A Depth Multiple Hierarchical StereoNet by Using Thermodynamic Color Guidance

Weiyi Kong¹, Menglong Yang², Qinzhen Huang¹*

¹School of Electrical information Engineering, Southwest Minzu University, Chengdu, Sichuan, 610041, China
²School of Aerospace Science and Engineering, Sichuan University, Chengdu, Sichuan, 610041, China
*Qinzhen Huang’s e-mail: 21500004@swun.edu.cn, veechiry@163.com

Abstract. Thermodynamics is devoted to the study of energy or materials, and few researchers have combined the laws of thermodynamics with stereoscopic vision. Thus, we propose Depth Multiple Hierarchical StereoNet (DMH-Net) by using thermodynamic color guidance, which present end-to-end deep architecture for real-time stereo matching, realize sub-pixel precision and produce high-quality refinement disparity on dataset benchmarks. At the same time, we use color thermodynamics to guide the image output of an accurate Hill disparity image. Our DMH-Net consists of four modules: the first module uses a deep spatial pooling method to extract multiple features that work on the initial left and right images. The second module uses color guidance to produce high-quality cost volume. We use the ahead of two modules to calculate depth disparity precision, which can segment coarse to fine to obtain background and edge of objects information. Thirdly module integrates the information about 3D cost volume and color channel to frame a 4D hierarchical cost volume. The fourth module is designed by depth hierarchical refinement to regress the final disparity. Finally, we achieve sub-pixel precision and real-time performance on benchmarks, which can produce high-quality thermodynamic disparity regression maps.

1. Introduction
In the recent development of computer vision, depth estimation, and disparity refinement are key researches[1, 2]. Since computer vision can be applied to unmanned driving, stereo image matching, and other fields, it is also the top priority of connecting the real world and the virtual world and the research on stereo vision matching has also attracted the attention of scholars[3]. For proper rendering and interaction between virtual and real-world objects, the depth is expected to be both dense and correct around objects’ boundaries. Depth sensors are often used to produce infrared vision. However, the sensors do not work well in the sunlight environment or other infrared sources[4], on the other hand, in indoors and in sunlight environment, there is a strong advantage from stereo vision systems. These systems use passive image data that does not interfere with the environment materials. Moreover, passive stereo systems can use structured light depth sensors, which is better than sparse patterns. Stereo vision also has capabilities to produce accurate object boundaries. However, some disadvantages are inevitable, such as on less textured or chaotic textured surfaces[5]. While traditional methods to deal with the ambiguous regions, they usually choose working on an entire image that used handcrafted functions. However, this method has a disadvantage about complex and time consuming, and cannot work well on the treatment of complex and ambiguous texture areas. With the rapid development of
machine learning, the use of machine learning functions to process data and the training of a network model to find information in the shaded area has become a new development direction.

In the stereo vision field, stereo matching is a key step and worthy attention. In recent years, the research on stereo matching has gained satisfying achievement on KITTI stereo benchmarks and Middlebury datasets[6]. Besides, cost aggregation is also important in stereo matching research. Recent research work on global and local support Windows for each pixel and the average cost area is based on image filtering methods. Although the Gaussian smoothing method adopted by the researchers is efficient and easy to operate, the results are not good, especially in the fuzzy depth boundary region. Yoon and Kweon first proposed to adopt a joint bilateral filter method to produce the cost volume[7, 8]. Since then, some researchers have proposed an edge-aware filtering method to explore cost aggregation. For example, geodesic weight and segment support, in which, the elapsed time of the filter guided image method is independent of the kernel size[9, 10]. Therefore, both stereo methods have shown leading speed and accuracy performance. However, stereo systems in the traditional algorithm have some disadvantages, such as predict ambiguous regions in complexity less context texture or confusing surfaces. Stereo matching and refinement steps cannot improve their performance at the same time, nor can they be completely combined, which will lead to slow processing speed and poor disparity matching effect in stereo matching. To solve these problems, we propose a depth of multiple hierarchical stereonet.

Our main contributions include:

- We improve disparity matching accuracy by deep spatial multiple pooling module and innovated propose to produce deep cost volume with the color guide.
- We demonstrate a hierarchical framework that aims to produce 4D refine disparity to regression.
- We use the thermodynamic color guiding principle as the guiding factor for Hill disparity in the network and output hill disparity matching image.
- Our hierarchical refinement network that achieves real-time performance and produces high-quality sub-pixel disparity.

2. Related Work

An interface separates the object to be studied from the surrounding environment, and this artificial separation is called a thermal system[36, 37, 38, 39]. The research of the thermodynamic system is generally based on energy conversion and is rarely combined with image research[40]. However, Stereo depth has a large number of applications in many fields. So scholars attach great importance to research in the field of stereo depth, which is included four steps. Respectively, namely cost calculation, cost matching, aggregation, optimization, and refinement disparity. The more famous ones, for example, belief propagation graph partitioning technology was applied to stereo study that is also a form of discrete tag. We refer interested readers to surveys and methods described here[11, 12, 13]. With the stereo research development, the state-of-the-art methods use CNNs in the disparity estimation step and use global or semi-global matching to estimate consistent disparity. Among them, the representative algorithms DispNet demonstrate an end-to-end neural network with a correlation layer for volume[14], this method is superior to the traditional stereo algorithm in performance.

The techniques develop about disparity estimation, there are several algorithms presenting advantage performance, such as MC-CC, which introduced a Siamese network to compare the left and right images. To predict consistent disparity estimation, the semi-global matching method was used in[15]. DispNet was also famous in disparity estimation research that demonstrates a real-time network with a correlation layer, in which, used dot product of feature for stereo construction. Liang et.al improved DispNet by introducing a novel iterative filtering process[16]. For building filter cost, GC-Net adopts 3d cost filtering method and soft Argmax process to regression depth. Based on GC-Net, PSMNet was improved using pyramid process to enrich features. The cost volume calculation step is important that can improve the performance of disparity estimation, which can also obtain a better cost volume. It has been demonstrated by several works that the matching cost calculation performance can be improved using convolution neural networks. For example, Park and Lee worked on enlarging the receptive field of features and outperformed with the convolutional layers[17]. The performance of PSMNet is better than
that of GCNet, because PSM deepens the extraction of different levels to improve the feature recognition rate.

Our work is aimed at the imprecision of predicting disparity in stereo matching, we propose depth multiple extracts features methods to build a 4D cost volume. we also creatively according to the second law of thermodynamics and express the energy of different parts of the stereo disparity in different levels of color utilizing color thermal guidance, which uses parallel strategies can realize real-time speed. Rather than traditional stereo matching methods[18, 19], our work obtains more hierarchical information in predicted disparity. Deserve to be mentioned, our algorithm finally achieves sub-pixel accuracy and can match the output stereo disparity in real-time, which is superior to other advanced stereo researches in both performance and accuracy.

3. Depth multiple hierarchical network by Thermally Guide

3.1. Network architecture
Instead of using traditional CNNs methods[17, 20], we break our algorithm into depth feature predictions, cost volume with thermodynamic color guided and hierarchical refinement procedures. The flow of the algorithm is shown in Figure 1.

![Figure 1. Thermally guided depth multiple hierarchical framework.](image)

3.2. Depth spatial multiple pooling module
We aim to train a real-time disparity prediction framework by given pairs of input images, this method of operation is inspired by relevant algorithms[21, 22]. Depth spatial multiple pooling module used to extract the multiple-image information, the principle of main characteristics derived from coarse to fine, the image of the one-dimensional information using global spatial pyramid pooling that works on all rough information to extract the whole image, and then is closely connected to the image of the refinement on the second pooling to extract information, the second layer to connect the third layer of refinement to extract the texture information in the main character is the one-dimensional image by pooling mapped to high dimension space, the refinement of the original image texture information of all step by step, to find different levels of disparity and eventually obtain high precision of prediction of disparity maps. To ensure the denseness of connections in each layer when extracting features, we adopted a multi-level shrinkage ratio of 1, 2, and 4. It can also be understood that this operation is a residual learning method of shrinkage in proportion. This can also maximize the extraction of one-dimensional features from the original image, which can be used to generate the predicted disparity step below to ensure the accuracy of the predicted disparity.

3.3. Depth cost volume with thermodynamic color guided
We use the predicted value of the step output disparity produce depth cost volume, using color information as a thermodynamic guide, closely integrated into the depth of the multiple disparity cost volume, through the left at every level of disparity will feature maps and right together, forming a volume of the 4D deep space cost module by involving different levels of characteristics to promote the
stereo matching, combined with a thermodynamic color difference as a guide, which can obtain high-quality depth cost volume. In our work, depth the cost volume process by $3 \times 3 \times 3$ convolution along the height, width, and depth dimensions, and then reduce the resulted cost by using convolution. Next, followed by convolution with dilation 1, 2, and 4 in parallel, and the stride is 2. The step of depth convolution is used to combine the information from more receptive fields with thermodynamic color guided. Residual learning is very effective in the disparity refinement process so we propose a cascade of such block to iteratively improve the quality of our disparity prediction. In Figure 2 we show the initial input image, the background prediction image, and the specific image using the thermodynamic color as a supervisory guide.

![Figure 2](image_url)

Figure 2. This set of images shows the original image, the background segmentation prediction map, and the thermodynamic color segmentation image from top to bottom. The thermodynamic energy distribution is used to represent color energy disparity in different regions. In the process of disparity prediction, we use the background image as the supervision and guidance factor.

3.4. 4D hierarchical convolution network
Considering differences between pairs of potential matching descriptors along horizontal scanlines, we obtain hierarchical features in left and right images by integrating depth spatial multiple pooling and depth cost volume by color guided to construct a 4D feature volume. In a hierarchical pyramid of 4 volumes, every volume has increasing spatial resolution and increasing disparity resolution. In recently methods, cost volumes are 3D included height by width by disparity, our hierarchical cost volume includes 4th dimension, increases for later layers that represent feature channels C.

3.4.1. Feature volume decode
Because of the difference between scanlines matching and description pairs, we design multiple feature volume filters to realize low dimension features reflect the high dimension, which can obtain more accurate and effective feature information. We define two connected three-dimensional convolution, use dilated convolution processing to obtain more detailed information, which proved the dilated convolution method is feasible. We design six Conv-3D modules to filter feature volume. To improve
the ability to extract features, we apply depth spatial multiple pooling to generate features that capture sufficient global context for high-resolution inputs.

Figure 3. This figure has shown the depth of multiple hierarchical refinement regression processes, which included three layers. The disparity images of regression reconstruction are obtained and finally output.

3.5. Depth Hierarchical refinement regression
Multi-hierarchical dense connection of texture features at different levels can greatly improve disparity estimation and performance. The following comparison experiment will give a comparison between multi-hierarchical and single-tiered performance to prove the effectiveness of multi-level and dense connection (EPE is reduced from 2.55 to 2.50). In the multi-hierarchical, the first convolutional layer outputs high-frequency information, resulting in a large reconstruction error in the region with a large change in object boundary and color. Among them, if the region of the object surface is far away from the boundary, a very accurate initial disparity estimation will be generated, that is, the real reconstruction error is very small. However, texture changes will lead to color changes, so the reconstruction error of the first convolution layer is not accurate. To solve this problem, we propose a multi-hierarchical dense connection method, which aims to improve the accuracy of reconstruction error, improve the stability of computing features, obtain more context information and enrich the sensing field. In the previous study of stereo disparity refinement, residual learning was proved to be effective. Different from previous studies, we proposed a densely linked multiple-depth residual module to improve the quality of disparity prediction through multi-layer iteration. To estimation the continuous disparity map, we use disparity regression based on Siamese framework, detail framework shown in Figure 3. We use each disparity probabilities \( d \) to calculate and use the predicted cost \( c_d \) utilizing the softmax operation. The predicted disparity denotes in equation (1). And softmax function represents each disparity \( d \) weighted by its probability, we denote it as equation (2). We use soft argmax that to have a differentiable but also efficient. For each pixel \( i \) the regressed disparity estimation as equation (3).

\[
\hat{d} = \sum_{d=0}^{D_{max}} d \times \sigma(-c_d)
\]

(1)
\[
\sigma (-c_d) = \frac{\sum_{d'=1}^{N} e^{-c_i(d')}}{\sum_{d'=1}^{N} e^{-c_i(d')}},
\]
\[
d_i = \frac{\sum_{d=1}^{N} e^{-c_i(d)}}{\sum_{d'=1}^{N} e^{-c_i(d')}}.
\]

4. Fusion loss function

In this section, we calculated depth hierarchical net refinement disparity, which included approximate smooth loss \(L_1\) and three refine loss. The refine loss is defined as below equation (4) and equation (5). We use the multi-scale loss to ensure the refinement disparity accurate, which can output one dimension disparity to any for pyramid layers and can also regularize the overall network. \(L_1\) represents the loss on the finest level, and \(L_4\) represents the loss on the most coarse level. According to the depth spatial multiple pooling method, we obtain features including the dilated ratio about 1, 2, and 4. And then we use these multi-features integrate to the depth cost volume. The output result of the system is the 4D hierarchical module input, in which, we use a 4D hierarchical convolutional ratio as 2, 4, and 6 that can connect multiple hierarchical loss functions. The total loss denotes as below equation (6), which realizes to calculate disparity from coarse to fine.

\[
L(d, \hat{d}) = \frac{1}{N} \sum_{i=1}^{N} smooth_{L_1}(d_i - \hat{d}_i),
\]

\[
smooth_{L_1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases}
\]

\[
L = L_1 + \frac{1}{2} L_2 + \frac{1}{2^2} L_3 + \frac{1}{2^6} L_4.
\]

5. Experiments

We evaluate our algorithm on stereo benchmarks such as KITTI 2012, KITTI 2015, and Sceneflow datasets[23, 24]. In this section, we demonstrate the detail in our experiment and also visualized the output of our results, as shown in Figure 4. We tested our framework rectified stereo datasets including Sceneflow[25], KITTI 2012[26], KITTI 2015[26]. We achieve high-quality results at a part of the computational cost required by the state of the art.

5.1. Data sets

In our work, we use the public data sets, including Scene Flow, KITTI 2012, and KITTI 2015. The detail about them as follows. Scene Flow: A large-scale composite data set containing 35, 454 training images and 4,370 test images, of which the height and width of the images are 540 and 960 respectively. This data set sets the difference graph between density and rate as ground truth[23]. In our experiment, some pixels with large disparity were excluded from the loss calculation regardless of whether the disparity was greater than the set limit. KITTI 2012: A real-world data set for driving a car street View, where the image size is 376 high and 1240 wide. It consists of 194 training stereo image pairs with a sparse true difference and 195 test image pairs without ground true difference obtained with LiDAR. The whole training data is divided into the training set and the verification set about 5:1. KITTI 2015: A real data set from the driving car Street View, where the image size is the same as KITTI 2012 above. It includes
200 training stereo image pairs with sparse true difference obtained by LiDAR and 200 test image pairs without ground true difference[29]. In our work, we divide the entire training data into an 80% training set and a 20% validation set.

Training: We using PyTorch to implement our architecture. In our work, we trained network end to end with Adam set beta1 = 0.9 and beta2 = 0.999. For data preprocessing, we adopted color normalization on the whole data set. And then, cropped images randomly to size H = 256 and W = 512, set the max-disparity as 192. During training, we used the Scene Flow data set with a constant learning rate of 0.001 for 10 epochs. Next process, we trained our model on Scene Flow directly for testing and used the model trained with Scene Flow data after fine-tuning on the KITTI training set for 300 epochs. The fine-tuning learning rate began at 0.001 for the first 200 epochs and 0.0001 for the remaining 100 epochs. In terms of equipment, we trained on an Nvidia Ge Force Titan RTX GPU with a batch size of 2, and the parameters were initialized from random.

5.2. Sub-pixel cost precision
In this section, we introduce a crucial variable about choosing the right technology in the precision depth system. In the triangulation system, we set the baseline, the focal length as f and a sub-pixel precision <5 has an error e, increases quadratically with the distance. Recent stereo matching methods perform a discrete search and then retrieve the accurate disparity use a parabola interpolation[30], which leads to a sub-pixel precision 0.25 pixels, that roughly correspond to 4.5 cm error at 3m distance for a system with a 55 cm baseline. Here we show that our algorithm architectures are a breakthrough in terms of sub-pixel precision, therefore, compete with other technologies not only for short distances but also as well as in long. Each operation of our algorithm architecture, including the running time of each expanded convolution layer and refinement layer, and the accuracy of EPE matching pixels. It proves the effective function of the refining network. With each step of refinement disparity matching network, the running time and error of matching get lower and lower. We assess the precision of our algorithm and evaluate Scene Flow then compute the average error that is accurately matched at integer locations. The average of over a hundred million pixels correspond results.
5.3. Benchmark performance

In this section, we compare our algorithm with some advanced algorithms to prove the effectiveness of our work. Our depth multi-hierarchical model has obtained the best performance of disparity estimation on KITTI 2012 and 2015 data sets. Table 1 is the results of comparison on KITTI 2012 and KITTI 2015 respectively. In the evaluation experiment of KITTI 2012, the following evaluation indexes are used: Out-Noc represents the percentage of error pixels in the non-blocking area, Out-All represents the percentage of error pixels in the total area, Avg-Noc represents the average disparity in the non-occlusion area, Avg-All represents the endpoint error Run-time represents the speed of various algorithms in KITTI 2012. In the KITTI 2015 assessment experiment, the percentage of the six error pixels (both greater than 3 pixels or 5 percent of disparity errors) were used, and b-g indicated the static background pixel, f-g represents the dynamic pixels, all is the sum of the two. All represent a total proportion, Noc represents the average disparity. The detail is shown in Table 1. The results of the comprehensive table one and table two show that our depth refining module improves the overall disparity of the algorithm and improves the general ability. In different scenarios, our depth multiple hierarchical models achieve the best disparity estimation performance on both KITTI 2012 and 2015 data sets.

Table 1. The evaluation of KITTI is compared with other existing algorithms.

| index      | GC-Net[31] | DRR[34] | SGM-Net[33] | MC-CNN[12] | DispNet[35] | Ours       |
|------------|------------|---------|-------------|------------|-------------|------------|
| D1-bg-All  | 2.21 %     | 2.58 %  | 2.66 %      | 2.89 %     | 4.32 %      | **2.20%**  |
| D1-fg-All  | 6.61 %     | 6.04 %  | 8.64 %      | 8.88 %     | 4.41 %      | **3.40%**  |
6. Conclusion
In this paper, we propose an innovative depth multi-hierarchical stereo network by using thermal color guidance. The characteristics of the original image at different levels are extracted utilizing the deep multi-level pyramid method, and the traditional three-dimensional cost volume with thermodynamic color guided are fused in the disparity matching stage to construct a four-dimensional cost volume, which contains rich and accurate image disparity information. In the step of disparity refinement regression, we use a multi-level and densely connected hourglass module to carry out accurate disparity regression and finally output a high-precision disparity map. For other stereo methods, we have achieved sub-pixel accuracy and state-of-the-art results. Moreover, our algorithm is superior to other methods in various evaluation indexes.

References
[1] Faugeras, O. Three-dimensional computer vision: a geometric viewpoint. 29 (1993).
[2] Bleyer, M., Rhemann, C. & Rother, C. Patchmatch stereo - stereo matching with slanted support windows. 1-11 (2011).
[3] Arteta, C., Lempitsky, V., Noble, J. A. & Zisserman, A. Learning to detect partially overlapping instances. 3230-3237 (2013).
[4] Geiger, A., Ziegler, J. & Stiller, C. Stereoscan: Dense 3d reconstruction in real-time. 963-968 (2011).
[5] Geiger, A., Roser, M. & Urtasun, R. Efficient large-scale stereo matching. 25-38 (2010).
[6] Goesele, M., Snavely, N., Curless, B., Hoppe, H. & Seitz, S. M. Multi-view stereo for community photo collections. 1-8 (2007).
[7] Geiger, A., Lenz, P. & Urtasun, R. Are we ready for autonomous driving? the kitti vision benchmark suite. 3354-3361 (2012).
[8] Geiger, A., Lenz, P., Stiller, C. & Urtasun, R. Vision meets robotics: The kitti dataset. The Int. J. Robotics Res. 32, 1231-1237 (2013).
[9] Hirschmuller, H. Accurate and efficient stereo processing by semi-global matching and mutual information. 2, 807-814 (2005).
[10] Golemati, S., Stoitsis, J., Sifakis, E. G., Balkizas, T. & Nikita, K. S. Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery. Ultrasound Medicine Biol. 33, 1918-1932 (2007).
[11] Seitz, S. M., Curless, B., Diebel, J., Scharstein, D. & Szeliski, R. A comparison and evaluation of multi-view stereo reconstruction algorithms. 1, 519-528 (2006).
[12] Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis Mach. Intell. 30, 328-341 (2008).
[13] Grossi, E., Lops, M. & Venturino, L. A novel dynamic programming algorithm for track-before-detect in radar systems. IEEE Transactions on Signal Process. 61, 2608-2619 (2013).
[14] Hosni, A., Rhemann, C., Bleyer, M., Rother, C. & Gelautz, M. Fast cost-volume filtering for visual correspondence and beyond. IEEE Transactions on Pattern Analysis Mach. Intell. 35, 504-511 (2013).
[15] Klaus, A., Sormann, M. & Karner, K. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. 3, 15-18 (2006).
[16] Lang, M. et al. Nonlinear disparity mapping for stereoscopic 3d. 29, 75 (2010).
[17] Martin, D., Fowlkes, C. C. & Malik, J. Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Transactions on Pattern Analysis Mach. Intell. 26, 530-549 (2004).
[18] Lucas, B. D. & Kanade, T. An iterative image registration technique with an application to stereo vision. 674-679 (1981).

[19] Marr, D. & Poggio, T. A computational theory of human stereo vision. Proc. The Royal Soc. B: Biol. Sci. 204, 301-328 (1979).

[20] Luo, W., Schwing, A. G. & Urtasun, R. Efficient deep learning for stereo matching. 5695-5703 (2016).

[21] Mei, X. et al. On building an accurate stereo matching system on graphics hardware. 467X74 (2011).

[22] Rusu, R. B., Bradski, G., Thibaux, R. & Hsu, J. Fast 3d recognition and pose using the viewpoint feature histogram. 2155-2162 (2010).

[23] Zbontar, J. & LeCun, Y. Computing the stereo matching cost with a convolutional neural network. 1592-1599 (2015).

[24] Sun, J., Zheng, N. & Shum, H. Stereo matching using belief propagation. IEEE Transactions on Pattern Analysis Mach. Intell. 25, 787-800 (2003).

[25] Scharstein, D., Szeliski, R. & Zabih, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Comput. Vis. 47, 7-42 (2001).

[26] Scharstein, D. & Szeliski, R. High-accuracy stereo depth maps using structured light. 1, 195-202 (2003).

[27] Zhao, T., Aggarwal, M., Kumar, R. & Sawhney, H. Real-time wide area multi-camera stereo tracking. 1, 976-983 (2005).

[28] Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. (2014).

[29] Strasdat, H., Davison, A. J., Montiel, J. M. M. & Konolige, K. Double window optimisation for constant time visual slam. 2352-2359 (2011).

[30] Zhu, J., Wang, L., Yang, R. & Davis, J. Fusion of time-of-flight depth and stereo for high accuracy depth maps. 1-8 (2008).

[31] Gehrig, S., Eberli, F. & Meyer, T. A real-time low-power stereo vision engine using semi-global matching. 134-143 (2009).

[32] Besse, F., Rother, C., Fitzgibbon, A. & Kautz, J. Pmbp: Patchmatch belief propagation for correspondence field estimation. Int. J. Comput. Vis. 110, 2-13 (2014).

[33] Yang, Q. A non-local cost aggregation method for stereo matching. 1402-1409 (2012).

[34] Hu, X. & Mordohai, P. A quantitative evaluation of confidence measures for stereo vision. IEEE Transactions on Pattern Analysis Mach. Intell. 34, 2121-2133 (2012).

[35] Tippetts, B., Lee, D., Lillwythe, K. & Archibald, J. Review of stereo vision algorithms and their suitability for resource-limited systems. J. Real-time Image Process. 11, 5-25 (2016).

[36] Ibarrabahena J, Romero R J, Velazquezavelar L, et al. Experimental thermodynamic evaluation for a single stage heat transformer prototype build with commercial PHEs[J]. Applied Thermal Engineering, 2015: 1262-1270.

[37] Lima E C, Hosseinibandegharaei A, Morenopirajan J C, et al. A critical review of the estimation of the thermodynamic parameters on adsorption equilibria. Wrong use of equilibrium constant in the Van't Hoof equation for calculation of thermodynamic parameters of adsorption[J]. Journal of Molecular Liquids, 2019: 425-434.

[38] Nagabhushana G P, Shivaramaiah R, Navrotsky A, et al. Direct calorimetric verification of thermodynamic instability of lead halide hybrid perovskites.[J]. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113(28): 7717-7721.

[39] Tran H N, You S, Chao H, et al. Thermodynamic parameters of cadmium adsorption onto orange peel calculated from various methods: A comparison study[J]. Journal of environmental chemical engineering, 2016, 4(3): 2671-2682.

[40] Wang K, He Y. Thermodynamic analysis and optimization of a molten salt solar power tower integrated with a recompression supercritical CO2 Brayton cycle based on integrated modeling[J]. Energy Conversion and Management, 2017: 336-350.