# Adaptive Constraint K-Segment Principal Curves for Intelligent Transportation Systems

Junping Zhang, Dewang Chen

Shanghai Key Lab. of Intelligent Information Processing
School of Computer Science
Fudan University
jpzhang@fudan.edu.cn

Nov 8th, 2008

# Outline

1. Survey
2. ACKPCs
   - Theoretical Foundation
3. Applications
   - Freeway Traffic Stream Modeling
   - Learning High-precision GPS
4. Theoretical Analysis
5. Conclusion

# Outline

# Outline

1. Survey
2. ACKPCs
   - Theoretical Foundation
3. Applications
   - Freeway Traffic Stream Modeling
   - Learning High-precision GPS
4. Theoretical Analysis
5. Conclusion

# Outline

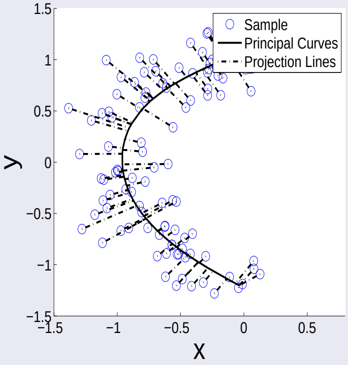
# Outline

# Principal Curves

## Motivations

- To discover intrinsic structure hidden in the data
- Geometrically intuitive
- One way is to search curves across the "middle" of data distribution.

## Example: semi-circle-shape distribution



# What is Principal Curves?

**Definition [Hastie, 1988]**

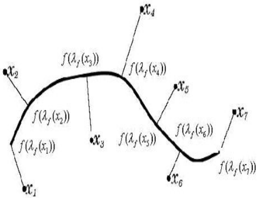
The smooth curve $f(\lambda)$ is a principal curve if the following conditions are satisfied:

1. $f$ does not intersect itself
2. $f$ has finite length inside any bounded subset of $\mathbb{R}^d$
3. $f$ is self-consistent, that is

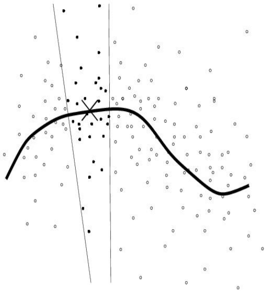
$$E(X|\lambda_f(X) = \lambda) = f(\lambda) \quad \forall \lambda \in \mathbb{R}^1 \tag{1}$$

# Illustration

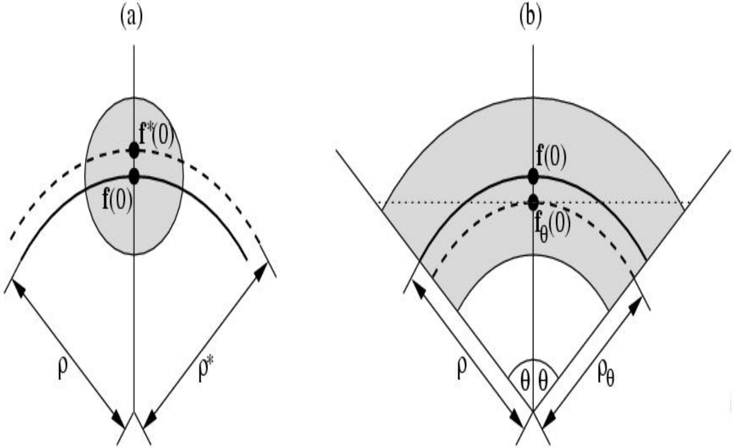
## Projection




## Self-Consistency [Kégl, 2002]



# Model Bias and Estimation Bias

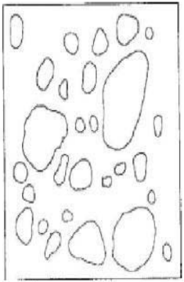


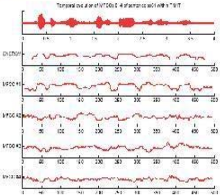
# Survey on Principal Curves

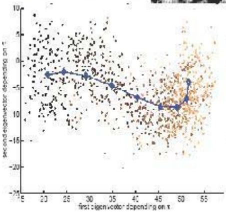
## Brief Survey

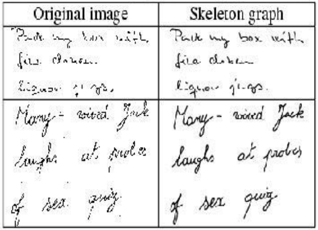
- HS Principal Curves [Hastie and Stuetzle, 1988]
- BR Principal Curves: Closed-shape and estimation bias problem [Banfield and Raftery, 1992].
- T Principal Curves: model bias[Tibshirani, 1992].
- Principal Curves of Principal Oriented Points (PCOP) [Delicado, 2001]
- K-segment Principal curves (KPCs)[Kégl, 2000, Sandilya, 2002]
- Twinned Principal Curves [Koetsier, 2004]
- Elastic Maps [Gorban, 2005]
- Local Principal Curves [Einbeck, 2007]
- ....

# Some Examples of Applications with PCs

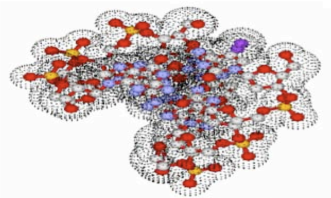


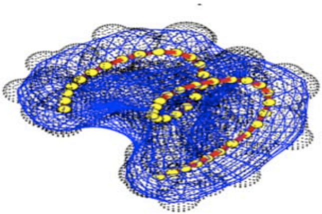


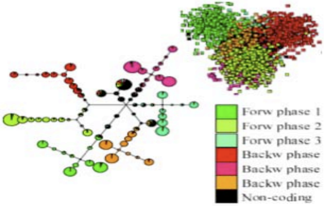




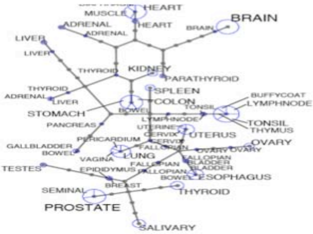
# Some Examples of Applications with PCs








*Bacillus halodurans*



# Several Issues

## Several Issues

1. Uniqueness and Existence
2. Unobservable Region and Prior Knowledge.
3. Projection Approach.
4. Vertex Optimization.
5. Abnormal Data
6. Parameter Sensitivity

# The KPCs algroithm

## Existence Theorem

Assume that $E\|X\|^2 < \infty$. Then for any $L > 0$ there exists a curve $f^*$ with $l(f^*) \leq L$ such that

$$\Delta(f^*) = \inf\{\Delta(f) : l(f) \leq L\} \tag{2}$$

## Convergence Rate Theorem

The expected squared loss of the KPCs, as $n \to \infty$, to the squared loss of the principal curve of length $L$ at a rate

$$\begin{aligned} J(f_{k,n}) &= (\Delta(f_{k,n}) - \Delta(f_k^*)) + (\Delta(f_k^*) - \Delta(f^*)) \\ &= O(n^{-1/3}) \\ &\leq \sqrt{\frac{kC(L,D,d)}{n}} + \frac{DL+2}{k} + O(n^{-1/2}) \end{aligned} \tag{3}$$

# Unobservable Region or Prior Knowledge

## Approaches

- Utilizing the knowledge from the unobservable region $\bar{A}$

$$PCs(\mathbf{A}) \neq PCs(\mathbf{A} \cup \bar{\mathbf{A}}). \quad (4)$$

- Under some practical condition, priori information from the “true” principal curves can be approximately obtained.

## Redefinition

$$\min \quad Err(x,f) \quad (5)$$
$$s.t. \quad E(X|\lambda_f(X) = \lambda) = f(\lambda) \quad \forall \lambda \in \mathbb{R}^1 \quad (6)$$
$$s.t. \quad p1 = StartPoint,\ p2 = Endpoint. \quad (7)$$

# Refinements on vertex optimization

## Verifying and removing the abnormal vertices

$$R(v_i) = \begin{cases} 1, & \text{if } (l_{s_{i-1}} \text{ or } l_{s_i}) > r \text{ and } \frac{N_{v_i}}{n} < 0.01 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Where

$$\begin{aligned} l_{s_{i-1}} &= \|\boldsymbol{v}_i - \boldsymbol{v}_{i-1}\| \quad 2 \le i \le K+1. && (9) \\ l_{s_i} &= \|\boldsymbol{v}_{i+1} - \boldsymbol{v}_i\| \quad 1 \le i \le K. && (10) \\ r &= \max_{\boldsymbol{x} \in X} \|\boldsymbol{x} - \frac{1}{n} \sum_{\boldsymbol{y} \in X} \boldsymbol{y}\| && (11) \end{aligned}$$

and

$$N_{\boldsymbol{v}_i} = \begin{cases} n_{\boldsymbol{s}_{i-1}} + n_{\boldsymbol{v}_i} + n_{\boldsymbol{s}_i}, & \text{if } 2 \le i \le K \\ n_{\boldsymbol{s}_i} + n_{\boldsymbol{v}_i}, & \text{if } i = 1 \\ n_{\boldsymbol{v}_i} + n_{\boldsymbol{s}_{i-1}}, & \text{if } i = K+1 \end{cases} \quad (12)$$

# Projection Refinement

## Projection Formula

$$f_K^{(m)}(\boldsymbol{x}_{\lambda_f}) = \begin{cases} \boldsymbol{v}_l, & \text{if } \lambda_f = 0 \\ \frac{(\boldsymbol{v}_{l+1}-\boldsymbol{v}_l)(\boldsymbol{v}_{l+1}-\boldsymbol{v}_l)'\boldsymbol{x}}{\|\boldsymbol{v}_{l+1}-\boldsymbol{v}_l\|^2}, & \text{if } 0 < \lambda_f < 1 \\ \boldsymbol{v}_{l+1}, & \text{if } \lambda_f = 1 \end{cases} \quad (13)$$

Where

$$\lambda_f(\boldsymbol{x}) = \frac{< \boldsymbol{x}, (\boldsymbol{v}_{l+1} - \boldsymbol{v}_l) >}{\|\boldsymbol{v}_{l+1} - \boldsymbol{v}_l\|^2} \quad (14)$$

# Vertex Optimization

$\nabla_{\boldsymbol{v}_i} G_n(\boldsymbol{f}_K^{(m)})$ of each vertex $\boldsymbol{v}_i$ is defined as:

$$\nabla_{\boldsymbol{v}_i} G_n(\boldsymbol{f}_K^{(m)}) = \nabla_{\boldsymbol{v}_i}(\Delta_n(\boldsymbol{f}_K^{(m)}) + \zeta P_{\boldsymbol{v}_i}(\boldsymbol{f}_K^{(m)})) \tag{15}$$

where

$$P_{\boldsymbol{v}_i}(\boldsymbol{f}_K^{(m)}) = \begin{cases} \frac{1}{K+1}(P_V(\boldsymbol{v}_i) + P_V(\boldsymbol{v}_{i+1})), & \text{if } i = 1 \\ \frac{1}{K+1}(P_V(\boldsymbol{v}_{i-1}) + P_V(\boldsymbol{v}_i) + P_V(\boldsymbol{v}_{i+1})), & \text{if } 1 < i < K+1 \\ \frac{1}{K+1}(P_V(\boldsymbol{v}_{i-1}) + P_V(\boldsymbol{v}_i)), & \text{if } i = K+1 \end{cases} \tag{16}$$

and

$$P_V(\boldsymbol{v}_i) = r^2(1 + \cos\gamma_i) \quad \text{if } 1 < i < K+1 \tag{17}$$

$$\zeta = \zeta' \cdot \frac{k}{n^{1/3}} \frac{\sqrt{\Delta_n(f_{k,n})}}{r} \tag{18}$$

# The Inner and Outer Loop

## Inner Loop

$$|1 - \frac{G_n(\boldsymbol{f}_k^{(m)})}{G_n(\boldsymbol{f}_k^{(m-1)})}| \leq \delta. \quad (19)$$

Where $\delta$ is equal to $1e-3$ without loss of generality.

## Outer Loop

$$c(n, \Delta_n(\boldsymbol{f}_K)) = \beta n^{1/3} \frac{r}{\sqrt{\Delta_n(\boldsymbol{f}_K)}}. \quad (20)$$

where $\beta$ is an experimental parameter which is set to be a constant value $0.3$ [Kegl, TPAMI]

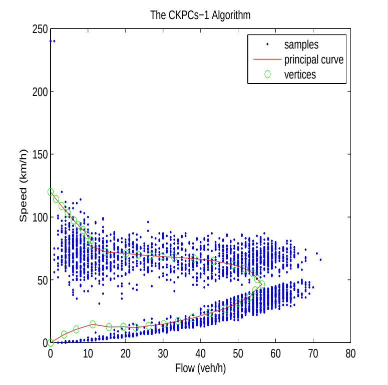
# Adding Vertices

## Adding Vertices

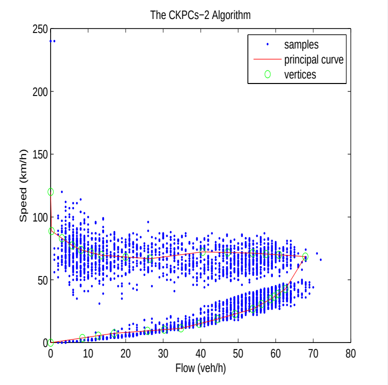
- The segment has the longest length.
- The segment has minimal average squared distance from samples which projected in this segment and corresponding projection locations.
- The segment has the maximal number of samples projected into there.
- Simultaneous partitioning all the segments without any constraint.

# Example

## ACKPC-1



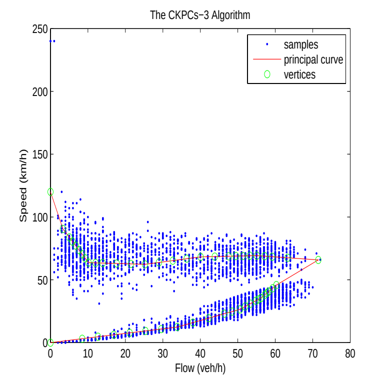

## ACKPC-2



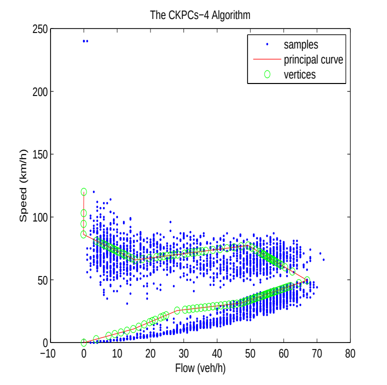
# Example



## ACKPC-3




## ACKPC-4



# Quantitative Results

Table: Quantitative comparisons of the four data sets; Table entries are of the form A$\pm B(C)$, where A denotes AveSE, B represents STD, and C refers to MaxSE.

| | Distorted Half Circle | Distorted S-Shape |
|---|---|---|
| | From $X$ to PC | From $X$ to PC |
| KPC-1 | $0.064 \pm 0.045(0.189)$ | $0.024 \pm 0.019(0.095)$ |
| KPC-2 | $0.048 \pm 0.037(0.177)$ | $0.027 \pm 0.020(0.087)$ |
| KPC-3 | $\mathbf{0.044} \pm \mathbf{0.035}(\mathbf{0.162})$ | $0.030 \pm 0.024(0.109)$ |
| KPC-4 | $0.057 \pm 0.037(0.185)$ | $0.035 \pm 0.029(0.129)$ |
| ACKPC-1 | $0.050 \pm 0.035(0.161)$ | $0.023 \pm 0.019(0.091)$ |
| ACKPC-2 | $0.053 \pm 0.034(0.164)$ | $\mathbf{0.023} \pm \mathbf{0.019}(\mathbf{0.088})$ |
| ACKPC-3 | $0.050 \pm 0.035(0.162)$ | $0.026 \pm 0.021(0.106)$ |
| ACKPC-4 | $0.050 \pm 0.035(0.159)$ | $0.029 \pm 0.025(0.120)$ |

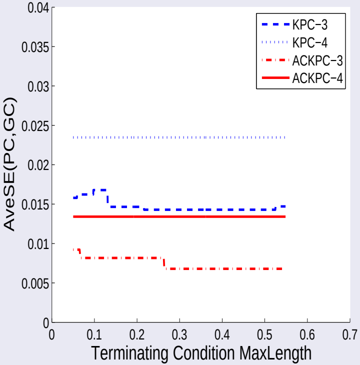
# Quantitative Results

Table: Quantitative comparisons of the four data sets; Table entries are of the form A$\pm B(C)$, where A denotes AveSE, B represents STD, and C refers to MaxSE.

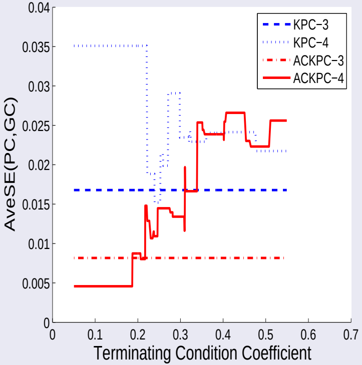
| | Distorted Half Circle | Distorted S-Shape |
|---|---|---|
| | From PC to GC | From PC to GC |
| KPC-1 | $0.037 \pm 0.023(0.078)$ | $0.012 \pm 0.009(0.033)$ |
| KPC-2 | $0.029 \pm 0.014(0.062)$ | $0.017 \pm 0.010(0.038)$ |
| KPC-3 | $0.019 \pm 0.015(0.056)$ | $0.017 \pm 0.010(0.036)$ |
| KPC-4 | $0.033 \pm 0.022(0.075)$ | $0.023 \pm 0.017(0.059)$ |
| ACKPC-1 | $0.031 \pm 0.010(0.030)$ | $\mathbf{0.008} \pm \mathbf{0.008}(\mathbf{0.032})$ |
| ACKPC-2 | $0.020 \pm 0.019(0.070)$ | $0.009 \pm 0.009(0.039)$ |
| ACKPC-3 | $0.014 \pm 0.010(0.030)$ | $0.008 \pm 0.010(0.039)$ |
| ACKPC-4 | $\mathbf{0.013} \pm \mathbf{0.008}(\mathbf{0.024})$ | $0.013 \pm 0.013(0.040)$ |

# Sensitive Analysis

## TermCondMaxLen

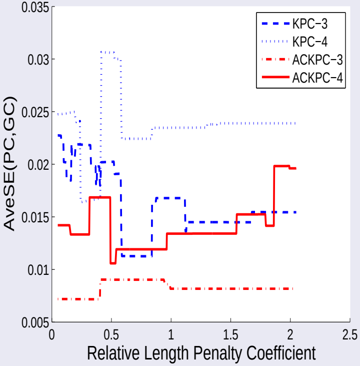



## TermCondCoef

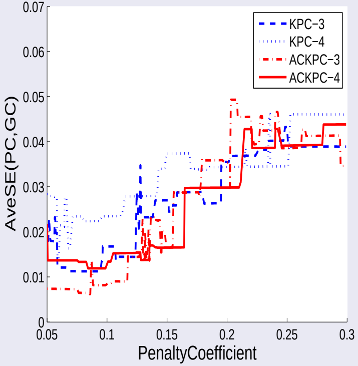


# Sensitive Analysis

## RelaLenPenaltyCoeff




## PenaltyCoeff



# Properties of Traffic Stream Data

## Application: Three Traffic Stream Models

- The density-speed model is the basic model and the fundamental diagram of traffic flow theory.
- The occupancy (or density)-flow model, is very important to ramp metering.
- The flow-speed model is often used to judge the level of service (LOS) of freeways.

# Properties of Traffic Stream Data



## Some key points with apparent physical meanings

1. Vehicles stop within the detection zone, which means flow$=0$ vehicles per hours (Veh/h), speed$=0$ kilometer per hours (Km/h), occupancy$=100\%$.
2. When there are so few vehicles, the drivers can choose the speed freely. In the limit sense, it means occupancy=$0$, flow$=0$ and speed=maxSpeed (=120 Km/h).

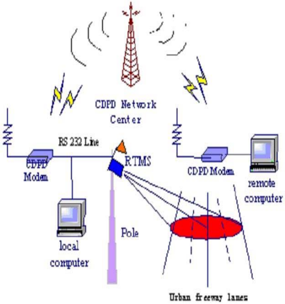
# Data Collection

- Collection Location: at third ring road in Beijing, where a total of $72$ RTMS (remote traffic microwave sensor) were set up.
- Collection time: Apr $1$rd, $2005 \longrightarrow$ April $5$th, $2005$.
- Sensor Installation: One detector in the outer third ring road were installed to collect data and each detector monitored three lanes, the inner lane $1$, middle lane $2$ and outer lane $3$.
- Sampling Cycle: per $2$ minutes and a total of $5055$ sets of data were obtained.
- Collection Route: All the field data are transmitted to Beiijing Traffic Administration Bureau (BTAB) through CDPD (Cellular Digital Packet Data) wireless network.

# An illustration of sampling traffic stream data

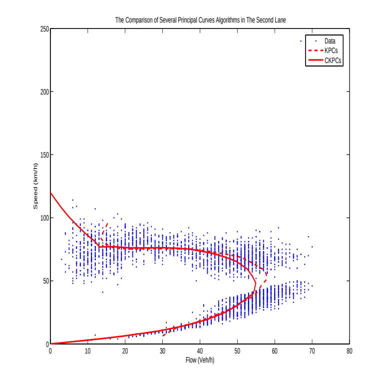
## Two detectors' locations


## Data collection scheme

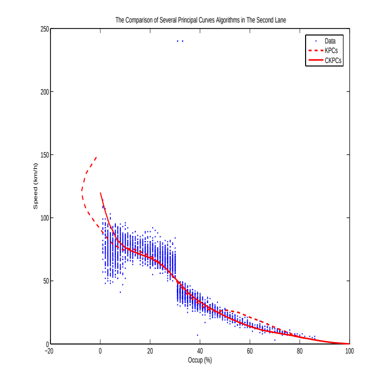


# Results

## Flow Vs Speed-2

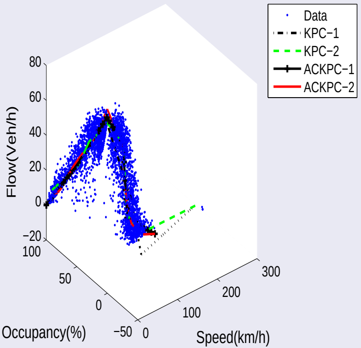



## Occupancy Vs Speed-2

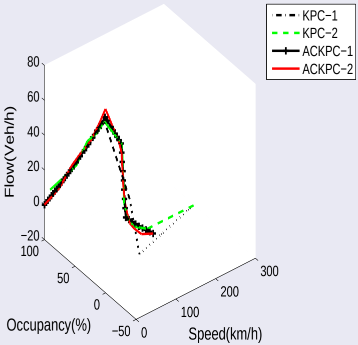


# 3-D Results

## The First Lane with Data




## The First Lane without Data



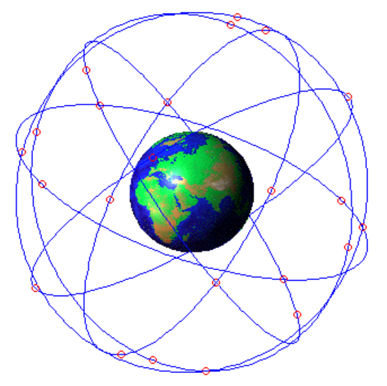
# Quantitative Analysis: ACKPCs Vs KPCs

Table: Quantitative comparisons of the three freeway traffic stream data; Table entries are of the form A$\pm B(C)$, where A denotes AveSE, B represents STD, and C refers to MaxSE.

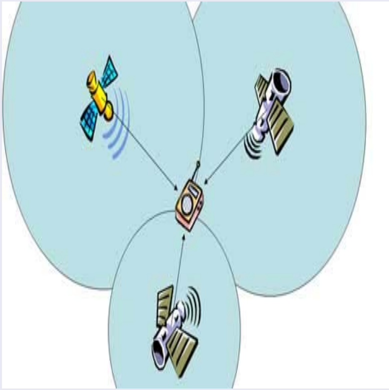
| | Lane-1 | Lane-2 | Lane-3 |
|---|---|---|---|
| | From $X$ to PC | | |
| KPC-1 | $8.191 \pm 6.375(40.236)$ | $6.014 \pm 7.035(163.200)$ | $7.040 \pm 6.435(150.520)$ |
| KPC-2 | $\mathbf{5.715} \pm \mathbf{4.553}(\mathbf{32.751})$ | $6.431 \pm 7.335(168.700)$ | $7.570 \pm 6.280(144.070)$ |
| KPC-3 | $5.568 \pm 4.483(34.087)$ | $8.008 \pm 7.769(171.720)$ | $7.151 \pm 6.265(162.54)$ |
| KPC-4 | $7.351 \pm 4.730(33.281)$ | $7.950 \pm 7.701(171.530)$ | $9.436 \pm 7.701(171.530)$ |
| ACKPC-1 | $5.959 \pm 5.979(120.010)$ | $\mathbf{5.878} \pm \mathbf{6.222}(\mathbf{143.280})$ | $\mathbf{6.816} \pm \mathbf{5.998}(\mathbf{146.880})$ |
| ACKPC-2 | $6.194 \pm 6.192(120.010)$ | $6.178 \pm 6.231(143.280)$ | $7.401 \pm 5.973(146.880)$ |
| ACKPC-3 | $7.181 \pm 6.425(120.010)$ | $6.294 \pm 6.158(143.280)$ | $7.002 \pm 5.753(146.880)$ |
| ACKPC-4 | $6.186 \pm 6.375(120.010)$ | $6.402 \pm 6.722(143.280)$ | $7.316 \pm 6.504(146.880)$ |

# Motivation: What is GPS?

## Global Positioning System

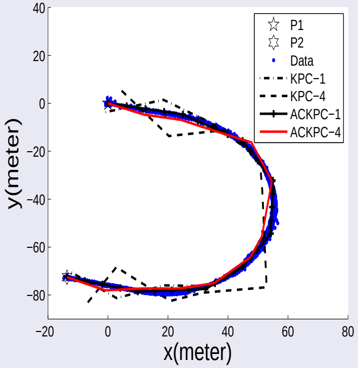


## How GPS Works

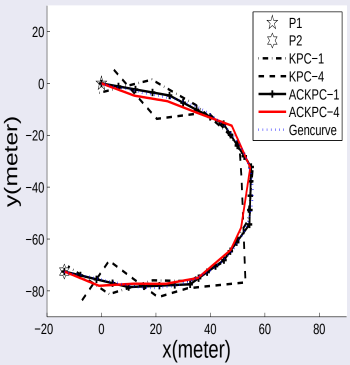


# Illustration-1

## Data and the constraint K-segments principal curves




## High-precision GPS data and the ACKPCs



# Quantitative Analysis

Table: Quantitative comparisons of improved low-precision GPS position estimates; Table entries are of the form A$\pm B(C)$, where A denotes AveSE, B represents STD, and C refers to MaxSE.

| | From Data to PC | From PC to GC |
|---|---|---|
| KPC-1 | $1.553 \pm 1.198(5.511)$ | $1.621 \pm 1.393(5.619)$ |
| KPC-2 | $1.927 \pm 1.578(8.219)$ | $7.980 \pm 14.331(60.67)$ |
| KPC-3 | $1.376 \pm 0.845(3.994)$ | $1.537 \pm 1.048(4.010)$ |
| KPC-4 | $3.920 \pm 2.913(12.375)$ | $3.752 \pm 2.712(11.820)$ |
| ACKPC-1 | $\mathbf{0.512} \pm \mathbf{0.405}(\mathbf{2.745})$ | $\mathbf{0.505} \pm \mathbf{0.297}(\mathbf{1.202})$ |
| ACKPC-2 | $0.872 \pm 0.577(3.079)$ | $0.765 \pm 0.583(2.618)$ |
| ACKPC-3 | $1.000 \pm 0.912(4.972)$ | $1.940 \pm 3.480(13.812)$ |
| ACKPC-4 | $1.210 \pm 0.790(4.465)$ | $1.126 \pm 0.752(2.664)$ |

$$
\begin{aligned}
&\Delta\left(\boldsymbol{f}_2(\cdot),\boldsymbol{f}^*(\cdot)\right)-\Delta\left(\boldsymbol{f}_1(\cdot),\boldsymbol{f}^*(\cdot)\right)\\
=\;&\underbrace{\frac{1}{K}\sum_{i=1}^{j}\left\|\boldsymbol{v}_i'-\boldsymbol{f}^*\left(\lambda_{f^*}\left(\boldsymbol{v}_i'\right)\right)\right\|}_{=0}-\frac{1}{K}\sum_{i=1}^{j}\left\|\boldsymbol{v}_i-\boldsymbol{f}^*\left(\lambda_{f^*}\left(\boldsymbol{v}_i\right)\right)\right\|\\
+\;&\underbrace{\frac{1}{K}\sum_{i=j+1}^{K}\left\|\boldsymbol{v}_i-\boldsymbol{f}^*\left(\lambda_{f^*}\left(\boldsymbol{v}_i\right)\right)\right\|-\frac{1}{K}\sum_{i=j+1}^{K}\left\|\boldsymbol{v}_i-\boldsymbol{f}^*\left(\lambda_{f^*}\left(\boldsymbol{v}_i\right)\right)\right\|}_{=0}\\
=\;&\underbrace{-\frac{1}{K}\sum_{i=1}^{j}\left\|\boldsymbol{v}_i-\boldsymbol{f}^*\left(\lambda_{f^*}\left(\boldsymbol{v}_i\right)\right)\right\|}_{<0} \qquad (21)
\end{aligned}
$$

# Analysis on Learning High-precision GPS



Assume that

$$\boldsymbol{x} = \boldsymbol{f}_{highGPS}(\lambda) + \boldsymbol{\eta} \tag{22}$$

Therefore, the mathematical expectation of $\bar{\boldsymbol{x}}$ is:

$$E\left\{\bar{\boldsymbol{x}}\right\} = \frac{1}{\omega}\sum_{i=1}^{\omega} E\left\{\boldsymbol{f}_{highGPS}^{(\omega)}(\lambda)\right\} + \frac{1}{\omega}\sum_{i=1}^{\omega} E\left\{\boldsymbol{\eta}_i\right\} \tag{23}$$

Then, we have:

$$\sigma_{\bar{x}}^2 = \frac{\omega}{\omega^2}\sigma_\eta^2 = \frac{1}{\omega}\sigma_\eta^2 \tag{24}$$

# Conclusion

- In this paper, we propose the ACKPCs algorithm for automatical removing unexpected vertices in which vertex optimization of the KPCs algorithm fails
- Introducing samples in the unobservable region or prior knowledge for the improvement of PCs.
- The ACKPC algorithm is less sensitive to outliers and the setting of parameters than the KPC algorithms.
- Experiments show that the ACKPCs algorithms are of potential to the practical field such as ITS.

# Further Research

## Further Research

- Further enhancing the accuracy of the ACKPCs algorithm from both theoretical and algorithmic aspects.
- Developing some new principal curves algorithms.
- Potential Application: GPS-based Electronic Map

# For Further Reading I


- **Junping Zhang**, Dewang Chen, Uwe Kruger, Adaptive Constraint K-Semgent Principal Curves for Intelligent Transportation Systems, IEEE Transactions on Intelligent Transportation Systems, vol. 9, no. 4, pp. 1524-1535, 2008
- Uwe Kruger, **Junping Zhang**, Lei Xie, Developments and Applications of Nonlinear Principal Component Analysis - a Review, In: Principal Manifolds for Data Visualization and Dimension Reduction, Lecture Notes in Computational Science and Engineering , Vol. 58, pp. 1-44, Springer, Berlin-Heidelberg-New York, 2007
- Dewang Chen, **Junping Zhang**, Shuming Tang, and Jue Wang, Freeway Traffic Stream Modelling based on Principal Curves and its Analysis,IEEE Transactions on Intelligent Transportation Systems, Vol 5, No. 4, pp: 246-258, 2004
- **Junping Zhang**, Jue Wang, An Overview about Principal Curves (in chinese), Chinese Journal of Computers, 26(2), pp. 137-148, 2003