An adaptive image segmentation method with automatic selection of optimal scale for extracting cropland parcels in smallholder farming systems

Zhiwen Cai [1], Qiong Hu [2], Xinyu Zhang [1], Jingya Yang [1], Haodong Wei [1], Zhen He [2], Qian Song [3], Cong Wang [2], Gaofei Yin [4], Baodong Xu [1,5,*]

[1] Macro Agriculture Research Institute, College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China; zhiwen.cai@webmail.hzau.edu.cn (Z.C.); zxy0119@webmail.hzau.edu.cn (X.Z.); jingya.yang@webmail.hzau.edu.cn (J.Y.); hzaursaa@webmail.hzau.edu.cn (H.W.)
[2] Key Laboratory for Geographical Process Analysis & Simulation of Hubei Province/College of Urban and Environmental Sciences, Central China Normal University, Wuhan 430079, China; huqiong@mail.ccnu.edu.cn (Q.H.); hzhen@mails.ccnu.edu.cn (Z.H.); wangcong@mail.ccnu.edu.cn (C.W.)
[3] Key Laboratory of Agricultural Remote Sensing (AGRIRS), Ministry of Agriculture and Rural Affairs/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China; songqian01@caas.cn
[4] Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 610031, China; yingf@swjtu.edu.cn
[5] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Aerospace Information Research Institute, Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, China
* Correspondence: xubaodong@mail.hzau.edu.cn

Part 1: The statistical result of cropland parcels over 8 test tiles.

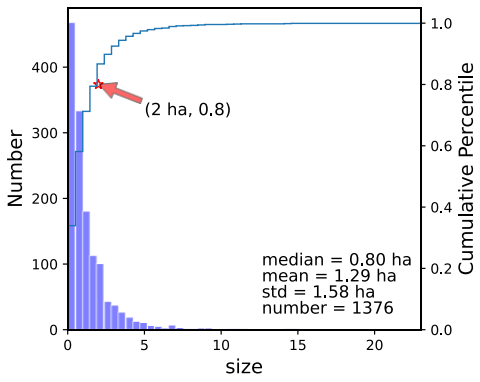

Figure S1. The distribution of cropland parcel size in 8 test tiles.

Part 2: The scale parameter estimation of multi-resolution segmentation.

The candidate scale parameters of multi-resolution segmentation were selected based on the Estimation of Scale Parameter (ESP) tool and visual assessments. First, we used the ESP tools to obtain the candidate optimal scales for each test tile, as shown in Figure S2. Then, the scales of greater than 200 or less than 100 were discarded due to the extremely under- and over-segmentation errors evaluated by the visual assessment. Finally, the final 5 optimal segmentation scales were selected based on their presence frequency ranking at 8 test tiles.

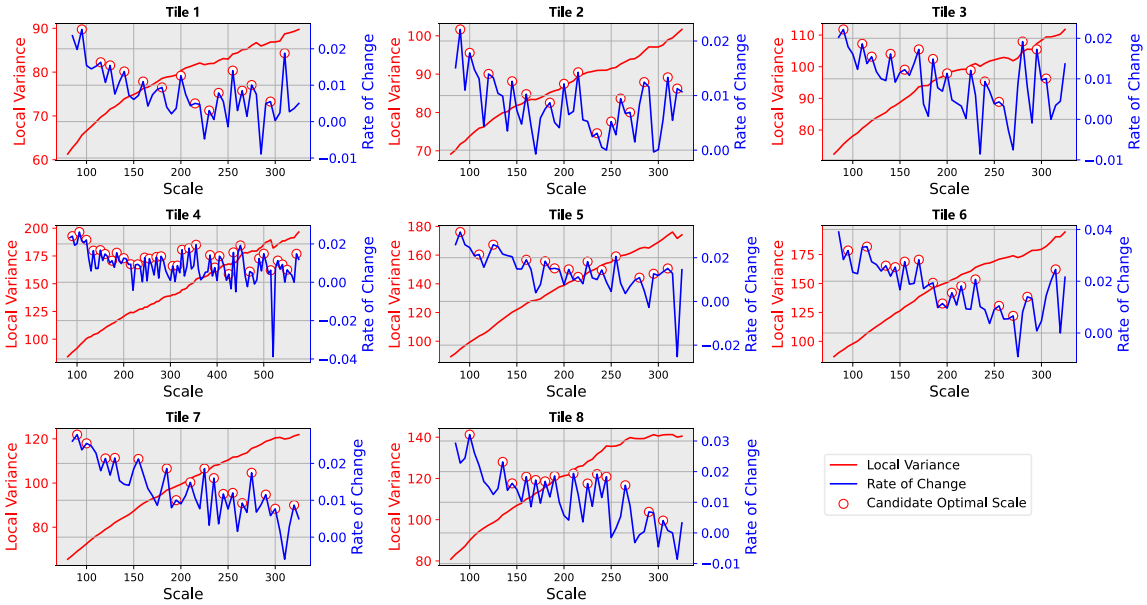

Figure S2. The candidate scale parameters of multi-resolution segmentation over 8 test tiles using the Estimation of Scale Parameter (ESP) tool.

<u>Part 3:</u> The selection of the clusters number for k-means method.

We explored the performance of k-means clustering with different cluster numbers over two tiles, as shown in Figure S3. It can be observed that the clustering results tend to be fragmented with the increase of cluster numbers. In detail, the k-means would generate fragmented objects and introduce the over-segmentation errors as long as the cluster number was overestimated. Since the coarse segmentation tended to derive the under-segmented objects that can be further input to the fine segmentation, the number of clusters in k-means was set as 2 in our study.

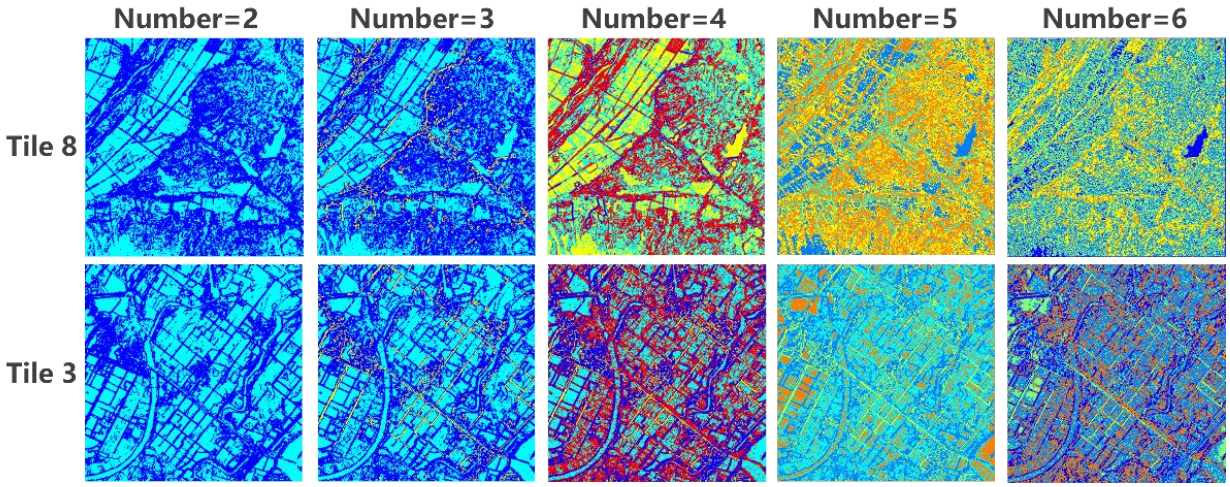

Figure S3. The clustering results of k-means with different clusters numbers.

<u>Part 4:</u> The correlation analysis between different texture features.

We calculated the correlation coefficients (r) between each two candidate texture features, as shown in Figure S4. It can be found that Dissimilarity (Dis) was highly correlated to Variance (Var) and Homogeneity (Hom) with r of 0.73 and -0.78, respectively. This means that the information of Dis could be substituted by Var and Hom. Similarly, Angular second moment (Asm) showed high consistency with Ent (R = -0.91). Therefore, we believe that there was no need for selecting Dis and Asm when the Var, Hom and Entropy (Ent) features were selected. To avoid the confusion, we have revised several sentences in section 4.1 and added the correlation analysis in part 2 of supplementary material.

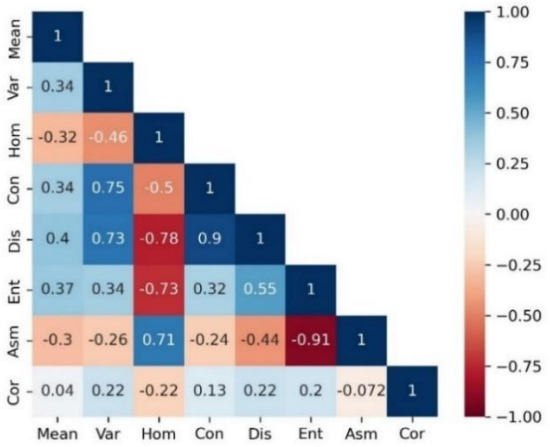

Figure S4. The correlation coefficient between each two texture features

<u>Part 5:</u> Analysis of different co-occurrence shifts of the GLCM effect.

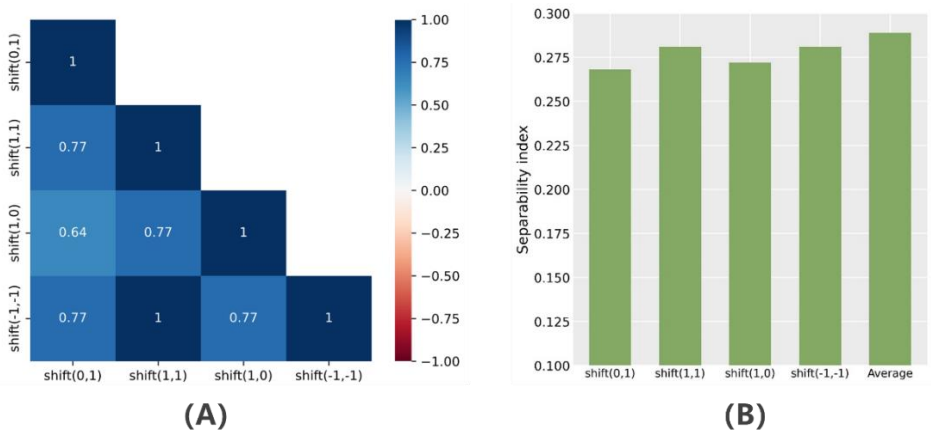

Figure S5. (A) The correlation coefficient between each two co-occurrence shifts; (B) The separability index of *Homogeneity* feature extracted by different shifts.

Part 6: Analyzing the effect of the thresholds of a and b in the bandwidth selection.

The thresholds of a and b were determined based on the relationship between h_s and the average local variance, as shown in Figure S6. When the thresholds of a and b decreased, the estimated spatial bandwidth (h_s) and the corresponding local variance would be smaller, which would decrease the spectral bandwidth (h_r) or texture bandwidth (h_t). As a result, the smaller bandwidth would derive smaller segments. To perform the fine segmentation within the coarse segmentation objects and reduce too many over-segmented objects simultaneously, the thresholds of a and b can be derived as long as FOALV and SOALV start to increase significantly over the study area. If the MSAOS was implemented to other regions, the thresholds of a and b can also be calculated based on the characteristics of the study area.

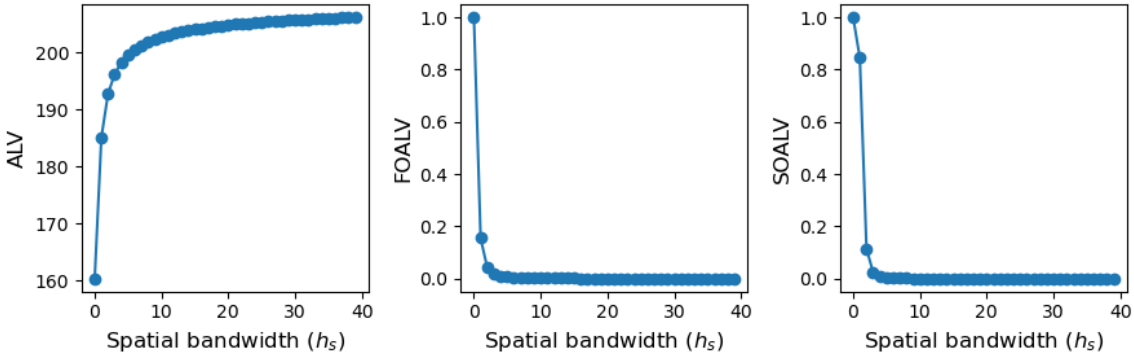

Figure S6. The relationship between spatial bandwidth and ALV, FOALV and SOALV over the study area.

Part 7: The computational efficiency of the MSAOS as well as the basic mean shift algorithm

The main difference between the MSAOS and the basic mean shift was the automatic selection of segmentation bandwidth. Specifically, the computational efficiency of basic mean shift method was primarily affected by the spatial bandwidth (h_s), which determined the number of adjacent pixels for calculation at each iteration. Since different h_s would require different computational resources, we fixed other bandwidths and increased the spatial bandwidth from 5 to 20 with the interval of 5 for simulating the bandwidth selection process. The sum of running time at each bandwidth was regarded as the running time of the basic mean shift segmentation, as shown in Table S1. It can be observed that the computational efficiency of MSAOS was much better than the basic mean shift algorithm, which can be attributed to the great benefits of the coarse segmentation and optimal scale selection.

Table S1. The running time of the MSAOS over 8 test tiles.

	Basic mean shift segmentation (sec)					MSAOS without region merging(sec)			Region merging (sec)
	h_s=5	h_s=10	h_s=15	h_s=20	Sum	Coarse segmentation	Fine segmentation with optimal scale	Sum	
Tile 1	18.6	53.4	124.3	201.8	398.1	47.9	64.9	112.8	84.7
Tile 5	19.4	59.6	132.6	213.6	425.2	45.8	59.4	105.2	90.1
Tile 2	18.1	54.2	126.5	199.3	398.1	46.5	63.3	109.8	43.7
Tile 3	19.2	58.1	133.1	210.8	421.2	46.9	64.9	111.8	79.2
Tile 4	18.4	55.4	128.6	208.2	410.6	46.8	63.9	110.7	76.1
Tile 6	19.4	58.5	136.8	216.2	430.9	45.3	50.9	96.2	109.0
Tile 7	19.2	58.3	136.1	219.5	433.1	50.3	64.7	115.0	94.9
Tile 8	19.8	59.4	135.1	214.4	428.7	46.8	57.4	104.2	98.0