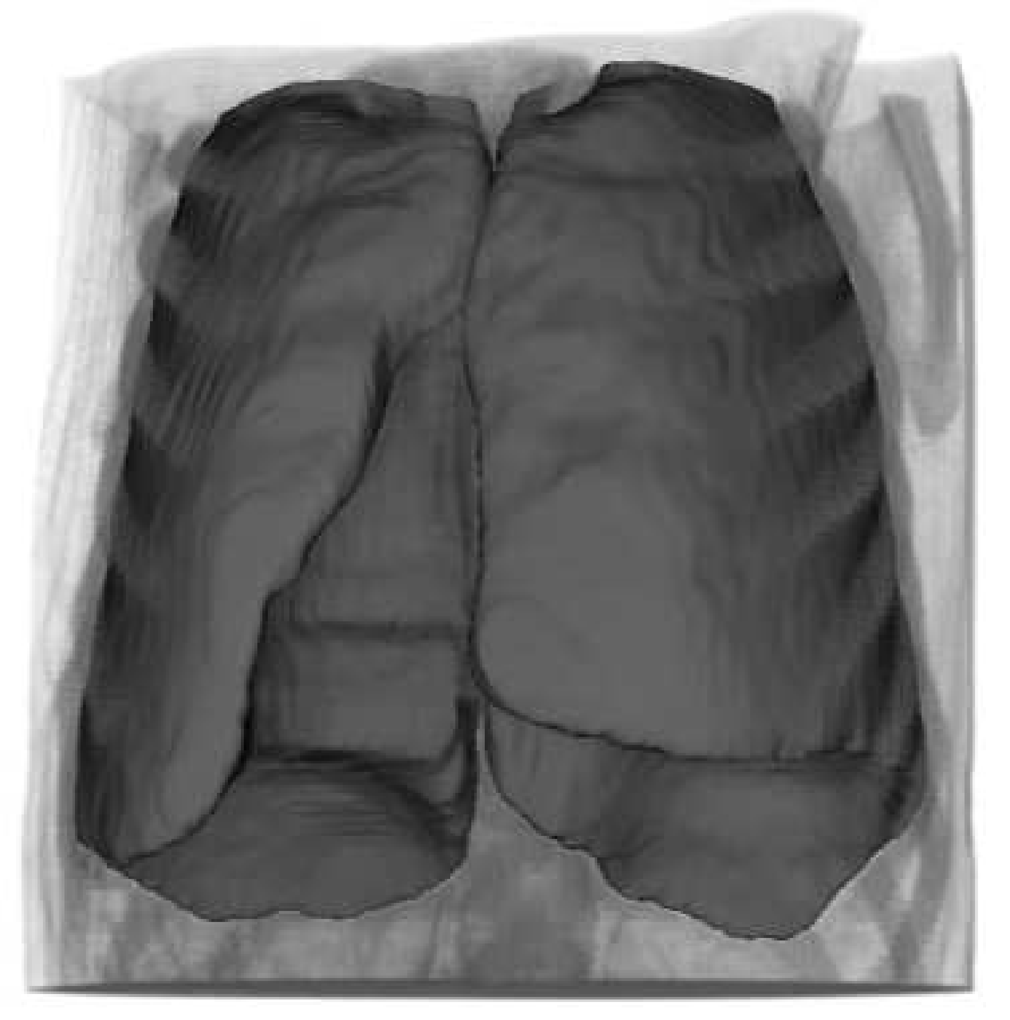


Level-sets segmentation

3-D results (II)

Discrete GC

3-D results (II)

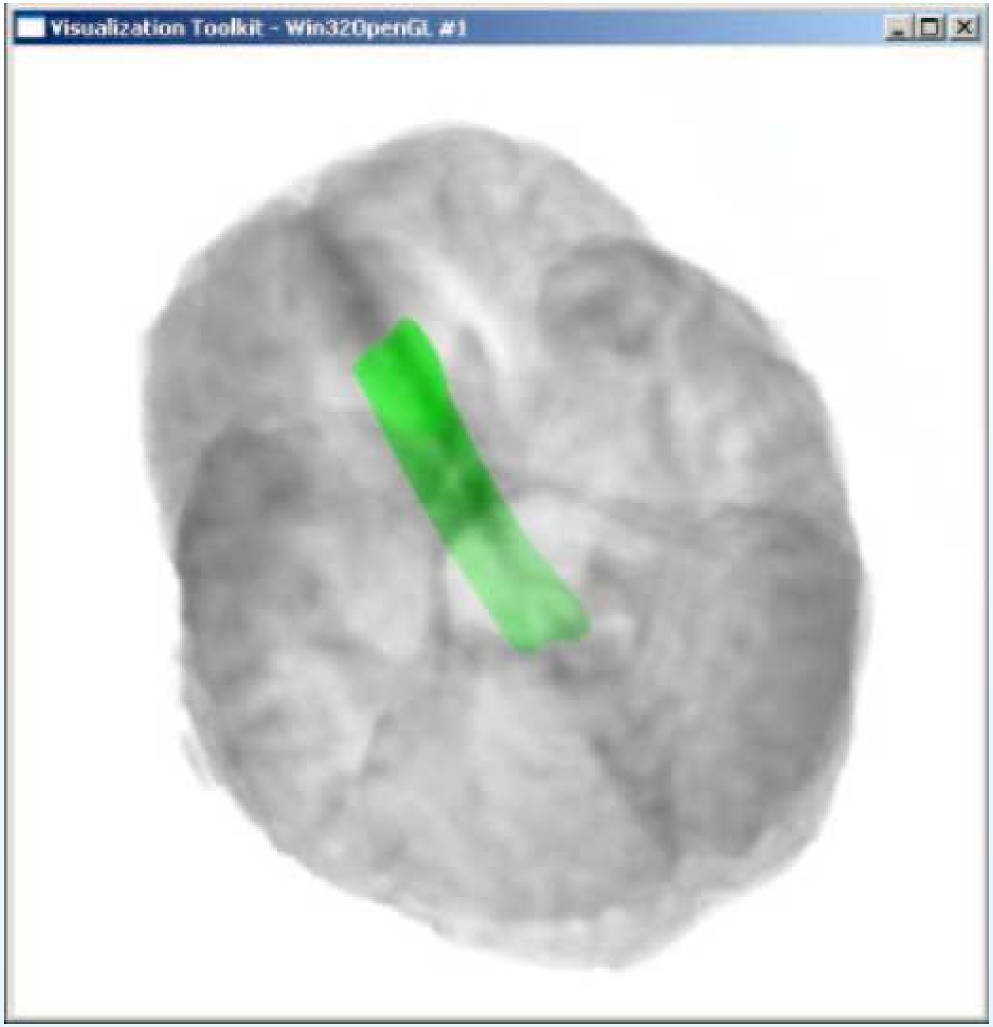


Continuous maxflow segmentation

Brain segmentation

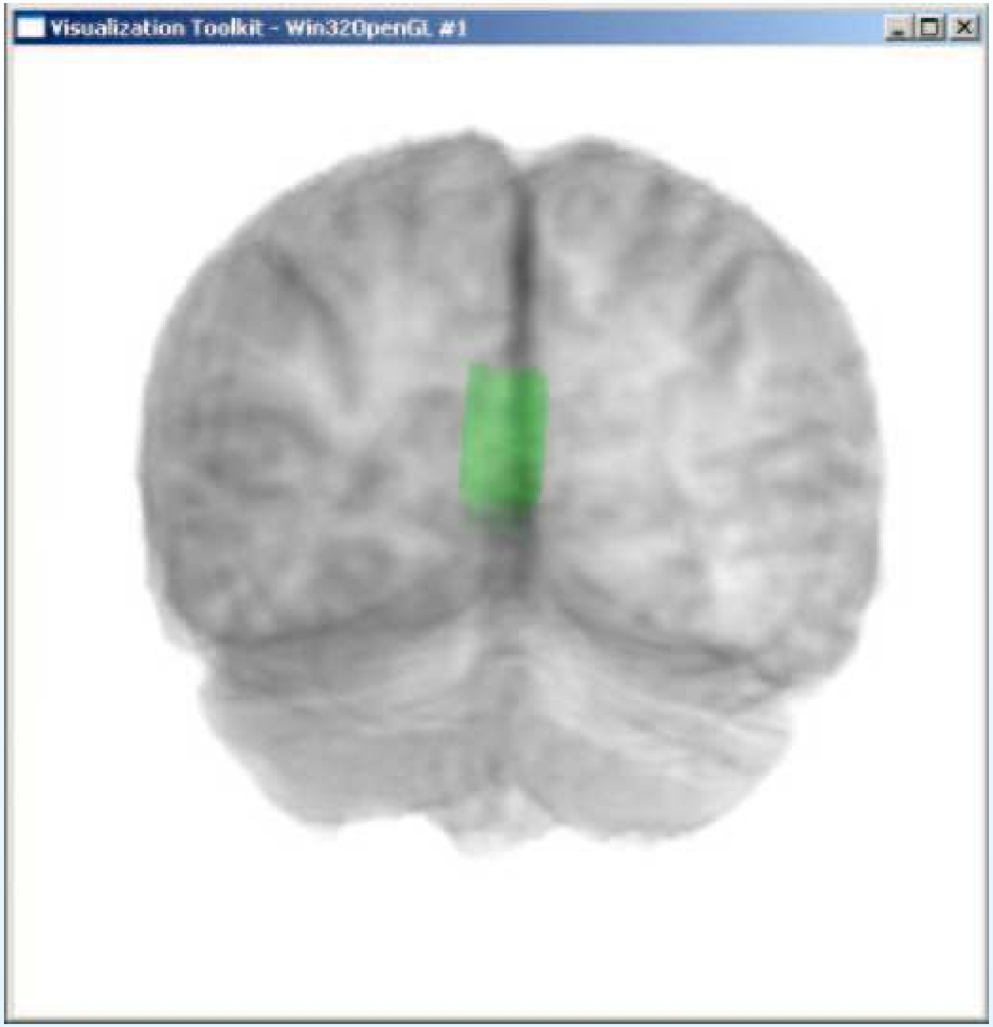
Corpus callosum interactive seed and boundary selections.

Brain segmentation



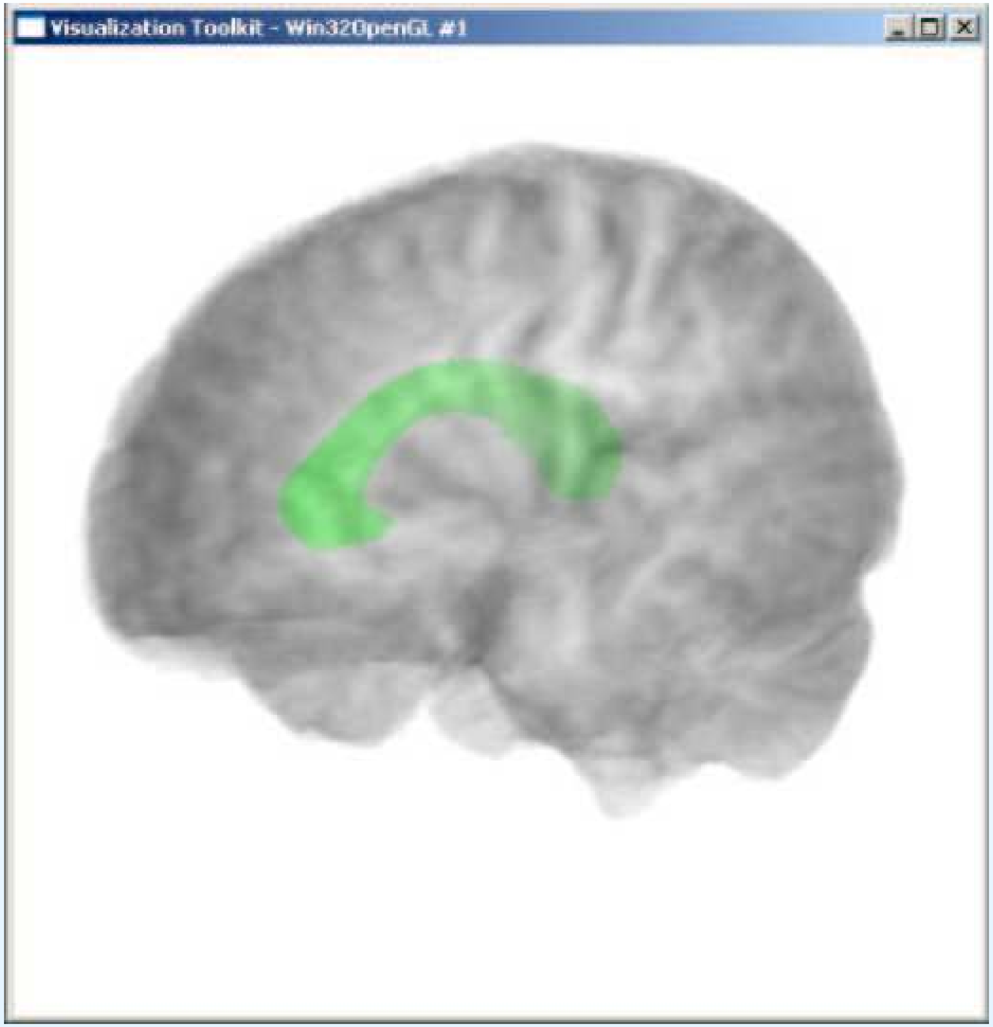
3D segmentation, below.

Brain segmentation



3D segmentation, rear.

Brain segmentation



3D segmentation, side.

Stereo

Figure 1: Stereo reconstruction, parcmeter scene ; discrete GC and CMF solutions

Speed

On the lung segmentation, $200 \times 160 \times 90$

- Level set GAC: 279s
- Graph cut: 128s
- Continuous maximal flow: 28.8s

Conclusion

- Correct solution to the minimal surface problem in 3D (and more)
- Better results than GAC or discrete maximal flow
- Surprisingly faster implementation.
- Limited to binary problems
- An optimal solution does not mean a pleasing solution.

Future work:

- Proof of convergence
- Texture and motion

References (I)

- R. Adams and L. Bischof. Seeded Region Growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6): 641–647, 1994.
- B. Appleton and H. Talbot. Globally optimal geodesic active contours. *Journal of Mathematical Imaging and Vision*, (23) 67–86, 2005.
- B. Appleton and H. Talbot. Globally minimal surfaces by continuous maximal flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):106–118, 2006.
- S. Beucher and C. Lantuéjoul. Use of watersheds in contour detection. In *Int. Workshop on Image Processing*, Rennes, France, Sept. 1979. CCETT/IRISA.
- J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6): 679–698, 1986.
- V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal on Computer Vision*, (1): 61–79, 1997a.
- V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. Minimal surfaces based object segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19: 394–398, 1997b.

References (II)

- Laurent D. Cohen and Ron Kimmel. Global minimum for active contour models: A minimal path approach. *International Journal of Computer Vision*, 24(1): 57–78, August 1997.
- J. L. R. Ford and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.
- A.V. Goldberg and R.E. Tarjan. A new approach to the maximum-flow problem. *Journal of the ACM*, 35: 921–940, 1988.
- R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast geodesic active contours. *IEEE Trans. On Image Processing*, 10(10): 1467–1475, 2001.
- M. Iri. *Survey of Mathematical Programming*. North-Holland, Amsterdam, 1979.
- M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.
- R. Sedgewick. *Algorithms in C*. Number 5. Addison-Wesley, third edition, 2002.
- G. Strang. Maximal flow through a domain. *Mathematical Programming*, 26: 123–143, 1983.