Technical Report 2001-10-31
Internetworking and Media Communications Research Laboratories
Department of Math & Computer Science, Kent State University
http://medianet.kent.edu/technicalreports.html

SYSTEM OVERVIEW OF MEDIANET BIRD-OF-FLOCK TRACKING TRANSCODER

Zhong Guo and Javed I. Khan

Technical Report 2001-10-31

Internetworking and Media Communications Research Laboratories
Department of Math & Computer Science
Kent State University, 233 MSB, Kent, OH 44242
(Revised October 31, 2001)
javed@kent.edu

1 Object-based Video Transcoding

Object based encoding plays an increasingly important role for carrying perceptually pleasant video at lower rates. In this research we are studying how perhaps perceptual object based coding concept can be incorporated into transcoding for extreme rate transformation required in splicing asymmetric segments of the Internet. Object based video encoding is an open problem. It is much more difficult to detect an object than to compute a wavelet or DCT co-efficient. Indeed MPEG-4 objects are currently thought to have best chance for computer model generated synthetic video where objects do not need to be detected or in limited domain (such as head and shoulder) small format video. Even in these special situations object detection algorithms are generally quite computation intensive. Consequently, these techniques have been pursued primarily for first stage video encoding. Because most first stage encoding scenarios (except for live video) allow for off-line processing.

In this context, the transcoding scenario has several additional challenges from conventional first stage encoding. First, the original frame images are no longer explicitly available. The transcoder receives an encoded video stream. Secondly, the object detection has to be performed extremely fast at the rate of the stream. Thirdly, it already contains some refined information (such as motion vectors). Consequently, techniques for transcoding can generally be used in first stage encoding but the reverse in not always possible.

In this research we have developed a low computation based real-time object tracking algorithm suitable for full motion focus region based perceptual video recompression called **Bird-of-Flock Motion Tracking Object Transcoder** (BOF Transcoder). BOF Transcoder technique accepts some logical and high-level initial description of the video objects in terms of initial position, and shape. It then automatically tracks the region covered by this object for subsequent perceptual encoding based on real-time motion analysis. We have restricted the problem that no pixel level decoding of DCT or image components is allowed to conform to the constraints of the transcoding scenario.

This transcoder has been developed based on the self-organizing Network adaptive transcoder (SONET) system developed at MEDIANET Lab distribution, which has additional features of motion-vector bypass.

In this document we present the system description of this concept transcoder that we have recently implemented using this tracking algorithm. This distribution contains the second version of KENT MEDIANET LAB implementation of an ISO-/IEC 13818-2 stream subnet transcoding system. The system can perform dynamic video rate adaptation, and perceptual bit-allocation.

The work has been funded by DARPA Research Grant F30602-99-1-0515 under it's Active Network initiative.

2 How to install BOF Transcoder?

You can download latest distribution release from http://www.medianet.kent.edu. After download, use gunzip command to extracts the distribution zip file. This should give a tar file. Untaring the tar file will generate all components, each in a separate directory below a directory named ‘demo’. Demo will have subdirectories named svr, xdec, xenc, player, and test. To generate the executable, run make in the top demo directory. It should generate all the required executables.

These system has been tested on Redhat Linux 6.2 and 7.1, but basically you can run it on any UNIX like operating systems.

3 How to test run the BOF Transcoder?

Before we can start the transcoding we need to provide the network configuration to the system via the file “net.par” in the demo directory. You can run the server, transcoder and player from three terminal windows on one machine, or each on a different machine of a LAN. You must provide the IP addresses and port numbers for svr and xcoder in “net.par”, and the two port numbers can not be same.

To play the video, you need to use serverrun, xcoderrun and playerrun commands in server, transcoder and player windows, respectively. Video will be shown on the 32bit MPEG-2 player.

4 How to serve your own video?

In svr directory, you can find a sample MPEG2 video stream file “svr.m2v”. It can easily be replaced by any other MPEG-2 video stream. The default encoding sequence GOP size N = 15 and M = 3. If a video stream with different sequence is used the “test.par” file should be modified in the test subdirectory.

5 Transcoder bit-rate control

The overall outgoing bit-rate of the transcoder can be adjusted by specify the number of bits/second in the “rate.par” file in the test subdirectory.

6 Motion vector bypass control

Motion vector bypass can be enabled by modify the content of “bypass.par” file in the test subdirectory to “1 1”. You can disable it using “1 0”. With bypass, encoder will use the motion vectors from the decoding process directly, which results in fast encoding but sometimes less quality.

7 How to track objects?

All the parameters for tracking objects are in “track.par” file in the test subdirectory. Currently, up to two objects can be tracked simultaneously. Whenever an object is being tracked, a window will be displayed by the player to show the macroblocks under tracking. Green window corresponds to object one and red to object two.

The “track.par” file is read by the transcoder for every frame. So one can make modification any time during execution. All parameter values in this file should be integers.

The first 8 parameters in the list are related to object one and the following 8 parameters repeat everything for object two. One can start tracking an object by specifying its location in the frame as a rectangle area. The TL_X, TL_Y, BR_X and BR_Y are integers from 0 to 100. They specify the coordinates of top-left and bottom-right corners of that rectangle area using percentage relative to the picture size with 0 percent being the top-left corner of the picture. One can also start tracking by giving the BIRTH_SIZE and BIRTH_SPEED values. When an object with size greater than BIRTH_SIZE macroblocks and speed greater that BIRTH_SPEED pixels/frame is found, the system will start tracking it. But the first way of starting tracking take high priority. The second way of starting tracking will be suppressed once you specified an rectangle area. You have to set all four coordinates to 0 if you want to use the second way. When an object being tracked has a size less that or equal to DEATH_SIZE macroblocks, or has a speed less than or equal to DEATH_SPEED pixels/second, system will stop tracking it.

8 Flock-of-Bird Parameters

All other parameters apply to both objects unless specified by its name:

DEACT_MV_THRESHOLD: 0 ~ 100 percent, if the motion value of a macroblock is greater than this percentage of the object motion value, then this macroblock is considered not following the object movement.

DEACT_TIME_THRESHOLD: if a macroblock has been not following the object movement for consecutively more than this number of frames, it will be discarded from this object.

DEACT_MB_NUM_LIMIT: the maximum percentage of macroblocks of an object that can be discarded in one frame.

ACT_MV_THRESHOLD: 0 ~ 100 percent, if the motion value of a macroblock is less than this percentage of the object motion value, then this macroblock is considered to be following the object movement.

ACT_TIME_THRESHOLD: if a macroblock has been following the object movement for consecutively more than this number of frames, it will be taken into this object.

MONITOR_SET_WIDTH: macroblocks within this number of macroblocks away from an object will be monitored.

DISPLAY_MARGIN: the number of macroblocks that the tracking window will extend out of the object.

The last 3 parameters set the quality of corresponding areas in the output video. There values range from –14 to 14, with –14 indicating worst quality and 14 best quality.

9 References

[1] Bhattacharjee, S., Kenneth L. Calvert and Ellen W. Zegura. An Architecture for Active Networking. High Performance Networking' 97 ,White Plains, NY, April 1997. [also available at http://www.cc.gatech.edu/projects/canes/papers/anarch.ps.gz, October 98]

[2] Haskell B. G., Atul Puri and Arun Netravali, Digital Video: An Introduction to MPEG-2, Chapman and Hall, NY, 1997.

[3] Information Technology- Generic Coding of Moving Pictures and Associated Audio Information: Video, ISO/IEC International Standard 13818-2, June 1996.

[4] Keesman, Gertjan; Hellinghuizen, Robert; Hoeksema, Fokke; Heideman, Geert, Transcoding of MPEG bitstreams Signal Processing: Image Communication, Volume: 8, Issue: 6, pp. 481-500, September 1996,

[5] Tennenhouse, D. L., J. Smith, D. Sincoskie, D. Wetherall & G. Minden., "A Survey of Active Network Research", IEEE Communications Magazine, Vol. 35, No. 1, Jan 97, pp 80-86

[6] Wetherall, Guttag, Tennenhouse, "ANTS: A Tool kit for Building and Dynamically Deploying Network Protocols", IEEE OPENARCH' 98, San Francisco, April 1998. Available at: http://www.tns.lcs.mit.edu/publications/openarch98.html

[7] Youn, J, M.T. Sun, and J. Xin, "Video Transcoder Architectures for Bit Rate Scaling of H.263 Bit Streams," will beappeared to 'ACM Multimedia 1999', Nov., 1999. pp243-250.