Emotional video mashup

Riccardo Leonardi @ NEM Summit

2011

Sergio Benini and Riccardo Leonardi

Department of Information Engineering
DII-SCL, University of Brescia
Via Branze, 38 - 25123 Brescia - ITALY
New forms of creativity

- Semiprofessional tools for multimedia production are now within the reach of common users
- New forms of creativity are emerging

**VIDEO MASHUP**

The practice of combining multiple audiovisual sources into a derivative work whose semantics could be very different compared to the one of the original videos
Video mashup and...

- Different types of mashup
  - Movie trailer
  - Political
  - Advertising
  - Music video
  - ...

Interactive video mashup based on emotional identity
...The need of material

- Loads of multimedia = loads of needs: classify, retrieve, summarize, ...

**Understanding the Semantics (*)**

Cognitive level
(is there a man depicted?, which is the location?, ...)

Emotional level
(mood of the scene, which emotions does the scene convey?, ...)

Aim of the work

• Interactive framework for new content generation
• Material: set of 90 feature movies

Novelties

• Mixing task focuses on the emotive sphere
• Emotional coherence VS cognitive coherence
Emotional space for movies

- 3D space (*) whose dimensions are related to low level features eligible to have an emotional impact

- The trajectory evolution over time provides a strong characterisation of the multimedia item

Emotional axes

• Consider a movie as a design object (*)

<table>
<thead>
<tr>
<th>“Design” concept</th>
<th>Object Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural</td>
<td>Does it convey <em>passion</em> or <em>reflection</em>?</td>
</tr>
<tr>
<td>Temporal</td>
<td>Turned to the <em>past</em> or projected toward the <em>future</em>?</td>
</tr>
<tr>
<td>Energetic</td>
<td>Is it seductive and <em>energetic</em> or flat and <em>basic</em>?</td>
</tr>
</tbody>
</table>

Our Transposition

<table>
<thead>
<tr>
<th>High level concept</th>
<th>Dichotomic adjectives</th>
<th>Low level feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural</td>
<td>Warm / cold</td>
<td>Illuminant of the scene</td>
</tr>
<tr>
<td>Temporal</td>
<td>Dynamic / slow</td>
<td>Motion dynamics</td>
</tr>
<tr>
<td>Energetic</td>
<td>Energetic / minimal</td>
<td>Sound energy</td>
</tr>
</tbody>
</table>

Emotional space overview

- **低级特征**
  1. 光源
  2. 运动动态
  3. 声音能量

- **二元形容词**
  1. 温暖/寒冷
  2. 动态/缓慢
  3. 能量/简约

- **语义维度**
  1. 自然
  2. 时间
  3. 能量

交互式视频拼接基于情感身份
Natural dimension

- **Illuminant**
  - Used by directors to underline the mood of a scene
  - Suitable for the natural axis ~ warm/cold colours\(^(*)\)
  - Estimated by using a “white patch” algorithm\(^(**)\)

- **Mapped on the natural axis**
  - Non uniform quantization law


• **Motion dynamics**
  
  - Used by directors to stress the identity of a scene
    - ex: speed and dynamism VS calm and tranquility
    - achieved with shot type/length, amount of motion...
  
  - Defined as a combination of
    - shot length (sequences of short shots ~ high narration pace)
    - motion activity (MPEG7 descriptor) (*)

Energetic dimension

- Audio dynamics
  - Scene semantics is usually stressed thanks to audio
    - ex: gentle for romantic moments VS aggressive for actions
    - audio: speech + soundtrack + ...
  - Audio energy
    - Down-sampling at 8 kHz single-channel
    - Average log energy of samples > threshold
Drawing movies & mixing points

Movie trajectory
• Shot $\rightarrow$ point $\bar{S}_i = \{N_i, T_i, E_i\}$
• Trajectory $\rightarrow$ 3D cubic spline

Mashup: candidate selection
• Mixing point candidates
• Minimum distance criteria
  • on shot coordinates
  • on movie trajectories
Mashup framework

Interactive video mashup based on emotional identity
Mixing point candidates

• When the user decides to switch movie, the system looks for a restricted number of shots belonging to other movies

• Minimum distance criteria
  – on shot coordinates (for retrieval)...

\[
d (\bar{S}_i, \bar{S}_j) = \left[ (N_i - N_j)^2 + (T_i - T_j)^2 + (E_i - E_j)^2 \right]^{\frac{1}{2}}
\]

  – …and movie trajectories (for ranking)

\[
D (\xi_A, \xi_B) = \sum_{k \in N_h} d (\bar{S}_{i+k}, \bar{S}_{j+k})
\]
Assisted Mashup (AM)

Composition: quick, intuitive, and suitable for beginners

1. The user selects a movie
2. On a specific user command, the user stops the movie
3. The highest shot in the rank is proposed as the best candidate for the continuation of the mashup
4. In case the user approves it as a suitable mixing point, the playback is resumed from the new inserted shot
5. If not, the next shot from the list is submitted for user approval
Creative Mashup (CM)

More in depth control on the mixing activities

• The user can focus on a particular sub-volume of the emotional space to obtain a product with a very specific emotional identity (e.g., fast rhythm, cold lighting and loud music)

• In the AM module
  – only one shot is proposed at a time

• In the CM
  – the whole pool of browsable candidates for selection is shown at once
Graphical User Interface

Interactive video mashup based on emotional identity
User validation and examples

• 8 users, 90 movies, CM module
• Create a mashup (2-3 mins long) trying to convey a particular mood

  – The Matrix

  – Raiders of the lost ark

  – The war of worlds
Another example

- A beautiful mind
- The Patriot
- All or nothing

Interactive video mashup based on emotional identity
Evaluation

• Questionnaire for rating users’ satisfaction with
  – the quality of the generated mashups
  – the respondent to their own expectations

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The colour atmosphere does not change across the transitions between mashup segments</td>
</tr>
<tr>
<td>2.</td>
<td>The sensation of motion does not change across the transitions between mashup segments</td>
</tr>
<tr>
<td>3.</td>
<td>There is no brisk change in audio between different mashup segments</td>
</tr>
<tr>
<td>4.</td>
<td>The general mood of the whole mashup is coherent</td>
</tr>
<tr>
<td>5.</td>
<td>Rate the whole experience of creating the mashup</td>
</tr>
<tr>
<td>6.</td>
<td>Using such a tool stimulated my creativity</td>
</tr>
<tr>
<td>7.</td>
<td>The final product responds to my expectations</td>
</tr>
</tbody>
</table>

![Bar chart showing average marks for each statement]
Conclusions

• Interactive mashup framework
  – editing and remixing pre-existing material from movies

• The system provides a selection of emotionally similar scenes
  – Continuity of the evoked mood

• All users judged the whole experience
  – As a positive recreational activity
  – Stimulating for their creativity
Thank you