

時系列データに対する評価指標形成支援のための視覚的分析フレームワークの提案

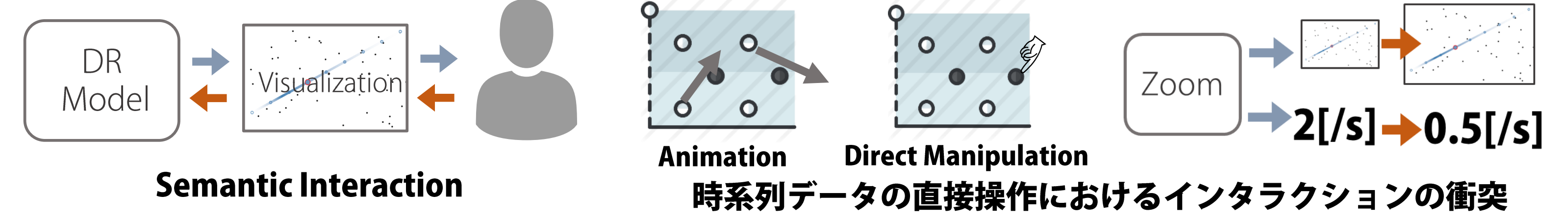
高見 玲, 高間 康史 (首都大学東京 システムデザイン研究科 情報科学域)

研究概要/研究背景

目的: 視覚的分析による時系列データの評価指標形成支援

- 多次元 + 時系列データの普及 [Aigner et al. 07] e.g. センサログ, 医療, 金融
- 機械学習への活用 -> 評価指標 (外れ値の判定基準など) が必要
 - 大規模データ: 次元削減などの前処理 -> 可視化
 - ドメイン専門家: モデル/パラメータ調整の困難性
- ドメイン専門家の分析作業: 視覚的分析インタフェース [Keim et al. 08]
 - 視覚的洞察 -> 仮説形成/検証 -> 知識 (e.g. 評価指標) 形成 [Sacha et al. 14]

- 人と計算機の協調による分析アプローチ e.g. **Semantic interaction** [Endert et al. 12]
 - > 時系列性考慮の必要 + 多次元時系列データの適切な可視化の困難性

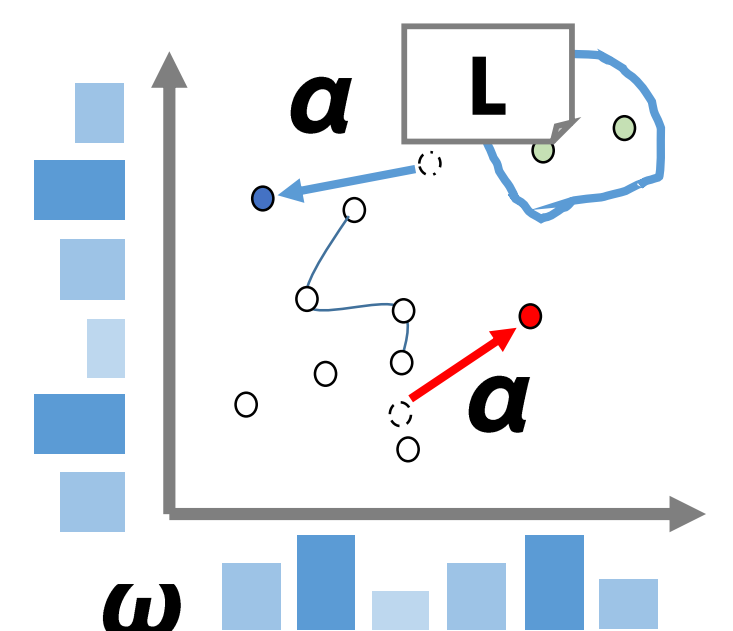


- 時系列データへの適用のためのフレームワーク提案/実装
- **次元削減 + 軌跡表現**を用いて可視化, 分析/指標形成を支援

提案フレームワーク

分析手法: 事後解析(post hoc analysis), 対象データ: 離散的データ

- e.g. 医療データ, スポーツ, 金融
- 投影の操作: パラメータ α, ω, τ で表現
 - α : 各データ点の移動度 (バイアス)
 - ω : 低次元投影軸 (主成分) における各属性の係数
 - τ : 操作適用先の時点
- ラベル L (任意): オブジェクトの所属グループ

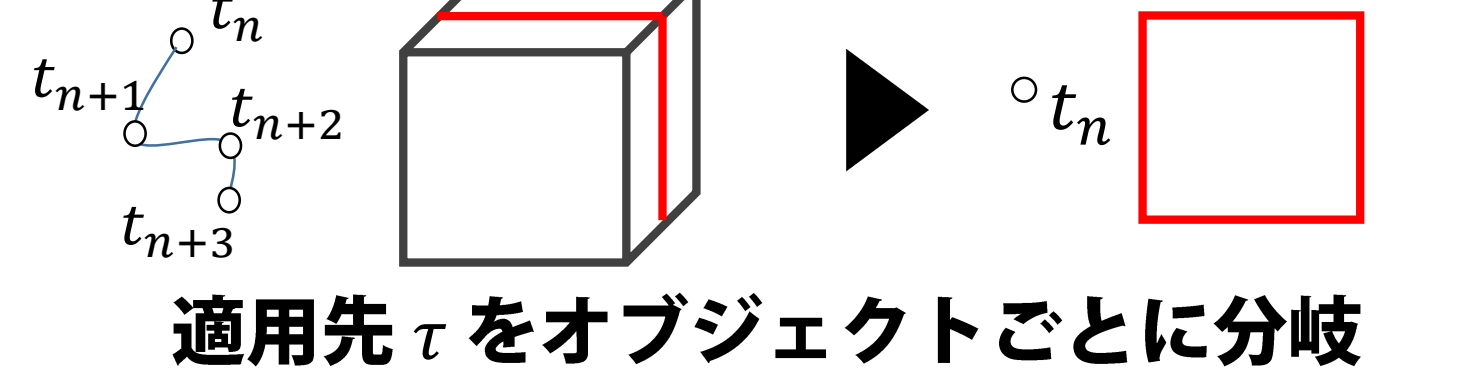


操作対象 (データ点, 軌跡, 凸包): 3次元キューブ上にモデル構築システム: 散布図へのユーザ操作を解釈 -> α, ω を変更 / ω を指標として出力

- 時系列データ $d_{tm} \in \mathbb{R}^M$ の座標 $P_{tm} = (X_{tm}, Y_{tm})$
 - 時点 $\tau \in \{1, 2, \dots, T\}$, データ $n \in \{1, 2, \dots, N\}$, 属性 $m \in \{1, 2, \dots, M\}$

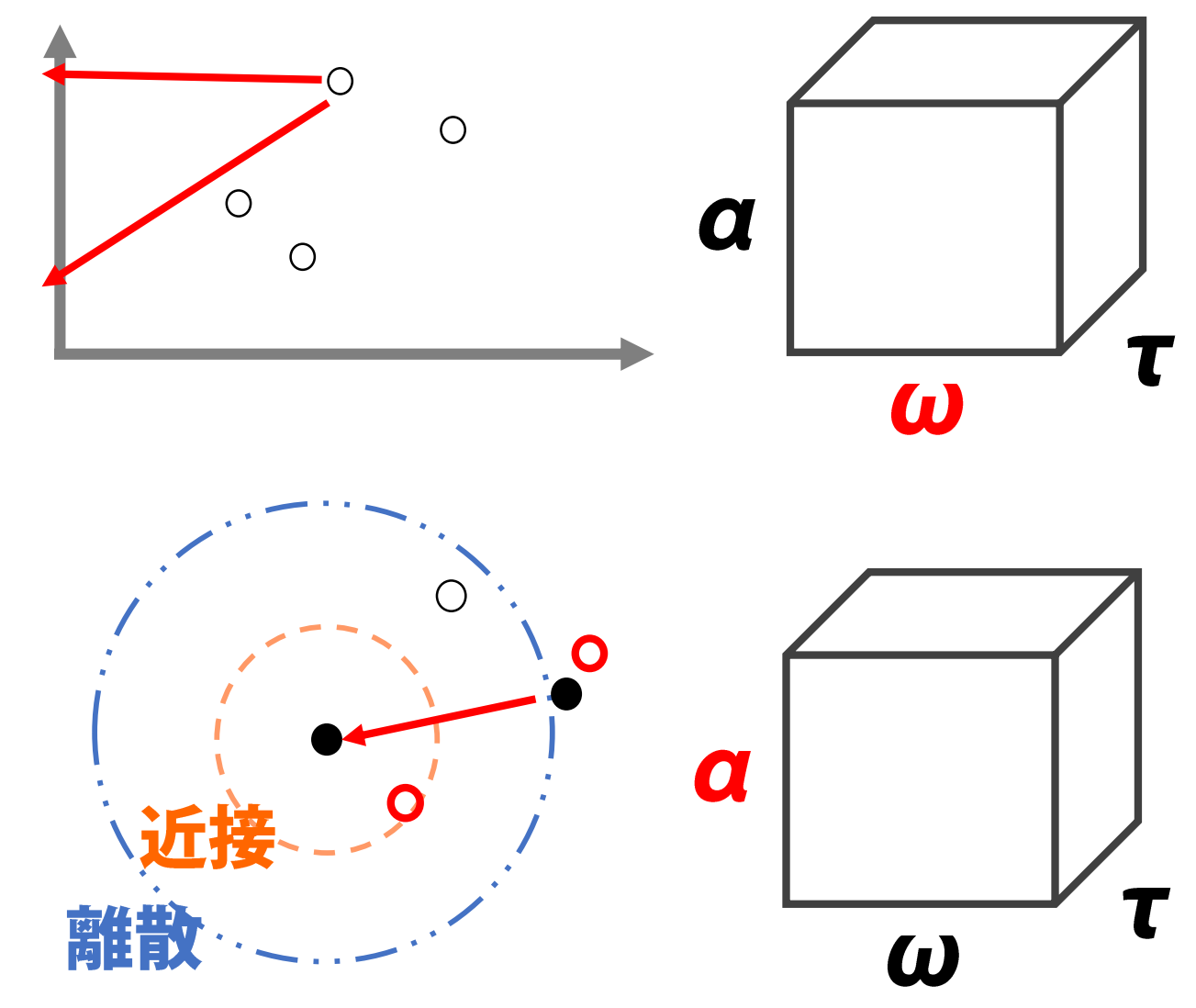
$$X_{tm} = \sum_{m=1}^M d_{tm} \omega_{tm}^x + \alpha_{tm}^x, Y_{tm} = \sum_{m=1}^M d_{tm} \omega_{tm}^y + \alpha_{tm}^y$$

- データ点毎の重み: $\alpha_{\tau} \in \mathbb{R}^{2 \times N}$
- 属性ごとの重み: $\omega_{\tau} \in \mathbb{R}^{2 \times M}$
- ラベル: $L \in \mathbb{R}^N, l \in C^N, C = \{l_1, l_2, \dots\}$



ユーザ操作の解釈手法

- ① 絶対的操作
 - 投影基準の**大局的**な変更
 - 軸上下への対象の移動 (ω の変更)
 - > 投影の各属性の ω 調整: 調整結果を最適化してフィードバック
- ② 相対的操作
 - 投影の**局所的**な微調整
 - 散布図内での制御点の移動 (α の変更)
 - α はデータへのバイアス: 最適化計算を用いて ω -> 投影に反映



探索過程: ω, α を交互に修正し, 最終的に ω を指標として活用



凸包(オブジェクト集合)の操作

- 軌跡, データ点の集合 -> トレンド把握, グループング
- 凸包の作成: ラベルの指定/変更
- 凸包の直接操作 -> メンバシップ更新, グループ結合

インタフェースとしての実装

相補的マルチビュー

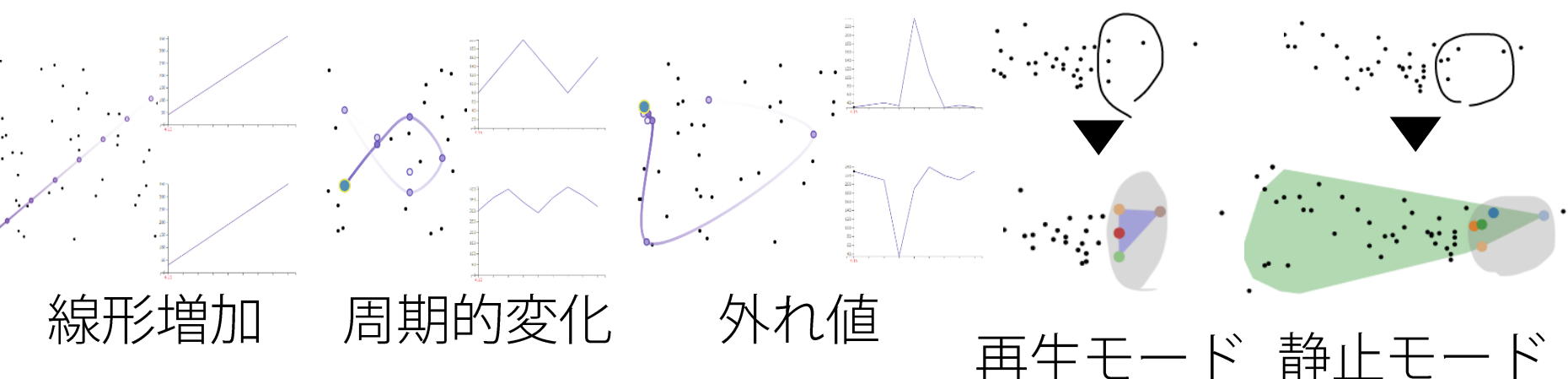
- **散布図**: 軌跡の直接操作 -> 視覚的傾向, 仮説形成
- **詳細**: 仮説の検証
- **連携** -> 探索・知識形成支援, 一貫性保持

探索モードの導入 -> 各可視化手法を活用

- **再生** (アニメーション): データの変化**特性**の把握
- **静止** (軌跡の探索): 知識形成のための**詳細な探索**

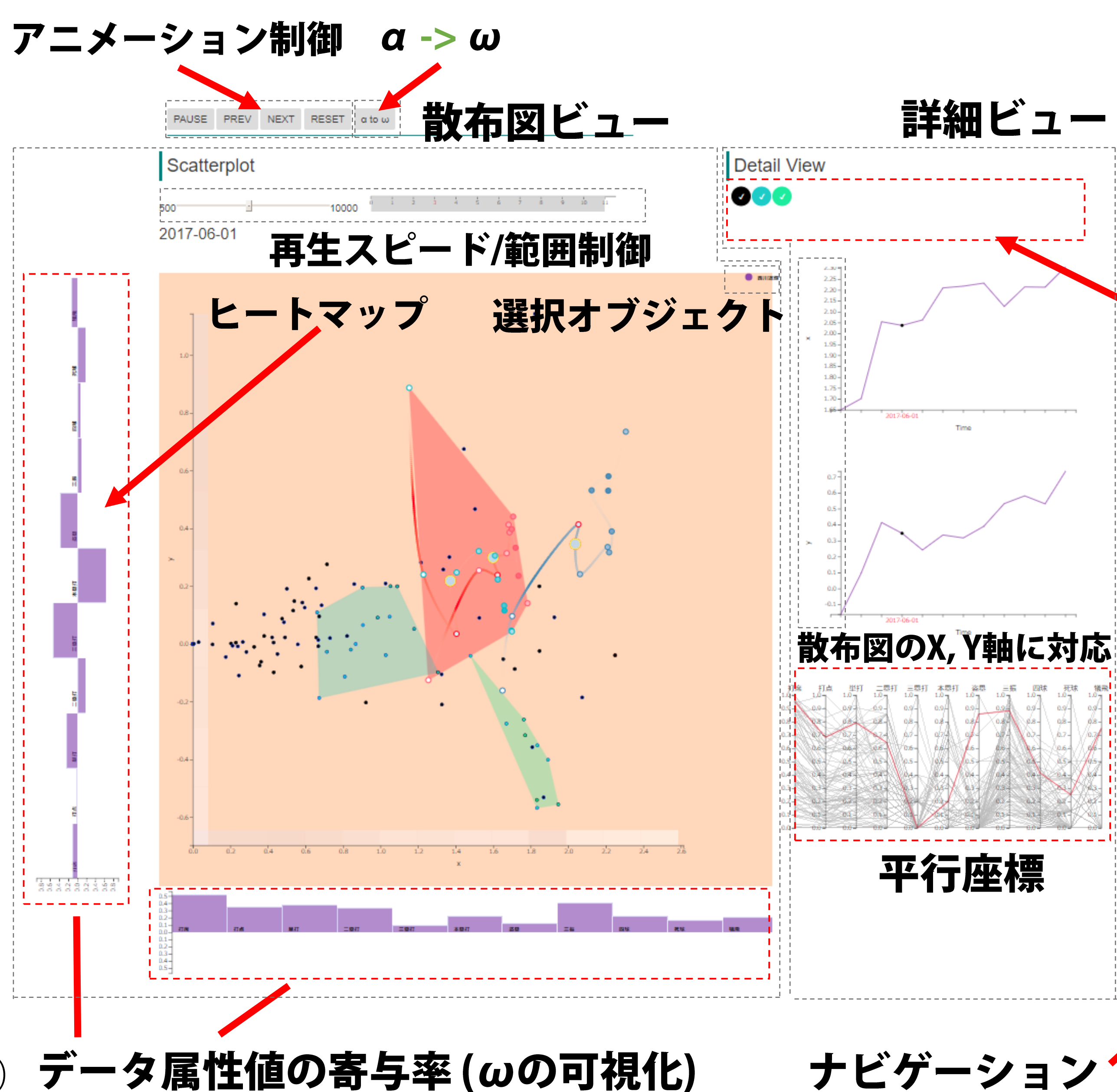
時系列データの**分析過程**: 再生 -> 静止

軌跡/凸包形状 -> データの変化概要の把握



軌跡による時間的ナビゲーション

- (1) 軌跡の各時点のマウスオーバー
- (2) その時点の散布図を重畳表示
- (3) クリックで全体に反映 (時間的遡及)



① 直接操作対象変更

- 現在の対象をアイコンで表現

② 探索モード(再生/静止)変更

- パレットによるスケッチベース入力
- 結果は凸包で表現
- 表示/非表示の制御が可能

④ オブジェクト削除

- 現在の操作対象 (軌跡 or 凸包) 削除

⑤ スプラインエディタ

- 用途: 軌跡形状の微修正
- 適用結果は α に反映

ω の変更

- X, Y軸へドロップされたデータが**強調される**軸を構成

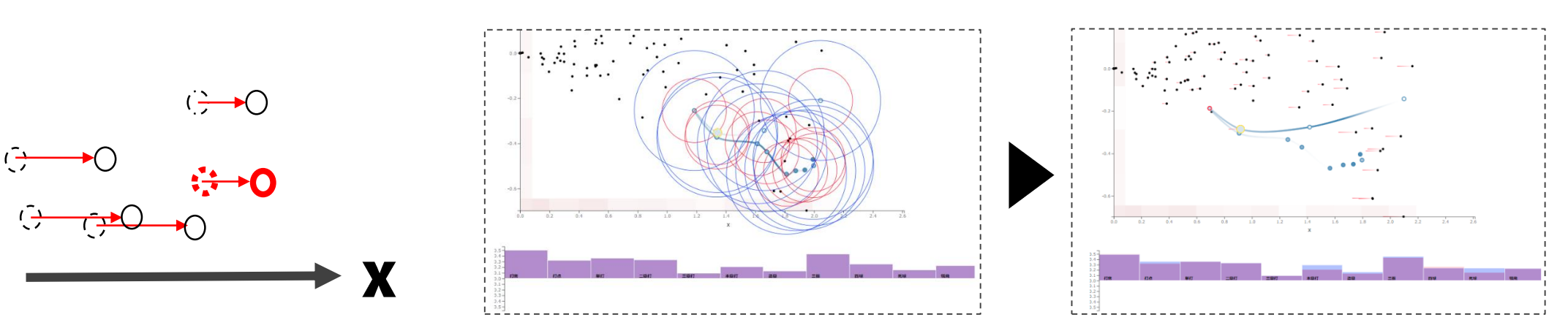
1. ドロップされたデータ属性値に基づき更新 -> ω'

$$\omega = \frac{d_{tm} \text{ と属性 } m \text{ の中央値の差}}{\text{ドロップ位置係数 } c} \times \omega$$

(ドロップ位置 ÷ X or Y軸の値域)

2. ω' を元に新規座標 $P'_{tm} = (X'_{tm}, Y'_{tm})$ を算出

- α が変更されていた場合, ω への還元後に実施



α の変更

- データのドロップ位置に基づき算出: $P \rightarrow P'$
- 移動対象 + 近接 or 離散点も考慮

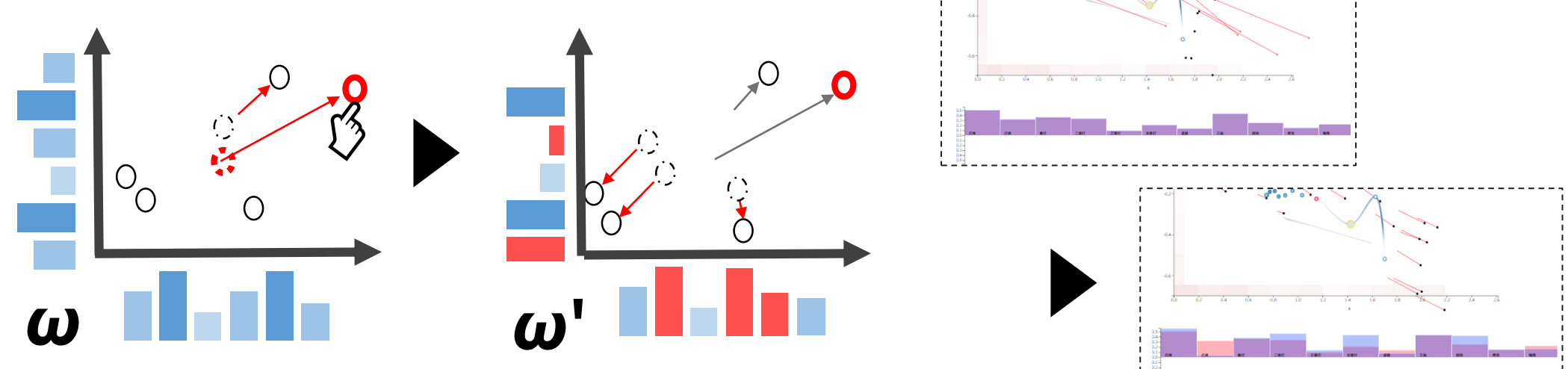
1. α -> 新規座標 $P'_{tm} = (X'_{tm}, Y'_{tm})$ を算出

$$X'_{tm} = X_{tm} + \alpha_{tm}^x, Y'_{tm} = Y_{tm} + \alpha_{tm}^y$$

2. 投影間 (P と P') の差を最小化, α を ω に還元

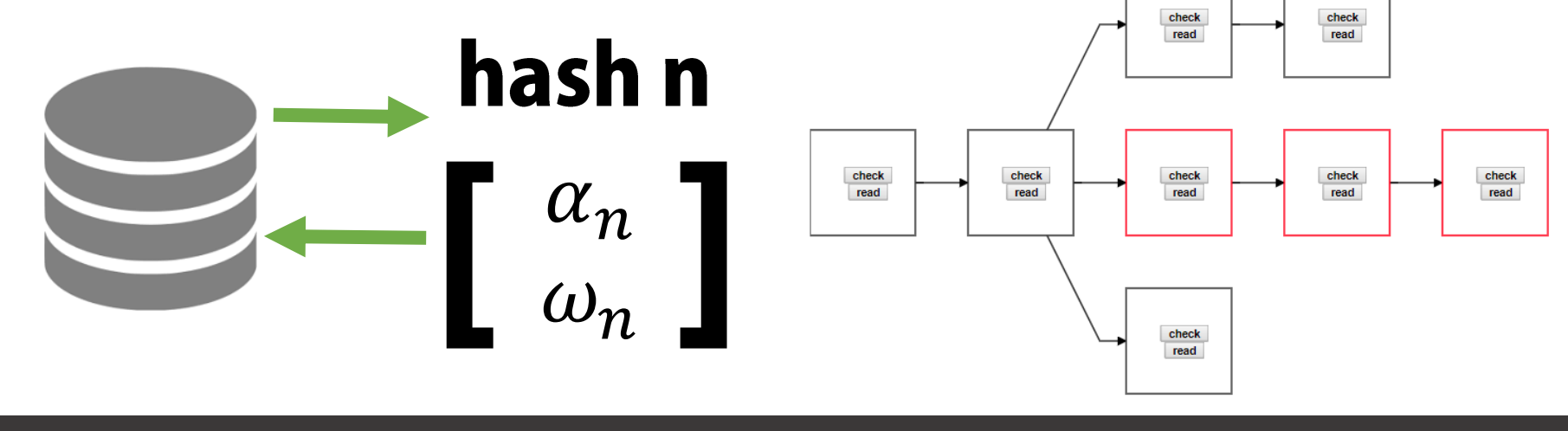
- 逐次二次計画法で ω を決定
- $\text{argmin}_{\omega_{\tau}} \|d_{\tau} \omega_{\tau} - P'_{\tau}\|_2 + \lambda \|\omega_{\tau} - \omega'_{\tau}\|$

-> データ座標計算, 再投影



履歴管理機能

- 洞察形成支援: 作業過程の**外在化**が重要 -> 分析支援, ストーリーテリング, 発想支援
- バックエンドで α, ω 格納 -> 指標再利用



今後の展望

インタフェースの評価と改善

- **ケーススタディ**, 定性的な有効性検証:
- ドメイン専門家による使用
- 定量的実験: 指標形成 -> 主観/客観的に評価