Partnership for International Research and Education
A Global Living Laboratory for Cyberinfrastructure Application Enablement

Project Title:
Data Mining for Video Encoding (low complexity H.264)

Student: Paula A. Carrillo, Master student, FAU
FIU/FAU Advisor: Dr. Hari Kalva, FAU
PIRE International Partner Advisor: *Dr. Tao Pin, Tsinghua University*

National Science Foundation

I. Research Overview and Outcome

OVERVIEW

- H.264 standard video encoder offers great compression with high quality but at the expense of high computational complexity.
- The proposed method uses data mining for exploiting the structural similarity in video in order to make optimal mode decisions instead of doing the exhaustive mode evaluation.
- Research has been done on low complexity H.264. However, use of data mining approach has not been explored.
- The proposed solution was implemented in the Intel® H.264 IPP video encoder.

PROPOSE SOLUTION

- Machine learning using J.48 WEKA classifier.
- Inter frame mode decision speed up based on trained trees.
- Different metrics and decision trees were implemented and evaluated.

PROCESS OF TRAINNING AND IMPLEMENTATION

The process of obtaining data for training is done offline. In this supervised learning approach, we used the data of the first four frames of the flower.yuv video in CIF format for training. Full search was the search algorithm used in order to get the motion vectors (MV) and residual metrics are calculated based on the MVs. The residual and current macro block metrics and the macro block mode selected by standard Intel® IPP H.264, are saved in an ARRF file. Next, using WEKA tool, mode decision trees are discovered through J48 classifier algorithm. Then, these trees are implemented as if-else statements in the Intel® H.264 encoder. The trees replace the original complex Inter mode decision.

TEST ENVIRONMENT

- Processor: Intel® Core™ 2 CPU 6600 @ 2.4 GHz
- Videos in Cif (352x288) format, 240 frames, IPPP, GOP (Group of Pictures)=100
- Bitrate range: 100Kbps – 2.5Mbps
- Search algorithm: Full search, range of search = 15.
- The proposed solution was tested against Full Search with Split =0, Full Search with Split =1 and Full Search with Split =2.
- Split=0 Implies evaluation only of Inter mode 16x16, Split=1 Implies evaluation of Inter mode 16x16, 8x16, 16x8 and 8x8, Split=2 Implies evaluation of Inter 16x16, 8x16, 16x8, 8x8 and sub-modes (8x4, 4x8 and 4x4)

OUTCOME

Our method classifies Inter modes between {16x16, 16x8, 8x16, and 8x8} macro block shapes. Split 1 option has the same Inter mode classification. Therefore, we primarily compare against Split 1

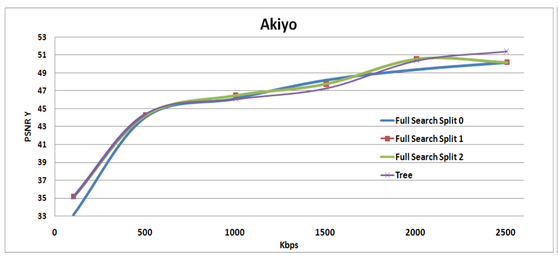
- Encoded speed up comparing our method with the standard Intel® Full Search Split 1 method varies between 42.33% and 228.64% !!!!
- Average increase in PSNR and Bitrate comparing our method with Full Search Split 1 are negligible. The PSNR difference compared to the full complexity mode is between -0.05 (better PSNR!!) and 0.57. In Bitrate the range is between -0.35% (reduced Bitrate!!!) and 1.41%.

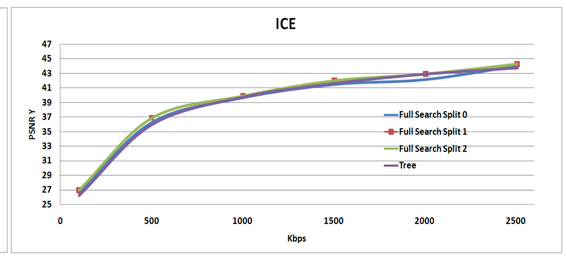
Notes: These ranges depend primarily on the video context, texture and format.

TABLES AND GRAPHICS RESULTS

Our solution was tested on Akiyo, Ice, Harbour and Football YUV sequences.

Video	% Increase FPS comparing with Full Search Split 0	% Increase BitRate comparing with Full Search Split 0	Δ decrease PSNR comparing with Full Search Split 0	% Increase FPS comparing with Full Search Split 1	% Increase BitRate comparing with Full Search Split 1	Δ decrease PSNR comparing with Full Search Split 1	% Increase FPS comparing with Full Search Split 2	% Increase BitRate comparing with Full Search Split 2	Δ decrease PSNR comparing with Full Search Split 2	% Speed up Encoder time comparing with Full Search Split 1	% Speed up Encoder time comparing with Full Search Split 2
Akiyo	149.59	0.00	-0.61	228.64	-0.35	-0.05	260.69	-0.35	-0.05	69.40%	90.45%
Ice	34.94	-0.01	0.00	76.12	-0.01	0.43	455.16	-0.01	0.43	42.92%	81.83%
Harbour	3.55	-0.01	0.20	34.03	0.00	0.22	81.55	0.00	0.22	23.18%	75.37%
Football	10.37	1.28	0.30	42.33	1.41	0.57	71.29	1.41	0.57	29.31%	76.21%





Note: FPS - Frames Per Second; PSNR - Peak Signal Noise Ratio

In Akiyo graphic we can see that the curve "Tree" has a better performance in low bitrates than the others.
In Ice graphic the "Tree" curve behavior is extremely close that the rest of the curves.

Machine learning method shows a close quality performance comparing with the standard Intel® IPP H.264 encoder but with significant time saving

FUTURE WORK

- It is necessary to explore a more adaptive mode decision depending on the video context and format using multiple pre-trained trees which switch between trees according with movie changes. For example scenes changes, increase/decrease of visual details, light and/or movement)
- On-line training is a computational expensive solution. However, learning online could generate more accurate trees decisions and it is an interesting area to explore.

II. International Experience

TSINGHUA UNIVERSITY

Tshighua University is the perfect research environment.

This project was the result of a close collaboration between the Human Computer Interaction and Multimedia group at Tshinghua university and MLAB (Multimedia Laboratory) at FAU. The collaboration was not only in the form of access to the latest technology/servers used for our simulations and tests but also, through exchange of ideas in regular group meetings. The feedback of the group helped us to explore more options and work on our weaknesses during the whole process.
During my experience in Tsinghua university my research advanced faster, It was the pedal to the metal !!

BEIJING, CHINA

It is just a beautiful city with too much history and future and too much places for sightseeing and enjoying

III. Acknowledgement

The material presented in this poster is based upon the work supported by the National Science Foundation under Grant No. OISE-0730065. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.