Uniformization, the Monge-Kantorovich Problem and Medical Imaging

Sigurd Angenent & Steve Haker & Ron Kikinis & Allen Tannenbaum

12 October, 2005
Medical Imaging produces pictures like these

PET

MRI

CT
Medical Imaging produces pictures like these

PET

MRI

CT

(These are actually 2D slices of 3D images)
“Segmentation”
Extracting relevant geometric objects from an image
“Segmentation”
Extracting relevant geometric objects from an image
Putting together all the slices leads to a surface.

The segmented surface is only minimally smooth.
Registration:

Find a “canonical method” to map the given surface onto a standard surface, such as the unit sphere $S^2$. 
Registration:

Find a "canonical method" to map the given surface onto a standard surface, such as the unit sphere $S^2$. 

\[ \text{?} \rightarrow S^2 \]
Registration:

Find a “canonical method” to map the given surface onto a standard surface, such as the unit sphere $S^2$.
Which map has the smallest distortion?
Which map has the smallest distortion?

An **isometry**, which preserves all intrinsic distances on the surface would have the least imaginable distortion. However, given two surfaces $S$ and $\Sigma$, there is usually no isometry from $S$ to $\Sigma$. 
Which map has the smallest distortion?

An isometry, which preserves all intrinsic distances on the surface would have the least imaginable distortion. However, given two surfaces $S$ and $\Sigma$, there is usually no isometry from $S$ to $\Sigma$.

Options: instead of looking for an isometry, try to find
Which map has the smallest distortion?

An isometry, which preserves all intrinsic distances on the surface would have the least imaginable distortion. However, given two surfaces $S$ and $\Sigma$, there is usually no isometry from $S$ to $\Sigma$.

Options: instead of looking for an isometry, try to find

conformal maps, or area preserving maps, or . . .
**Definition.** A map $f : S \to \Sigma$ is conformal if “$f$ preserves angles,” or, if there is some positive function $\lambda : S \to \mathbb{R}_+$ such that

$$\langle df_p \cdot \vec{v}, df_p \cdot \vec{w} \rangle = \lambda(p) \langle \vec{v}, \vec{w} \rangle$$

holds for all points $p \in S$ and all tangent vectors $\vec{v}, \vec{w} \in T_p S$. 
Definition. A map $f: S \to \Sigma$ is conformal if “$f$ preserves angles,” or, if there is some positive function $\lambda: S \to \mathbb{R}_+$ such that

$$\langle df_p \cdot \vec{v}, df_p \cdot \vec{w} \rangle = \lambda(p) \langle \vec{v}, \vec{w} \rangle$$

holds for all points $p \in S$ and all tangent vectors $\vec{v}, \vec{w} \in T_pS$.

Some examples...
Theorem. (Gauss) A sufficiently small neighborhood of any point $p$ on a smooth surface $\Sigma \subset \mathbb{R}^3$ has a conformal parametrization $X : U \rightarrow \Sigma$ ($U \subset \mathbb{R}^2$ open).

$$|X_u|^2 = |X_v|^2 = \lambda(u, v), \quad X_u \cdot X_v = 0.$$
Theorem. (Gauss) A sufficiently small neighborhood of any point $p$ on a smooth surface $\Sigma \subset \mathbb{R}^3$ has a conformal parametrization $X : U \rightarrow \Sigma$ ($U \subset \mathbb{R}^2$ open).

$$|X_u|^2 = |X_v|^2 = \lambda(u, v), \quad X_u \cdot X_v = 0.$$ 

The inverse of $X$ provides conformal coordinates $u, v$ on the neighborhood $X(U)$ of $p$. 
**Theorem. (Gauss)** A sufficiently small neighborhood of any point $p$ on a smooth surface $\Sigma \subset \mathbb{R}^3$ has a conformal parametrization $X: U \to \Sigma$ ($U \subset \mathbb{R}^2$ open).

$$|X_u|^2 = |X_v|^2 = \lambda(u, v), \quad X_u \cdot X_v = 0.$$  

The inverse of $X$ provides conformal coordinates $u, v$ on the neighborhood $X(U)$ of $p$.

In conformal coordinates the metric is given by

$$(ds)^2 = \lambda(u, v)((du)^2 + (dv)^2).$$
Example – Stereographic projection.

\[ X(u, v) = \begin{pmatrix} \frac{2u}{1 + u^2 + v^2} \\ \frac{2v}{1 + u^2 + v^2} \\ \frac{1 - u^2 - v^2}{(1 + u^2 + v^2)} \end{pmatrix} \]
Example – Stereographic projection.

\[ X(u, v) = \begin{pmatrix} \frac{2u}{1 + u^2 + v^2} \\ \frac{2v}{1 + u^2 + v^2} \\ \frac{(1 - u^2 - v^2)}{(1 + u^2 + v^2)} \end{pmatrix} \]
Example – Mercator’s map. The exponential map $z \mapsto e^z$ is a conformal map from the complex plane to itself. The composition of stereographic projection and the exponential map is therefore also conformal.
Example – Mercator’s map. The exponential map $z \mapsto e^z$ is a conformal map from the complex plane to itself. The composition of stereographic projection and the exponential map is therefore also conformal.
Existence of conformal maps in general. If $\Sigma \subset \mathbb{R}^3$ is a simply connected surface, then the uniformization theorem implies that there exists a conformal map $f : \Sigma \to S^2$. 
Existence of conformal maps in general. If $\Sigma \subset \mathbb{R}^3$ is a simply connected surface, then the uniformization theorem implies that there exists a conformal map $f : \Sigma \to S^2$.

Uniqueness: Choose a point $p \in \Sigma$, a unit tangent vector $\vec{v} \in T_p \Sigma$ and a scale $\lambda > 0$. Then there is only one conformal map $f : \Sigma \to S^2$ such that

- $f(p)$ is the northpole $N$, 

Existence of conformal maps in general. If $\Sigma \subset \mathbb{R}^3$ is a simply connected surface, then the uniformization theorem implies that there exists a conformal map $f : \Sigma \rightarrow S^2$.

**Uniqueness:** Choose a point $p \in \Sigma$, a unit tangent vector $\vec{v} \in T_p \Sigma$ and a scale $\lambda > 0$. Then there is only one conformal map $f : \Sigma \rightarrow S^2$ such that

- $f(p)$ is the northpole $N$,
- $df_p \cdot \vec{v}$ is tangent to the Greenwich meridian, and
Existence of conformal maps in general. If $\Sigma \subset \mathbb{R}^3$ is a simply connected surface, then the uniformization theorem implies that there exists a conformal map $f : \Sigma \rightarrow S^2$.

**Uniqueness:** Choose a point $p \in \Sigma$, a unit tangent vector $\vec{v} \in T_p \Sigma$ and a scale $\lambda > 0$. Then there is only one conformal map $f : \Sigma \rightarrow S^2$ such that

- $f(p)$ is the northpole $N$,
- $df_p \cdot \vec{v}$ is tangent to the Greenwich meridian, and
- $\|df_p\| = \sqrt{\lambda}$.
Generalized stereographic projection. We call the composition

$$F : \Sigma \setminus \{p\} \xrightarrow{f} S^2 \setminus \{N\} \xrightarrow{X} \mathbb{R}^2 = \mathbb{C}$$

a generalized stereographic projection.
Generalized stereographic projection. We call the composition

\[ F : \Sigma \setminus \{p\} \xrightarrow{f} S^2 \setminus \{N\} \xrightarrow{X} \mathbb{R}^2 = \mathbb{C} \]

a generalized stereographic projection.

\( F \) is smooth on \( \Sigma \setminus \{p\} \), and satisfies a PDE,

\[ \Delta_\Sigma F = 0, \]

where \( \Delta_\Sigma \) is the Laplace-Beltrami operator.
Generalized stereographic projection. We call the composition

\[ F : \Sigma \setminus \{p\} \xrightarrow{f} S^2 \setminus \{N\} \xrightarrow{X} \mathbb{R}^2 = \mathbb{C} \]

a generalized stereographic projection.

\( F \) is smooth on \( \Sigma \setminus \{p\} \), and satisfies a PDE,

\[ \Delta_\Sigma F = 0, \]

where \( \Delta_\Sigma \) is the Laplace-Beltrami operator.

In conformal coordinates \( \Delta_\Sigma F = \frac{1}{\lambda}(F_{uu} + F_{vv}) \).
Generalized stereographic projection. We call the composition
\[ F : \Sigma \setminus \{p\} \xrightarrow{f} S^2 \setminus \{N\} \xrightarrow{\chi} \mathbb{R}^2 = \mathbb{C} \]
a generalized stereographic projection.

\( F \) is smooth on \( \Sigma \setminus \{p\} \), and satisfies a PDE,
\[ \Delta_{\Sigma} F = 0, \]
where \( \Delta_{\Sigma} \) is the Laplace-Beltrami operator.

In conformal coordinates \( \Delta_{\Sigma} F = \frac{1}{\chi}(F_{uu} + F_{vv}) \).

Boundary condition near \( p \)?
The PDE for $F$. Given a point $p \in \Sigma$, a scale $\lambda > 0$ and a unit vector $\vec{v} \in T_p \Sigma$.

Choose $\vec{w} \in T_p \Sigma$ so $\{\vec{v}, \vec{w}\}$ is orthonormal. Then $F$ satisfies

$$\Delta_\Sigma F = \lambda \nabla_{\vec{v} + i\vec{w}} \delta_p(u, v)$$

(1)

in the sense of distributions.

Proof: In conformal coordinates $z = u + iv$ near $p$ one has

$$F(u, v) = \frac{A}{u + iv} + \mathcal{O}(1)$$

for some constant $0 \neq A \in \mathbb{C}$. 

The PDE for $F$. Given a point $p \in \Sigma$, a scale $\lambda > 0$ and a unit vector $\vec{v} \in T_p \Sigma$.

Choose $\vec{w} \in T_p \Sigma$ so $\{\vec{v}, \vec{w}\}$ is orthonormal. Then $F$ satisfies

$$\Delta_\Sigma F = \lambda \nabla_{\vec{v} + i\vec{w}} \{\delta_p(u, v)\} \tag{1}$$

in the sense of distributions.

Proof: In conformal coordinates $z = u + iv$ near $p$ one has

$$F(u, v) = \frac{A}{u + iv} + O(1)$$

for some constant $0 \neq A \in \mathbb{C}$. 
Solution by finite elements. For every $\phi \in C^\infty(\Sigma)$ one has

$$\int \int_\Sigma \nabla \phi \cdot \nabla F \, dA = \langle -\Delta_\Sigma F, \phi \rangle$$

$$= \langle -\lambda \nabla \vec{v}_+ i \vec{w} \delta_p, \phi \rangle$$

$$= \lambda \left[ \nabla \vec{v}_+ i \vec{w} \right](p) \right\} \tag{2}$$
Solution by finite elements. For every $\phi \in C^\infty(\Sigma)$ one has

$$\int \int_{\Sigma} \nabla \phi \cdot \nabla F \, dA = \langle -\Delta_{\Sigma} F, \phi \rangle$$

$$= \langle -\lambda \nabla \vec{v} + i \vec{w}, \delta_p, \phi \rangle$$

$$= \lambda \left[ \nabla \vec{v} + i \vec{w} \phi \right](p) \tag{2}$$

Since the surface $\Sigma$ is approximately given through a triangulation, one can consider only piecewise linear functions on $\Sigma$. Then (2) leads to a (large but sparse) system of linear equations which can be solved by, e.g. the conjugate gradient method.
The result:
Synthetic image
Other applications – virtual colonoscopy
Other applications – virtual colonoscopy

This surface has the topology of a cylinder rather than a sphere.
Given two images
find a canonical map between them
A map from the rectangle \( R \) to itself:
What is an image?
What is an image?

Let $R = [0, 1] \times [0, 1]$.

1. An image could be a grayscale on $R$, i.e. a measurable function $m : R \to [0, 1]$.
What is an image?

Let $\mathbb{R} = [0, 1] \times [0, 1]$.

1. An image could be a grayscale on $\mathbb{R}$, i.e. a measurable function $m : \mathbb{R} \to [0, 1]$

2. An image could be a density, i.e. a Borel measure $\mu$ on $\mathbb{R}$ with density $m$: $d\mu = m(x, y)\,dL$. 
What is an image?

Let $R = [0, 1] \times [0, 1]$.

1 – An image could be a grayscale on $R$, i.e. a measurable function $m : R \rightarrow [0, 1]$

2 – An image could be a density, i.e. a Borel measure $\mu$ on $R$ with density $m$: $d\mu = m(x,y) dL$.

The difference is in the mapping behaviour
A map $\phi : \mathbb{R} \to \mathbb{R}$

- acts on a grayscale by composition, $(\phi, m) \mapsto m \circ \phi$
A map $\phi : \mathbb{R} \to \mathbb{R}$

- acts on a grayscale by composition, $(\phi, m) \mapsto m \circ \phi$
- acts on densities by pull-back, $(\phi, \mu) \mapsto \phi^* \mu$

where $(\phi^* \mu)(E) = \mu[\phi(E)]$. 
A map \( \phi : \mathbb{R} \to \mathbb{R} \)

- acts on a grayscale by composition, \((\phi, m) \mapsto m \circ \phi\)
- acts on densities by pull-back, \((\phi, \mu) \mapsto \phi^* \mu\)

where \((\phi^* \mu)(E) = \mu[\phi(E)]\).

If \( \mu = m(x) dL \) then

\[ \phi^* \mu = \det(d\phi) \ m \circ \phi \ dL. \]
Monge’s transportation cost. Let $C : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a given function. Then Monge (1781) defined the cost of transporting the measure $\mu$ to $\phi^*\mu$ to be

$$M(\phi) = \int_{\mathbb{R}} C(x, \phi(x)) \, d\mu(x)$$
Monge’s transportation cost. Let $C : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a given function. Then Monge (1781) defined the cost of transporting the measure $\mu$ to $\phi^*\mu$ to be

$$M(\phi) = \int_{\mathbb{R}} C(x, \phi(x)) \, d\mu(x)$$

Monge’s choice was $C(x, y) = |x - y|$. We will choose

$$C(x, y) = \frac{1}{2} |x - y|^2.$$
The Kantorovich solution. Given two measures $\mu_{1,2}$ on $\mathbb{R}$ with the same total mass, any measure preserving transformation $\phi : \mathbb{R} \rightarrow \mathbb{R}$ (i.e. $\mu_1 = \phi^*(\mu_2)$) determines a reallocation measure $\gamma$ on $\mathbb{R} \times \mathbb{R}$ by the formula

$$\gamma(E) = \int_{\mathbb{R} \times \mathbb{R}} \chi_E(x, \phi(x)) \, d\mu_1(x).$$
The Kantorovich solution. Given two measures $\mu_{1,2}$ on $\mathbb{R}$ with the same total mass, any measure preserving transformation $\phi : \mathbb{R} \to \mathbb{R}$ (i.e. $\mu_1 = \phi^*(\mu_2)$) determines a reallocation measure $\gamma$ on $\mathbb{R} \times \mathbb{R}$ by the formula

$$
\gamma(E) = \int_{\mathbb{R} \times \mathbb{R}} \chi_E(x, \phi(x)) d\mu_1(x).
$$

The measure $\gamma$ is supported on the graph of $\phi$, and satisfies

$$
\gamma \geq 0, \quad p_1^*(\gamma) = \mu_1, \quad p_2^*(\gamma) = \mu_2.
$$
The Kantorovich solution. Given two measures $\mu_{1,2}$ on $\mathbb{R}$ with the same total mass, any measure preserving transformation $\phi : \mathbb{R} \to \mathbb{R}$ (i.e. $\mu_1 = \phi^*(\mu_2)$) determines a reallocation measure $\gamma$ on $\mathbb{R} \times \mathbb{R}$ by the formula

$$\gamma(E) = \int_{\mathbb{R} \times \mathbb{R}} \chi_E(x, \phi(x)) \, d\mu_1(x).$$

The measure $\gamma$ is supported on the graph of $\phi$, and satisfies

$$\gamma \geq 0, \quad p_1^*(\gamma) = \mu_1, \quad p_2^*(\gamma) = \mu_2. \quad (3)$$

The Monge transportation cost of $\phi$ is

$$M(\phi) = \int_{\mathbb{R} \times \mathbb{R}} C(x, x') \, d\gamma(x, x') = \langle \gamma, C \rangle.$$
The Kantorovich solution. Given two measures $\mu_{1,2}$ on $\mathbb{R}$ with the same total mass, any measure preserving transformation $\phi : \mathbb{R} \to \mathbb{R}$ (i.e. $\mu_1 = \phi^*(\mu_2)$) determines a reallocation measure $\gamma$ on $\mathbb{R} \times \mathbb{R}$ by the formula

$$\gamma(E) = \int_{\mathbb{R} \times \mathbb{R}} \chi_E(x, \phi(x))d\mu_1(x).$$

The measure $\gamma$ is supported on the graph of $\phi$, and satisfies

$$\gamma \geq 0, \quad p_1^*(\gamma) = \mu_1, \quad p_2^*(\gamma) = \mu_2. \quad (3)$$

The Monge transportation cost of $\phi$ is

$$M(\phi) = \int_{\mathbb{R} \times \mathbb{R}} C(x, x')d\gamma(x, x') = \langle \gamma, C \rangle. \quad (4)$$
Minimization of $M(\gamma)$ with constraints (3) is a classical example of a linear programming problem.
Existence & uniqueness theorems.

A minimizing reallocation measure always exists.

— (Kantorovich)
Existence & uniqueness theorems.

A minimizing reallocation measure always exists.
— (Kantorovich)

If the cost functional is “nice” then the minimizing measure $\gamma$ is supported on the graph of a measure preserving transformation $\phi$ (Gangbo, McCann, Feldman, . . .)
Existence & uniqueness theorems.

A minimizing reallocation measure always exists.  
— (Kantorovich)

If the cost functional is “nice” then the minimizing measure $\gamma$ is supported on the graph of a measure preserving transformation $\phi$ (Gangbo, McCann, Feldman, . . .)

How to compute $\phi$?
Steepest descent for Monge’s cost. To reduce the transportation cost of a given measure preserving map \( \phi_0 : (\mathbb{R}, \mu_1) \rightarrow (\mathbb{R}, \mu_2) \), by deforming it through a family of m.p. maps \( \phi_t : (\mathbb{R}, \mu_1) \rightarrow (\mathbb{R}, \mu_2) \), write the maps as

\[
\phi_0 = \phi_t \circ s_t,
\]

where \( s_t : \mathbb{R} \rightarrow \mathbb{R} \) preserve \( \mu_1 \).
Steepest descent for Monge’s cost. To reduce the transportation cost of a given measure preserving map $\phi_0 : (\mathbb{R}, \mu_1) \to (\mathbb{R}, \mu_2)$, by deforming it through a family of m.p. maps $\phi_t : (\mathbb{R}, \mu_1) \to (\mathbb{R}, \mu_2)$, write the maps as

$$\phi_0 = \phi_t \circ s_t,$$

where $s_t : \mathbb{R} \to \mathbb{R}$ preserve $\mu_1$.

The $s_t$ are determined by their velocity fields,

$$\frac{\partial s_t}{\partial t} = v_t \circ s_t.$$

We will choose the maps $s_t$ by chosing the velocities $\tilde{v}_t$. 
\[
\frac{dM(\phi_t)}{dt} = \frac{d}{dt} \int_{\mathbb{R}} \frac{1}{2} |x - \phi_t(x)|^2 m_1(x) \, dx
\]
(set \(x = s_t(\xi)\))
\[
\frac{dM(\phi_t)}{dt} = \frac{d}{dt} \int_{\mathbb{R}} \frac{1}{2} |x - \phi_t(x)|^2 m_1(x) \, dx \quad (\text{set } x = s_t(\xi))
\]
\[
= \frac{d}{dt} \int_{\mathbb{R}} \frac{1}{2} |s_t(\xi) - \phi_0(\xi)|^2 m_1(\xi) \, d\xi.
\]
\[
\frac{dM(\phi_t)}{dt} = \frac{d}{dt} \int_\mathbb{R} \frac{1}{2} |x - \phi_t(x)|^2 m_1(x) \, dx \\
= \frac{d}{dt} \int_\mathbb{R} \frac{1}{2} |s_t(\xi) - \phi_0(\xi)|^2 m_1(\xi) \, d\xi \\
= \int_\mathbb{R} (s_t(\xi) - \phi_0(\xi)) \cdot \vec{\nu}_t(s_t(\xi)) m_1(\xi) \, d\xi
\]
\[
\frac{dM(\phi_t)}{dt} = \frac{d}{dt} \int_R \frac{1}{2} |x - \phi_t(x)|^2 m_1(x) \, dx 
\quad \text{(set} \ x = s_t(\xi)\text{)}
\]

\[
= \frac{d}{dt} \int_R \frac{1}{2} |s_t(\xi) - \phi_0(\xi)|^2 m_1(\xi) \, d\xi,
\]

\[
= \int_R (s_t(\xi) - \phi_0(\xi)) \cdot \vec{v}_t(s_t(\xi)) m_1(\xi) \, d\xi,
\]

\[
= \int_R (s_t(\xi) - \phi_t(s_t(\xi))) \cdot \vec{v}_t(s_t(\xi)) m_1(\xi) \, d\xi,
\]
\[ \frac{dM(\phi_t)}{dt} = \frac{d}{dt} \int_\mathbb{R} \frac{1}{2} |x - \phi_t(x)|^2 m_1(x) \, dx \quad \text{(set } x = s_t(\xi)) \]

\[ = \frac{d}{dt} \int_\mathbb{R} \frac{1}{2} |s_t(\xi) - \phi_0(\xi)|^2 m_1(\xi) \, d\xi, \]

\[ = \int_\mathbb{R} (s_t(\xi) - \phi_0(\xi)) \cdot \vec{v}_t(s_t(\xi)) m_1(\xi) \, d\xi, \]

\[ = \int_\mathbb{R} (s_t(\xi) - \phi_t(s_t(\xi))) \cdot \vec{v}_t(s_t(\xi)) m_1(\xi) \, d\xi, \]

\[ = \int_\mathbb{R} \vec{v}_t(x) \cdot (x - \phi_t(x)) \, d\mu_1(x) \]
Thus
\[
\frac{dM(\phi_t)}{dt} = \int_{\mathbb{R}} \tilde{v}_t(x) \cdot (x - \phi_t(x)) \, m_1(x) \, dx
\]

Since \( s_t \) preserves \( \mu_1 = m_1(x) \, dL(x) \), one has
\[
\text{div}(m_1(x) \tilde{v}) = 0. \quad (5)
\]
Thus
\[
\frac{dM(\phi_t)}{dt} = \int_{\mathbb{R}} \vec{v}_t(x) \cdot (x - \phi_t(x)) \ m_1(x) dx
\]

Since \( s_t \) preserves \( \mu_1 = m_1(x) dL(x) \), one has
\[
\text{div}(m_1(x)\vec{v}) = 0. \tag{5}
\]

We choose \( \vec{v}_t \) so that
\[
m_1(x)\vec{v}_t = \phi_t(x) - x + \nabla p_t(x)
\]

where \( \Delta p_t = -\Delta (\phi_t(x) - x) \). Then
\[
\frac{dM(\phi_t)}{dt} = -\int_{\mathbb{R}} |x - \phi_t(x) + \nabla p|^2 \ dx
\]
Differentiate $\phi_0 = \phi_t \circ s_t$ w.r.t. time to get the transport equation

$$\frac{\partial \phi_t}{\partial t} + \vec{v}_t \cdot \nabla \phi_t = 0.$$
Summary: Steepest descent is achieved when

\[ m_1(x)\vec{v}_t = \phi_t(x) - x + \nabla p_t(x) \]
\[ \Delta p_t = \Delta(x - \phi_t(x)) \]
\[ \frac{\partial \phi_t}{\partial t} + \vec{v}_t \cdot \nabla \phi_t = 0 \]

and \( \vec{v} \) is tangential to \( \partial R \).
Summary: Steepest descent is achieved when

\[ m_1(x) \vec{v}_t = \phi_t(x) - x + \nabla p_t(x) \]

\[ \Delta p_t = \Delta(x - \phi_t(x)) \]

\[ \frac{\partial \phi_t}{\partial t} + \vec{v}_t \cdot \nabla \phi_t = 0 \]

and \( \vec{v} \) is tangential to \( \partial R \).

Theorems.

- Smooth solutions: short-time existence and uniqueness,
- Weak solutions: global existence and convergence for \( t \rightarrow \infty \) in the context of reallocation measures.

——— (show the movie clip) ———