The Cumulative Distribution Transform For Data Analysis And Machine Learning

Akram Aldroubi

Vanderbilt University

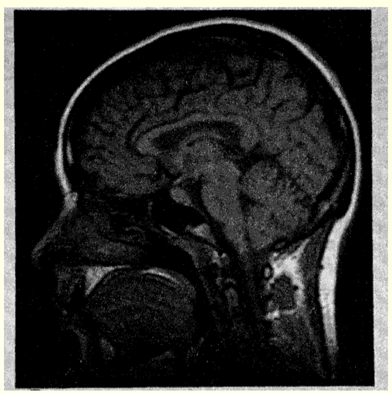
Jubilee of Fourier Analysis and Applications
In Celebration of John Benedetto 80th Birthday

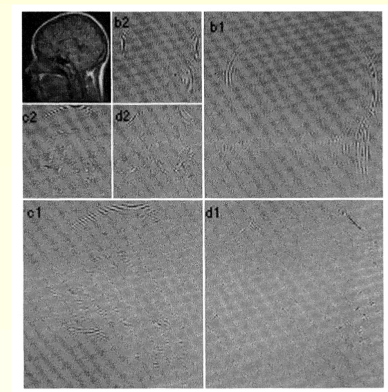
Supported by NIH Grant (Gustavo Rohde PI)

Gustavo Rohde

Transports Transform

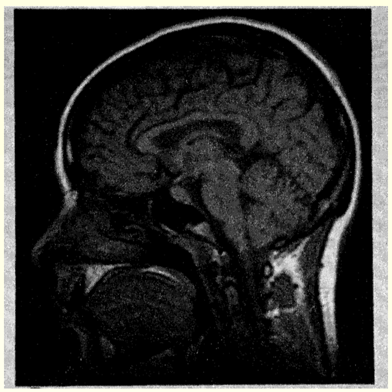
Transforms: Fourier Transform, Wavelet transform, Zak Transform, Shearlets, Scattering “transform,” ...

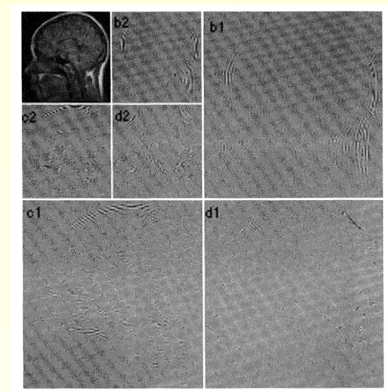




Transports Transform

Transforms: Fourier Transform, Wavelet transform, Zak Transform, Shearlets, Scattering “transform,” ...





Transport Transforms: Non-linear transforms based on transport theory: Monge and Katorovich Transport theory

The Monge Problem (1781)

The Monge Problem Let μ be a pile of sand on $X \subset \mathbb{R}^n$, find the "most efficient way" to transport it to the hole in the ground ν on $X \subset \mathbb{R}^n$.

The Monge Problem (1781)

The Monge Problem Let μ be a pile of sand on $X \subset \mathbb{R}^n$, find the "most efficient way" to transport it to the hole in the ground ν on $X \subset \mathbb{R}^n$.

Let μ, ν be probability measures on $\mathbb{R}^n$ find a map $T^\dagger : \mathbb{R}^n \to \mathbb{R}^n$ such

$$T^\dagger = \arg \min_{\nu = T_{\#}\mu} \int_{\mathbb{R}^n} \|x - T(x)\|^2 d\mu(x),$$

where $\nu(B) = \mu(T^{-1}(B))$ for all measurable measurable sets B.

The Monge Problem (1781)

The Monge Problem Let μ be a pile of sand on $X \subset \mathbb{R}^n$, find the "most efficient way" to transport it to the hole in the ground ν on $X \subset \mathbb{R}^n$.

Let μ, ν be probability measures on $\mathbb{R}^n$ find a map $T^\dagger : \mathbb{R}^n \to \mathbb{R}^n$ such

$$T^\dagger = \arg \min_{\nu = T_{\#}\mu} \int_{\mathbb{R}^n} \|x - T(x)\|^2 d\mu(x),$$

where $\nu(B) = \mu(T^{-1}(B))$ for all measurable measurable sets B.

Weak version of the Monge problem: The Kantorovich problem (1939).

Brenier's Theorem (1991)

(Brenier's Theorem) Let μ, ν be two probability measures on $\mathbb{R}^n$ (finite 2nd moments) that are absolutely continuous w.r.t Lebesgue measure. Then there exists a map $T^\dagger : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T^\dagger = \arg \min_{\nu = T_{\#}\mu} \int_{\mathbb{R}^n} \|x - T(x)\|^2 d\mu(x),$$

where $\nu(B) = \mu(T^{-1}(B))$ for all measurable measurable sets B. Moreover, $T^\dagger$ is unique. (Generalization by Gangbo and McCann 1996, Villani 2006)

Brenier's Theorem (1991)

(Brenier's Theorem) Let μ, ν be two probability measures on $\mathbb{R}^n$ (finite 2nd moments) that are absolutely continuous w.r.t Lebesgue measure. Then there exists a map $T^\dagger : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T^\dagger = \arg \min_{\nu = T_{\#}\mu} \int_{\mathbb{R}^n} \|x - T(x)\|^2 d\mu(x),$$

where $\nu(B) = \mu(T^{-1}(B))$ for all measurable measurable sets B. Moreover, $T^\dagger$ is unique. (Generalization by Gangbo and McCann 1996, Villani 2006)

The transport transform: Let $d\mu(x) = s(x)dx$ and $d\nu(x) = s_0(x)dx$, where r is a fixed reference signal, then the transform $\tilde{s}$ of s is the unique solution to the Monge problem above, i.e., $\tilde{s} = T^\dagger$.

The CDT transform

Let s a smooth probability density function on $[0,1] \subset \mathbb{R}$, and s_0 a reference probability density function on $[0,1] \subset \mathbb{R}$.

The Cumulative Distribution Transform

$$\int_0^{\tilde{s}(x)} s(\xi) d\xi = \int_0^{x} s_0(\xi) d\xi, \quad x \in [0,1].$$

The CDT transform

Let s a smooth probability density function on $[0,1] \subset \mathbb{R}$, and s_0 a reference probability density function on $[0,1] \subset \mathbb{R}$.

The Cumulative Distribution Transform

$$\int_0^{\tilde{s}(x)} s(\xi)d\xi = \int_0^x s_0(\xi)d\xi, \quad x \in [0,1].$$

Typically $s_0 = \chi[0,1]$.

The CDT transform

Let s a smooth probability density function on $[0,1] \subset \mathbb{R}$, and s_0 a reference probability density function on $[0,1] \subset \mathbb{R}$.

The Cumulative Distribution Transform

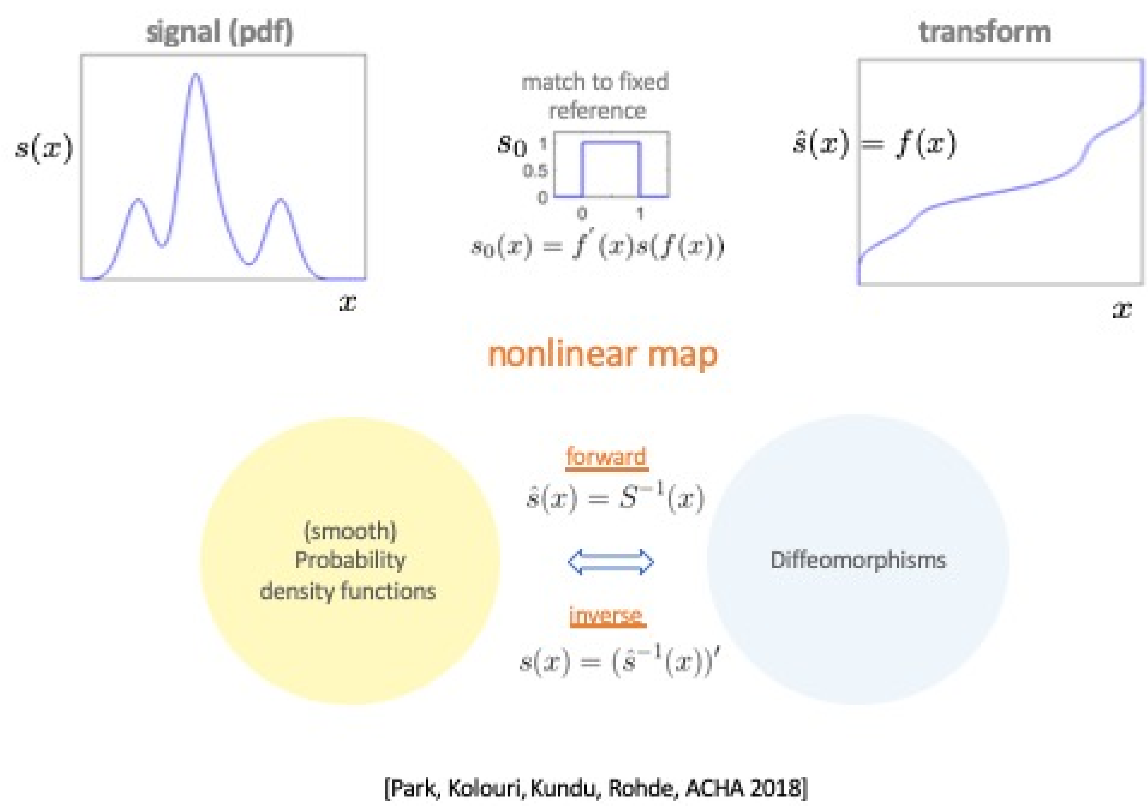
$$\int_0^{\tilde{s}(x)} s(\xi) d\xi = \int_0^x s_0(\xi) d\xi, \quad x \in [0,1].$$

Typically $s_0 = \chi[0,1]$.

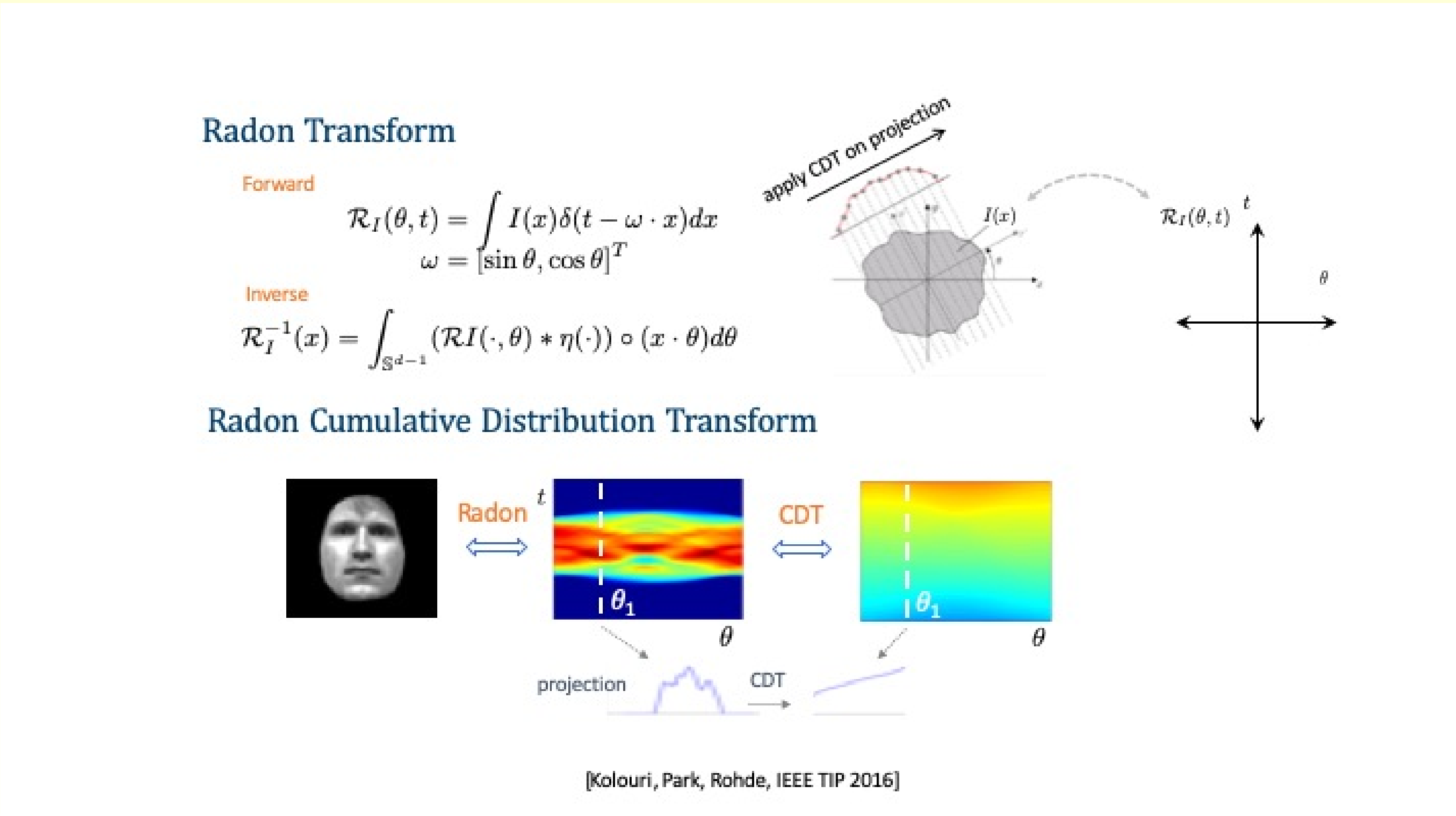
Inverse Transform:

$$\tilde{s}'(x) s(\tilde{s}(x)) = s_0(x).$$

CDT and its inverse

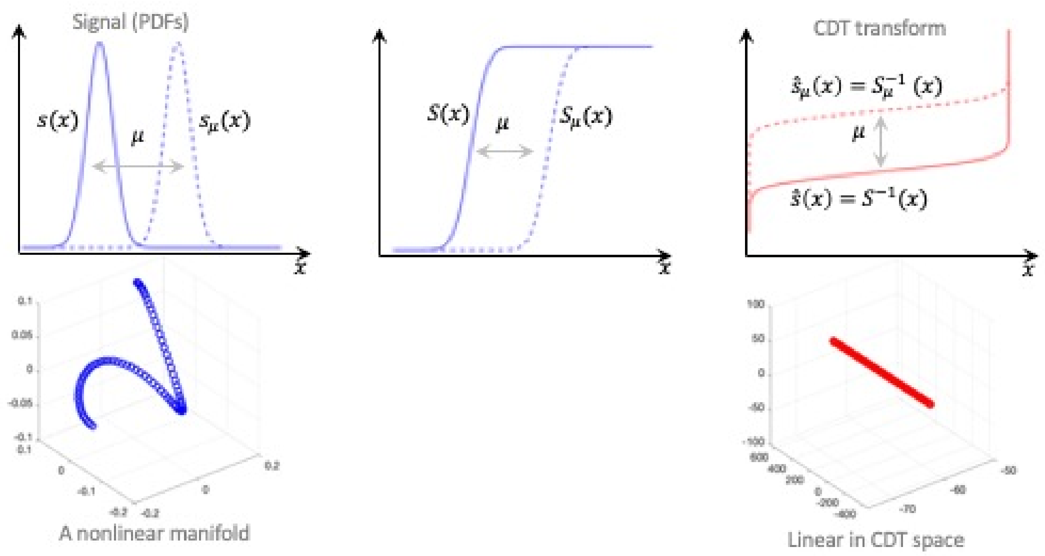


Radon-CDT

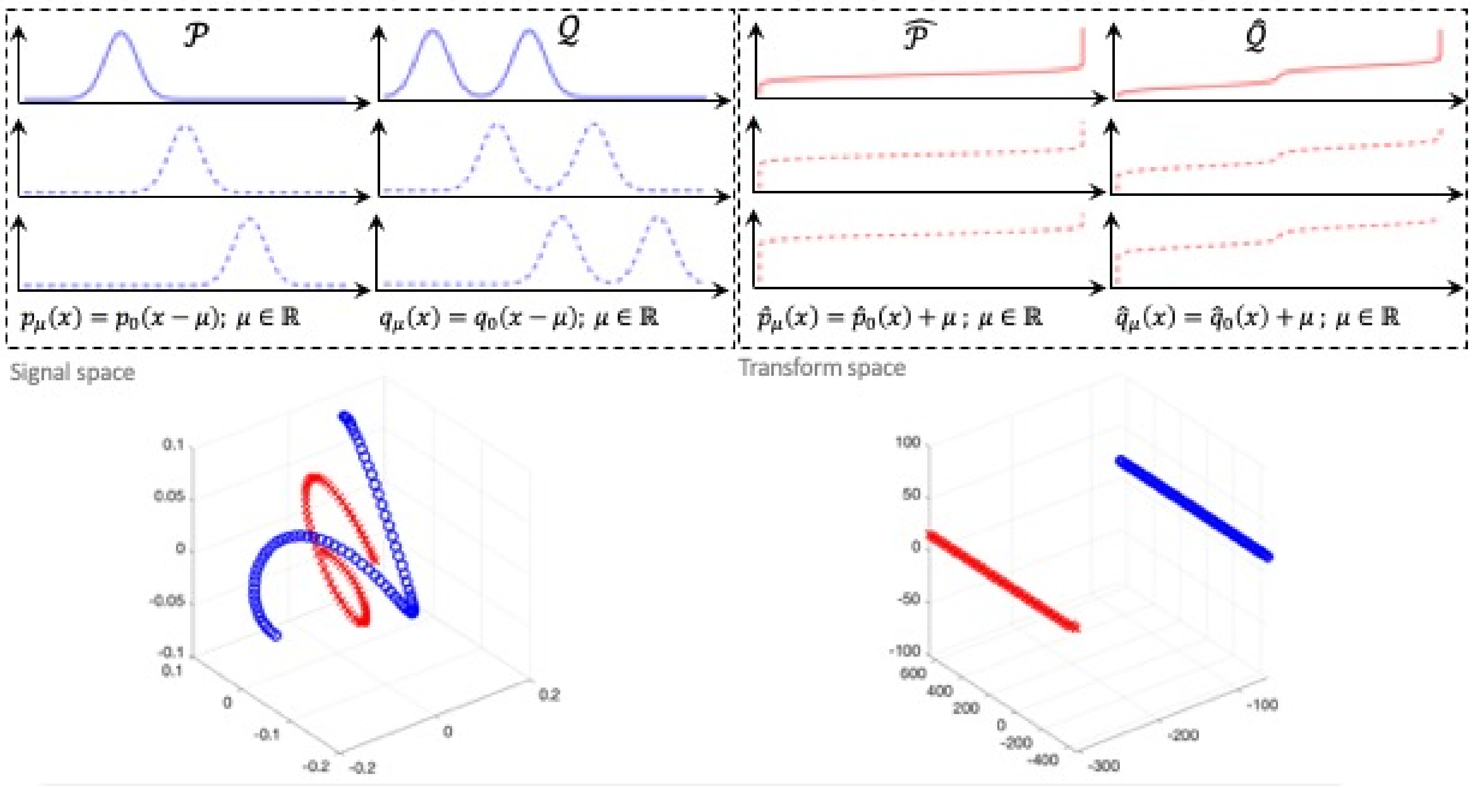

Radon Transform
Forward
$\mathcal{R}_I(\theta, t) = \int I(x)\delta(t - \omega \cdot x)dx$
$\omega = [\sin\theta, \cos\theta]^T$
Inverse
$\mathcal{R}_I^{-1}(x) = \int_{\mathbb{S}^{d-1}} (\mathcal{R}I(\cdot,\theta) * \eta(\cdot)) \circ (x \cdot \theta) d\theta$
apply CDT on projection
$I(x)$
$\mathcal{R}_I(\theta, t)$
t
θ
Radon Cumulative Distribution Transform
Radon
CDT
t
θ_1
θ
θ_1
θ
projection
CDT
[Kolouri, Park, Rohde, IEEE TIP 2016]

Convexification properties

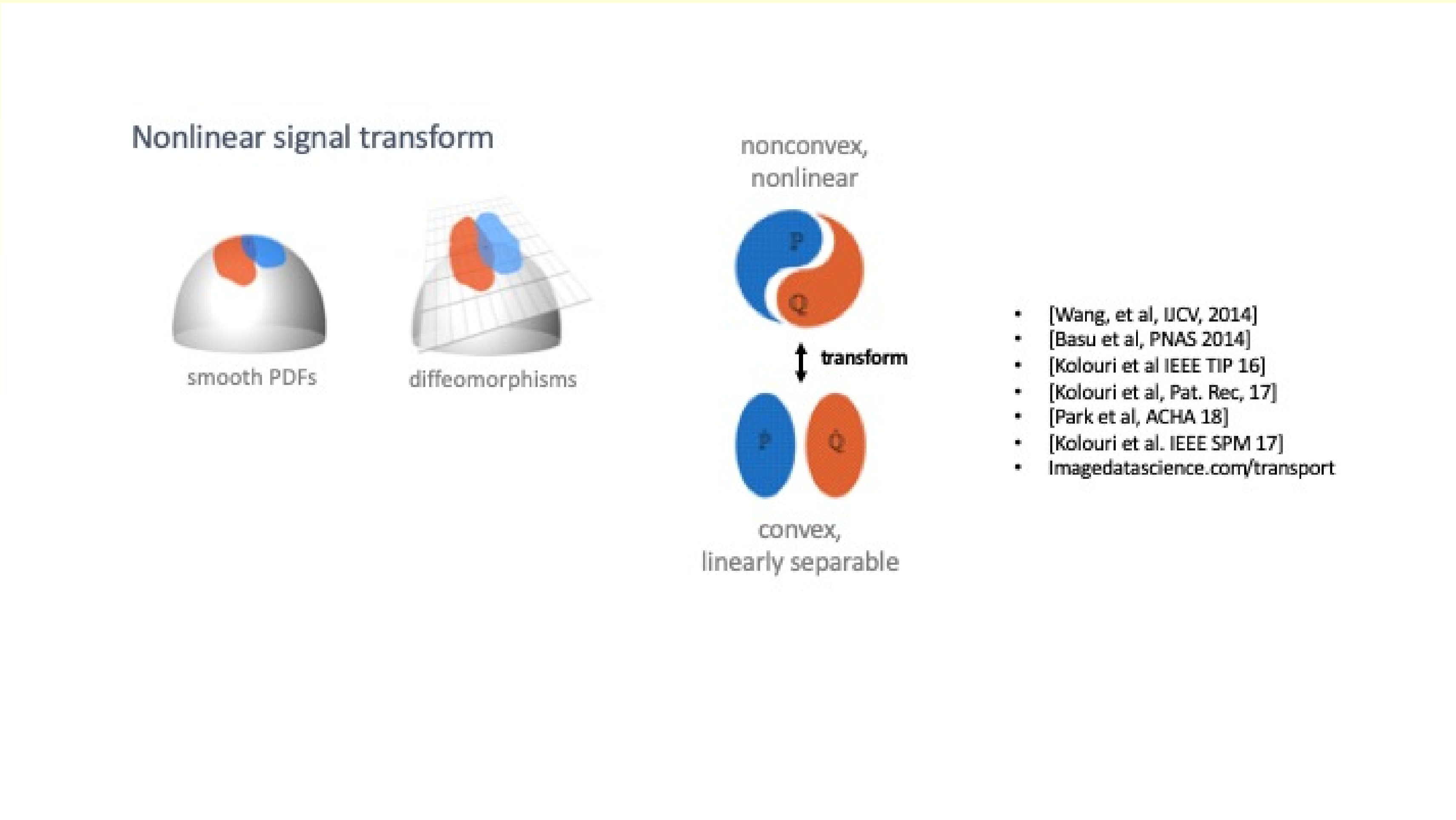
$T_s : \mathbb{R} \to L^2(\mathbb{R})$: $T_s(\mu)(x) = s_\mu(x) = s(x-\mu)$.



Convexification properties



Convexification properties



Wasserstein distance between two measures

Let $\Pi(\mu, \nu)$ be the set of all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν.

2-Wasserstein distance between μ and ν:

$$W_2^2(\mu, \nu) = \min_{\pi \in \Pi} \int_{\mathbb{R}^n} \|x - y\|^2 d\pi(x, y).$$

Wasserstein distance between two measures

Let $\Pi(\mu, \nu)$ be the set of all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν.

2-Wasserstein distance between μ and ν:

$$W_2^2(\mu, \nu) = \min_{\pi \in \Pi} \int_{\mathbb{R}^n} \|x - y\|^2 d\pi(x, y).$$

Existence of a minimizer is due to Kantorovich.

Wasserstein distance between two measures

Let $\Pi(\mu,\nu)$ be the set of all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν.

2-Wasserstein distance between μ and ν:

$$W_2^2(\mu,\nu) = \min_{\pi\in\Pi} \int_{\mathbb{R}^n} \|x-y\|^2 d\pi(x,y).$$

Existence of a minimizer is due to Kantorovich.

The solution $T^\dagger$ of the transport map of the Monge problem give rise to the minimizer to the Kantorovich problem: $d(\mu(x)\delta(y=T^\dagger(x)) \in \Pi$ and minimizes $W_2(\mu,\nu)$.

Wasserstein distance between two measures

Theorem (Kolouri, Rohde) Let s_1 and s_2 be two signals (PDFs) and $\tilde{s}_1, \tilde{s}_2$ be their CDT transform with respect to fixed reference s_0. Then

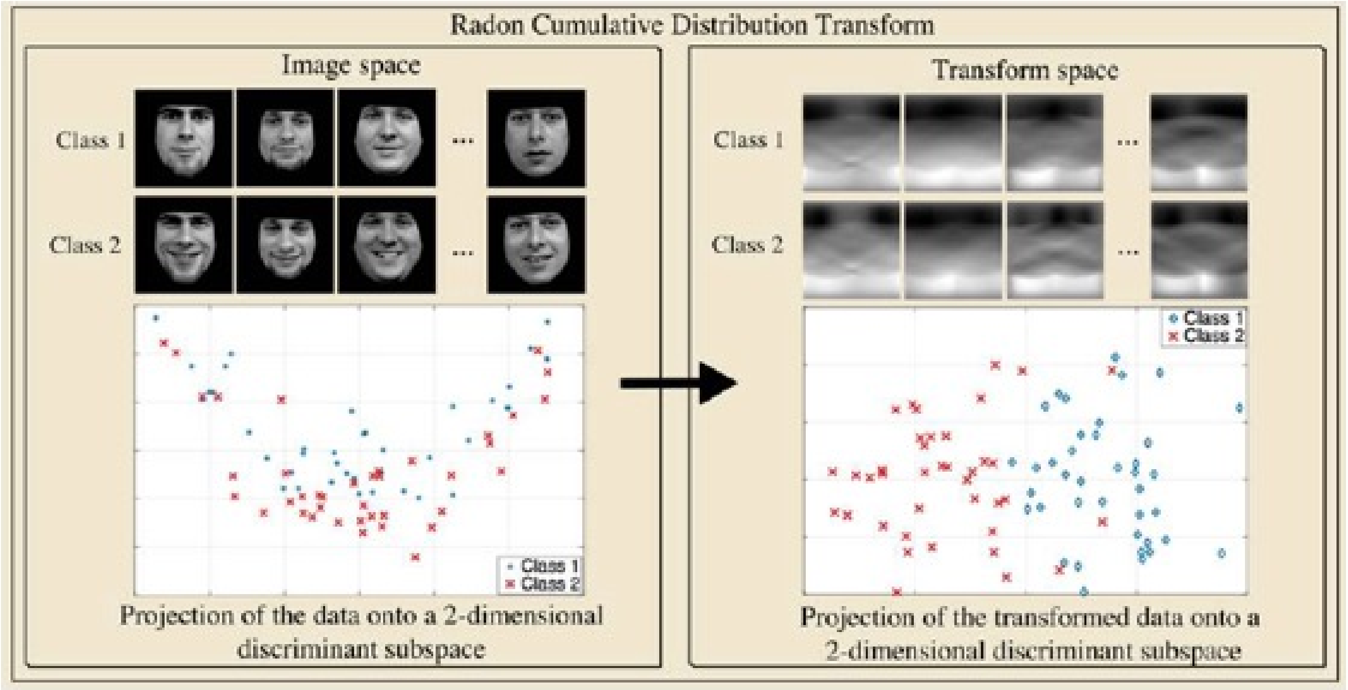
$$\|\tilde{s}_2 - \tilde{s}_1\|^2_{L^2(\mathbb{R}^n)} = W_2(\mu, \nu) = \min_{\pi \in \Pi} \int_{\mathbb{R}^n} \|x - y\|^2 d\pi(x, y)$$

where $d\mu(x) = s_1(x)dx$ and $d\nu = s_2(x)dx$.

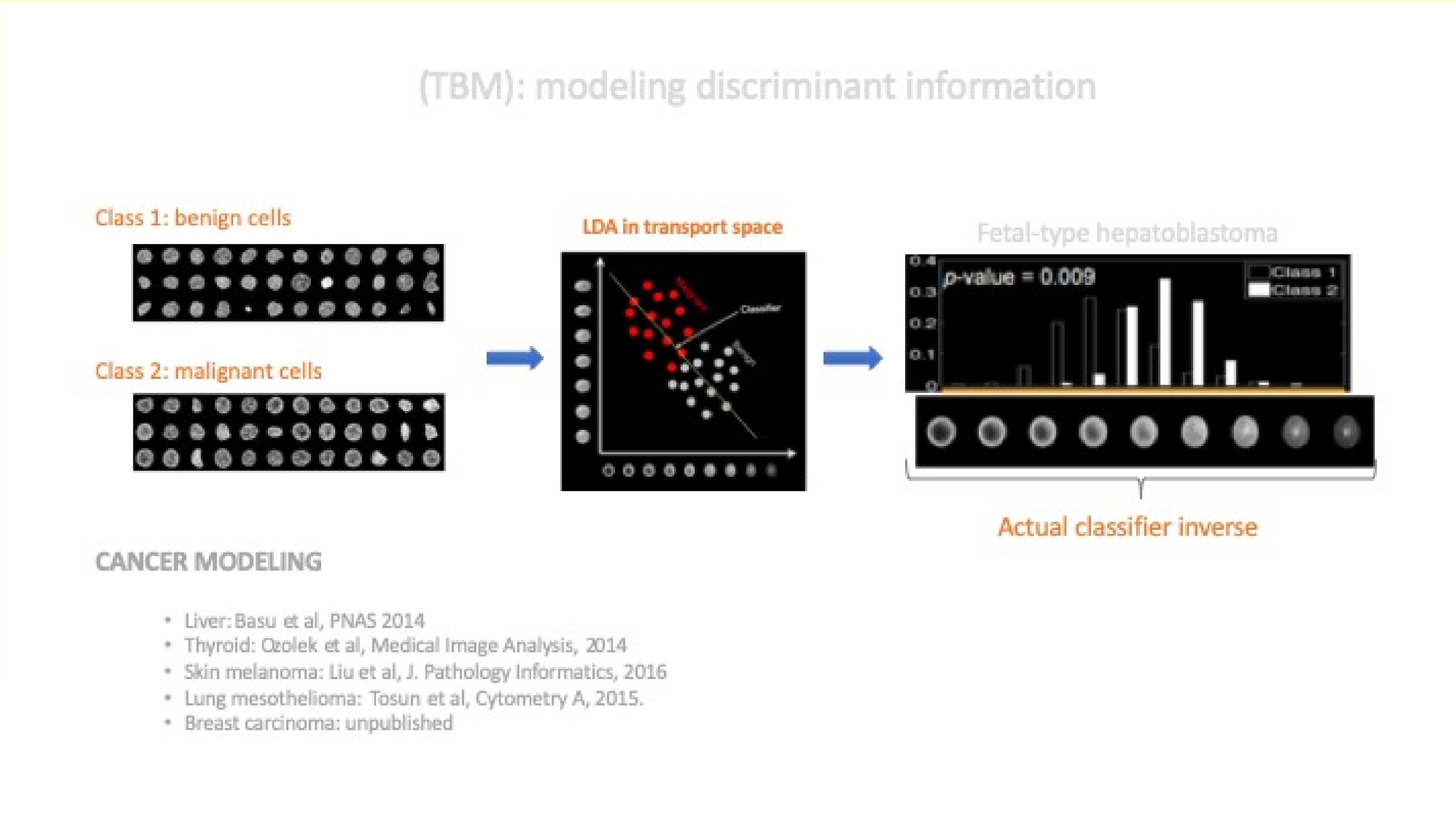

signal space,
Wasserstein
geodesics
transform space
Euclidean geodesics
[Kolouri, Rohde, CVPR 16]
1

Applications

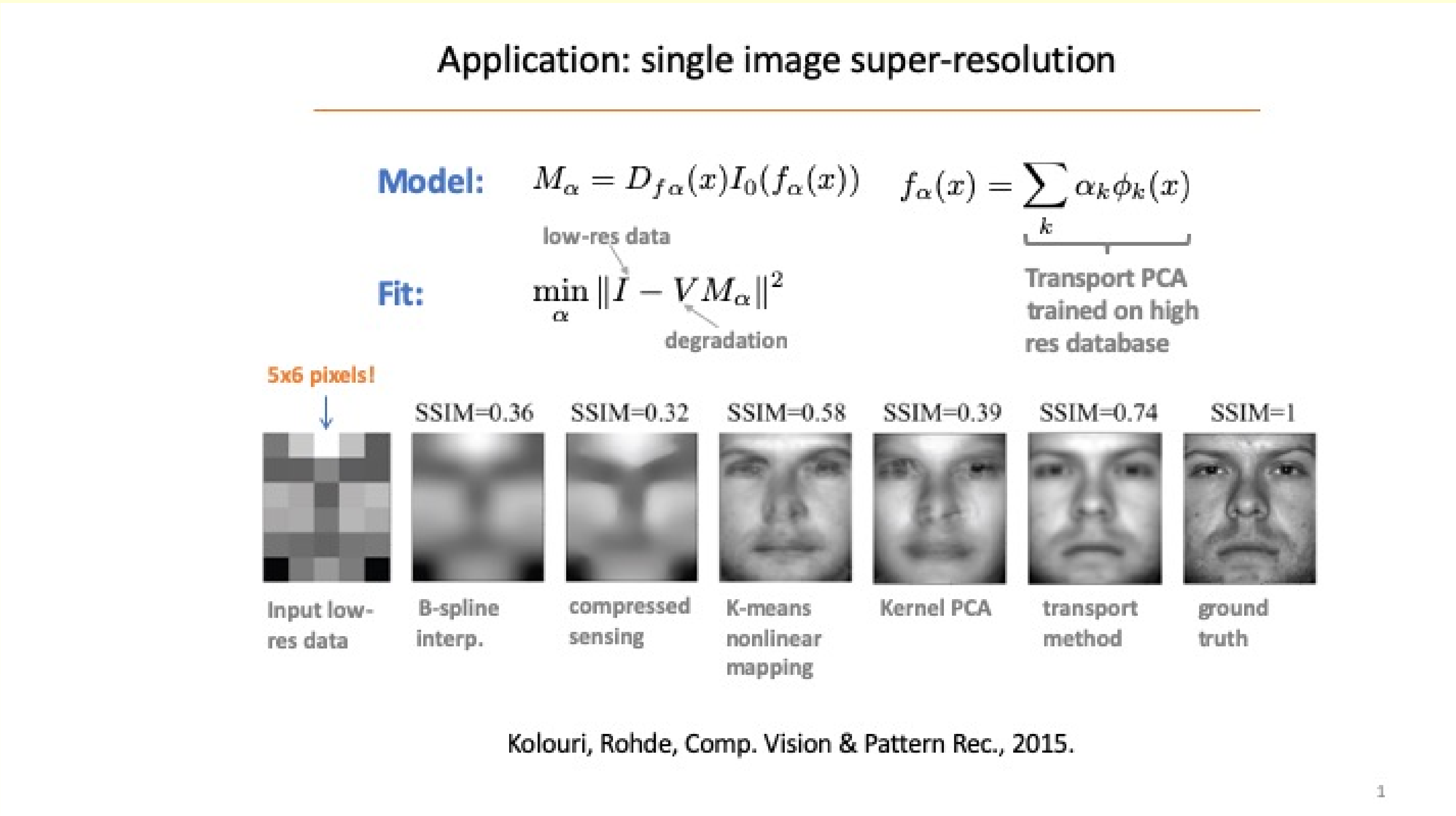
Example: classifying facial expressions



Applications



Applications



Happy Birthday Caro