bradscholars

Semantic content analysis for effective video segmentation, summarisation and retrieval.

Item Type	Thesis
Authors	Ren, Jinchang
Rights	 The University of Bradford theses are licenced under a Creative Commons Licence.
Download date	2025-06-13 10:17:13
Link to Item	http://hdl.handle.net/10454/4251

SEMANTIC CONTENT ANALYSIS FOR EFFECTIVE VIDEO SEGMENTATION, SUMMARISATION AND RETRIEVAL

PhD

Jinchang Ren (UB06027594)

Supervisors: Prof. Jianmin Jiang and Dr. Stan S. Ipson

ABSTRACT

This thesis focuses on four main research themes namely shot boundary detection, fast frame alignment, activity-driven video summarisation, and highlights based video annotation and retrieval. A number of novel algorithms have been proposed to address these issues, which can be highlighted as follows.

Firstly, accurate and robust shot boundary detection is achieved through modelling of cuts into sub-categories and appearance based modelling of several gradual transitions, along with some novel features extracted from compressed video. Secondly, fast and robust frame alignment is achieved via the proposed subspace phase correlation (SPC) and an improved sub-pixel strategy. The SPC is proved to be insensitive to zero-mean-noise, and its gradient-based extension is even robust to non-zero-mean noise and can be used to deal with non-overlapped regions for robust image registration. Thirdly, hierarchical modelling of rush videos using formal language techniques is proposed, which can guide the modelling and removal of several kinds of junk frames as well as adaptive clustering of retakes. With an extracted activity level measurement, shot and sub-shot are detected for content-adaptive video summarisation. Fourthly, highlights based video annotation and

retrieval is achieved, in which statistical modelling of skin pixel colours, knowledge-based shot detection, and improved determination of camera motion patterns are employed.

Within these proposed techniques, one important principle is to integrate various kinds of feature evidence and to incorporate prior knowledge in modelling the given problems. High-level hierarchical representation is extracted from the original linear structure for effective management and content-based retrieval of video data. As most of the work is implemented in the compressed domain, one additional benefit is the achieved high efficiency, which will be useful for many online applications.

KEYWORDS:

Semantic content analysis, shot boundary detection, video segmentation, subspace phase correlation, frame alignment, video summarisation, hierarchical modelling, adaptive clustering, content-based retrieval, automatic annotation, TRECVID, digital video processing.