Clustering analysis of visible spectra of asteroid Ryugu and its preliminary global spectrum map. R. Honda[1], Y. Yokota[2,1], E. Tatsumi[3], R. Hayashi[1], A. Barucci[4], D. Perna[4,5], M. Matsuoka[2], D. Domingue[6], T. Morota[7], S. Kameda[8], T. Kouyama[9], H. Suzuki[10], M. Yamada[11], N. Sakatani[2], C. Honda[12], L. LeCorre[6], M. Hayakawa[2] K. Yoshioka[3], Y. Cho[3], Y. Yamamoto[2], N. Hirata[12], N. Hirata[13], Y. Fujii[1], T. Nakamura[14], T. Hiroi[15], H. Sawada[2] and S. Sugita[3], [1]Kochi University (2-5-1 Akebono-cyo, Kochi, Japan, honda@is.kochi-u.ac.jp), [2]ISAS/JAXA, Japan, [3]University of Tokyo, Japan, [4]Paris Obs., France, [5]INAF, Italy, [6]Planetary Science Institute, USA, [7]Nagoya University, Japan, [8]Rikkyo University, [9]National Institute of Advanced Industrial Science and Technology, Japan, [10]Meiji University, Japan, [11]Chiba Institute of Technology, Japan, [12]University of Aizu, Japan, [13]Kobe University, Japan, [14]Tohoku University, Japan, [15]Brown University, USA.

Introduction: The Japanese asteroid exploration spacecraft Hayabusa2 arrived at asteroid 162173 Ryugu on June 27th, 2018, and began observing its morphology and distribution of material across its surface. One of the optical navigation camera suites onboard Hayabusa2, ONC-T, has obtained multi-band spectra by imaging Ryugu with seven broadband filters ranging from 0.39- 095μm. In this study, spectral clustering analysis was conducted to extract representative reflectance spectra across Ryugu's surface aiming to examine the spatial distribution of spectral properties for selecting landing site candidates. In this presentation, technical evaluation of the clustering analysis for the multi-band spectra, including noise, and the preliminary results are presented along with a global spectral map of Ryugu.

Specification of ONC-T: The ONC-T is a telescopic, charge-coupled-device (CCD) camera with its line of sight along the –Z direction in the spacecraft coordinate system. The CCD image sensor has 1024 × 1024 pixels with a 6.27° × 6.27° field of view (FOV). From the home-position (HP) altitude of 20 km, a 2 km × 2 km of area can be imaged with a pixel resolution of 2 m/pix. The ONC-T has a 7-color filter wheel with a panchromatic window[1][2]. Based on the ONC-T's observations, Ryugu's spectra have been found to be basically flat and there are only small deviations of up to several percent between the bands.

Test of Clustering Methods Using Artificial Data: 7-band spectra are regarded as the data points in 7-dimension attribute space, the spectrum groups are automatically found by a clustering method using the Euclidian distance between the spectra as the similarity measure. Multiple clustering methods, K-means[3], K-means++[4], Gaussian Mixture Model via EM (expectation and maximization algorithm) (GM+EM), and X-means[5] are examined for artificial test dataset, focusing on checking the ability to select the optimal number of components. X-means searches the space of cluster locations and number of clusters to optimize Bayes information criterion (BIC) in a hierarchical way, while the number of clusters is given as an input in the other methodologies.

Thus the adequate number of clusters is determined by using the following method. For K-means, K-means++, and Gaussian mixture model, we conduct trials with various numbers of clusters and calculate BIC for each case assuming each of the obtained clusters are Gaussian and then selected the number of the clusters with the least BIC. In the case of X-means, most frequently obtained number of clusters among 100 trials are determined as the most probable number of clusters.

As artificial input test data, we selected four "Ryugoid" spectra created for landing site selection dry-dun training as the base spectra and added an offset noise $N_{offset}(\mu_0, \sigma_0)$ and random noise at each band $N_{band}(\mu, \sigma)$. As result of the validation using 100 data sets and 100 trials, the X-means method obtained correct numbers of clusters most stably. The use of BICs is also effective in K-means, K-means++ and Gaussian mixture model. Even for X-means, it becomes difficult to find the correct number of clusters when the standard deviation of the offset noise exceeds 7 %. Based on this result, we selected X-means as the standard clustering method for examining Ryugu's spectra.

Ryugu's spectra preparation: Two types of datasets are prepared for clustering analysis of Ryugu's spectra. The first dataset is normalized I/F created from the multi-band images obtained in the Box-A observations conducted at 20km altitude on July 12th, 2018. Raw images are converted into the irradiance after flat-field correction and hardware correction, and then converted into I/F by dividing the obtained irradiance with the solar irradiance. The images are then spatially co-registered to the v-band images and normalized with v-band values. The second dataset is the reflectance at the standard viewing geometry of phase angle of 0°incident angle of 0° and emission angle of 30° created by photometricaly correcting [6] the first dataset.

The obtained spectra are highly noisy due to the effect of surface slope shades. Thus we selected the data in the central area of FOV by selecting the pixels with the longitude difference from S/C nadir point ranges from -30° to 30° aiming to sample as many reliable data as possible including both poles. In addition, the

data with the emission angle less than 50° and the incidence angle less than 50° are also added to include wider areas in the polar region such as both sides of Otohime Saxum in the south-pole area. Then the mean value of 5 x 5 pixels (about 10 m x 10 m) boxes are calculated to reduce noise. Assuming the data with I/F values of v-band less than 0.01 are effective pixels, only the average of bins with the ratio of effective pixels is more than 60% are selected as the valid data.

Clustering Results: We used the first dataset (I/F spectra normalized at v-band) for the preliminary analysis, instead of the normalized reflectance.

Nine groups are found as a result of the clustering of 7-band spectra of normalized I/F using X-means. Figure 1 shows the mean spectra of each cluster. The resulting mean spectra show that the slope of the spectra is the most dominant factor. Comparing the standard deviation of each cluster and the difference of mean values, clusters #4, #5, #6, #7, #8, #0 overlapped each other and clusters #1, #2, #3 are relatively isolated. Moreover, the average v-band reflectance has the negative correlation with the slope of the mean spectra.

Figure 2 shows the global spatial distribution of cluster ID together with the mosaicked v-band reference image. The relatively flat and bright spectra (#1) is distributed around the equatorial Ryujin Dorsum, Tokoyo Fossa, and at the bottom of some craters. Reddish material, such as cluster #4, #5, #6, is distributed in the other areas, and bluish material is distributed around both poles, in particular at Otohime boulder. The slight difference between the western and eastern hemispheres and north-south asymmetry are also seen in the global cluster map. The spatial distribution obtained by other method G-mode [7] shows similar distributions [8] for similar number of clusters.

The calibration method and its accuracy are still under examination, in particular for ul-bands. If we eliminate the ul-band data and conduct clustering with 6 bands, the number of clusters is reduced to 7 and the north-south asymmetry is weakened. We are planning to evaluate the validity of the north-south asymmetry using new data obtained after solar conjunction.

Interpretation: Since the validity of ul-band is under examination, we compared 6 band spectra without ul–band with meteorite spectra. Based on comparison with laboratory-measured meteorites focusing on visible spectral slope, clusters #4, #5, #6, #7, #8, #0, #1, #2 are similar to moderately heated or shocked CC-like material, while cluster #3 is similar to heated CI-like material. The possible causes of gradation of clusters #4, #5, #6, #7, #8, #0, #1, #2 are heating, space weathering, higher organic content (flattening), larger grain size and larger porosity effect of mixing of end members via binning. The depth at w-band is about 1% except for cluster #3. The validity of the absorption must be examined by accumulating the data and increasing the accuracies of calibration and the method of analysis. Further discussion and the result with the data after the photometric correction will be shown in the presentation.

Acknowledgement: The authors are thanks to Y. Miki for test with artificial data. This study was supported by JSPS International Planetary Network.

References: [1] Suzuki, H. et. al. (2018) *Icarus*, 300, 15, 341-359, [2] Kameda, S. et. al. (2017) *SSR* 208, 17–31, [3] MacQueen, J. B. (1967) *Proc. of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 281–297, [4] Arthur, D. and Vassilvitskii, S. (2007) *Proc. eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 1027-1035. [5] Pelleg, D. and Moore, A. W. (2000), *ICML* 1, 727-734, [6] Domingue, D. et al. (2019) *LPSC*, [7] Barucci A. et al. (1987) *Icarus* 72, 304, [8] Barucci A. et al. (2019) *LPSC*.

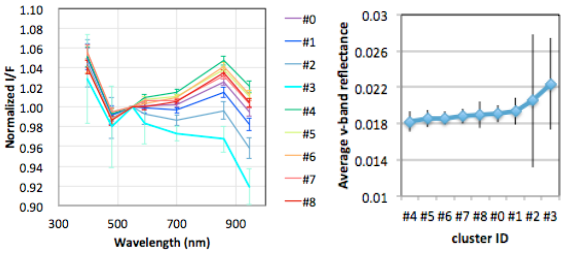

Figure 1. The mean spectra (I/F normalized with v band value) (left) and average v-band reflectance(right) of obtained clusters. Bars of each data point is standard deviation.

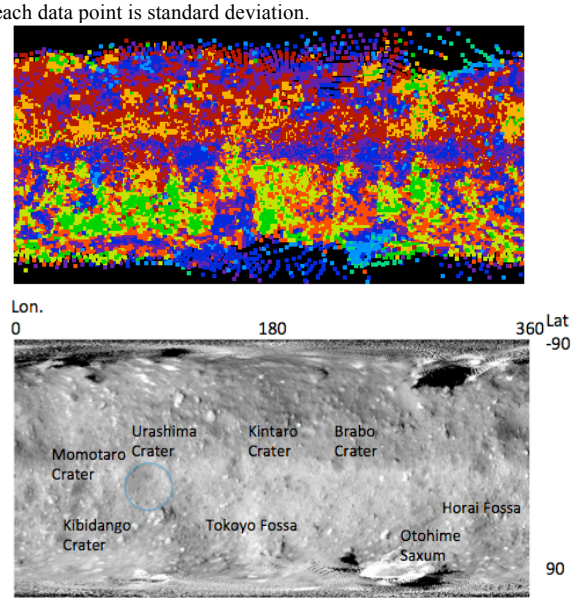

Figure 2. The spatial distribution of obtained cluster IDs (top) and the mosaicked index map (bottom). Colors in the cluster map is the same with ones in Figure 1.