



# University of HUDDERSFIELD

## University of Huddersfield Repository

Wang, Jing, Xu, Zhijie, Pickering, Jonathan and Mishra, Rakesh

Spatio-Temporal Volume Based Video Event Detection

### Original Citation

Wang, Jing, Xu, Zhijie, Pickering, Jonathan and Mishra, Rakesh (2009) Spatio-Temporal Volume Based Video Event Detection. In: University of Huddersfield Research Festival, 23rd March - 2nd April 2009, University of Huddersfield. (Unpublished)

This version is available at <http://eprints.hud.ac.uk/id/eprint/5234/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>



# Spatio-Temporal Volume Based Video Event Detection

Jing Wang, Zhijie Xu, Jonathan Pickering and Rakesh Mishra  
School of Computing and Engineering, University of Huddersfield  
Queensgate, Huddersfield HD1 3DH, UK

**AIMS OF THIS PROJECT**  
The aim of this project is to develop techniques for the automatic classification of events in a video using efficient volumetric video structures. Based on developing 3D volume processing and pattern analysis methods such as filter, segmentation, edge/surface detection, morphology and template mapping, the technique can automate recognition of specific events in a video which can be applied in video indexing and searching, surveillance, and human-computer interactions.

**EVENT DETECTION IN A VIDEO**  
A video event is something that happens in a given space during a period of time in a video, both the spatial and temporal signals can be continuous or discrete. Video event detection is the task of identifying specific events in videos. The process relies on some pre-defined event templates as models to compare with the input video streams. The goal of event detection is to identify and localize the closest matches depending on certain criteria.

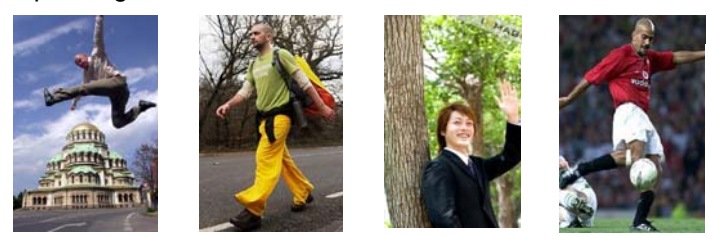


Figure 1 Sample events such as jumping, walking, waving and kicking

**SPATIO TEMPORAL VOLUME (STV)**  
STV is a 3D volume structure which extends the traditional video frames by adding a time-axis. The frames are organized as slices in a volume data according to the time order of the video.

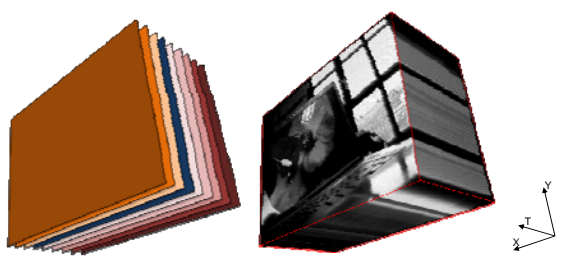
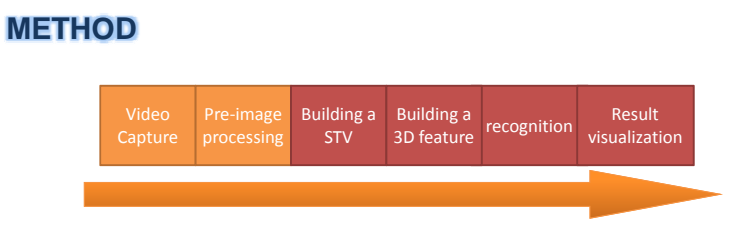


Figure 2 The STV structure. The left one is the framework of the STV and the right one is a real 3D video STV model. Notice the bar-like texture, which is shaped by the static camera and invariant frame information.

**STV-BASED EVENT DETECTION IN A VIDEO**  
Inherited from existing image processing techniques, the 3D STV processing methods devised in this research so far are mainly based on extending 2D image processing approaches, such as filtering and morphological processing into a 3D space.



**STEP1. CONSTRUCTING 3D STV TEMPLATES**

For example, the left 3D model can be used for event analysis or template matching in order to detect the waving action in a video.

**STEP2. PATTERN ANALYSIS AND RECOGNITION**  
The additional time domain generates new space-time volume entities that do not exist in the spatial domain only. Based on the 3D volume analysis methods, features such as a space-time "stick," "plate," and "ball" denotes different events in a video.

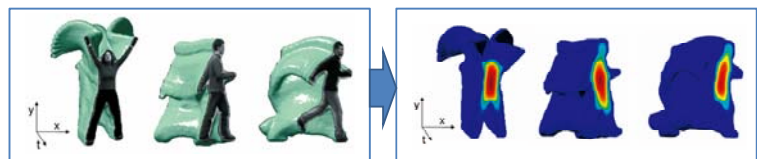


Figure 3 One sophisticated analytical measure of the STV is by adopting the Poisson equation. The figures at right is the Poisson distance of the STV events shown at left.

**CONCLUSIONS AND FUTURE WORK**  
It is envisaged that event detections can be benefited by the existing and matured 2D pattern analysis and machine intelligence researches and in turn to facilitate 3D data processing and feature/pattern analysis.

Next phase will see the completion of the literature review and prototype design in this research. At PhD stage, a library of event templates from different video footages will be formed with theoretical research into the innovation of 3D pattern analysis techniques based on the STV models.

switch on to

