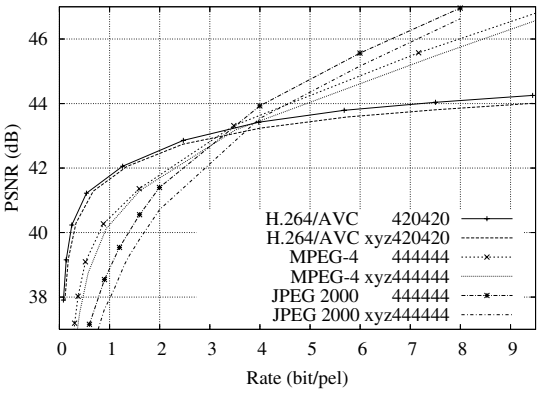
Multiband Video Coding Using H.264/AVC, MPEG-4 Studio Profile and JPEG 2000

Seishi TAKAMURA and Yoshiyuki YASHIMA
NTT Cyber Space Laboratories, NTT Corporation
{*takamura.seishi, yashima.yoshiyuki*}*@lab.ntt.co.jp*

Multispectral imaging, which is common in satellite image acquisition, is now being used in consumer-oriented visual imaging systems such as printers, digital/analog cameras, still/video cameras, projectors, and even displays. The reason is the variety of functionalities it provides. As the amount of image data is proportional to the number of bands, efficient compression of such data sets, particularly for video data, is necessary.

This paper provides a basic investigation of multispectral video coding techniques, such as interband decorrelation, visible/invisible separation, and non-linear signal transform prior to encoding. We use a six-band video sequence (48 bit/pel) as the test data. The video/image codecs of JPEG 2000, MPEG-4 Studio Profile, as well as the newly developed standard H.264/AVC, are compared in a coding experiment.

The results are shown in the figure on the right. In the figure, 420420 denotes the 4:2:0 subsampled format (because of the restriction of H.264/AVC), 444444 the full resolution for each band. xyz*X* denotes visible/invisible separated format, where *X* is replaced by either 420420 or 444444. It is observed that at bit rates of around 3.5 bit/pel, all schemes have similar PSNR values of 43 dB. H.264/AVC and JPEG 2000 offer better PSNR for lower and higher bit rates than 3.5 bit/pel, respectively. The difference between xyz*X* and *X*, i.e., the difference between with and without visible/invisible separation, is less than 0.4 dB for H.264/AVC and MPEG-4, and up to 1.0 dB for JPEG 2000.



Our H.264/AVC coding has some redundancy because it handles the two 4:2:0 sequences separately, particularly with regard to motion vector information. However, the performance benefits simply overwhelm this handicap at lower bit rates. Modification of the H.264/AVC codec to support more than three bands with full spatial resolution such as 444444 is a future task. We will also investigate the optimal encoding parameters, optimal bit allocation among components, subjective evaluation of decoded images, and optimal spatial resolution of each image band.

Acknowledgment
The authors would like to express their appreciation of Mr. N. Fujikawa and Mr. T. Uchiyama from Natural Vision Research Center, Telecommunication Advancement Organization, for supplying the multispectral video data.