RECFusion: Automatic Scene Clustering and Tracking in Videos from Multiple Sources

F.L.M. Milotta¹, S. Battiato¹, F. Stanco¹, V. D’Amico², G. Torrisi², L. Addesso²

¹ Department of Mathematics and Computer Science, University of Catania, Italy
² Telecom Italia JOL WAVE, Catania, Italy
Overview

1. Motivations
2. Main Aims
3. Notation
4. Intra-flow Classification & Inter-flow Classification
5. Cluster Tracking
   1. Cluster Log
   2. Initialization
   3. Tracking
   4. Update
6. Selection of the Most Representative Frame for Output
7. Demo
8. Conclusion and Future Works
Motivations

- The automatic processing of video data from many devices, as smartphones, tablets, webcams, surveillance cameras, etc., in the real-time context is not a trivial issue.
Main Aims

1. Analysis of video streams from multi-source multi-device context

2. Identification of the *scenes of interest* through clustering of video sequences

3. Time tracking of the computed scenes clusters
Notation

- **Denotes a frame from video stream (device) $v_k$ at time-slot $t_n$**

- **Output of Intra-flow classification**: denotes a scene identified by a $Scene_{ID}$ from video stream (device) $v_k$ at time-slot $t_n$

- **Output of Inter-flow classification**: denotes a scene identified by a $Cluster_{ID}$ from video stream (device) $v_k$ at time-slot $t_n$

- **Output of Cluster Tracking**: denotes a cluster identified by a $LoggedCluster_{ID}$ from video stream (device) $v_k$ at time-slot $t_n$
Intra- and Inter-flow Classifications

**INTRA-FLOW CLASSIFICATION OF $d_6$**

<table>
<thead>
<tr>
<th></th>
<th>$S_{1.6.1}$</th>
<th>$S_{1.6.2}$</th>
<th>$S_{2.6.3}$</th>
<th>$S_{1.6.4}$</th>
<th>$S_{3.6.5}$</th>
<th>$N_{6.6}$</th>
<th>$S_{2.6.7}$</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>$f_{1.1}$</td>
<td>$f_{1.2}$</td>
<td>$f_{1.3}$</td>
<td>$f_{1.4}$</td>
<td>$f_{1.5}$</td>
<td>$f_{1.6}$</td>
<td>$f_{1.7}$</td>
<td>$f_{1.8}$</td>
</tr>
<tr>
<td>$t_2$</td>
<td>$f_{2.1}$</td>
<td>$f_{2.2}$</td>
<td>$f_{2.3}$</td>
<td>$f_{2.4}$</td>
<td>$f_{2.5}$</td>
<td>$f_{2.6}$</td>
<td>$f_{2.7}$</td>
<td>$f_{2.8}$</td>
</tr>
<tr>
<td>$t_3$</td>
<td>$f_{3.1}$</td>
<td>$f_{3.2}$</td>
<td>$f_{3.3}$</td>
<td>$f_{3.4}$</td>
<td>$f_{3.5}$</td>
<td>$f_{3.6}$</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$t_4$</td>
<td>$f_{4.1}$</td>
<td>$f_{4.2}$</td>
<td>$f_{4.3}$</td>
<td>$f_{4.4}$</td>
<td>$f_{4.5}$</td>
<td>$f_{4.6}$</td>
<td>$f_{4.7}$</td>
<td>$f_{4.8}$</td>
</tr>
<tr>
<td>$t_5$</td>
<td>$f_{5.1}$</td>
<td>$f_{5.2}$</td>
<td>$f_{5.3}$</td>
<td>$f_{5.4}$</td>
<td>$f_{5.5}$</td>
<td>$f_{5.6}$</td>
<td>$f_{5.7}$</td>
<td>$f_{5.8}$</td>
</tr>
<tr>
<td>$t_6$</td>
<td>$f_{6.1}$</td>
<td>$f_{6.2}$</td>
<td>$f_{6.3}$</td>
<td>$f_{6.4}$</td>
<td>$f_{6.5}$</td>
<td>$f_{6.6}$</td>
<td>$f_{6.7}$</td>
<td>...</td>
</tr>
<tr>
<td>$t_7$</td>
<td>$f_{7.1}$</td>
<td>$f_{7.2}$</td>
<td>$f_{7.3}$</td>
<td>$f_{7.4}$</td>
<td>$f_{7.5}$</td>
<td>$f_{7.6}$</td>
<td>$f_{7.7}$</td>
<td>$f_{7.8}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**FROM INTRA- TO INTER-FLOW CLASSIFICATION AT $t_5$**

|   | $S_{1.1.5}$ | $S_{1.2.5}$ | $S_{2.3.5}$ | $S_{1.4.5}$ | $S_{3.6.5}$ | $S_{2.7.5}$ | $C_{1.1.5}$ | $C_{2.2.5}$ | $C_{2.3.5}$ | $C_{3.4.5}$ | $C_{1.5.5}$ | $C_{1.6.5}$ | $C_{2.7.5}$ | ... |
|---|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-----|
| $N_{5.5}$ | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
Why Cluster Tracking?

• **Intra-flow Classification:**
  - The $Scene_{ID}$s are related to a single video stream. Frames labeled with the same $Scene_{ID}$ but taken from different video streams could definitely represent different scenes!

• **Inter-flow Classification:**
  - The $Cluster_{ID}$s are related to a single time-slot. Frames labeled with the same $Cluster_{ID}$ but taken from different time-slots could definitely represent different scenes!

• **Cluster Tracking:**
  - $LoggedCluster_{ID}$s are related to the whole dataset of video streams: they represent the same scene in every time-slot!
Cluster Tracking (1)

The image contains a diagram that illustrates the concept of cluster tracking. The diagram shows a grid with rows labeled \( v_1, v_2, v_3, \ldots \) and columns labeled \( t_2, t_3, t_4, t_5, \ldots \). Each cell in the grid is labeled with a combination of a cluster \( C_{i,j,k} \) and a time slot \( f_1, f_2, \ldots \). The clusters and time slots are connected to nodes labeled \( C_{1,1,3}, C_{2,2,3}, \ldots \) and \( v_1, v_2, \ldots \) respectively, indicating a tracking relationship between them.
Cluster Tracking (2) – Cluster Log

For each cluster a centroid is logged and keep updated

Tracking of the cluster (Re-labeling)

\[ C_{ID,v_k,t_n} \rightarrow L_{ID,v_k,t_n} \]
Cluster Tracking (3) – Initialization

\[ C_{1,1,1}, C_{1,2,1}, C_{2,3,1}, C_{3,4,1}, C_{2,5,1}, C_{1,6,1}, C_{3,7,1}, \ldots \]

\[ t_1 \]

\[ C_{1,1,1}, C_{1,2,1}, C_{1,6,1}, C_{2,3,1}, C_{2,5,1}, C_{2,5,1}, \ldots \]
Cluster Tracking (4.1) – Tracking

Centroid definition
Match with logged clusters
Re-labeling (Tracking)

Minimum Distance + Thresholding
Cluster Tracking (4.2) – Tracking

COMPUTATION OF MINIMUM DISTANCE

THRESHOLDING

>= T

INSERT NEW LOG

< T

UPDATE

CL

L₁  L₂  L₃  ...

C₁

COMPUTATION OF MINIMUM DISTANCE

< T

UPDATE

T=0,15

INSERT NEW LOG

THRESHOLDING

>= T

INSERT NEW LOG
Cluster Tracking (5.1) – Update

\[ C_{1,1,3} \]
\[ C_{2,2,3} \]
\[ C_{2,3,3} \]
\[ C_{3,4,3} \]
\[ C_{1,5,3} \]
\[ C_{1,6,3} \]
\[ C_{2,7,3} \]

...
Cluster Tracking (5.2) – Update

*Weigthed Update*:

- Logs become even more stable after each update

\[ L'_a = \frac{u}{u + 1} L_a + \frac{1}{u + 1} C_b \]

- Where \( u \) is the number of updates performed, increased by 1 to take into account also the insertion
Cluster Tracking (5.3) – Update

#Cluster <= #LoggedCluster

\[ C_{1,1,t} \quad L_{1,1,t} \quad L_{3,2,t} \quad L_{3,3,t} \]

Update:

New log(s) could be defined

#Cluster > #LoggedCluster

\[ C_{1,1,t} \quad L_{2,1,t} \quad L_{1,2,t} \quad L_{3,3,t} \]

Update:

New log(s) will be defined

New log(s) could be defined
Most Representative Frame of a Cluster at time $t$: Frame with minimum distance from Cluster Centroid

Frame with minimum distance from Cluster Centroid

Centroid of Cluster 1

MIN

Video Output of Cluster 1

$C_{1,1,t}$ $C_{1,2,t}$ $C_{1,3,t}$

$C_{1,1,t}$ $C_{1,2,t}$ $C_{1,3,t}$

$C_1$

$d_1$ $d_2$ $d_3$

$C_1$
Selection of most representative frame: Behaviour of previous RECfusion version

For each time slot $t_n$ a single representative frame is computed from the most popular (the biggest) cluster.
Selection of most representative frame (1): RECfusion + Cluster Tracking

The most popular Cluster

$C_{MP}$

$\text{CL}$

$L_1$  $L_2$  $L_3$  ...

$C_1$  $C_2$  $C_3$  ...

18
Selection of most representative frame (2): RE Cfus ion + Cluster Tracking

http://recfusionproject.altervista.org/clustertracking.htm
Demo

http://recfusionproject.altervista.org/clustertracking.htm
Demo in a glance

http://recfusionproject.altervista.org/clustertracking.htm
Conclusion

• We have presented an extension to RECfusion:
  • **Cluster Tracking** has been added:
    • The different scenes within the collection are identified. They could be all selected and tracked in the whole collection of video-streams by multi-source devices, and might be used in the output video.
    • Previously, only the most popular cluster for each time-slot was used in the output video.
  • An evaluation of a **better threshold for LoggedClusters update** has been conducted:
    • Experimentations assess the goodness of the threshold equal to 0,15
Future Works

- We are planning to add some other functionalities:
  - **Assistive Technology:**
    - Exploiting wearable devices
    - ‘How much time have I spent in a specific room?’
  - **Security & Computer Forensics:**
    - it might give an advice on when the scenario is subject to some changes
RE Cfus ion related publications


• Further extension is in progress…

http://recoveryproject.altervista.org/
Thanks!
Any question?

http://recfusionproject.altervista.org/