Multimedia Event Detection

- There is a need for automatic tagging, indexing and searching of multimedia data, esp. automatic event detection in videos.
- Application: searching for a specific video content such as how to make a cake, how to repair an appliance, etc.

Challenges of Multimedia Event Detection
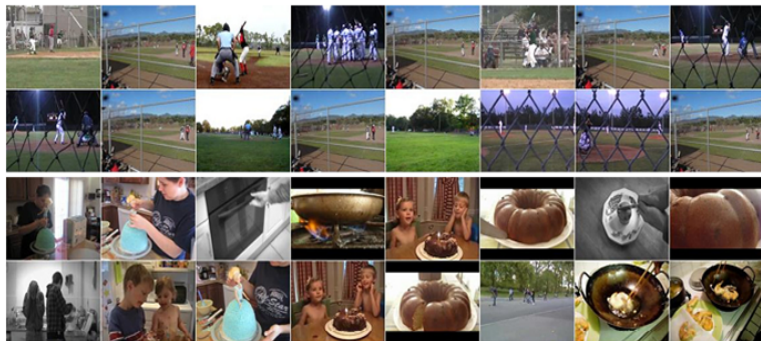

Making a cake   *Batting a run in*   *Assembling a shelter*

- Large content variation
  - ▶ Large number of events
  - ▶ Large number of background videos.

Challenges of Multimedia Event Detection

- Uncontrolled capturing conditions
  - Different time, location,
  - Clutter in the environment, camera motion, etc.

Challenges of Multimedia Event Detection

- Event name: Making a cake
- Event definition: One or more people make a cake
- Event description:
  - *scene*: indoors, typically a kitchen in a home, restaurant or other setting
  - *objects/people*: ingredients, bowls, spoons, mixers, etc.
  - *activities*: selecting ingredients, combining ingredients, pouring batter into pan, putting cake into oven,
- Few video examplars

- Searching for ad hoc events
  - Limited number of video examples
  - Textual descriptions are readily understood by humans, difficult to encode for machine learning approaches