A Solution of MV Translation for Interlaced MPEG2 to Progressive H.264 Video Transcoding

B. Pjetar, S. Očovaj, Ž. Lukač

Abstract — Translation of motion vectors for transcoding purposes is considered. MPEG2 interlaced video sequence is transcoded into progressive H.264 sequence. Besides converting MVs form half-pixel to quarter-pixel format, it is necessary to do additional modifications in order to adjust originally field prediction MVs to frame-based prediction.

Keywords — digital television, H.264, MPEG2, video transcoding

I. Introduction

MPEG2 video and H.264 are video compression standards. Video transcoding is the conversion of one video sequence to another. In this paper, conversion of MPEG2 interlaced video sequence into progressive H.264 video sequence is considered.

In progressive scan, video picture is captured or displayed line by line, whereas in interlaced scan, picture is divided in two fields – one field containing only even lines of the picture and the other field containing the odd lines.

The performed transcoding process can be described as follows. Original sequence is interlaced. It is encoded in two ways: as frame coded MPEG2 with frame and field predictions and as field coded MPEG2. The sequence is encoded for several bit rates: 2, 4 and 6 Mb/s. Relevant encoding parameters, such as MVs (motion vectors) and quantizer scale, are saved on MB (macroblock) level during complete sequence decoding. Decoded sequence is deinterlaced and encoded as H.264 using information obtained during MPEG2 decoding, i.e. side information on MB level is translated for H.264 encoding. Deinterlaced sequence is encoded as H.264 stream using only MB frame coding. The block diagram of the whole process is given in figure 1.

Work is partially supported by Ministry for science and environmental protection of Republic of Serbia in project TR-6136B.

B. Pjetar, MicronasNIT, Institut za informacione tehnologije, Novi Sad, Srbija (telefon +38-(0)21-4801-128, Fax +381-(0)21-450-721), email: branislav.pjetar@micronas.com).

S. Očovaj, Fakultet tehničkih nauka, Novi Sad, Srbija, (e-mail: stanislav.ocovaj@micronas.com).

Ž. Lukač, MicronasNIT, Institut za informacione tehnologije, Novi Sad, Srbija (e-mail: zeljko.lukac@micronas.com).

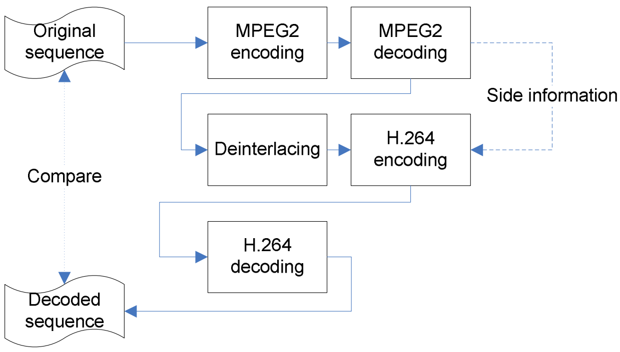

Fig. 1: Block diagram of transcoding flow

Quantizer scale of a MPEG2 MB is directly translated into corresponding H.264 QP (quantization parameter) by means of a table. The main object of this work was to find out how to translate MVs.

II. Implementation

A. Deinterlacing

A new frame is made of every original field. To retain the number of frames that was present in original sequence, every second frame is discarded. If the original sequence was top field first (TFF), only the frames containing lines of original top fields are preserved. For interlaced sequences that had bottom field first (BFF), the frames containing lines of original bottom fields are preserved.

B. Interlaced field to progressive sequence transcoding

Inter-prediction decision process of MBs belonging to a H.264 coded picture may choose between several previously decoded reference pictures for every prediction direction [1]. In MPEG2 standard, current field can use two previous reference fields for motion prediction [2]. B fields always use fields belonging to the same picture as reference, but for P frames, the situation is different. Under the assumption that sequence was coded as TFF, top field can use two fields belonging to the same picture as reference, whereas bottom field can use one field belonging to some of previously coded pictures, but it can also use the field that belongs to the same picture as current field. For transcoding purposes, only those MPEG2 MVs that were used by preserved fields (fields not discarded in deinterlacing) were used in H.264 coding. MVs of one MPEG2 MB are used for two vertically neighbouring H.264 MBs. If there is 16x8 prediction used in original MPEG2 MB, vectors of the upper MB part are used for the upper MB in H.264 MB pair, MVs of lower MPEG2 MB part are used for lower block. When 16x16 prediction is used in MPEG2, the same vectors are used by both vertically neighbouring H.264 MBs.

Some modifications were done on MVs. It was necessary to multiply MV horizontal and vertical component by two in order to convert MPEG2 half-pixel MVs into H.264 quarter-pixel MVs. Additional multiplication by two was needed for vertical component because MPEG2 MVs were in field format, whereas MVs in H.264 were meant to be in frame format. Besides that operation, additional adjustment of vertical component was needed. The reason is also change from field to frame referencing. Let's assume that MPEG2 sequence was coded as top field first. If MV of a MB in top field refers to pixels in bottom field and if vertical component of MV has value 0, that means that value 4 should be added to vertical component of H.264 MV in order to make that MV point to the same line of picture.

There is one more modification. The MVs that are pointing to interpolated fields (i.e. the vectors that in MPEG2 pointed to fields that were discarded in deinterlacing process) are modified to compensate (possible) difference between object position in original (discarded) field and new, interpolated field. The MV modification is done with respect to distance between the fields. For forward prediction, the distance is increased, whereas for backward prediction the distance is decreased by deinterlacing. Modification of MVs has shown improvement in experimental encodings.

Briefly, the following changes were made on MVs:

- Both horizontal and vertical component are multiplied by 2 to convert MV from half-pixel to quarter –pixel format
- Vertical component is additionally multiplied by 2 in order to convert MV from field to frame format
- Value 4 or -4 is added to vertical component of some MVs in order to make those MVs point to the proper line in frame
- Every MVs pointing to field discarded during deinterlacing process is modified in order to compensate (possible) position difference between referenced samples in original field and new (interpolated) field.

C. Interlaced frame to progressive sequence transcoding

When field prediction is used for a MPEG2 MB, each MPEG2 MB contains two halves – upper, containing samples of top field and lower containing samples of bottom field. Each 16x8 MB partition has its own motion vector(s). Vectors of one 16x8 partition will be adopted for entire H.264 MB. The choice of partition depends on which field was interpolated during deinterlacing – if top field was preserved, MV of top partition is used and vice versa.

Frame prediction for MB is a simpler case because there is no need to choose between MPEG2 MVs.

The modifications done on MVs of field predicted MBs are similar as for interlaced field to progressive transcoding.

III. EXPERIMENTAL RESULTS

All input streams had GOP (group of picture) size 24 and the number of successive B frame was 2. There was no additional motion search (refinement) after MV translation.

Transcoding was performed using different streams. For every MPEG2 quality, transcoding was performed for several resulting bit rates. Bit rate was changed by varying QP offset (a single offset was used for all MBs in a stream). Performance is observed on PSNR-bit rate curves (PSNR – Peak Signal to Noise Ratio). Several plots are given for streams "park run" and "pedestrian zone". On every plot, PSNR between original sequence and MPEG2 decoded sequence. PSNR loss is observed for the same bit rate.

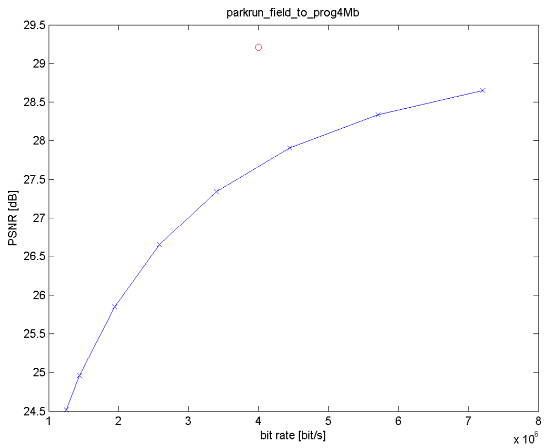

Fig. 2: Results of transcoding "park run" sequence, 4Mb/s MPEG2 quality, interlaced field to progressive transcoding

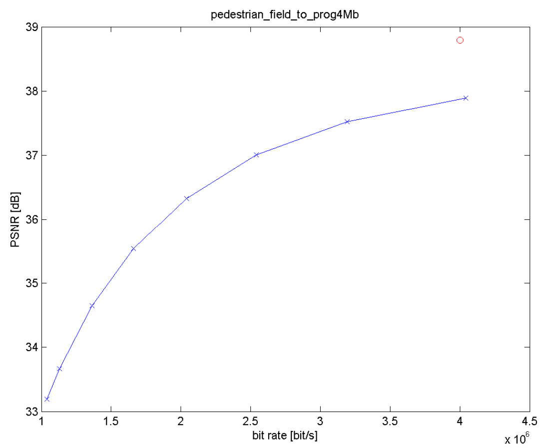

Fig. 3: Results of transcoding "pedestrian" sequence, 4Mb/s MPEG2 quality, interlaced field to progressive transcoding

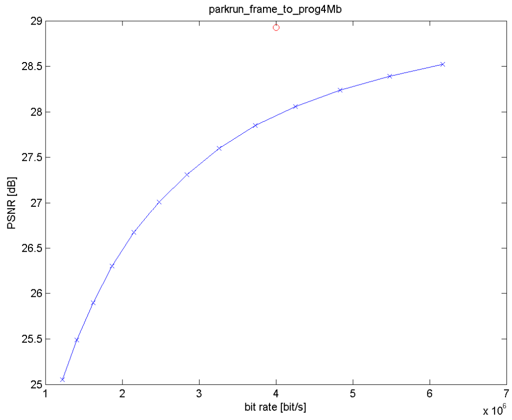

Fig. 4: Results of transcoding “park run” sequence, 4Mb/s MPEG2 quality, interlaced frame to progressive transcoding

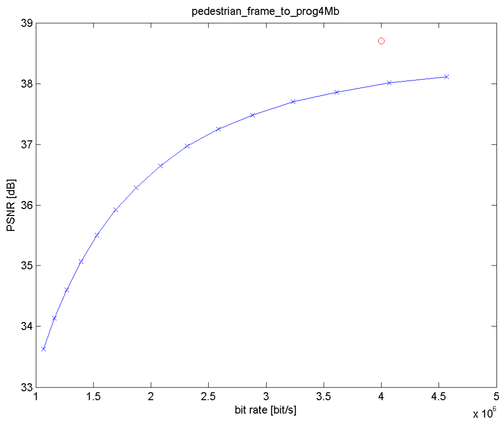

Fig. 5: Results of transcoding “pedestrian” sequence, 4Mb/s MPEG2 quality, interlaced frame to progressive transcoding

PSNR loss for original bit-rate was estimated using linear approximation of PSNR between H.264 decoded and original sequence. PSNR losses are shown in tables 1 and 2. PSNR losses for field to progressive transcoding are greater than those for interlaced frame to progressive transcoding. Presence of frame-predicted MBs in MPEG2 stream probably had influence on this difference – better coding efficiency is expected when a MB is frame-predicted in both MPEG2 and H.264 than when this is not the case.

TABLE 1:.PSNR LOSSES FOR INTERLACED FIELD TO PROGRESSIVE TRANSCODING

bitrate	*“parkrun”*	*“pedestrian”*	*“football”*
2Mb/s	1.41	0.60	0.56
4Mb/s	1.54	0.92	1.01
6Mb/s	1.49	1.12	1.07

TABLE 2: PSNR LOSSES FOR INTERLACED FRAME TO PROGRESSIVE TRANSCODING

Bitrate	*“parkrun”*	*“pedestrian”*	*“football”*
2Mb/s	0.87	0.37	0.25
4Mb/s	0.97	0.72	0.71
6Mb/s	0.95	0.94	0.83

IV. CONCLUSION

A method for translation of MVs form MPEG2 to H.264 stream is presented. PSNR losses varied for sequences and bit-rates. Generally, interlaced field to progressive transcoding gave greater PSNR losses, presumably because of additional loss introduced by greater amount of field coded MBs in MPEG2. Reference pictures used in MPEG2 are not those as in H.264 coding are not only different because of different block transforms and quantizations used in those two standards - additional difference is made by interlaced to progressive picture conversion. Transcoding with additional motion search would most probably give better results at the expense of increased transcoding complexity.

V. REFERENCES

[1] “Information technology – coding of audio visual objects – Part 10: advanced video coding,” ISO/IEC 14496-10:2003, Dec.2003.

[2] ITU-T Recommendation H.262, also ISO/IEC 13818-2, “Information Technology - generic coding of moving pictures and associated information”, 1995