



Diversified and Summarized Video Search System

Tingting Dong, Shoji Nishimura, Jianquan Liu
System Platform Research Laboratories, NEC Corporation, Japan

A novel video search system that can provide users with diversified and summarized search results automatically

1. Motivation and Concept

Query example:
Search by **face** + **clothing color**

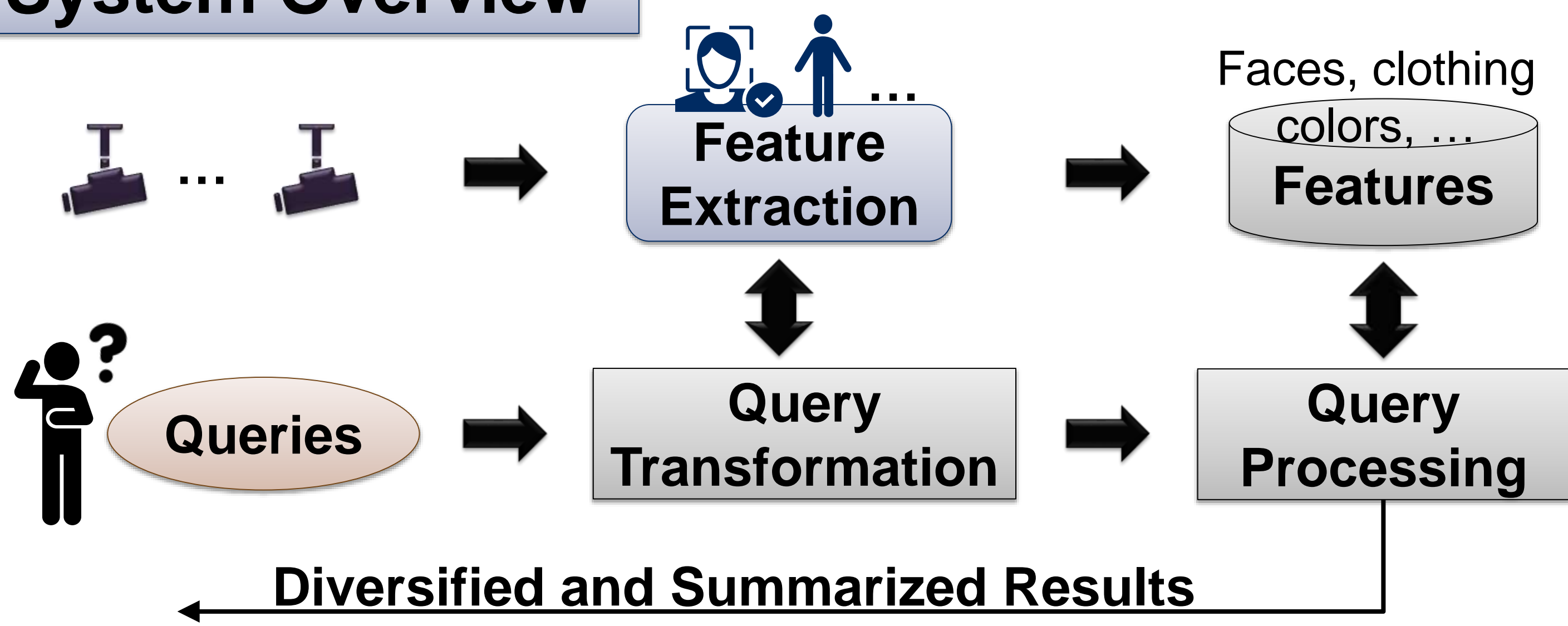


Diversified and Summarized Video Search



2. System Overview and Methods

System Overview



Methods

Objective:

- Provide users with **well-organized** and **intuitive** views of results

Diversification:

- Select objects with largest **diversity scores** as top results

Summarization:

- Assign each candidate to the most similar result

Ideas:

- Utilize statistical information inherent in underlying datasets
- Consider hidden relations (correlations) among data objects

Definitions:

- Diversity score (DIV)** of an object o_i w.r.t o_j combines **feature importance** and **independence**: **larger DIV** \rightarrow **more diverse**

$$DIV(o_i; o_j) = \sum_{m=1}^d RP(f_{m,i}) \cdot IP(f_{m,i}; f_j) \quad IP(f_{m,i}; f_j) = \prod_{u=1}^d IP(f_{m,i}; f_{u,j})$$

- Feature Importance**

- E.g. face is more **important** than clothing color in identifying a person

\rightarrow Representation Power (RP) of a feature $f_{m,i}$

- The possibility of using $f_{m,i}$ to identify an object o_i among all candidates in C

$$RP(f_{m,i}) = \frac{1}{\|\{o_p \mid Sim_{o_i, o_p}(f_m) \geq \tau(f_m), o_p \in C\}\|}$$

- Feature Independence**

- E.g. age has low **independence** w.r.t face because face can predict age

\rightarrow Independence Power (IP) of a feature $f_{a,i}$ w.r.t $f_{b,j}$

- The possibility of $f_{a,i}$ being not correlated by or independent of $f_{b,j}$

$$IP(f_{a,i}; f_{b,j}) = \frac{\|\{o_p \mid Sim_{o_i, o_p}(f_a) < \tau(f_a), o_p \in R_b\}\|}{\|R_b\|}, \quad R_b = \{o_p \mid Sim_{o_i, o_p}(f_b) \geq \tau(f_b), o_p \in C\}$$

3. Demo

Data: surveillance videos in three stores

- Three cameras in each store, multiple videos (outfits) by each camera
- Video recording time: 5-10 minutes
- Samples:



Convenient store



Electronics store



Bank ATM

Implementation of feature extraction:

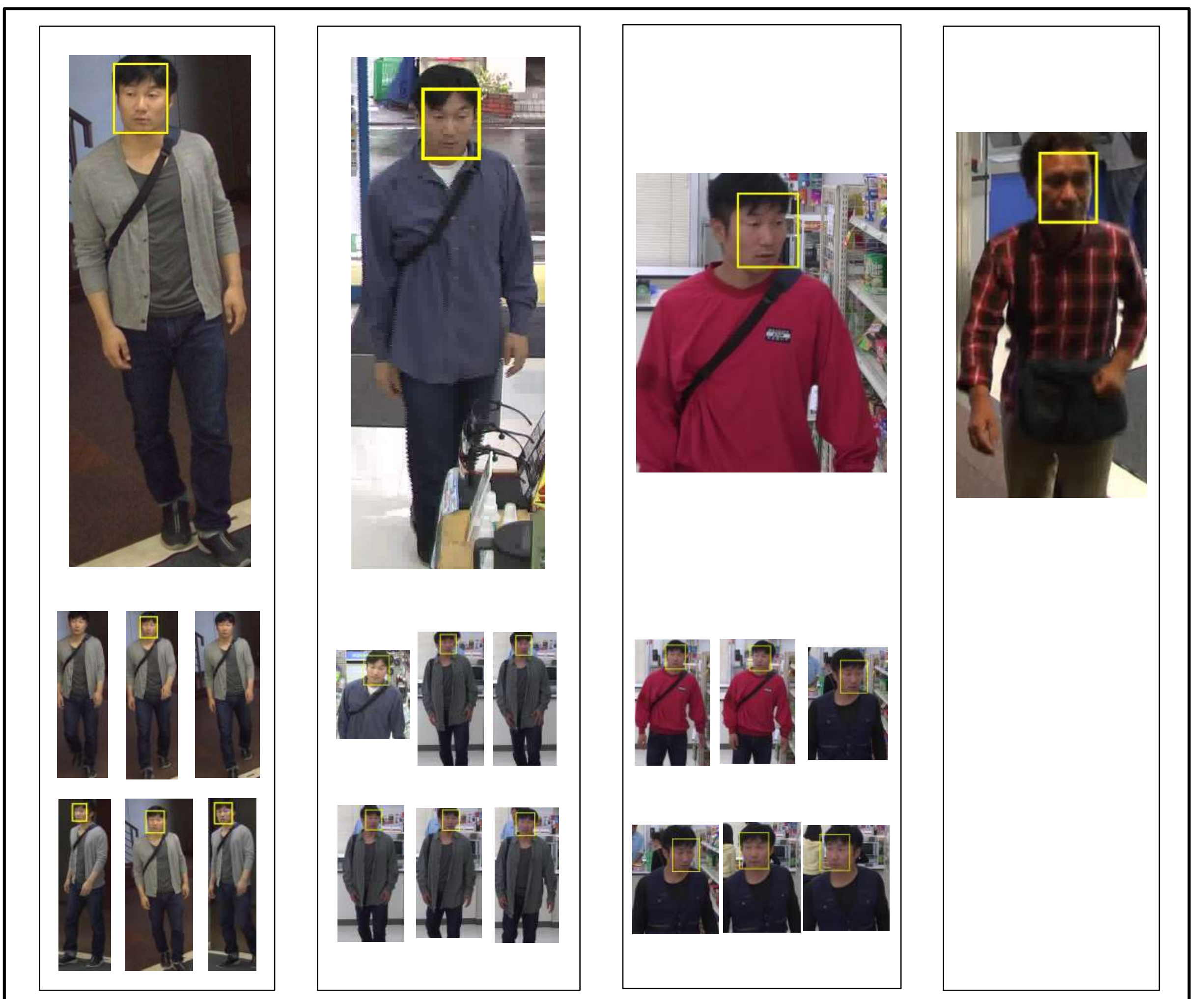
- Face features: **NeoFace®**
- Clothing color (HSV) features: **YOLO** + Clothing feature extraction library
- Other features (info): store location, camera ID, video ID, time, etc.

Queries:

- Search modes: **Face**, **Clothing color**, **Face + Clothing color**
- Default similarity thresholds: $\tau(\text{face}) = 0.8$, $\tau(\text{color}) = 0.8$, $\tau(\text{info}) = 0.8$

Example: search by Face + Clothing color

Query



4. References

- [1] Jianquan Liu, Shoji Nishimura, and Takuya Araki. Wally: A Scalable Distributed Automated Video Surveillance System with Rich Search Functionalities. MM 2014
- [2] Jianquan Liu, Shoji Nishimura, and Takuya Araki. AntiLoiter: A Loitering Discovery System for Longtime Videos across Multiple Surveillance Cameras. MM 2016
- [3] Jianquan Liu, Shoji Nishimura, and Takuya Araki. VisLoiter: a system to visualize loiterers discovered from surveillance videos. SIGGRAPH 2016
- [4] Jianquan Liu, Shoji Nishimura, Takuya Araki, and Yuichi Nakamura. A Loitering Discovery System Using Efficient Similarity Search Based on Similarity Hierarchy. IEICE Transactions 2017

Contact: Tingting Dong t-dong@ce.jp.nec.com