原 H.264 无参考视频质量评价方法 （使用了基于遗传编程方法的符号回归）

2013年10月21日 11:31:10 阅读数：4869

Nicolas Staelens 等人在《Constructing a No-Reference H.264/AVC Bitstream-based Video Quality Metric using Genetic Programming-based Symbolic Regression》论文中研究了H.264的视频质量评价方法。这篇论文我感觉真的是把无参考视频质量评价做到了很高的水平，很有必要记录一下其中的关键信息。

注：并不是特别了解基于遗传编程方法的符号回归，在此就不多讲述这方面的了。

文章首先回顾了一下客观视频质量评价算法：

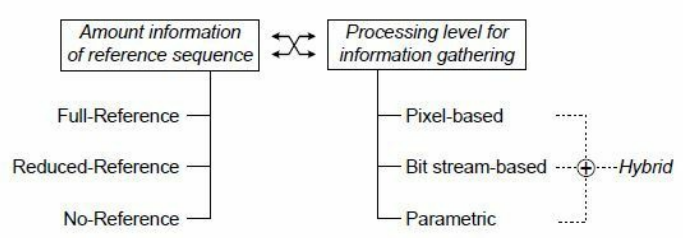

Fig. 1. Different categories of video quality metrics based on the amount of information which is used from the reference sequence or based on the processing level for extracting information in order to model perceived quality.

TABLE I
PARETOGP EXPERIMENTAL SETTINGS

Setting	Values
# replicates	5
# generations	310
population size	1000
archive size	100
crossover rate	0.9
subtree mutation rate	0.1
population tournament	5

选择的8个测试序列如下表所示。分别标明了来源以及描述。

TABLE II
CHARACTERISTICS OF THE EIGHT SELECTED TEST SEQUENCES.

Sequence	Source	Description
basketball	CDVL	Basketball game with score. Camera pans and zooms to follow the action.
BBB*	Big Buck Bunny	Computer-Generated Imagery. Close-up of a big rabbit. Slight camera pan while following a butterfly in front to the rabbit.
cheetah	CDVL	Cheetah walking in front of a chainlink fence. Camera pans to follow the cheetah.
ED*	Elephants Dream	Computer-Generated Imagery. Fixed camera focusing on two characters. Motion in the background.
foxbird3e	CDVL	Cartoon. Fox running towards a tree and falling in a hole. Fast camera pan with zoom.
purple4e	CDVL	Spinning purple collage of objects. Many small objects moving in a circular pattern.
rush hour	TUM	Rush hour in Munich city. Many cars moving slowly, high depth of focus. Fixed camera.
SSTB*	Sita Sings the Blues	Cartoon. Close-up of two characters talking. Slight camera zoom in.

8个测试序列的内容如下图所示。

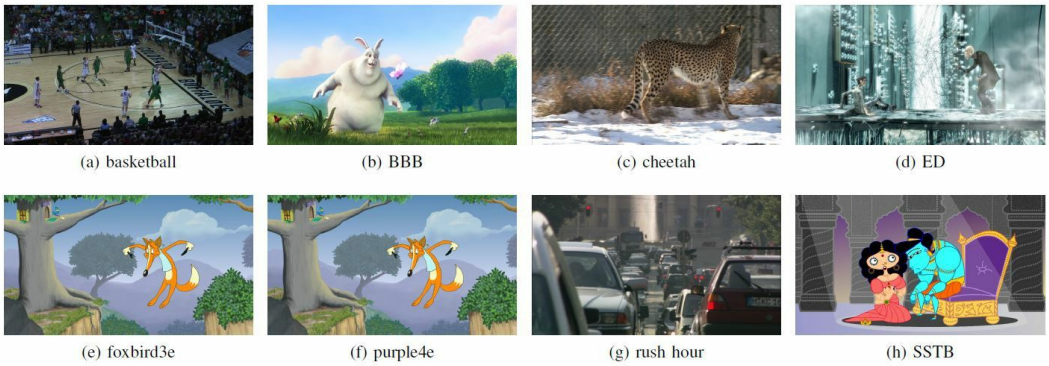

Fig. 2. Overview of the eight selected source video sequences, taken from open source movies, CDVL and TUM.

计算了8个测试序列的SI（空间复杂度）和TI（时间复杂度），并以散点图的形式画成如下图所示的图表。

注：有关SI（空间复杂度）和TI（时间复杂度）可以参考： 衡量视频序列特性的TI（时间信息）和SI（空间信息）

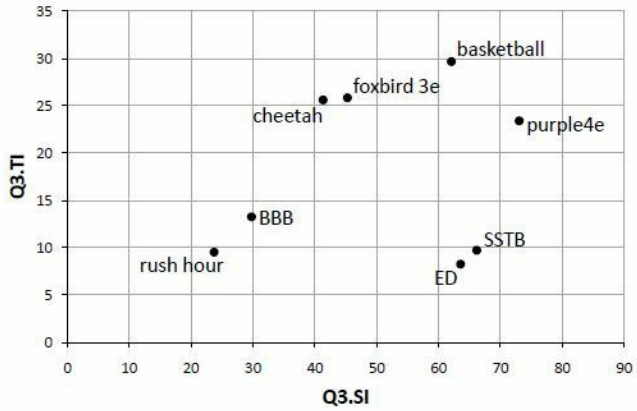

Fig. 3. Calculated Q3.SI and Q3.TI values for each sequence [41].

视频编码选项设定如下：

- Number of slices: 1, 4 and 8
- Number of B-pictures: 0, 1 and 2
- GOP size [42]: 15 (0 or 1 B-picture) or 16 (2 B-pictures)
- Closed GOP structure
- Bit rate: 15 Mbps

模拟丢包的时候，使用了名为nalu-drop classifier的工具。

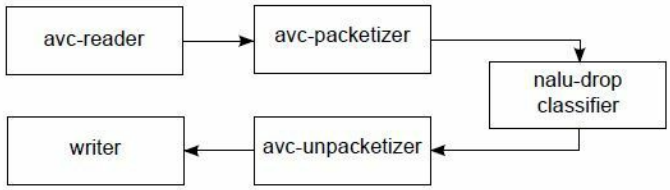

Fig. 4. RTP packets, which carry data from particular slices, are dropped using the *nalu-drop classifier* component. After unpacketizing, the resulting impaired sequence is saved to a new file.

本文打算从下列参数中选择可以用于建立视频质量评价模型的参数，备选参数数量真是大的惊人啊。

TABLE III

OVERVIEW OF PARAMETERS EXTRACTED FROM RECEIVED VIDEO BITSTREAM IN ORDER TO IDENTIFY LOCATION OF LOSS AND TO CHARACTERIZE VIDEO CONTENT. THE SUBSET OF 8 INFLUENTIAL PARAMETERS IS MARKED IN BOLD.

Parameter	Description
Encoder settings	
B-pictures, **slices**, GOP	Number of B-pictures, slices per picture and GOP size as specified during encoding.
Loss location and severity	
i_loss, **p_loss**, b_loss	Indication (1 or 0) whether the loss originates from an I-, P- or B-picture.
perc_pic_lost	Percentage of slices lost of the picture where the loss originates.
imp_in_gop_pos, imp_in_pic_pos	Temporal location within the GOP (begin, middle, end) and spatial location within the picture (top, bottom, middle) of the first lost slice.
imp_in_gop_idx, imp_in_pic_idx	Absolute position within the GOP and within the picture of the first lost slice.
imp_cons_slice_drops, imp_cons_b_slice_drops, imp_pic_drops	Number of consecutive slice drops, number of consecutive B-slice drops and number of entire picture drops.
drift	Temporal duration of the loss.
Video content characteristics	
perc_pb_4x4, perc_pb_8x8, perc_pb_16x16, perc_pb_8x16, perc_pb_16x8, perc_i_4x4, **perc_i_8x8**, perc_i_16x16	Percentage of I, P & B macroblocks of type 4x4, 8x8, 16x16, 8x16 and 16x8, averaged over the pictures in the GOP containing the loss.
perc_i_mb, perc_skip, perc_ipcm	Percentage of macroblocks encoded as I, skip and PCM, averaged over the pictures in the GOP containing the loss.
I_perc_4x4, **I_perc_8x8**, I_perc_16x16	Percentage of macroblocks of type 4x4, 8x8 and 16x16 in the first I or IDR picture of the GOP containing the loss.
abs_avg_coeff, avg_qp	Absolute average value of the macroblock coefficients and QP value, averaged over the P or B pictures in the GOP containing the loss.
I_abs_avg_coeff, I_avg_qp	Absolute average value of the macroblock coefficients and QP value in the first I or IDR picture of the GOP containing the loss.
perc_zero_coeff, I_perc_zero_coeff	Percentage of zero coefficients, averaged over the P or B pictures in the GOP containing the loss and average of zero coefficients in the first I or IDR picture of the GOP containing the loss.
avg_mv_x, avg_mv_y, stdev_mv_x, stdev_mv_y	Average absolute motion vector length and standard deviation in x- and y-direction, averaged over the P or B pictures in the GOP containing the loss. Motion vector magnitudes have quarter pixel precision.
avg_mv_xy, stdev_mv_xy	Average and standard deviation of the sum of the motion vector magnitudes in x- and y-direction, averaged over the P or B pictures in the GOP containing the loss. Motion vector magnitudes have quarter pixel precision.
perc_zero_mv	Average percentage of zero motion vectors, calculated over the P or B pictures in the GOP containing the loss.

经过计算后，得出了每个变量在预测视频质量这方面做出的贡献，如下图所示。

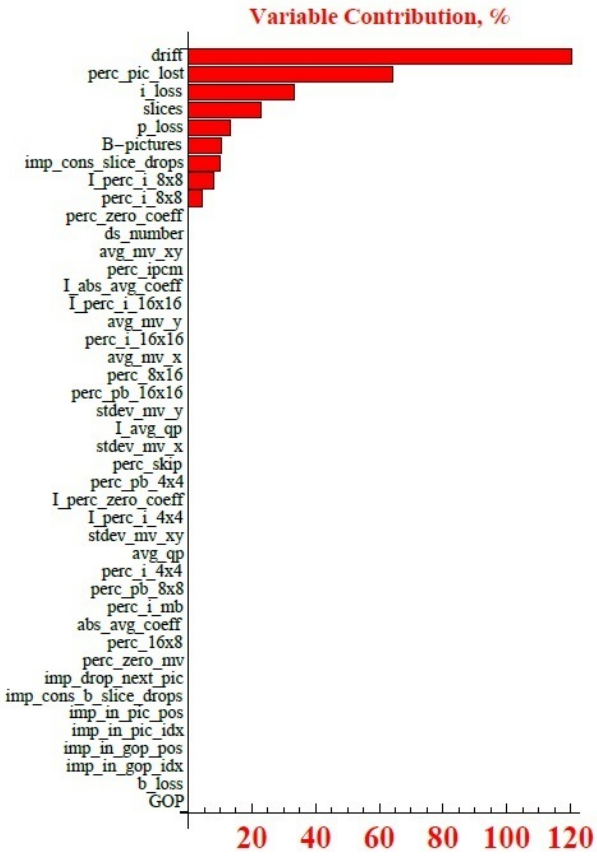

Fig. 8. Contribution of each variable into the prediction error of the regression models when removing that particular variable from the model.

最终选定了8个参数：perc_pic_lost, i_loss, slices, p_loss, B_pictures, imp_cons_slice_drops, I_perc_8x8 and perc_i_8x8。

有一些不明白的地方，先不多说了，看一看最终建立的模型，以树的形式显示如下图。

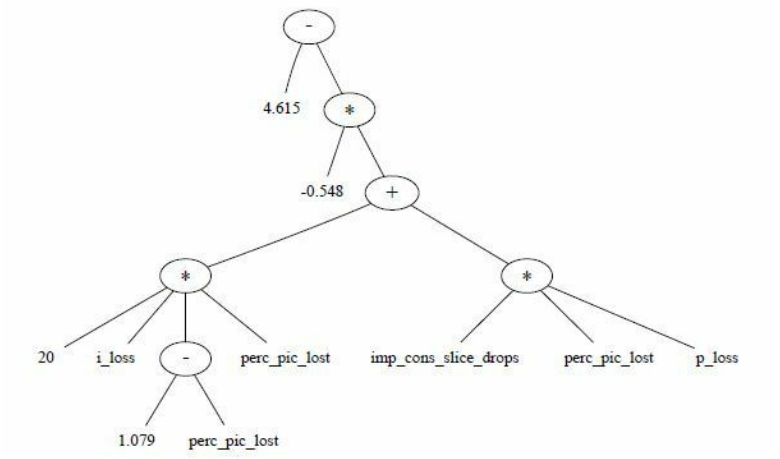

Fig. 12. Parse tree corresponding with the selected GP model indicated in Figure 10.

上图可以写成如下公式：

$$MOS_p = 4.615 - 0.548 \cdot (20 \cdot i\_loss \cdot (1.079 - perc\_pic\_lost) \cdot perc\_pic\_lost + imp\_cons\_slice\_drops \cdot perc\_pic\_lost \cdot p\_loss) \quad (9)$$

对此模型进行验证的结果如下表所示。作为对比，引入了两种视频质量评价算法：PSNR和VQM。非常令人震惊的是，该模型的性能竟然比这两种算法都要好。

注：PSNR介绍： http://blog.csdn.net/leixiaohua1020/article/details/11729289

VQM介绍： http://blog.csdn.net/leixiaohua1020/article/details/12685297

TABLE VI
PERFORMANCE EVALUATION OF OUR PROPOSED METRIC, PSNR AND VQM AGAINST THE EPFL-POLIMI VIDEO DATABASE.

	PLCC	SROCC	Pred. error
GP metric	0.8816	0.8830	0.2227
PSNR	0.7374	0.7463	0.4562
VQM	0.8127	0.8344	0.3395

文章标签： H.264 无参考 视频 质量评价

个人分类： 视频质量评价 视频编码

所属专栏： 视频质量评价