Spatio-temporal pre-processing methods for region-of-interest video coding

Linda S. Karlsson

Department of Information Technology and Media
Mid Sweden University

Licenciate Thesis No. 21
Sundsvall, Sweden
2007

Mittuniversitetet
Informationsteknologi och medier
SE-851 70 Sundsvall
SWEDEN

ISBN 978-91-85317-45-5
ISNN 1652-8948

Akademisk avhandling som med tillstånd av Mittuniversitetet framlägges till offentlig granskning för avläggande av teknologie licenciatexamen fredagen den 27 april 2007, 13.00 - 15.00, i L111, Mittuniversitetet, Holmgatan 10, Sundsvall.

Tryck: Tryckeriet Mittuniversitetet

Abstract

In video transmission at low bit rates the challenge is to compress the video with a minimal reduction of the percieved quality. The compression can be adapted to knowledge of which regions in the video sequence are of most interest to the viewer. Region of interest (ROI) video coding uses this information to control the allocation of bits to the background and the ROI. The aim is to increase the quality in the ROI at the expense of the quality in the background. In order for this to occur the typical content of an ROI for a particular application is firstly determined and the actual detection is performed based on this information. The allocation of bits can then be controlled based on the result of the detection.

In this licenciate thesis existing methods to control bit allocation in ROI video coding are investigated. In particular pre-processing methods that are applied independently of the codec or standard. This makes it possible to apply the method directly to the video sequence without modifications to the codec. Three filters are proposed in this thesis based on previous approaches. The spatial filter that only modifies the background within a single frame and the temporal filter that uses information from the previous frame. These two filters are also combined into a spatio-temporal filter. The abilities of these filters to reduce the number of bits necessary to encode the background and to successfully re-allocate these to the ROI are investigated. In addition the computational compexities of the algorithms are analysed.

The theoretical analysis is verified by quantitative tests. These include measuring the quality using both the PSNR of the ROI and the border of the background, as well as subjective tests with human test subjects and an analysis of motion vector statistics. The qualitative analysis shows that the spatio-temporal filter has a better coding efficiency than the other filters and it successfully re-allocates the bits from the foreground to the background. The spatio-temporal filter gives an improvement in $PSNR_{ROI,Avg}$ of more than 1.32 dB or a reduction in bitrate of 31 % compared to the encoding of the original sequence. This result is similar to or slightly better than the spatial filter. However, the spatio-temporal filter has a better performance, since its computational complexity is lower than that of the spatial filter.