IMPROVED ADVANCED MOTION VECTOR PREDICTION SCHEME FOR SURVEILLANCE VIDEO CODING

Yumei Zhang* Siwei Dong† Yaowei Wang* Yonghong Tian† Peiyin Xing†

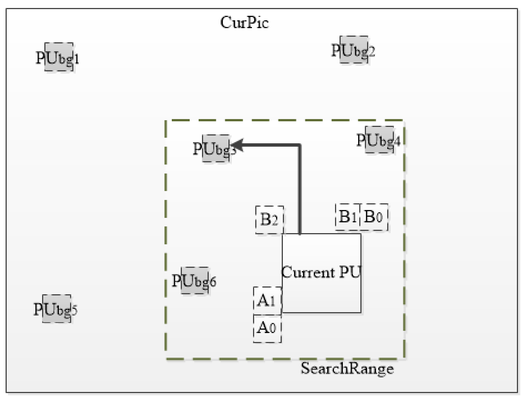
* Beijing Institute of Technology † Peking University

INTRODUCTION

- Background frame based video coding achieves a remarkable compress performance for surveillance video coding.
- We proposed an improved Advanced Motion Vector Prediction (AMVP) scheme to make Motion Vector Prediction (MVP) more accurate.
- The scheme applies to prediction units (PUs) which reference background (BG)-frame.

PROPOSED METHOD

- Analysis:
  - ✓ Case: current PU and PU_{bgi} references BG-frame, A_i and B_j reference normal frame.
  - ✓ **Before**: MVP of current PU is calculated by A_i and B_j.
  - ✓ **After**: MVP of current PU is calculated by PU_{bgi}.

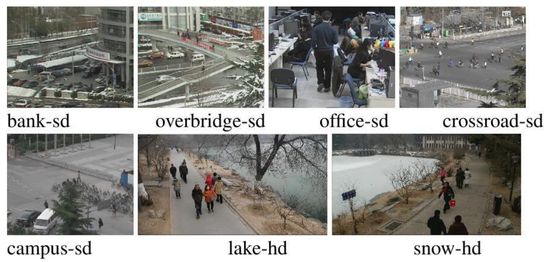


- Propose:
  - ✓ **Background Modeling and Updating.**
    - A background frame is modeled from original frames and then encoded as long-term reference[1][2].
  - ✓ **Recalculate MVP and Reconstruct MVP Candidate List.**
    - Record PUs which reference BG-frame(i.e., PU_{bgi}) .
    - Set a search range, find the best texture matching Pus(i.e., PU_{bg3}) in recorded PUs, and recalculate MVP.
    - Adopt zero MV and the new calculated MVP as final MVP candidates.

Input : The current PU is in the size of $2N \times 2N$ at position (x, y) of the current frame I. A(x,y) is the recorded MVs in 4×4 block level.
Output: MVP of the current PU.
for $m = (0\ to\ N/2 - 1)$ **do**
  for $n = (0\ to\ N/2 - 1)$ **do**
    $sx = (x + 4 \times m)$;
    $sy = (y + 4 \times n)$;
    for $Cq = ((x - R)/4\ to\ (x + R)/4)$ **do**
      for $p = ((x - R)/4\ to\ (x + R)/4)$ **do**
        if $Ax(q, p)! = 0\ or\ Ay(q, p)! = 0$ **then**
          $p = \sum_{i,j=1}^{4} |I(sx + i, sy + j) - I(4q + i, 4p + j)|$;
          if $s < Th$ **then**
            $Th = s$;
            $Mx = Ax(q, p)$;
            $My = Ay(q, p)$;
          end
          else
            $M = (0, 0)$;
          end
        end
        else
          $M = (0, 0)$;
        end
      end
    end
  end
end
$MVP_x = (\sum_{k=0}^{N^2/4} Mx_k) \times 4/N^2$;
$MVP_y = (\sum_{k=0}^{N^2/4} My_k) \times 4/N^2$;

EXPERIMENTAL RESULTS

- Experimental Settings:
  - ✓ Four 720x576(sd) and three 1920x1080(hd) videos[4][5].
  - ✓ Low-delay configuration with common test conditions[3].



- Experimental Results:

Sequence	Proposed vs HM12.0-S	
	BD-Rate	BD-PSNR
bank-sd	-0.35%	-0.003
overbridge-sd	-0.21%	0.004
office-sd	0.10%	-0.005
crossroad-sd	-0.15%	0.003
campus-sd	-0.36%	0.011
lake-hd	-0.12%	0.001
snow-hd	-0.10%	0.003
average	**-0.17%**	**0.002**

CONCLUSION

- An improved AMVP scheme for surveillance video coding is proposed.
- The scheme takes advantage of the correlations among the spatial PUs.
- Achieved **0.17%** bit saving compared to HM12.0-S.

ACKNOWLEDGEMENTS

- This work is partially supported by the National Basic Research Program of China under grant 2015CB351806, the National Natural Science Foundation of China under contract No.61471042 and No.61390515.

REFERENCES

[1] X. Zhang, Y. Tian, T. Huang, and W. Gao, "Low-complexity and high-efficiency background modeling for surveillance video coding." in VCIP, 2012, pp. 1–6

[2] X. Zhang, L. Liang, Q. Huang, Y. Liu, T. Huang, and W. Gao, "An efficient coding scheme for surveillance videos captured by stationary cameras," in Visual Communications and Image Processing 2010. International Society for Optics and Photonics, 2010, pp. 77442A–77442A.

[3] JCT-VC, "HM common test conditions and software reference configurations," in JCTVC-L1100, Jan.2013.

[4] X. Zhang, T. Huang, Y. Tian, and W. Gao, "Background-modeling based adaptive prediction for surveillance video coding," 2014.

[5] W. Gao, Y. Tian, T. Huang, S. Ma, and X. Zhang, "IEEE 1857 standard empowering smart video surveillance systems," 2013.