Acta Technica *62* No. 2A/2017, 165–176

3D reconstruction of tree and limb based on aerial image of UAV[1]

Tiebo Sun[2,3], Jianming Kan[2], Jinhao Liu[2], Qingqing[2], Kai Ma[2], Tingting Sun[2]

Abstract. To obtain more accurate and more detailed information, UAV (unmanned aerial vehicle) aerial in the application is inevitable going from two dimension to three dimension. While trees, due to their structure complexity, make the 3D reconstruction have more research significance. UAV aerial images, because of being affected by aerial height, the information contained in the tree crown texture and profile is poor. In terms of this problem, a feature extraction method based on watershed segmentation is proposed, so as to fully extract the feature points that can reflect the crown structure, and the region correlation coefficients are calculated to match the feature regions accurately. The results showed that in the tree modeling method, based on L-system, the structure of the growth of the trees is simulated. At last, it is concluded that in the case of less number of feature points, it well restores the limb structure trees, and constructs a complete model of trees.

Key words. 3D reconstruction, aerial image of UAV, watershed.

1. Introduction

In the UAV image processing, the target 3D reconstruction and localization have become one of the research directions in recent years by means of sequence image analysis [1]. The three-dimensional reconstruction problem based on UAV aerial images can theoretically be attributed to the reduction of the image scene and the 3D structure of the target from the sequence image. Applying the theory and technology of 3D reconstruction of sequence images, the 3D model of scene and target can be obtained by processing of unmanned aerial vehicle sequence images [2]. The three-dimensional structure of the target can be reconstructed and measured accurately based on the intrinsic constraints of the same scene on multiple images. But for

[1]The authors acknowledge the National Natural Science Foundation of China (Grant: 51578109), the National Natural Science Foundation of China (Grant: 51121005).

[2]School of Technology, Beijing Forestry University, 100083, Beijing, China

[3]Department of Mechanical and Electrical Engineering, Jiangsu Food & Pharmaceutical Science College, 223003, Jiangsu, China

the sequence image obtained from the UAV aerial photography, under the effect of height, the color, texture and contour information that the UAV aerial images contained are fuzzy.

In addition, the process of aerial shaking, offset, rotation and so on will cause the camera swing and rotation so that the camera imaging model has much complexity compared with the general camera imaging models. To determine the UAV aerial camera model is the main problem of UAV aerial image 3D reconstruction [3–4]. In the UAV navigation and agricultural applications, it often requires the measurement of a high degree of information on plants and trees, which created a necessity for 3D reconstruction of crown based on aerial images. As a result, 3D reconstruction of trees has become the focus of many scholars for the study [5–6]. However, trees, plants and other objects, influenced by their growth conditions (including photosynthesis, nutrient cycling, energy transmission and forest leaf light environment etc.) and some human intervention (such as pruning branches etc.), the uncertainty of morphology and structure is great. In consequence, it is difficult to establish the experience knowledge of tree morphology [7]. In addition, the surface texture of trees is very similar or almost no texture, so it is difficult to find the corresponding relationship between the surface points of trees in the images, so the difficulty of three-dimensional reconstruction of crown trees is increased. These factors make the extraction of crown feature points and matching and modeling of complex shape and structure of trees have become a key issue for the crown 3D reconstruction. It has very important theoretical guidance and practical significance in accurately using the aerial sequence images for the reconstruction of crown model.

2. Method

Watershed segmentation method is a segmentation method based on mathematical morphology and topology theory. Its basic idea, as the name suggests, is to regard the image as the topological landscape on geodesics. The gray value of each pixel in the image represents the altitude of the point, and every local minimum and the affected region are called the catchment basin, and the boundary of catchment basin forms the watershed. Watershed segmentation usually takes the gradient image as the input so that it can quickly get the closed area, but there is often over segmentation phenomenon [8]. Watershed transformation can locate the edge accurately, and it has the advantages of simple operation and easy parallel processing. In this paper, we hope that the segmentation method can sufficiently distinguish the light and shade areas of the crown image, and the appropriate over segmentation of the watershed segmentation algorithm can meet our requirements to some extent.

The concept and formation of watersheds can be explained by simulating the immersion processes. On the surface of every local minimum, a hole is pierced, and then the whole model is slowly immersed in water. With the immersion deepening, the influence domain of each local minimum slowly expands outward, to build a dam in the two water collecting basin confluence, which forms a watershed. The watershed computation process is an iterative annotation procedure. L.Vincent [9] proposed a more classic watershed segmentation algorithm, and the calculation is

divided into two steps: one is the sorting process, and the other is the flooding process. Firstly, the gray level of each pixel is ordered from low to high, then in the achieving submerged process from low to high, for each local minimum, in the effect domain in h order height, FIFO structure is used for the judgment and annotation.

What the watershed transformation obtained is the catchment basin of the input images, and the boundary between the catchment basin is the watershed. Obviously, the watershed represents the maximum point of the input image. Therefore, in order to obtain the edge information of the image, the gradient image is usually used as the input image, that is

$$g(x,y) = \text{grad}\ (f(x,y)) = \sqrt{[f\,(x,y) - f\,(x-1,y)]^2 + [f\,(x,y) - f\,(x,y-1)]^2}\,. \tag{1}$$

In the above formula, $f\,(x,y)$ represents the original image. Therefore, this paper uses gradient watershed segmentation method, and the algorithm process is as follows: first of all, the Sobel operator is used for longitudinal gradient operation; then, the gradient image is opened and closed for smoothing processing; at last, watershed segmentation is conducted for the smoothed images.

The watershed algorithm has a good response to the weak edge, and the phenomenon of over segmentation is caused by the noise in the image and the slight change of the gray level of the object surface. There are many minimum points in the image in general, and there is usually an over segmentation phenomenon. The gradient threshold segmentation improvement or labeling of watershed algorithm can be applied to connect multi-minimum value areas. But it should be seen that, the watershed algorithm has a good response to the weak edge, which is guaranteed by the closed continuous edge. In addition, the closed catchment basin obtained by the watershed algorithm provides the possibility of analyzing the regional features of the image.

Through the above watershed segmentation, the bright area and the dark area of the crown image can be distinguished well. For the area information recording the segmentation, seen from the image, the pixels the closer to the center, the more it represents the pixel characteristics of that area. By extracting the centroid $(\overline{u}_i, \overline{v}_i)$ of each region as feature points, this paper thinks that for the bright region, the centroid point is approximate to the brightest point; for the dark area, its centroid is similar to the darkest point, and therefore the calculation formulas of $\overline{u}_i$ and $\overline{v}_i$ are as follows.

$$\overline{u}_i = \frac{1}{n_i} \sum_{k=0}^{n_i} u_k\,, \tag{2}$$

$$\overline{v}_i = \frac{1}{n_i} \sum_{k=0}^{n_i} v_k\,. \tag{3}$$

The matching method based on region correlation coefficients only takes into account the local features of the image when matching, so the pixel pairs with high correlation coefficients are not necessarily the correct matching points. That is to say, if the two areas are the matching region pairs, then their correlation coefficient must

be relatively large. On the contrary, RGB regional correlation coefficient calculated by the two regions does not necessarily mean that these two regions are the regions matched to each other. Therefore, the matching algorithm of this paper cannot avoid the existence of some mismatching points.

To further improve the accuracy of the matching algorithm, this paper removes the wrong matching point by calculating the ratio of the most relevant one and sub related one [10]. Among them, the ratio calculation formula of the most relevant one and sub related one is:

$$K = D_{\text{nearest}}/D_{\text{hypo}-\text{nearest}}\,, \tag{4}$$

where D_{nearest} refers to the Euclidean distance between the most relevant matching point and the point to be matched, and $D_{\text{hypo}-\text{nearest}}$ suggests the Euclidean distance between the sub relevant matching point and the point to be matched.

According to the literature [11], the closest distance between the correct matching pairs should be significantly smaller than that of the wrong matching pairs. For the false matching points, due to the impact of high dimension feature space, there may be many other false matching points with similar distance. Thus, this correlation distance is taken as the error matching criteria to determine the individual fuzzy examples. Therefore, for a matching area, if the maximum correlation coefficient and the sub correlation coefficient are more closely, the ratio of them will be greater. It indicates that the matching area pairs are more likely to be the wrong matching examples. On the contrary, if the ratio of the most relevant and sub relevant is smaller, then it means that the matching is more likely to be the correct matching example.

So, a suitable threshold is selected for the proportion. If the proportion of K is greater than a threshold, it indicates that the most relevant matching and the sub relevant matching are very similar. In this paper, we can think the closest Euclidean distance matching might be the wrong matching, and eliminate the matching point.

3. 3D modeling

3.1. L-system

L-system is proposed by the American biologist A. Lindenmayer [12]. At the very beginning of this system is the method of simulating plant biology and cell morphology and growth, and later it is developed into the fractal method effectively simulating scenes of nature (especially vegetation type) in the computer graphics. L-system is a language prompt system consisting of several words, symbols and control parameters. It uses the right hand rectangular coordinate system to define the three-dimensional L-system. The direction is defined by the 3 vectors, H, L and U [13]. The 3D coordinate system of the L-system is shown in Fig. 1.

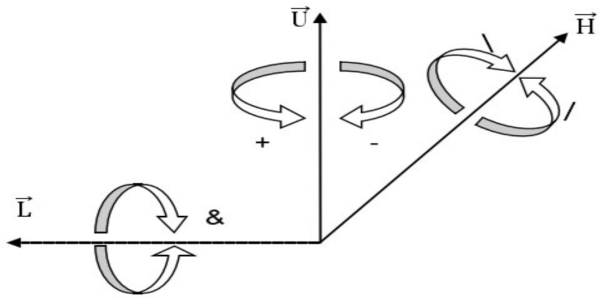

Fig. 1. 3D coordinate system of L-system

3.2. Canopy structure acquisition

In this paper, the skeleton information of the tree canopy is obtained by a certain method. Through the command of the L-system, a three-dimensional model of the crown is created by using Open GL. The modeling process is shown in Fig. 2.

Set a collection Trunk List of tree trunks, and the 3D scatter set contained in these tree blocks is the Include Points collection. The method for obtaining skeleton information is as follows:

(1) To find the highest point of the Z coordinate, and find dz of the highest point and the lowest point. The point dz is taken as the tree height and the vertical axis where the highest point locates is taken as the central trunk of trees. The depth is 1, the trunk set Trunk List is added, and the highest point is added in the tree trunk points set Include Points.

(2) To scan the feature points not in the Include Points collection, and to find the minimum point pi at which the current point is concentrated away from the tree trunk, and the corresponding trunk ti (the depth of the restricted ti is no more than 6, and the branch is no more than 3).

(3)To lead out a new tree trunk $tnew$ with angle of 30 degrees with ti from the tree trunk ti to the point pi. The depth of $tnew$ is the depth of ti (the maximum depth of restricted tree is no more than 4), which is inserted behind ti when adding Trunk List, and the point pi is added to the existing point set Include Points.

4. Results and discussion

Feature points extraction and matching:

In order to obtain the light and shade area in the crown, we make the watershed area segmentation of the crown region, to distinguish the light and shade areas of the crown. The contrast between the watershed segmentation and the original image is shown in Fig. 3. It can be seen from the graph that, the watershed segmentation well achieves the desired effect in this article.

After obtaining the light and shade area of the crown area, the volume of each area is recorded in pixels number, and the statistical region information is shown in

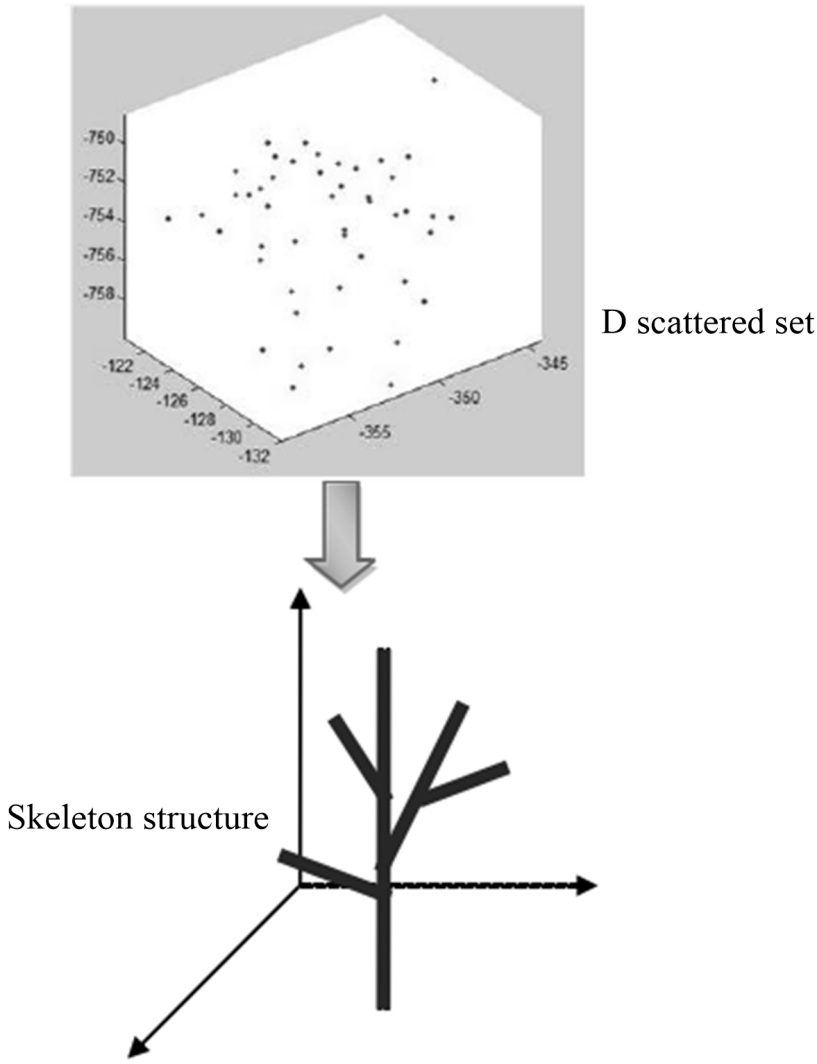

Fig. 2. Modeling process

Table 1. Analyzing the data, it can be found that, in the segmented regions, the size of most pixels is between 40 to 150. Although there is a larger area and a smaller area, it is after all in the minority. The number of areas in the intervals of [1,39] and [150181] has a total of 8. On the whole, the size of the region segmented by watershed is still relatively uniform.

By the statistics of segmentation region information, the size of each region is relatively uniform, so the centroid $(\overline{u}_i, \overline{v}_i)$ of each region is calculated as the feature point of the feature area, and then added to the feature points set of the image. The image feature points extracted are shown in Fig. 4, left part. It can be seen from the

figure that, watershed segmentation can fully extract the brighter and dark points in the tree crown, and these points can well reflect the crown branches information. In the watershed segmentation results, the feature points set position is shown in Fig. 4, right part.

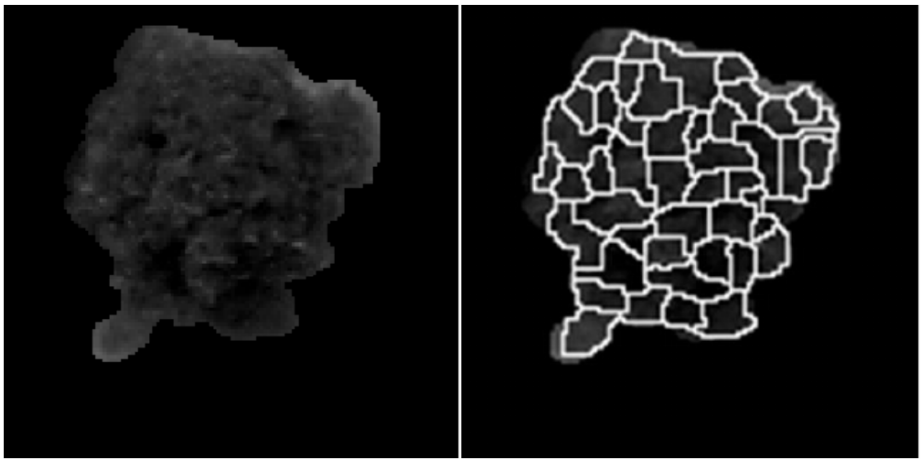

Fig. 3. Watershed segmentation results for tree crowns

Table 1. Results statistics of watershed segmentation

Items	Values	Ratio
The maximum value of areas, unit: pixel	181	-
The minimum value of areas, unit: pixel	1	-
The total number of areas	47	1.000
[1,39]	5	0.106
[40,99]	26	0.553
[100,149]	13	0.277
[150,181]	3	0.064

3D computation and 3D modeling:

In order to obtain the most accurate and complete canopy structure, we take different thresholds in the experiments and compare the corresponding results. Several representative threshold values are selected, the feature points set is extracted, and the skeleton structure and 3D model of the tree are constructed. Figure 5 shows the result when the threshold is 0.93, obtained 28 points. Due to the interference error points, the trunk effect is not beautiful. Figure 6 shows the experimental results when the threshold is 0.91, obtained 23 points, well removing the interference characteristics points, and the effect has been greatly improved. Figure 7 shows the experimental results when the threshold is 0.88. The number of feature points obtained by is only 17. When the false matching is removed, at the same time, the effective information is reduced, and the model constructed is not complete.

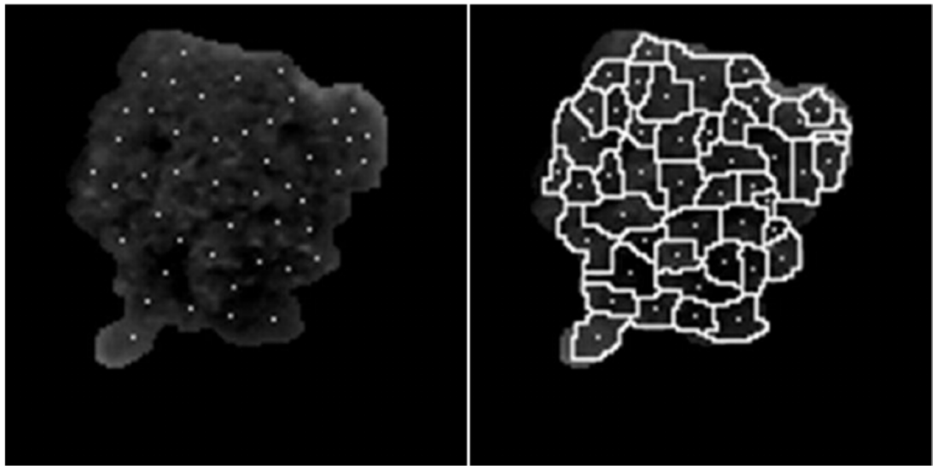

Fig. 4. Watershed segmentation of images and feature points extracted 3D computation and 3D modeling

It can be seen from the above results that, the K threshold is too large, and the removed false matching points are too small, then the left false matching points have interference on the results, thus affecting the 3D modeling effect. If the threshold of K is too small, then in full removal of false matching points, at the same time, it will also remove some correct matching points, and caused the loss of information.

In order to better and more intuitively compare, the proposed feature extraction and matching method are used at the same time for feature extraction and matching [14]. The mean geometric registration errors calculated are compared, and the results are shown in Table 2. As can be seen from the table, this method can match the feature points more accurately, and it takes the appropriate threshold K, and achieves better reconstruction results.

Table 2. Mean geometric registration error

Methods	Mean geometric registration error	Number of feature points
SIFT-KNN	1047.997	5
The proposed method ($K \leq 1$)	1041.99	47
The proposed method ($K \leq 0.8$)	1040.63	14
The proposed method ($K \leq 0.6$)	1039.91	9
The proposed method ($K \leq 0.4$)	1038.23	4

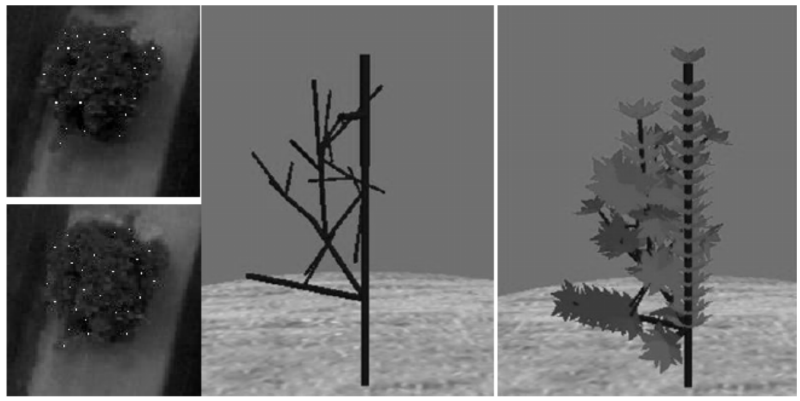

Fig. 5. Experimental results for threshold 0.93

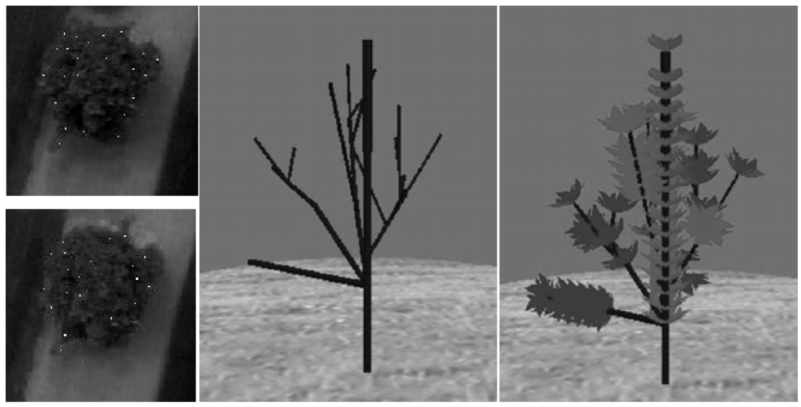

Fig. 6. Experimental results for threshold 0.91

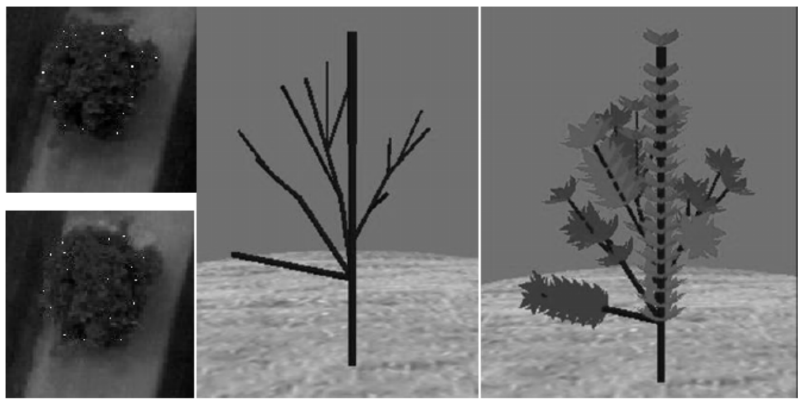

Fig. 7. Experimental results for threshold 0.88

5. Conclusion

In the feature extraction and matching, due to the complex structure of crown and features of UAV aerial image itself, a feature point extraction based on watershed region segmentation completely extracts the feature points that can reflect the crown structure. And the feature matching method based on the RGB region correlation coefficient, in the case of limited image information, can also accurately match the feature points, and then carry out 3D reconstruction. In the tree crown modeling method, based on L-system method, the structure of the growth of the trees is simulated. In the case of less number of feature points, it well restores the limb structure of trees, and constructs a complete model of trees.

References

[1] G. Caroti, I. Martínez-Espejo Zaragoza, A. Piemonte: *Accuracy assessment in Structure from Motion 3D reconstruction from uav-born images: The influence of the data processing methods.* ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences *XL-1/W4* (2015), 103–109.

[2] A. Eltner, D. Schneider: *Analysis of different methods for 3D reconstruction of natural surfaces from parallel-axes UAV images.* Photogrammetric Record *30* (2015), No. 151, 279–299.

[3] M. A. Yucel, R. Y. Turan: *Areal change detection and 3D modeling of mine lakes using high-resolution unmanned aerial vehicle images.* Arabian Journal for Science and Engineering *41* (2016), No. 12, 4867–4878.

[4] S. Harwin, A. Lucieer, J. Osborn: *The impact of the calibration method on the accuracy of point clouds derived using unmanned aerial vehicle multi-view stereopsis.* Remote Sensing *7* (2015), No. 9, 11933–11953.

[5] J. R. Alonso, A. Fernández, J. A. Ferrari: *Reconstruction of perspective shifts and refocusing of a three-dimensional scene from a multi-focus image stack.* Applied Optics *55* (2016), No. 9, 2380–2386.

[6] H. Claes, J. Soete, K. van Noten, H. El-Desouky, M. M. Erthal, F. Vanhaecke, M. Özkul, R. Swennen: *Sedimentology, three-dimensional geobody reconstruction and carbon dioxide origin of pleistocene travertine deposits in the Ballık area (south-west Turkey).* Sedimentology *62* (2015), No. 5, 1408–1445.

[7] J. A. Sereno, H. Lee: *The contribution of segmental and tonal information in Mandarin spoken word processing.* Language and Speech *58* (2015), No. 2, 131–151.

[8] G. Franchi, J. Angulo: *Bagging stochastic watershed on natural color image segmentation.* International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing (ISMM), 27–29 May, Reykjavik, Iceland, Springer Nature LNCS *9082* (2015), 422–433.

[9] C. Duval, M. de Tayrac, K. Michaud, F. Cabillic, C. Paquet, P. V. Gould, S. Saikali: *Automated analysis of 1p/19q status by FISH in oligodendroglial tumors: Rationale and proposal of an algorithm.* PloS one *10* (2015), No. 7, e0132125.

[10] C. Mouelhi, J. Saint-Pierre: *The most relevant value creation indicator under competitive dynamics of the firm.* SSRN (2013), Electronic copy available at: http://ssrn.com/abstract=2266050.

[11] V. Belmonti, A. Berthoz, G. Cioni, S. Fiori, A. Guzzetta: *Navigation strategies as revealed by error patterns on the Magic Carpet test in children with cerebral palsy.* Frontiers in Psychology (2015), No. 6, paper 880.

[12] E. Kwak, S. Ahn, J. Jaworski: *Microfabrication of custom collagen structures ca-*

pable of guiding cell morphology and alignment. Biomacromolecules *16* (2015), No. 6, 1761–1770.

[13] J. L. Junkins, H. Bang: *Maneuver and vibration control of hybrid coordinate systems using Lyapunov stability theory.* Journal of Guidance, Control, and Dynamics *16* (1993), No. 4, 668–676.

[14] N. D. Serej, A. Ahmadian, S. Kasaei, S. M. Sadrehosseini, P. Farnia: *A robust keypoint extraction and matching algorithm based on wavelet transform and information theory for point-based registration in endoscopic sinus cavity data.* Signal, Image and Video Processing *10* (2016), No. 5, 983–991.

Received July 12, 2017