Efficient screen content coding based on convolutional neural network guided by a large-scale database

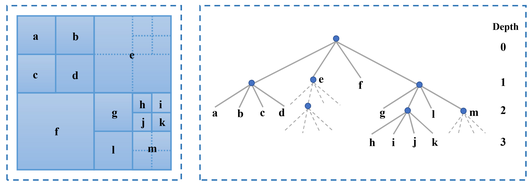
Lili Zhao, Zhiwen Wei, Weitong Cai, Wenyi Wang, Liaoyuan Zeng, Jianwen Chen

Email: zllmail@foxmail.com

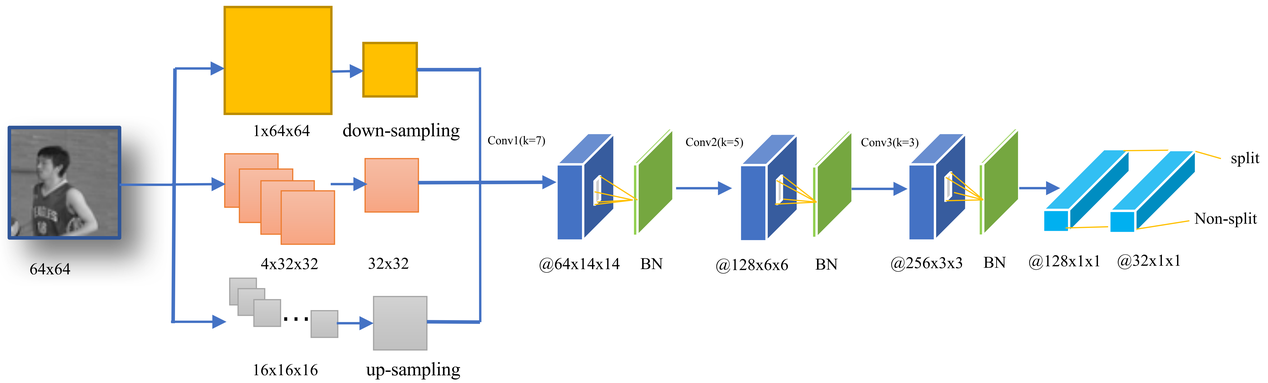
Introduction

Screen content videos (SCVs) are becoming popular in many applications. Compared with natural content videos (NCVs), the SCVs have different characteristics. Therefore, the screen content coding (SCC) based on HEVC adopts some new coding tools (intra block copy and palette mode etc.) to improve coding efficiency, but these tools increase the computational complexity as well. In this paper, we propose to predict the CU partition of the SCVs by a convolutional neural network (CNN) which is trained by the large-scale database that we firstly established for screen content coding. The proposed approach is implemented in SCC reference software SCM-6.1. Experimental results show that our proposed approach can save 53.2% encoding time with 2.67% BD-rate increase on average in All Intra (AI) configurations.

CU Partition



The Architecture of the Designed CNN



Method

The overall structure of the proposed convolutional neural network trained by the established database is shown in the figure above, which can make a quick decision. The CNN can learn the mapping relationship between the features of the current CU and its partition decision. Different from the existed approaches, the proposed method can handle various videos without the prediction result of neighboring CUs of the current CU and there is no need to extract fixed features manually for different content video sequences.

Database

The database is available online: https://github.com/MediaLabProject/SCC_Project. We provide 65 video sequences in our database. Specifically, there are 42 SCV sequences (Table.1), 12 NCV sequences we captured and 11 standard sequences [34] the Joint Collaborative Team on Video Coding (JCT-VC) provided. All sequences are divided into the classes of text and graphics with motion (TGM), mixed content (MIX), animation (AMT) and camera captured content (CAC). These sequences have different resolutions: 720p (1280x720), 1080p (1920x1080) and WQHD (2560x1440). All the sequences are encoded by the recent SCC reference software SCM version 6.1 under common test conditions (CTC) [14] with All-Intra configurations at different QPs (22, 27, 32 and 37). Then we obtained the CU partition decision of 3 depths for each sequence, which forms 12 sub-databases, corresponding to the 3 CU sizes of 3 depths and 4 QPs. The CUs to be split will be denoted as 1 and the others will be denoted as 0.

Conclusion

In this paper, we propose a novel deep learning-based approach for intra prediction in SCC. We also establish a largescale database for SCC intra prediction to give support to deep learning-based complexity reduction in SCC. The proposed approach reduces 53.2% encoding time with 2.67% BDBR and -0.27% BDPSNR loss on average. Moreover, the proposed approach can handle various content videos without any manual extracted feature. Considering that the development of hardware for deep learning is rapid, there is no doubt that the complexity of CNN-based algorithm will be further reduced.

Results

Approach	BDBR(%)	TS(%)
[1]	3.05	36.70
[2]	2.56	37.20
[3]	2.72	48.89
Proposed	2.67	53.21

References

1. Fanyi Duanmu, Zhan Ma, and YaoWang, “Fast cu partition decision using machine learning for screen content compression,” in IEEE International Conference on Image Processing , 2015, pp. 4972–4976.
2. Y. Kawakami, Gaoxing Chen and T. Ikenaga, "Content based mode and depth skipping with Sharp and Directional Edges for intra prediction in Screen Content Coding," 2016 IEEE 12th International Colloquium on Signal Processing & Its Applications (CSPA), Malacca City, 2016, pp. 46-49.
3. H. Yang, L. Shen and P. An, "An efficient intra coding algorithm based on statistical learning for screen content coding," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 2468-2472.

University of Electronic Science and Technology of China

School of Information and Communication Engineering

MediaLab, http://medialab.uestc.edu.cn/

Full Paper

Poster