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

Existing Method

Diversified and Summarized Video Search

Duplication results!

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:
  \[
  DIV(o_i; o_j) = \sum_{m=1}^{M} RP(f_{m,i}) \cdot IP(f_{m,i}; f_{j}) \cdot IP(f_{m,i}; f_{j}) = \prod_{m=1}^{M} IP(f_{m,i}; f_{j})
  \]
- **Feature Importance**
  - E.g., face is more important than clothing color in identifying a person
- **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 \( C \)
  \[
  RP(f_{m,i}) = \frac{\{o_j \mid Sim_n(o_i, f_{m,i}) \geq \tau(f_{m,i}), o_j \in C\}}{|C|}
  \]
- **Feature Independence**
  - E.g., face has low independence w.r.t face because face can predict age
  - **Independence Power (IP)** of a feature \( f_{m,i} \) w.r.t \( f_{m,j} \):
    - The possibility of \( f_{m,i} \) being not correlated by or independent of \( f_{m,j} \)
    \[
    IP(f_{m,i}; f_{m,j}) = \frac{\{o_j \mid Sim_n(o_i, f_{m,i}) \leq \tau(f_{m,j}), o_j \in C\}}{|C|}
    \]

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

**Examples:** search by Face + Clothing color

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: \( \tau(\text{face}) = 0.8, \tau(\text{color}) = 0.8, \tau(\text{info}) = 0.8 \)

4. References


Contact: Tingting Dong | dong@ece.jp.nec.com

ACM Multimedia 2017