2005 WSEAS Int. Conf. on REMOTE SENSING, Venice, Italy, November 2-4, 2005 (pp37-42)

Evaluation of Noise Removal of Radiance Data on Onboard Data Compression of Hyperspectral Imagery

Shen-En Qian[1], Josée Lévesque[2] and Robert A. Neville[3]
[1]Canadian Space Agency, 6767 Route de l'Aeroport, St-Hubert, Qc, J3Y 8Y9 CANADA
[2]Defence Research and Development Canada, 2459 Boul. Pie-XI N., Val-Bélair, Qc, G3J 1X5 CANADA
[3]Canada Centre for Remote Sensing, 588 Booth Street Ottawa, ON, K1A 0Y7 CANADA

Abstract: - This paper evaluates the impact of removing random noise of radiance data using a spectral-spatial smoothing approach on data compression onboard a hyperspectral satellite. A datacube acquired using a Short Wave Infrared Full Spectrum Imager II for target detection application of hyperspectral data was tested. The impact was evaluated using both the statistical based measures and a remote sensing application. The evaluation results show that compression on radiance data after removal of random noise produces better reconstruction fidelity and much higher evaluation scores for the remote sensing application than compression on radiance data without removal of random noise. The evaluation results indicate that random noise of radiance data should be removed before compression, if data compression is applied on radiance data onboard.

Key-Words: - Hyperspectral imagery, onboard data compression, noise removal, evaluation of impact

1 Introduction

The Canadian Space Agency (CSA) is developing an operational spaceborne Hyperspectral Environment and Resource Observer (HERO) mission [1]. HERO will be a Canadian optical Earth observation mission that will address the stewardship of natural resources for sustainable development within Canada and globally. Through targeted imaging, mapping and regular monitoring of the Earth's surface, HERO will acquire and deliver high-quality hyperspectral data that will support decision-making in the management of sensitive ecosystems and valuable natural resources. One of the challenges in the development of a spaceborne hyperspectral imager is the extremely high data rate due to the huge data volume generated on board, which exceeds the downlink capacity, and may quickly exhaust the onboard storage capacity. To deal with this challenge the CSA has been developing data compression technology for hyperspectral imagery for many years. Compression techniques for operational use have been developed [2-5]. Recently, two near lossless data compression techniques for hyperspectral imagery: Successive Approximation Multi-stage Vector Quantization (SAMVQ) and Hierarchical Self-Organizing Cluster Vector Quantization (HSOCVQ) have been developed and patented [6-7]. They well preserve spectral signature information for remote sensing applications [8-10], since the error introduced during the compression is less than the intrinsic noise of the original hyperspectral data caused by the instrument noise, and other error sources such as calibration and atmospheric correction [10]. The CSA would like to place a data compressor onboard HERO using these compression techniques to reduce the requirement for onboard storage and to better match the available downlink capacity.

This paper addresses the impact of removing random noise of radiance data using a spectral-spatial smoothing approach on SAMVQ and HSOCVQ data compression performance onboard a hyperspectral satellite to examine whether the compression should be applied on radiance data or on their noise removed version. This will help to decide whether or not the removal of random noise should be applied onboard before compression. The impact of the pre-processing and radiometric conversion of the raw data to radiance on data compression performance and ultimately on remote sensing applications has been carried out and reported in [11]. In this paper, we evaluate the impact of removal of random noise on data compression using both statistical based measures and a remote sensing application. The root mean squared error (RMSE), signal-to-noise ratio (SNR) and percentage error are used as the statistical based measures. A target detection application of hyperspectral data is used as the remote sensing application. In section 2, we first describe the test data. Then the data processing procedures of generating the remote sensing products for target detection is described in section 3. The evaluation results of the statistical measures are shown in section 4. The evaluation results of the target detection application are shown in section 5. The conclusion is presented in the last section.

2 Test Data Set

A hyperspectral datacube acquired using a Short Wave Infrared Full Spectrum Imager II (SFSI-II) was utilised

in this paper. It was acquired using the SFSI-II on June 7, 2002 flown at an altitude of 2900m with a ground sample size of 3.5m x 3.5m and 240 spectral bands between 1200 nm and 2450 nm with a band interval of 5 nm. The sky was clear with a few cirrus clouds. Man-made targets with five different materials, awnings, polythene, plastic tarp, cotton and vinyl mat, were deployed in the scene. Fig. 1 shows the man-made targets present in the scene of the test datacube. Seven pieces of awnings with varying sizes ranging from 12m x 12m to 0.2m x 0.2m, four pieces of polythene, four pieces of white tarp and four pieces of white cotton with varying sizes ranging from 6m to 0.5m were deployed. In addition, a 3m x 3m piece of white tarp was placed on a large vinyl turf mat (11m x 14m). The raw datacube is in 12-bit digital number and contains 140 along-track lines. The size of the datacube is 140 lines by 496 pixels per line by 240 bands.

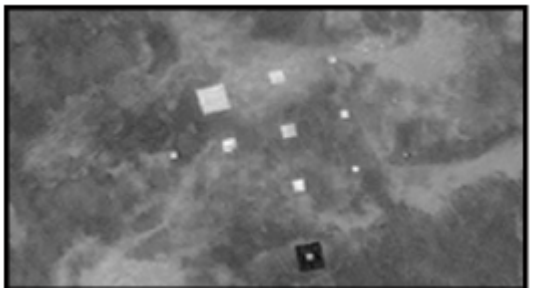

Fig. 1 Man-made targets present in the test SFSI-II datacube.

3 Data Processing Procedures

Fig. 2 shows the block diagram of the data processing procedures for target detection application using spectral un-mixing approach. There are three paths of processing procedures in the figure. In the upper path, which does not contain the onboard data compression, there are eight (8) processing steps to generate the fraction images from the acquired raw SFSI-II hyperspectral data. They are followings:

(1) Removal of periodic noise of raw data. Periodic noise varies from aircraft to aircraft. It is probably caused by electrical emission and electromagnetic interference of an aircraft. In order to remove the periodic noise, the data is transformed into Fourier frequencies and a notch filter is applied to remove the peaks that exceed 20 units.

(2) Subtraction of dark current. The average of the dark current recorded at the beginning and at the end of each flight line is subtracted from each pixel.

(3) Correction of smile: Correction of the spectral shift due to the slit curvature [12].

(4) Correction of keystone. Correction of the geometric distortion caused by misalignment of the lens in the focal plane [13].

(5) Radiometric conversion. A vicarious calibration is performed using calibration coefficients derived from a previous SFSI-II survey to convert the raw data into radiance. The radiance data is stored in 16-bit digital number.

(6) Removing noise of radiance data. Random noise of the radiance data is removed by applying a spectral-spatial smoothing based on the known noise characteristics of the SFSI-II sensor. This is done by averaging together spectra from a number of different pixels that are equivalent in light of the known noise characteristics of the sensor. This noise removal processing can significantly improve the data quality for remote sensing applications. Fig. 3 shows the typical radiance

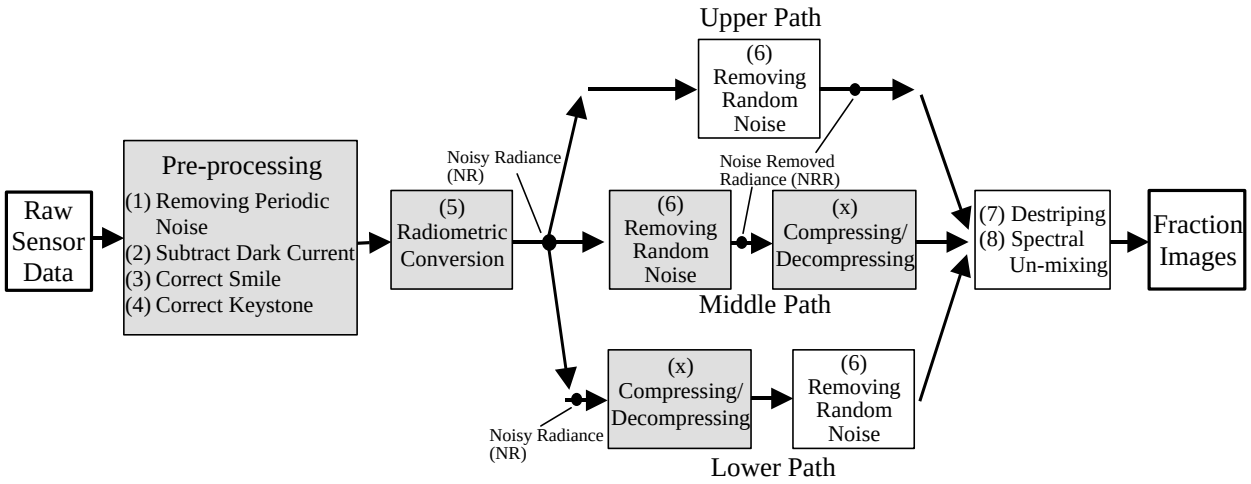

Fig. 2 Block diagram of the data processing procedures (shaded boxes denote onboard processing if an onboard data compression is placed)

spectra of the test SFSI-II datadube over the site before and after the noise removal. The H_2O band at 1470 nm and water vapour value cannot be properly estimated before the random noise is removed. They can be well estimated in the noise-removed radiance spectra.

(7) Destriping. Stripes in band images of radiance datacube are caused by the response difference between the adjacent detector elements. The intensity of the "bad" stripes is readjusted by applying a gain to them in order to correct them.

(8) Spectral unmixing. 9 endmembers and their corresponding regions-of-interest (ROI) are selected manually from the original radiance datacube. They consist of the five man-made targets and four ground features (forest, gravel road, sand and grass). Each ROI is then used to extract the corresponding endmember spectrum from a radiance datacube to be evaluated using constrained spectral unmixing. Fraction image of each endmember is produced for the radiance datacube and used to identify the targets.

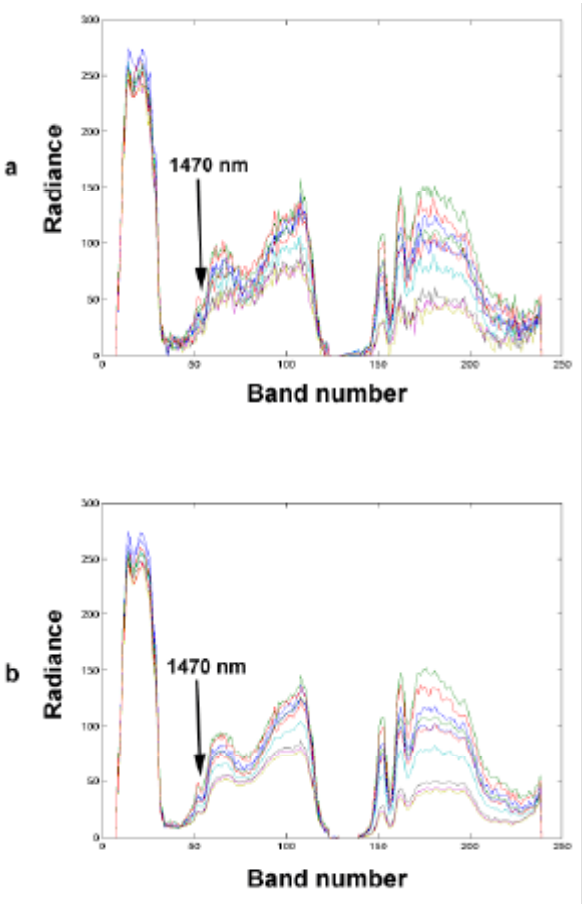

Fig. 3. Spectra of the SFSI-II radiance over the target site (a) before and (b) after removal of random noise. Band numbers start at 1100 nm (band 0) and end at 2500 nm (band 250).

Processing steps 1 to 4 are referred to as pre-processing in this paper. The evaluation of the processing procedures of the middle and lower paths in Fig. 2 is the subject of this paper. In the middle path processing procedure, the noisy radiance data are undergone the removal of random noise before they are sent to compression. The input of the compression is the random noise-removed radiance (NRR) data. In the lower path processing procedure, the noisy radiance data are sent to compression directly and undergone the removal of random noise after compression. The input of the compression is the noisy radiance (NR) data. The evaluation of the upper and middle path processing procedures (i.e. with and without onboard data compression on radiance data) has been carried out and reported in [10].

4 Evaluation Results of Statistical Measures

We evaluate the impact of removal of random noise on data compression using the statistical based measures first. Both the NRR and NR datacubes were compressed using SAMVQ at ratios 10:1 to 30:1 and HSOCVQ at ratios 10:1 to 20:1. The RMSE, SNR and Percentage Error (%E) used are defined as follows.

$$\mathrm{RMSE}=\sqrt{\left(\frac{1}{N_r \times N_c \times N_b}\right)\sum_{i=1}^{N_r}\sum_{j=1}^{N_c}\sum_{l=0}^{N_b}\left[x_{i,j}(l)-\hat{x}_{i,j}(l)\right]^2} \quad (1)$$

$$\mathrm{SNR}=10\log_{10}\frac{(\mathrm{Signal})^2}{\mathrm{MSE}} \quad (2)$$

$$\%\mathrm{E}=\left(\frac{1}{N_r \times N_c \times N_b}\right)\sum_{i=1}^{N_r}\sum_{j=1}^{N_c}\sum_{l=0}^{N_b}\left|\frac{x_{i,j}(l)-\hat{x}_{i,j}(l)}{x_{i,j}(l)}\right|\times 100\% \quad (3)$$

where $x_{i,j}(l)$ and $\hat{x}_{i,j}(l)$ are digital numbers of the radiance datacube and the reconstructed datacube respectively at spatial location (i, j) of band λ. N_r, N_c and N_b are the total number of lines, the total number of pixels per line and the total number of bands of the datacube respectively.

Table 1 lists the RMSE, SNR and %E of the reconstructed datacubes for both the NRR (shaded columns) and NR datacubes. It can be seen that compression on the NRR datacube produces better reconstruction fidelity than that on the NR datacube for the same compression algorithm and same compression ratio. The compression on the NRR datacube attains a SNR gain of between 7.4 and 9.7 dB and a percentage error reduction of between 1.8% and 2.3%. This is probably due to the removal of random noise of the radiance data. The NRR datacube is more compressible than the NR datacube. The higher

reconstruction fidelity of the NRR compressed datacubes are expected to produce more accurate fraction images for ground target detection application of hyperspectral data.

5 Evaluation Results Using a Remote Sensing Application

This section uses ground target detection of hyperspectral data as a remote sensing application to evaluate the impact of removing random noise of radiance data on data compression. The NRR and NR compressed datacubes using SAMVQ at ratios 10:1 to 30:1 and HSOCVQ at ratios 10:1 to 20:1 were decompressed to produce the reconstructed datacubes for evaluation. The reconstructed datacubes are of the same size and format as the original, but have undergone compression. A double blind test approach was used to reduce self-deception and bias in the evaluation of the impact. The compressed datacubes were named using an arbitrary number when sent back to the user. The user derived the fraction images for detecting the targets and evaluated the impact based on pre-defined criteria by comparing the products derived from the compressed datacubes with those derived from the original datacube. The user has no knowledge of the compression status of the datacubes evaluated. The following four criteria were pre-defined and used to assess the impact. 1 point was scored if a criterion was met.

(1) All targets that are present in the fraction images derived from the original datacube are present in the fraction images derived from the compressed datacubes,
(2) No targets other than the ones present in the fraction images derived from the original are present in the fraction images derived from the compressed datacubes,
(3) The T-test of the distribution of the fraction images derived from the compressed datacubes is not significantly different from that of the fraction images derived from the original datacube at a significance level of 0.01,
(4) The percentage standard error (SE%) of a target ROI is less than 5%. It is used to measure the relative average deviation of the fraction images derived from the compressed datacubes and is defined as follows:

$$SE\% = \frac{\frac{1}{n}\sqrt{\sum_{i=1}^{n}(f_i - \hat{f}_i)^2}}{\bar{f}} \times 100\% \quad (4)$$

where f_i is the fraction of an endmember for a pixel in an ROI derived from the original datacube, $\hat{f}_i$ is the fraction of the endmember for the same pixel derived from a compressed datacube, $\bar{f}$ is the mean fraction of the endmember for the ROI of the original datacube, and n is the size of the ROI.

The evaluation was performed on a ROI-by-ROI basis. The full score for a target ROI is 4. The full score for a compressed datacube is 20, as there are 5 targets. Table 2 lists the evaluation score per target, total score per datacube of the blind compressed datacubes with compression applied on the NRR datacube (shaded columns) and on the NR datacube.

When the compression was applied on the NRR datacube (i.e. the compression applied after removal of random noise), the compressed datacube SAMVQ 10:1 got a total score 18 out of the full score of 20 and was ranked #1. For targets polythene, cotton and vinyl mat, this datacube got a full score of 4. Both the compressed datacubes HSOCVQ 10:1 and 20:1 got a total score 15 out of 20 and were ranked #2. Compressed datacubes SAMVQ 20:1 and 30:1 got a total score 14 and 13 respectively. They were #3 and #4. The user accepted all the 5 compressed NRR datacubes based on the pre-defined criteria.

When the compression was applied on the NR datacube (i.e. the compression applied before removal of random noise), both the compressed datacubes at compression ratio 10:1 and 20:1 using SAMVQ got a total score 5 out of the full score of 20. The compressed datacube SAMVQ 30:1 got a total score 3. The compressed datacubes HSOCVQ 10:1 and 20:1 got a total score 6 and 7 respectively.

The total evaluation score of a NR compressed datacubes is much smaller than that of a NRR compressed datacube for the same compression ratio and the same compression algorithm. These evaluation scores are consistent with the statistical measures of the compressed datacubes shown in section 4. The evaluation results of this application show that removal of random noise of radiance data has significant impact on SAMVQ or HSOCVQ compression. Compression on the NRR radiance datacube by applying the removal of random noise in the radiance datacube before compression produces much higher evaluation scores and better user acceptability than that on the NR radiance datacube.

The noise removal algorithm was designed to remove random noise in SFSI-II radiance data. The SAMVQ and HSOCVQ algorithms are lossy data compression algorithms. They act like a low-pass filter, suppressing the high frequency noise during the compression [14]. The noise removal algorithm may not well remove the random noise in the reconstructed

SFSI-II radiance data after compression using SAMVQ and HSOCVQ, since the random noise in the reconstructed radiance data is different from that in the original (uncompressed) radiance data. The noise removal algorithm probably needs to be redesigned in order to effectively remove random noise in the compressed radiance data.

6 Conclusion

This paper evaluated the impact of removing random noise of radiance data using a spectral-spatial smoothing approach on data compression algorithms to be placed onboard a hyperspectral satellite to examine whether the compression should be applied on either radiance data or on their noise removed version. The evaluation result will help to decide whether or not the removal of random noise should be applied onboard before compression. An SFSI-II datacube acquired for target detection was tested. The impact of removing random noise on compression was evaluated using both statistical based measures and a remote sensing application. The root mean squared error, signal-to-noise ratio and percentage error were used as the statistical based measures. A target detection application of hyperspectral data was used as the remote sensing application. Both the statistical based measures and the remote sensing application-based measure show that removal of random noise of radiance data has significant impact on SAMVQ or HSOCVQ compression. Compression on the noise-removed radiance (NRR) datacube produces better reconstruction fidelity and much higher evaluation scores than that on the noisy radiance (NR) datacube. The evaluation results indicate that random noise of radiance data should be removed before compression, if data compression is applied on radiance data onboard.

Acknowledgements

The authors thank Lixin Sun at Dendron Resource Surveys Inc. (Ottawa, Canada) for developing the keystone detection algorithm and assistance to the development of the pre-processing algorithms.

Reference

[1] R. Buckingham, K. Staenz and A. B. Hollinger, “Review of Canadian airborne and apace activities in hyperspectral remote sensing”, *Canadian Aeronautics and Space Journal*, vol. 48, no. 1, pp.115-121, 2002.

[2] S.-E. Qian, A.B. Hollinger, D. Williams and D. Manak, "Fast 3D compression of hyperspectral imagery using vector quantization with spectral-feature-based binary coding", *Journal of Optical Engineering*, vol. 35, no. 11, pp.3242-3249, Nov. 1996.

[3] D. Manak, S.-E. Qian, A.B. Hollinger and D. Williams, "Efficient hyperspectral data compression using vector quantization and scene segmentation", *Canadian Journal of Remote Sensing*, vol. 24, pp.133-143, June 1998.

[4] S.-E. Qian, A.B. Hollinger, D. Williams and D. Manak, "Vector quantization using spectral index based multiple sub-codebooks for hyperspectral data compression", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 38, no. 3, pp.1183-1190, March 2000.

[5] S.-E. Qian, "Hyperspectral data compression using a fast vector quantization algorithm", *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no. 8, pp.1791-1798, Aug. 2004.

[6] S.-E. Qian and A.B. Hollinger, "System and method for encoding/ decoding multi-dimensional data using Successive Approximation Multi-stage Vector Quantization", U. S. Patent No. 6,701,021 B1, issued on March 2, 2004.

[7] S.-E. Qian and A.B. Hollinger "System and method for encoding multi-dimensional data using Hierarchical Self-Organizing Cluster Vector Quantization", U. S. Patent No. 6,724,940 B1, issued on April 20, 2004.

[8] C. Nadeau, G Jolly and H. Zwick, "Evaluation of user acceptance of compression of hyperspectral data cubes (Phase II)", final report of Canadian Government Contract No. 9F028-013456/001, July 24, 2003.

[9] B. Hu, S.-E. Qian, D. Haboudane, J. R. Miller, A.B. Hollinger and N. Tremblay, "Retrieval of crop chlorophyll content and leaf area index from decompressed hyperspectral data: the effects of data compression", *Remote Sensing of Environment*, vol. 92, pp.139-152, 2004.

[10] S.-E Qian, A. Hollinger, M. Bergeron, I. Cunningham, C. Nadeau, G. Jolly and H. Zwick, "A multi-disciplinary user acceptability study of hyperspectral data compressed using onboard near lossless vector quantization algorithm", *International Journal of Remote Sensing*, vol.26, no.10, pp.2163-2195, May 2005.

[11] S.-E. Qian, M. Bergeron, J. Levesque and A. Hollinger, "Impact of pre-processing and radiometric conversion on data compression onboard a hyperspectral satellite", *Proceedings IGARSS'2005 IEEE International Geoscience and Remote Sensing Symposium*, Seoul, South Korea July 2005.

[12] R.A. Neville, L. Sun and K. Staenz "Detection of spectral line curvature in imaging spectrometer data", *Proceedings SPIE*, vol. 5093, pp.144-154, 2003.

[13] R.A. Neville, L. Sun and K. Staenz "Detection of keystone in imaging spectrometer data", *Proceedings SPIE*, vol. 5425, 2004.

[14] S.-E. Qian, M. Bergeron and A. Hollinger, "Evaluation of Near Lossless Feature of Two On-board Data Compression Algorithms: SAMVQ and HSOCVQ," *WSEAS Trans. on Systems, Issue 5*, vol. 3, pp. 2153-2158, July 2004.

Table 1. Statistical Based Measures of Reconstructed Data with the Compression Applied on Noise-Removed Radiance (NRR) and Noisy Radiance (NR) Data

Compression Algorithm & Ratio	**RMSE**			**SNR (dB)**			**Percentage Error (%)**		
	NRR	**NR**	NRR-NR	**NRR**	**NR**	NRR-NR	**NRR**	**NR**	NRR-NR
SAMVQ - 10:1	8.93	21.92	-12.99	41.26	31.59	9.67	0.68	2.46	-1.78
SAMVQ - 20:1	12.92	28.00	-15.08	38.05	29.46	9.67	0.95	2.96	-1.87
SAMVQ - 30:1	15.22	33.35	-18.13	36.62	27.94	8.59	1.09	3.28	-2.19
HSOCVQ - 10:1	17.43	32.63	-15.20	35.45	28.08	8.68	1.33	3.64	-2.31
HSOCVQ - 20:1	20.85	36.05	-15.20	33.89	27.21	7.37	1.59	4.05	-2.46

Table 2 Application Evaluation Score Per Target, Total Score Per Compressed Datacube with Compression Applied on Noise-Removed Radiance (NRR) Datacube and Noisy Radiance (NR) Datacube

Compression Algorithm & Ratio	Score per Target										Total score per datacube	
	Awning (ROI 1 size=71)		Polythene (ROI 2 size=29)		Plastic tarp (ROI 3 size=29)		Cotton (ROI 4 size=28)		Vinyl mat (ROI 5 size=80)			
	NRR	NR	NRR	NR	NRR	NR	NRR	NR	NRR	NR	NRR	NR
SAMVQ - 10:1	3	2	4	0	3	1	4	1	4	1	18	5
SAMVQ - 20:1	3	2	3	0	2	1	3	1	3	1	14	5
SAMVQ - 30:1	3	1	2	0	2	0	3	1	3	1	13	3
HSOCVQ - 10:1	3	1	4	0	3	1	1	1	4	3	15	6
HSOCVQ - 20:1	3	1	4	0	2	1	3	2	3	3	15	7