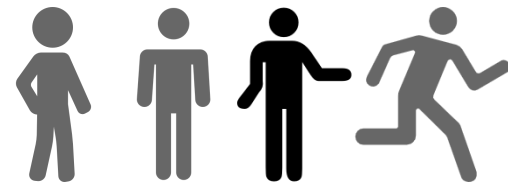




Towards Better Person Re-Identification through Interactive Visual Exploration and Incremental User Feedback



Reporter: Xiawang
Date: 2022/5/31



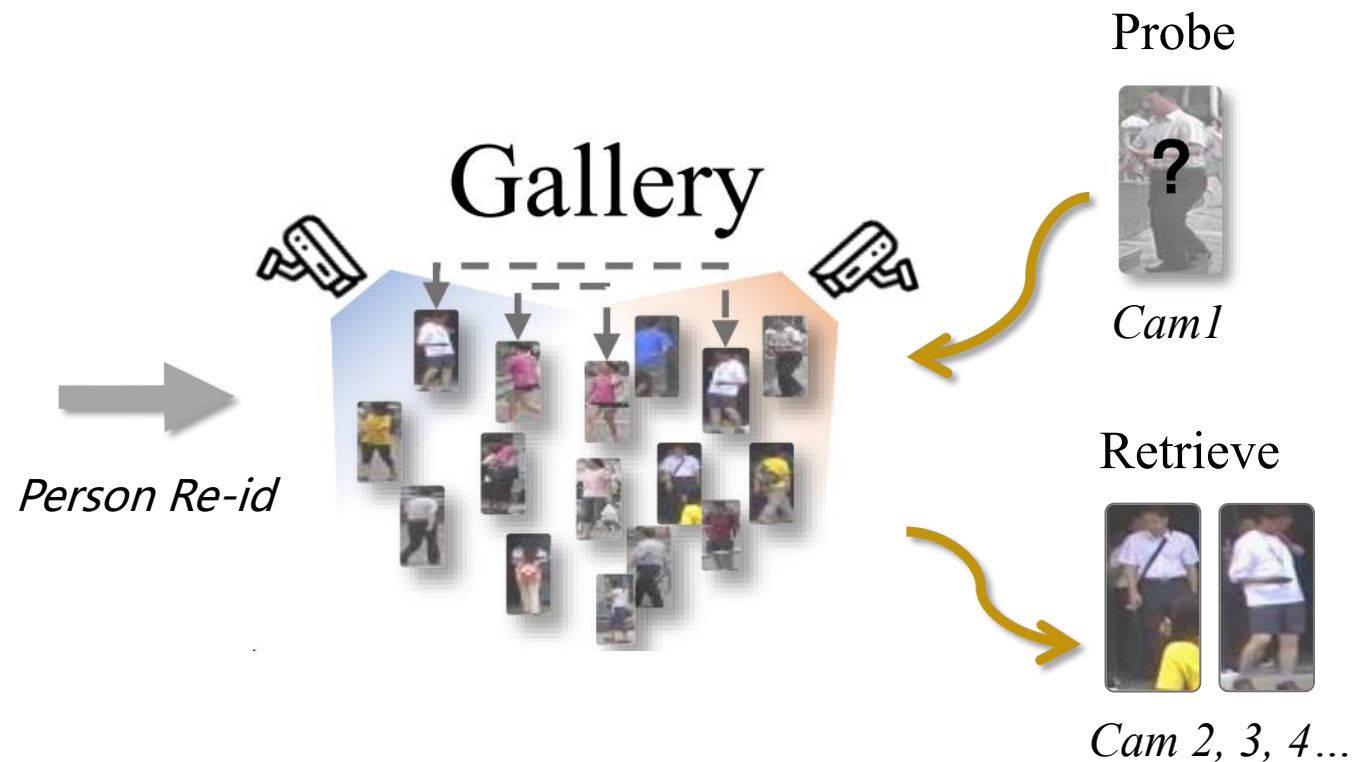
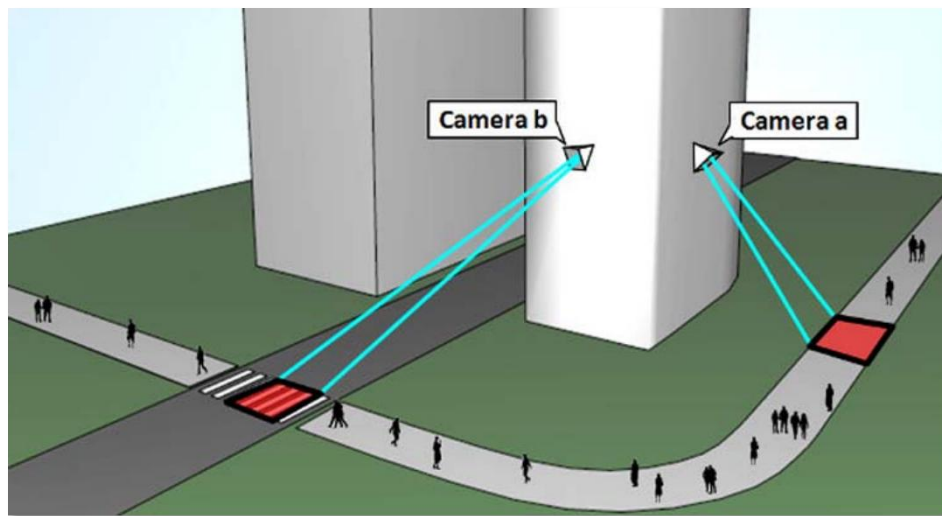
Content

- **Part1 : Background**
- Part2 : Challenges
- Part3 : Visual Analysis
 - Requirement
 - Data Extraction
 - User-feedback Mechanism
 - Visual Design
- Part4 : Case Study



Background : What is Person Re-ID?

- **Application scenarios:** Surveillance Video, Security Field
- **Target task:** identify the same pedestrian under different cameras



Background : Difficulties in Person Re-ID

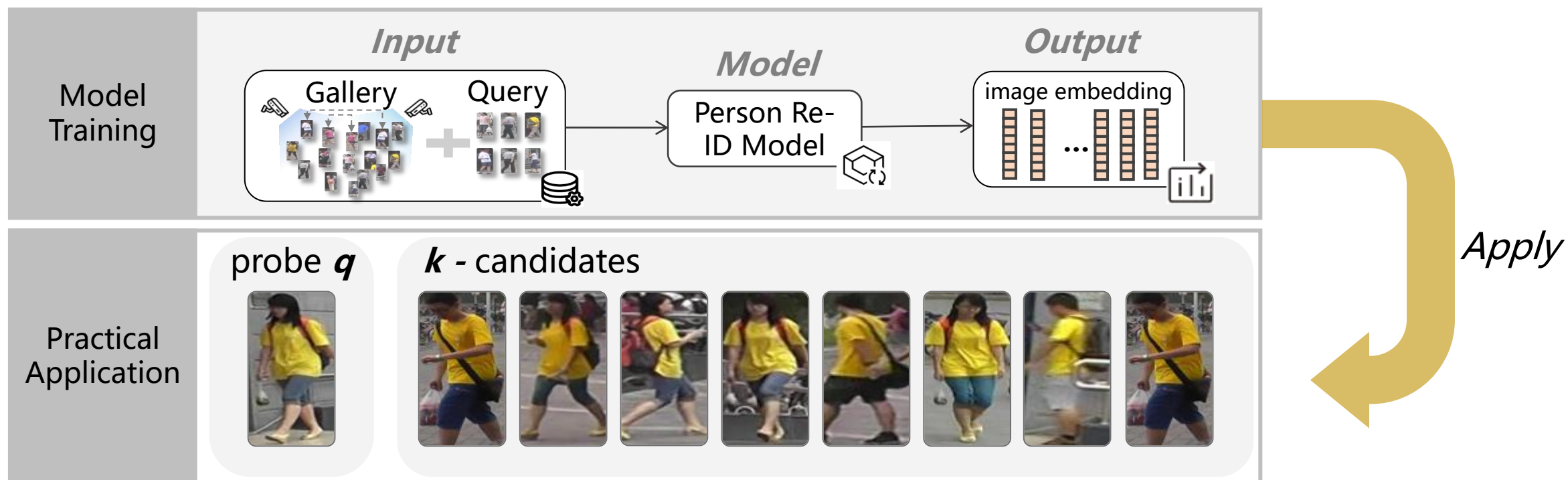
- **Challenge 1:** appearance variation of the same pedestrian is introduced due to viewing angle differences (**large inter-class differences**)
- **Challenge 2:** different pedestrians have similar appearances and require fine-grained distinction (**small intra-class differences**)



Does everyone like to wear yellow tops and dark pants?

Background: Person Re-ID in ML

- **Input-Output:** input data is *gallery* and *query*, output data is the *feature vector* of each image
- **Real-world deployment:** input probe q to the model, output k candidates sorted by similarity



Content

- Part1 : Background
- **Part2 : Challenges**
- Part3 : Visual Analysis
 - Requirement
 - Data Extraction
 - User-feedback Mechanism
 - Visual Design
- Part4 : Case Study



Challenges: Person Re-ID in real world

- Challenge 1: Person Re-ID models are difficult to apply in real-world deployments

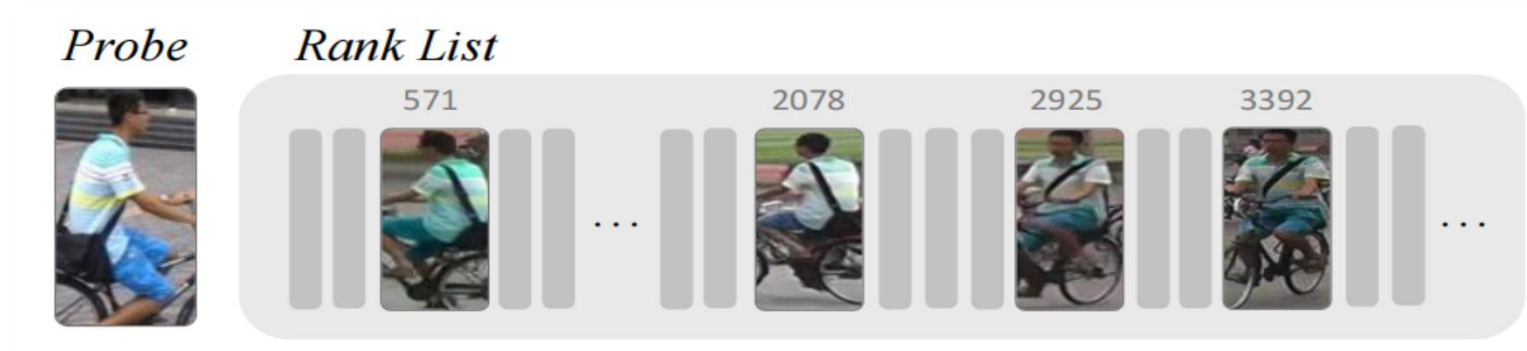


Cross-view appearance variations

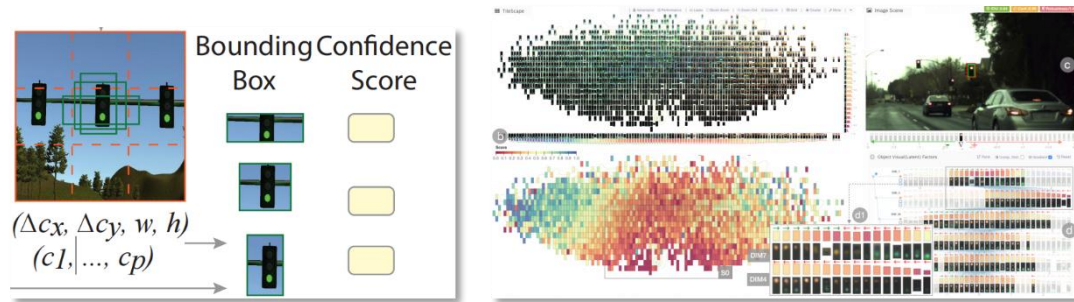


Similar appearance among different people

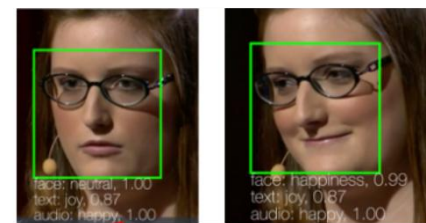
- Challenge 2: Automatic Person Re-ID algorithms may get bad results



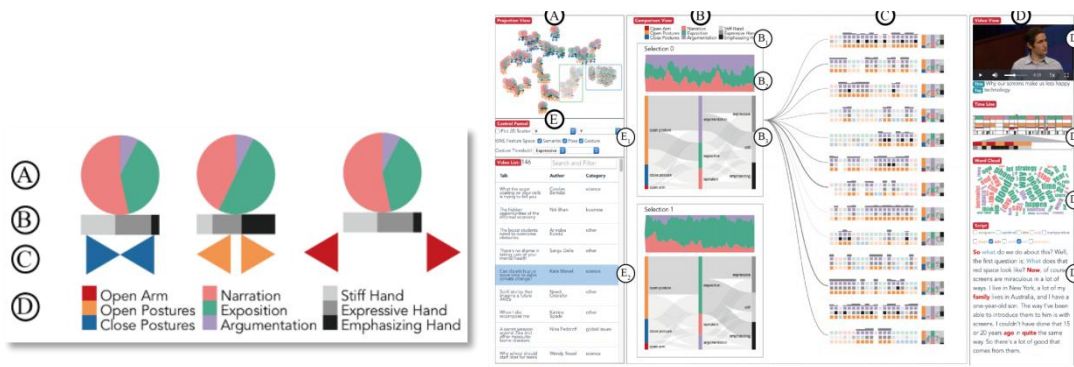
Related Works: Video X Vis



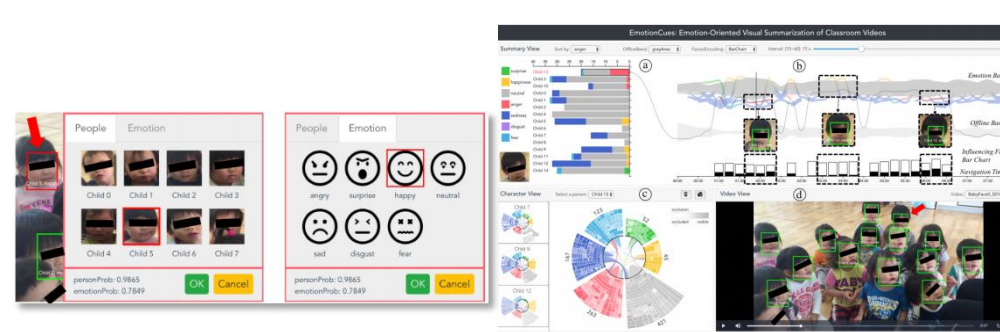
TVCG, 2020[1]



TVCG, 2019[3]



TCVG, 2018[2]



TVCG, 2020[4]

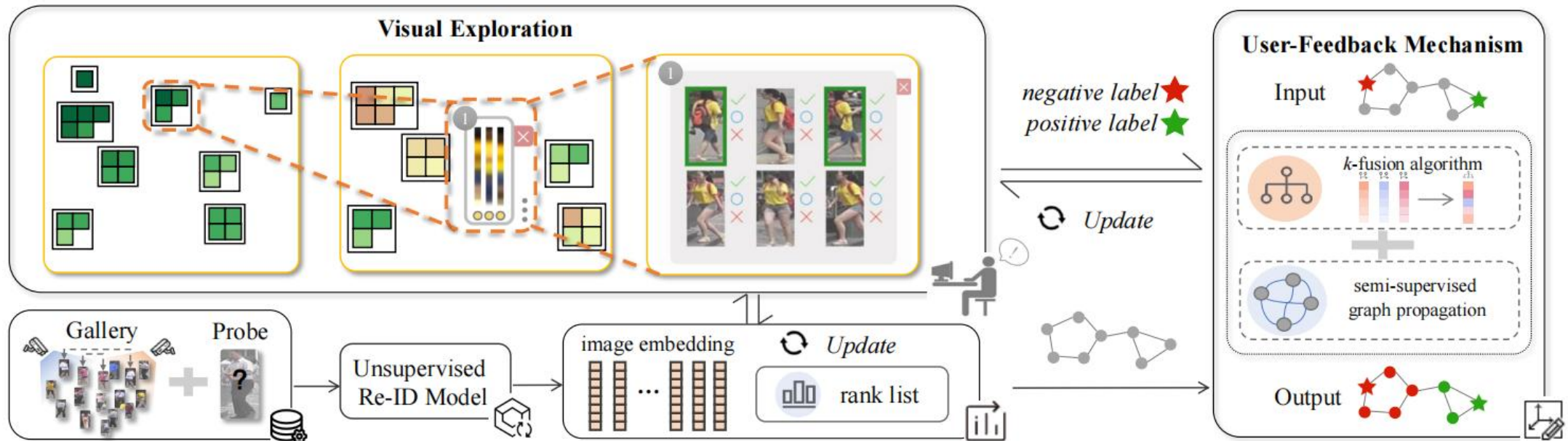
Content

- Part1 : Background
- Part2 : Challenges
- **Part3 : Visual Analysis**
 - Requirement
 - Data Extraction
 - User-feedback Mechanism
 - Visual Design
- Part4 : Case Study



Requirements: Instance-based target tasks

- **Task 1:** Incorporate human insights iteratively in the post-rank refinement
- **Task 2:** Summarize the retrieved samples given a probe
- **Task 3:** Guide user to retrieve the person-of-interest in search space
- **Task 4:** Ensure the content awareness of the retrieved person-of-interest



Demo

The screenshot displays the ReidVis interface with four main components highlighted by callouts:

- 1 Probe Panel:** Located in the top-left, it contains a 'Query' section with a 'probe' image of a person in a yellow shirt. Parameters include 'time : s1_001351', 'camera : c#01', 'rank : 500' (with a slider), and 'context : 0.5' (with a slider). There are also checkboxes for 'rank', 'route', 'image', and 'color'.
- 2 Search Space View:** The central area shows a large grid of small images, each with a grid overlay, representing the search space. A 'Rank summarization' bar at the top of this view shows a color gradient from 0 to 10.
- 3 Ranking List View:** At the bottom, it shows a horizontal list of images from the search results, each with a small grid overlay and a status indicator (e.g., 'x').
- 4 Spatio-temporal Information Panel:** Located in the bottom-left, it shows a vertical timeline with segments labeled 's3' and 's4', and columns labeled 'c1' through 'c6', displaying small images of the person in the yellow shirt.

1 Probe Panel

2 Search Space View

4 Spatio-temporal Information Panel

3 Ranking List View

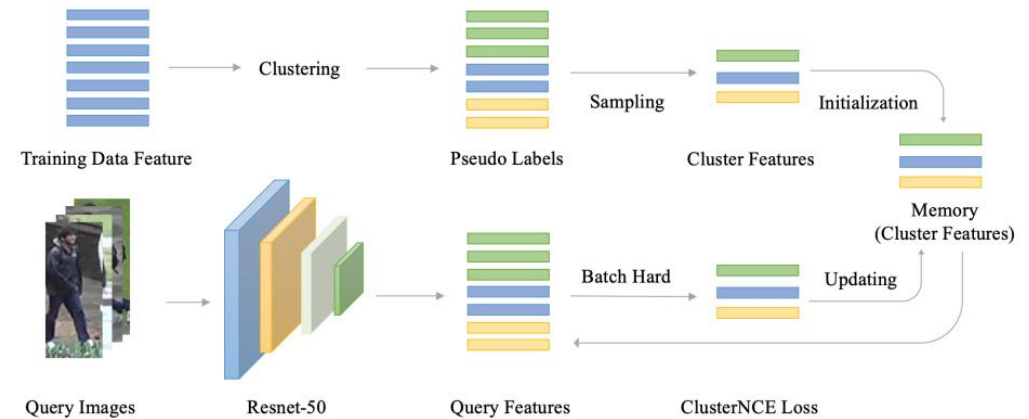
Data Extraction: Extract model results

□ **Dataset:** Market-1501

□ **Data Extraction**

- Step 1: acquire model generated ranking list
- Step 2: image feature extraction

□ **Data list**



[1]Dai Z, Wang G, Zhu S, et al. Cluster Contrast for Unsupervised Person Re-Identification. arXiv 2021[J]. arXiv preprint arXiv:2103.11568, 2021.

Varble	Description
pids	Person object ID
camids	camera ID
feature	512-dimensional vector
impath	image address
results	An array of results, used to match the Gallery image data

(1) Query

Varble	Description
pids	Person object ID
camids	camera ID
feature	512-dimensional vector
impath	image address

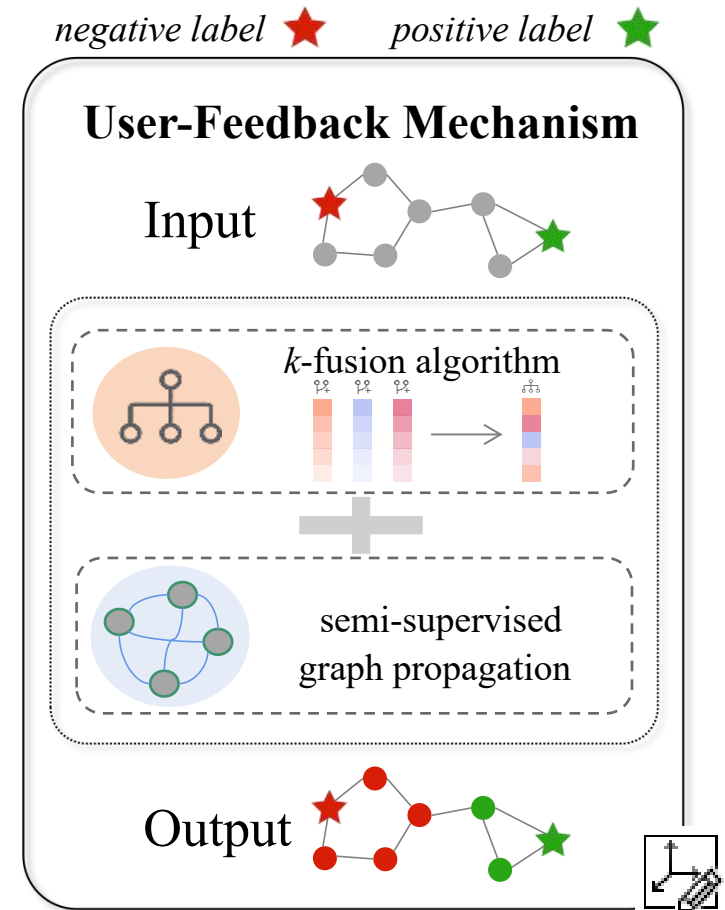
(2) Gallery

User-feedback Mechanism

- **Input:** Manual Label
- **Output:** Refined Ranking List
- **Algorithm:** Graph Propagation



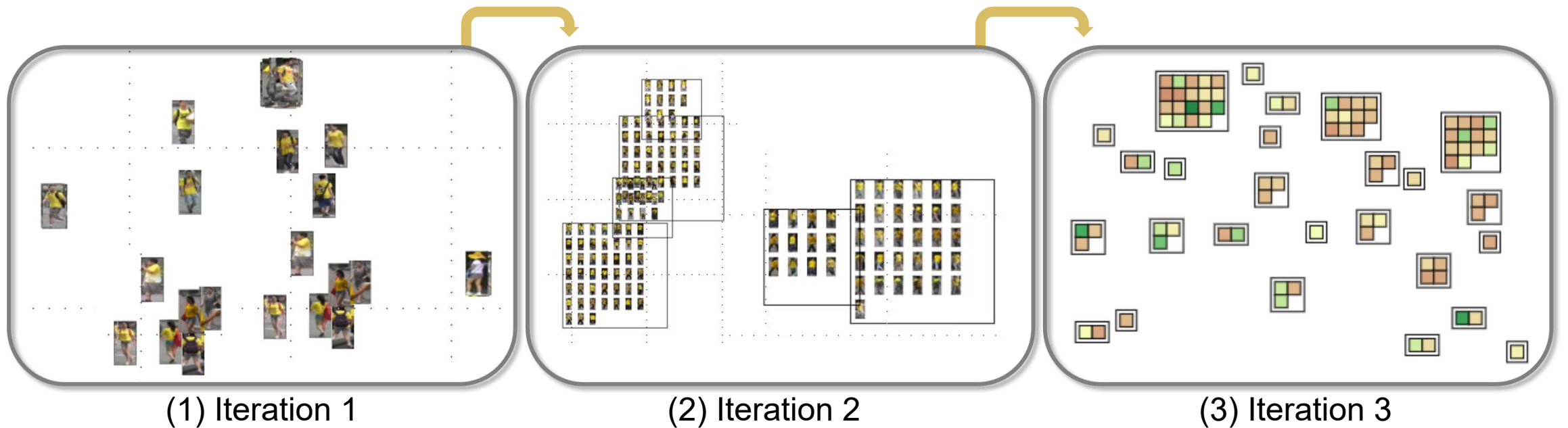
(1) Ranking List View



(2) User-feedback Mechanism

Visual Design: search space overview

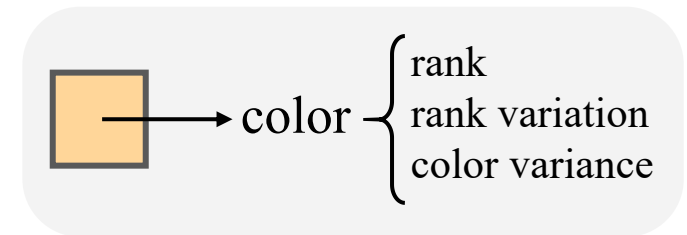
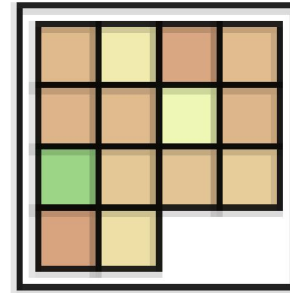
- **Step 1:** Dimensionality reduction
- **Step 2:** Clustering
- **Step 3:** Arrangement



Visual Design: search space design

Attribute-based visual encoding

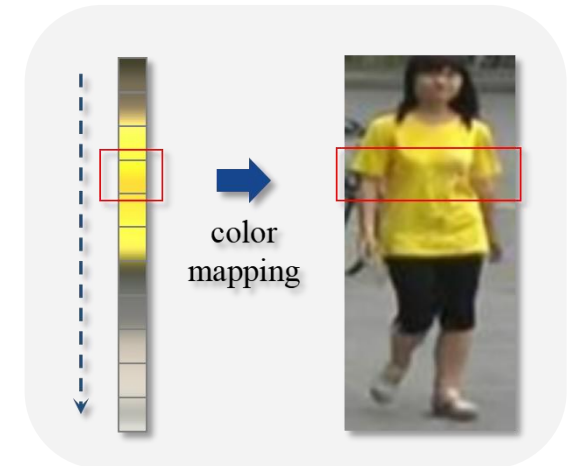
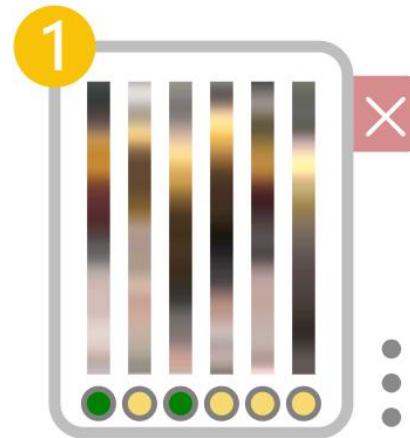
- rank-based colormap
- rank variation-based colormap
- color variance-based colormap



(1) Attribute-based visual encoding

Pixel-based visual encoding

- Compress the image horizontally
- Smooths image colors vertically

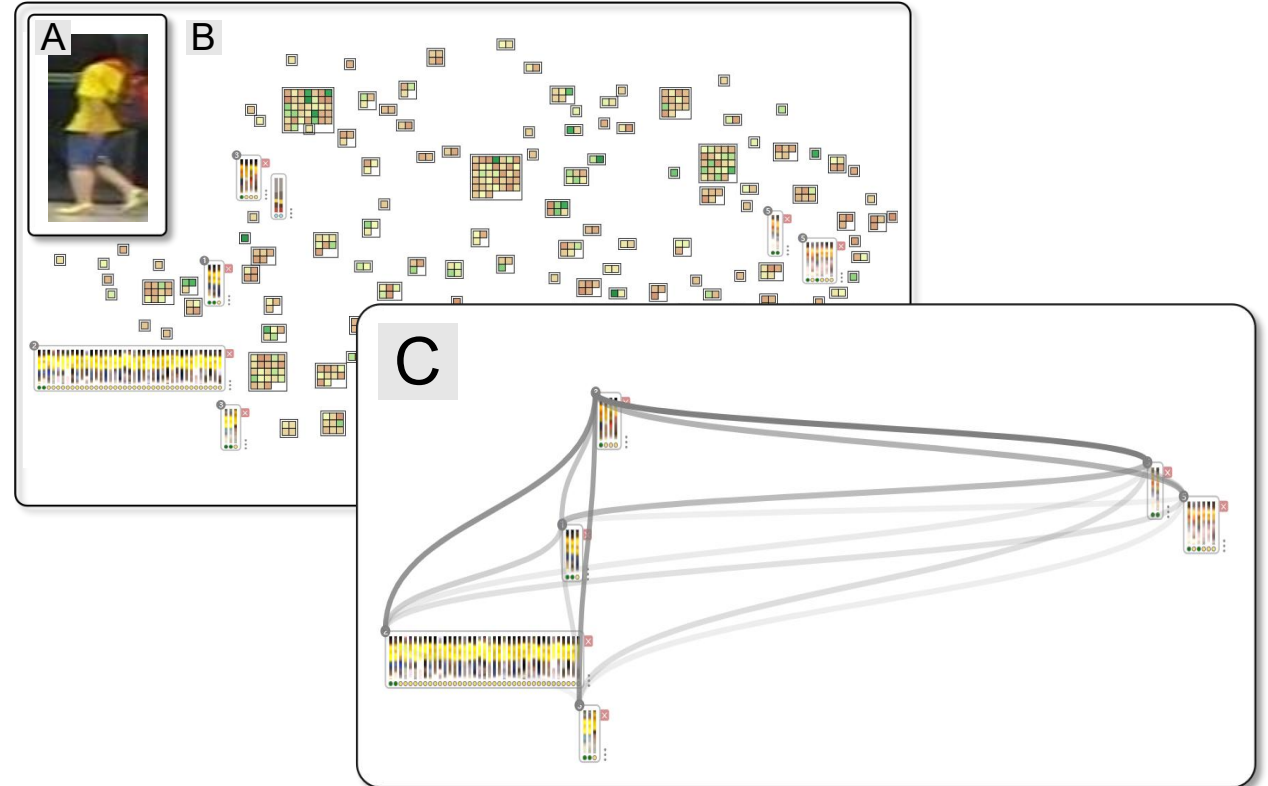


(2) Pixel-based visual encoding

Visual Design: search space design

□ Association-Based Visual Coding

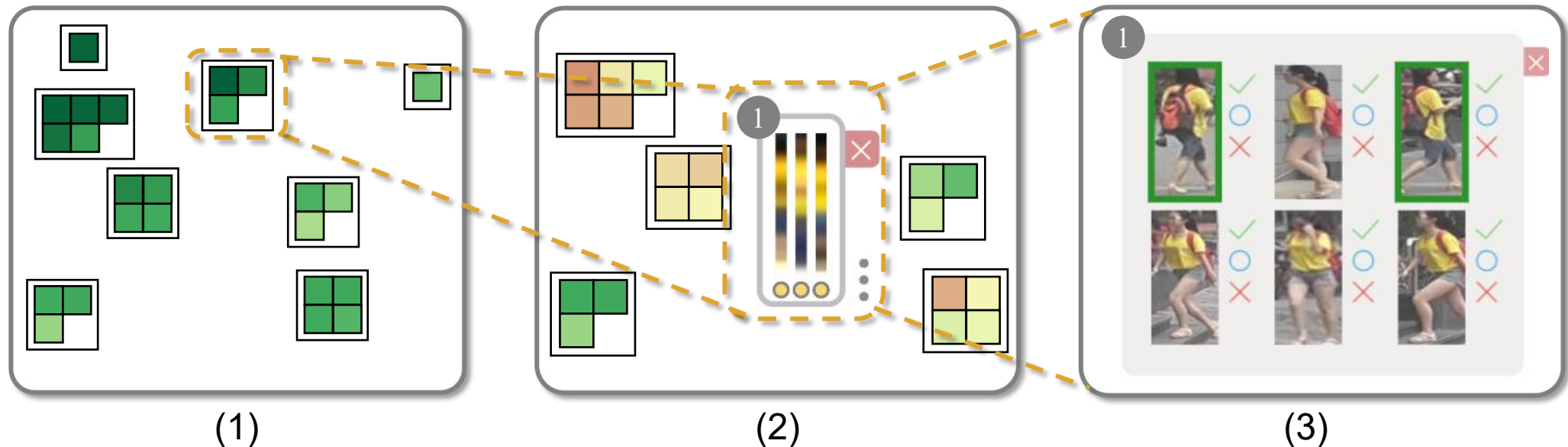
- show relationship in the historical feedback
- color transparency represents the color difference between nodes



Visual Design: interactions in search space

Multi-level exploration

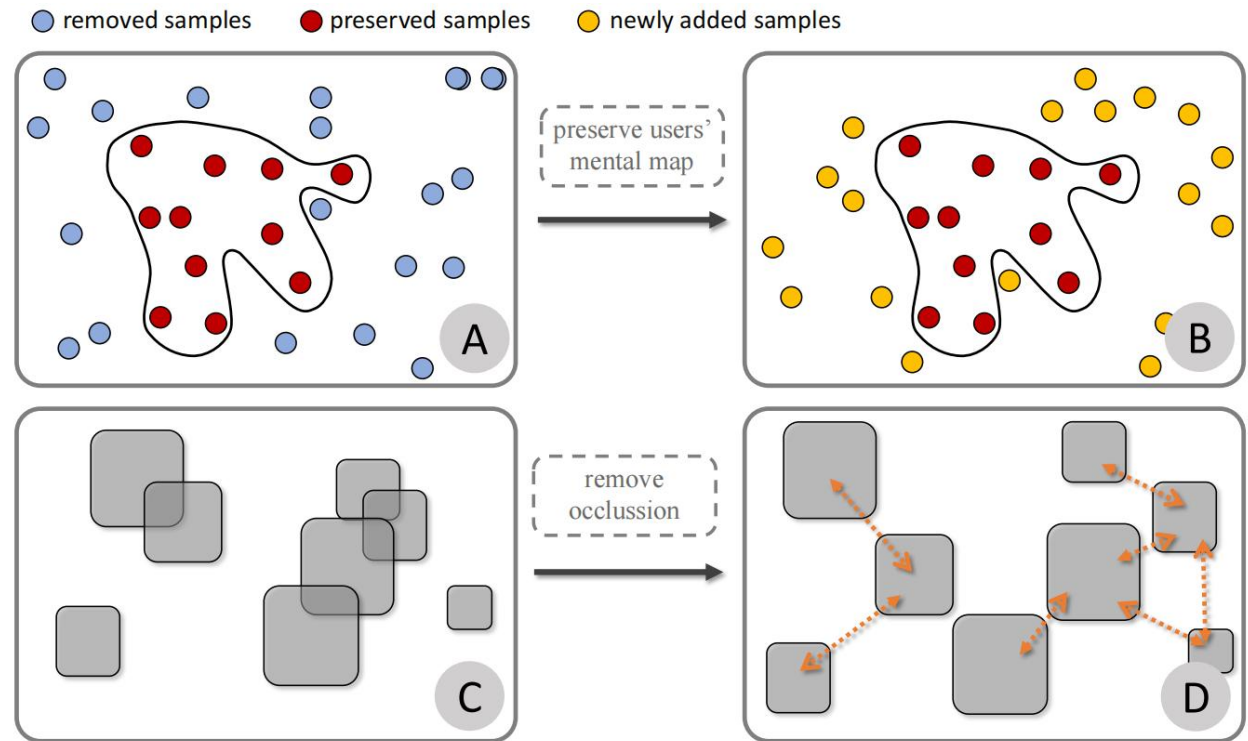
- Attribute-based visual encoding
- Pixel-based visual encoding
- image-based visual encoding



Visual Design: layout algorithm in search space

□ Optimization goals

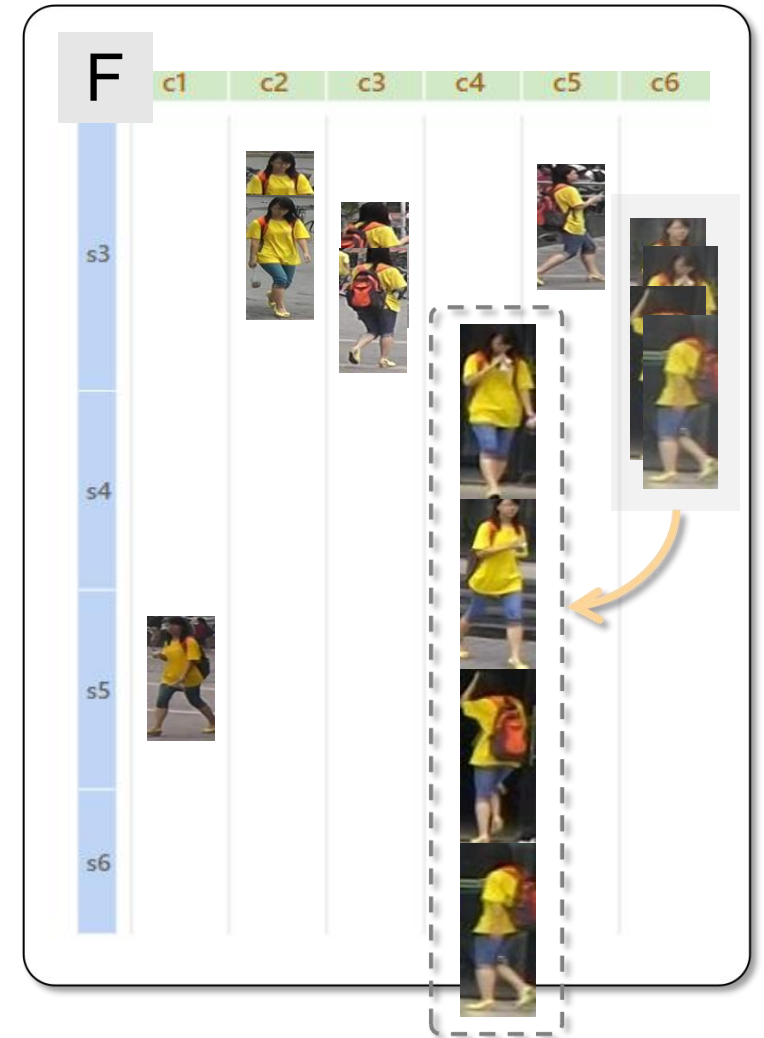
- Ensure the stability of the incremental layout
- Remove the occlusion between nodes



Visual Design: spatiotemporal panel

□ Spatio-temporal Information Panel

- Displays spatiotemporal information for each retrieved image in a 2D coordinate
- X-axis: cameras information
- Y-axis: time information

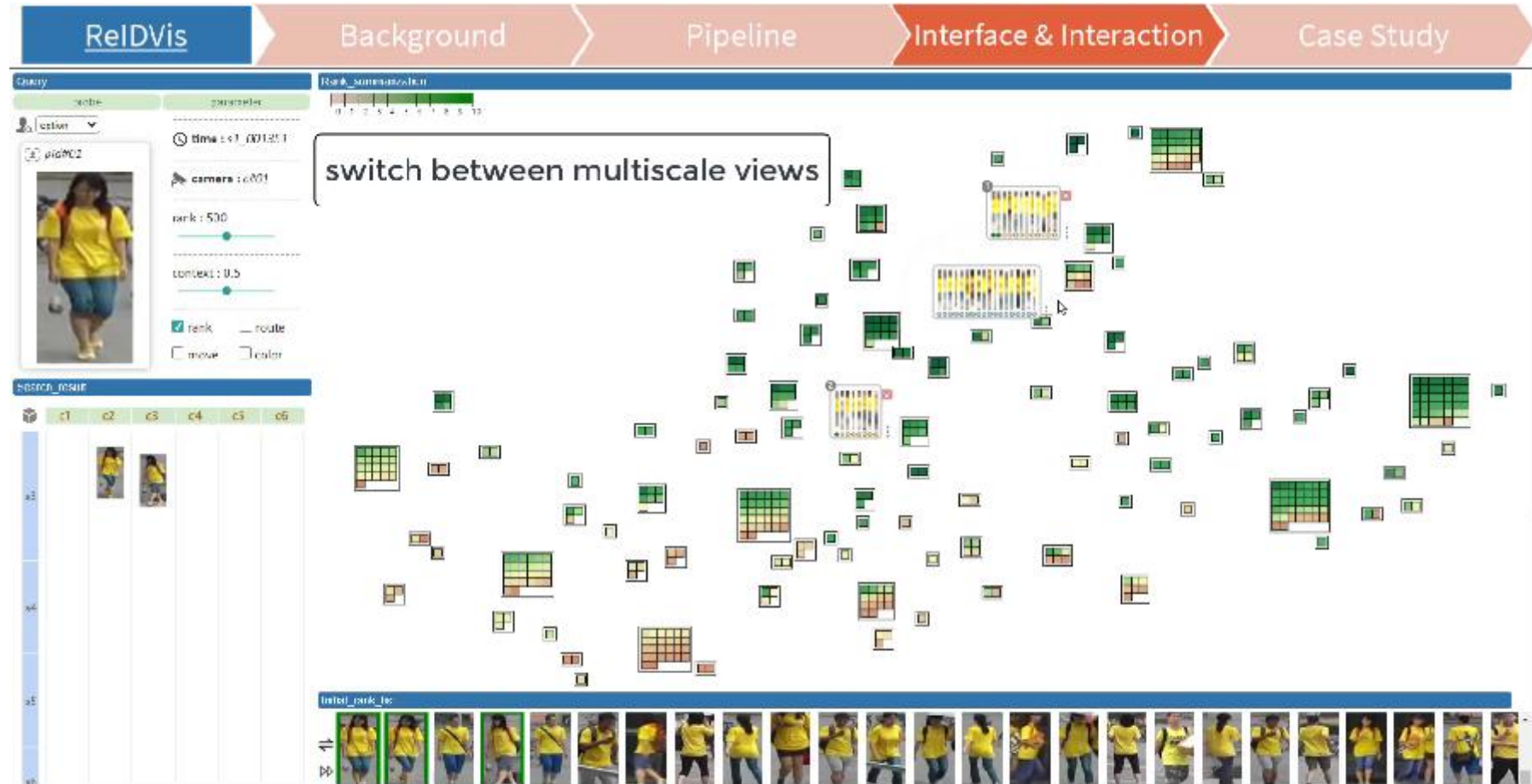


Content

- Part1 : Background
- Part2 : Challenges
- Part3 : Visual Analysis
 - Requirement
 - Data Extraction
 - User-feedback Mechanism
 - Visual Design
- **Part4 : Case Study**



Case study: video demo



Users can switch views in rank, route, move and color modes.

References

- [1] L. Guo, L. Zou, et al. "**VATLD**: A Visual Analytics System to Assess, Understand and Improve Traffic Light Detection." IEEE Transactions on Visualization and Computer Graphics. 2020.
- [2] H. Zeng, X. Wang, et al. "**EmoCo**: Visual Analysis of Emotion Coherence in Presentation Videos." IEEE Transactions on Visualization and Computer Graphics. 2019.
- [3] H. Zeng, X. Shu, et al. "**EmotionCues**: Emotion-oriented Visual Summarization of Classroom Videos." IEEE Transactions on Visualization and Computer Graphics. 2020.
- [4] A. Wu and H. Qu. "Multimodal Analysis of Video Collections: Visual Exploration of **Presentation Techniques** in TED Talks." IEEE Transactions on Visualization and Computer Graphics. 2018.

Thanks For Your Attention!