BUILDING EXTRACTION BASED ON DENSE STEREO MATCH AND EDISON ALGORITHM

BingXuan GUO [a], ZhiChao ZHANG [a], YuanZheng SHAO [a], Qi Li [a]

[a] State Key Lab of Information Engineering in Surveying Mapping and Remote Sensing,Wuhan University, China
-mobilemap@gmail.com

Commission WGS-PS, WG III/4

KEY WORDS: Extraction, Digital, Photogrammetry, Semi-automation, Dense Stereo Match, EDISON, Building Extraction

ABSTRACT:

Based on the high-resolution image data in urban areas, the paper investigate the key technologies in building detection and reconstruction, particularly focused on the Building Model Driving (BMD) algorithm for hip-roof building and box building. In this case, the boundaries of building rooftop are straight lines. The main methods, including generating the DSM data by a graph-cut based image dense stereo match under the energy minimization framework, constructing the coarse building model by the BMD algorithm, obtaining the 2D Building Edge Feature Vector (BEFV) by the EDISON algorithm. In order to refine the coarse model, projecting the coarse 3D building model to the original image and making a constraint to the edge of building rooftop by the BEFV. An integration has been made under all the above method to generate a semi-automatic building extracting system.

1. INTRODUCTION

Digital image acquisition of high resolution, accuracy and multi-spectrum and real-time imagery is coming true, how quickly and efficiently to survey the interested targets and obtain the information from these images has become an active research topic in digital photogrammetry. Till now, although many experts of diverse background have struggled for many years to create a automatic building extraction system, (Baillard, C., Schmid, C, 1999) there is still no one adaptive enough to handle most remotely sensed data. So building extraction semi-automatically from the high-resolution images is a more practical aspect. The pattern recognition ability of human in conjunction with the high process speed of computer can make a more practical building extraction system in the foreseeable future.

The semi-automatically extraction efficiency and accuracy of the building model from the images is limited for the lack of the excellent building model reconstruction algorithm. Previous works has been done in this field, a excellent review of detection of buildings from aerial space images has been made by Prof.Grun, etc (Grun et al, 1995; Robert W. Carroll), some more work is described in (Fua, 1996; Henricsson et al, 1996; Weidner, 1996). In recent years, more and more new kinds of inputs, such as range images and spectrum images are used. In this paper, we still focus on the use of stereo images.

In this paper, the reconstruction of buildings is split into three steps: in the first step the origin stereopair are preprocessed, including DSM generation, BMD algorithm model extraction. Then a coarse model is recovered, which consists of the three-dimensional edges of the building and the height of the building. In the final step, the 3D edges of the coarse model are reprojected to the original images to refine the model. In the following, the four steps are described in detail.

In Sections 2 the related technologies and concept , such as EDISON algorithm, disparity space image and dense stereo match, are considered. Section 3 describes the strategy of the Building Model Driving (BMD) algorithm. Section 4 presents the systematic integration which illustrates the full process. A proving experiment is also performed in section 5 and the conclusion is drawn in section 6.

2. BACKGROUND

2.1 EDISON algorithm

As the edge detection is one of the bases of our buildings extraction algorithm, we made much more efforts on it. After many tests and comparisons, we introduce EDISON algorithm to detect edge of the images. (P. Meer, B. Georgescu, 2001)

With few exceptions, the fundamental assumption of all step edge detectors is that the regions on either side of an edge are constant in colors or intensity. Much more effort has gone into making them robust to noise, but the noise is assumed to have statistically simple properties.

Convolution masks are ideal for realizing such assumption because the sign of the weight at a pixel tells us what side of the edge it is hypothesized to be on. We can think of a convolution as finding the weighted mean of each side and then computing the distance between the two means (Eric N. Mortensen, 2001; Mark A. Ruzon, 1999).While this assumption holds well enough for many applications, it does not hold in all cases. For instance, as scale increases, it is more likely that the weighted mean of each side will not be meaningful because an operator will include image features unrelated to the edge. This observation is even truer of color images. When only intensities are involved, the average over a large window is still perceptually meaningful because intensities are totally ordered. In color images, there is no such ordering, so the “mean color” of a large window may have little perceptual similarity to any of the colors in it.

we introduced EDISON detection algorithm to color remote sensing images. As reported recently, EDISON works well in edge detection of images used in other fields. This algorithm proposes an image segmentation and confidence based edge detection model.

Besides the above, this edge model has other advantages. The first is a lack of false negatives compared to other models false negatives result from a failure to “take into account all possible intensity variations that might accompany a step edge in practice”. Almost all of these variations are modelled implicitly. In homogeneities can be uncorrelated (due to noise) or correlated (due to texture) without affecting performance. The second benefit is that using distributions creates a unifying framework for edge detection in binary, grey-scale, color, or multi-spectral images, so long as a meaningful ground distance is defined.

2.2 Dense Stereo Match

In order to generate DSM, under a energy minimization framework, a dense stereo match between stereopair is taken. In the case of dense stereo match. The concept of a disparity space image or DSI is important. In general, a DSI is any image or function defined over a continuous or discretized version of disparity space (x, y, d). (Scharstein, D. and Szeliski, 2002)In practice, the DSI usually represents the confidence or log likelihood (i.e., cost) of a particular match implied by d(x, y).

The goal of a stereo correspondence algorithm is then to produce a univalued function in disparity space d(x, y) that best describes the shape of the surfaces in the scene. This can be viewed as finding a surface embedded in the disparity space image that has some optimality property, such as lowest cost and best (piecewise) smoothness. So a stereo correspondence can be formed to a energey minimization framework.

Under the DSI, we test various energy minimization algorithm and find among the common ones, graph cut works best but causes most computation cost. In order to get the best DSM data, we use In this case, we preprocess the images to generate DSM and save to DSM database. Therefore, the huge computation cost has less effect on the efficiency of our system.

3. BUILDING MODEL DRIVING (BMD) ALGORITHM

3.1 Geometric building model

For the task of building reconstruction, we first of all have to define the geometric building model. This model defines the hip-roof building and box building, which can be dealt with. (Brenner) comes to the conclusion, that the combined parametric models were the most suitable ones for automatic or semiautomatic building reconstruction. Both provide a good compromise between the reconstruction effort and the number of supplied real world building types. Combined parametric models imply geometric and topologic characteristics. They are based on constructive solid geometry (CSG). A large number of different building types can be modeled, especially buildings with different eave heights. In this paper, due to the low quantity of the DSM, we only describe the hip-roof building and box building model.

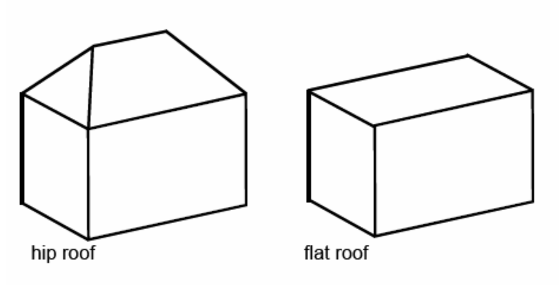

Figure 1: Geometric modal of the two kinds of building in this paper

Object recognition or reconstruction, in general, presumes knowledge about the perceived objects by some kind of object model. These object models can be considered as abstractions of real world objects. As model definition it is important to find balance between correctness and tractability i.e. the results given by the model must be adequate both in terms of the solution attained and the cost to attain the solution (Streilein, 1996). A priori knowledge or in other words constraints can be introduced by applying a very rigid building model. A rigid building model restricts the search space, which has to be examined to find a solution. On the other hand these models limit the number of possible building types which can be represented by a single model. In order to deal with the large architectural variations of building shapes, the utilized model should be as general as possible. Since most buildings are bounded by a set of planar surfaces and straight lines, in our approach a building is represented by a general polyhedron. Additional constraints are defined by the assumption that the coordinates of the given ground plan are correct and the borders of the roof are exactly defined by this ground plan. This provides sufficient restrictions to enable the reconstruction of buildings without loosing the possibility to deal with very complex buildings.

Two types of representation are feasible to describe the reconstructed buildings. The boundary representation (BRep) is probably the most widespread type of 3D representation. Many algorithms are available for computing physical properties or visualizations from that representation. The object is represented by its surface, which is decomposed into a set of faces, edges and vertices. The topology is additionally described by a set of relations which indicate how the faces, edges and vertices are connected to each other. In constructive solid geometry (CSG) simple primitives are combined by means of Boolean set operators. A CSG representation always results in valid 3D objects, i.e. in contrast to a BRep no topological check has to be performed in order to guarantee the closeness of the object surface.

3.2 Model Selection and Parameter Estimation

Each building is described by a combination of one or more basic building primitives. Each of them consists of a cuboid element with different roof types flat roof and hip roof. This type of representation is similar to the one used by (Englert and G¨ulch, 1996). In order to reconstruct more complex buildings, first the complete building has to be decomposed into these basic structures. This step can be realized automatically by the analysis of the given ground plan. Figure 6 shows the result of a ground plan decomposition into rectangular structures. Every rectangle defines one building primitive. Since position, orientation and horizontal extension of each cuboid is already defined by each rectangle, only the height of every cuboid as well as roof type and roof slope have to be determined as remaining parameters for the building primitives.

The parameters of the building primitives are estimated by a least squares adjustment, which minimizes the distances between the DSM surface and the corresponding building primitive, i.e. the building primitives are fit to the DSM surface. In order to apply the least squares adjustment first the appropriate model has to be selected. Additionally roof regions which do not fit to the selected model have to be excluded from the least squares adjustment to avoid gross errors of the estimated parameters. For both tasks the result of a segmentation of the DSM are used. This DSM segmentation into planar surfaces is supported by introducing ground plan information. Of course the given ground plan restricts the extension of the DSM area which has to be examined.More important, the implemented segmentation within each ground plan area can be based on the direction of the surface normals of the DSM, since possible orientations of planar surfaces to be extracted are predefined by the outline of the building. This is motivated by the observation that the direction of the unit normal vector of a possible roof plane emerging from an element of the ground plan has to be perpendicular to this segment. Hence, the different segments of the ground plan polygon are used to trigger the segmentation of a planar surface with a projected normal vector perpendicular to this element. A more detailed description of the segmentation process can be found in (Haala et al., 1997).

4. SYSTEM INTEGRATION

We develop a system to integrate all the above method, The reconstruction of buildings is split into four steps:
in the first step the origin stereopair are preprocessed to generate the DSM data. Then a detection of the image edge is carried out. In the third step, a coarse model is recovered, which consists of the three-dimensional edges of the building. In the final step, the coarse model is projected to the original images to refine the model. In the following the four steps are described in detail.

4.1 Stereopair DSM generation

Because of the huge dense match cost of graph cut algorithm, we provide a batch process on the DSM generation of all the images. constructing the Disparity Space Image (DSI), then the energy minimization framework, minimized by the graph-cut algorithm, is taken to generate the corresponding DSM data. The generated DSM data contains the relative initial three-dimensional coordinates of the building.

4.2 Edge detection of building rooftop

The second step, by selecting the appropriate parameters, the image segmentation is applied with the EDISON algorithm, which separates the building areas from the background areas. Then the EDISON algorithm is used again to detect the building edges of the stereopair. After removing the noise of edge segment, line fitting and end points competition are introduced to optimize the edge. The BEFV is generated after the optimization.

4.3 Coarse model extraction

In this step, the coarse building models are generated according to the DSM data acquired from the first step. The left image of the original stereopair is chosen as a part of the outer interface to the user, so that the user can clearly recognize the building area. In the course of the practical implementation, on the outer interface to the user, the only operation that the user needed to do is to click the mouse in the interior of the building areas. On the inner interface, the corresponding areas in DSM data is selected, then based on the selected areas, the BMD algorithm is applied to produce the coarse building model. The property of the BMD algorithm is described as follows: Automatic building roof planar separation algorithm. Normal of each point in DSM is used as voting to detect planar. Flowing by isolate point assignment, planar joining can optimize the detected planar. And multi-thresholds are used to make algorithm more robust. Planar edge points competition strategy is an important mechanism which can eliminate the bad influence caused by the order of planar detected out. The ridge lines are finally detected by intersection of two planar. The coarse building models possess the initial attributes of the building, such as the length, width, height, footprint, the ridge line of the hip-roof building.

4.4 Refine coarse model

The final step, using the BEFV generated in the second step, the coarse model could be refined. project the edge of the coarse three-dimensional model to the original stereopair image to produce a line buffer in the corresponding BEFV, the WLDC algorithm is proposed to find the corresponding line segment in the stereopair, the WLDC algorithm is described as follows: To get the direction of the optimal line segment, the length and the direction of each line segment in the line buffer is taken as the parameters of a weight function: W=f(l, d), the line segment with bigger W has bigger influence on the optimal line segment's direction. A pair of corresponding optimal line segment is generated after executing the WLDC algorithm. Then the precise three-dimensional coordinates of the building edges can be derived by a space intersection, the coarse model of the building can be refined to a more precise three-dimensional model.

5. TEST

We have tested our integrated system .A complete process introduced in section four has been made on a stereopair.
The stereopair in Figure 1 is taken as the input data to test the integrated system.

For the stereopair in Figure 2, we preprocess it under the energy minimization framework using graph cut algorithm, a huge but acceptable computation will be carried out. Then we display it in a disparity map image type in Figure 3. The result of the edge detection using Edison algorithm is in Figure 4. After these, project the dsm data into the edge map to refine the coarse model generated in 4.3.

Figure 2: stereopair including box building and hip-roof building.(taking from the URL:www.middlebury.edu/stereo.)

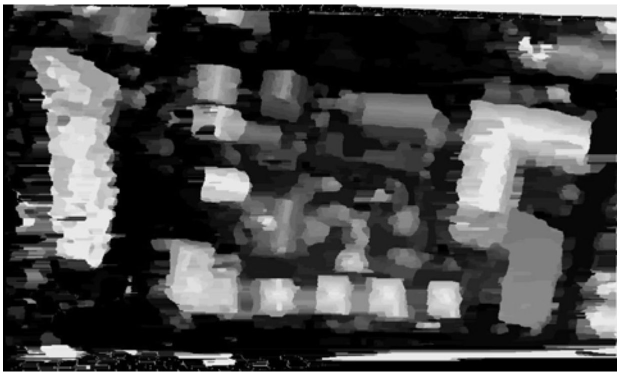

Figure 3: disparity map image generated using graph-cut algorithm dense stereo match.

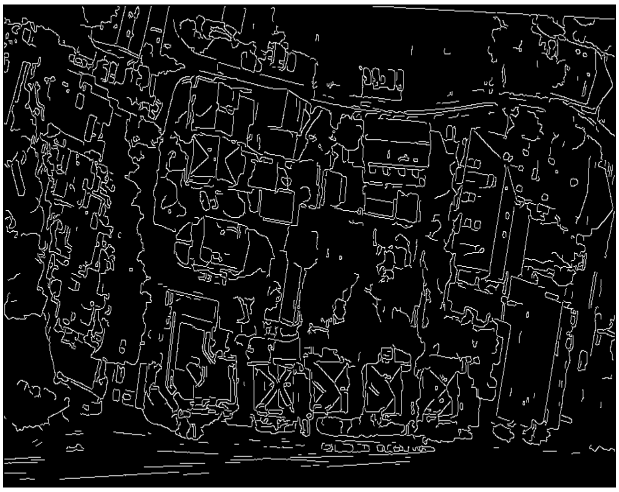

Figure 4: image edge map generated by EDISON algorithm

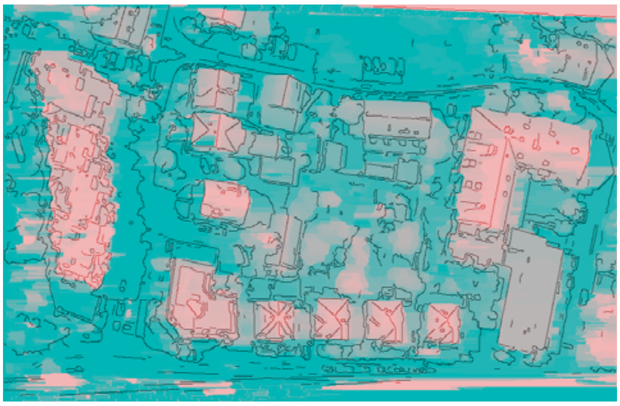

Figure 5: project the dsm data to the left (or right)image edge map to refine the edge fo the coarse model

6. CONCLUSIONS

By using the algorithm and the processes proposed in the paper, we design an experiment to prove the efficiency and the increased success rate of the building extraction. Currently only the box building and hip-roof building been extracted. Taking the same stereopair as the input images, compared with the ground truth, the accuracy and efficiency of our result is comparable to the result of the commercial digital photogrammetric workstation , and what’s more, the degree of automation is higher than the above two software.

Due to the self-occlusion and at step edges.(Brenner, Haala 1998) Thus, the lateral dimension of the building can not derive high accuracy, especially in urban areas. The future work will be focus on the range image because its high accuracy.

Based on the generation of the DSM data from the stereopair, the precise building models are reconstructed by using the BMD algorithm. The experiment showed that our method of building extraction posses three advantages: the high degree of automation, the high success rate and the comparable accuracy. Our method can be competed with the optimal algorithm and software by virtue of the above advantages.

REFERENCES

Baillard, C., Schmid, C., Zisserman, A. and Fitzgibbon, A., 1999. Automatic line matching and 3D reconstruction of buildings from multiple views. In: *IAPRS*, Vol. 32 Part 3-2W5.

Baltsavias, E., Gr¨un, A., and VanGool, L. (eds), 2001. *Automatic Extraction of Man-Made Objects from Aerial and Satellite Images (III)*. Balkema Publishers, Netherlands.

Bauer, J., Karner, K., Schindler, K., Klaus, A. and Zach,C., 2003. Segmentation of Building Models from Dense 3D Point-Clouds. In: *27th Workshop of the Austrian Association for Pattern Recognition*.

Brenner, C., 2000. Towards Fully Automatatic Generation of City Models. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. (33) B3/1, pp. 85–92.

Brenner, C. and Haala, N., 1998. Rapid acquisition of virtual reality city models from multiple data sources. In: *H. Chikatsu and E. Shimizu (eds)*, IAPRS Vol. 32 Part 5, pp. 323.330.

Bruzzone, L., and D. Prieto, 2000. Automatic analysis of the ifference image for unsupervised change detection, *IEEE ransactions on Geoscience and Remote Sensing*, 38(3):pp. 171-1182.

Bruzzone, L., and S. Serpico, 1997. An iterative technique for he detection of land-cover transitions in multitemporal remote ensing images, *IEEE Transactions on Geoscience and Remote* Sensing, 35(4):pp. 858-867.

Canny, J., 1986. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine intelligence*, 6(6):pp. 679-698.

C. Christoudias, B. Georgescu, P. Meer, 2002. Synergism in low-level vision. *16th International Conference on Pattern Recognition*, Quebec City, Canada, vol. IV, 150-155.

Dick, A., Torr, P. and Cipolla, R., 2002. A Bayesian Estimation of Building Shape Using MCMC. In: *Seventh European Conference on Computer Vision*, Vol. II, pp. 852–866.

Dowman, I., 1998. Automated procedures for integration of satellite images and map data for change detection: The archangel project, *GIS-between visions and applications. IAPRS*, 32(4):pp. 162-169.

Estes, J., 1992. Technology and policy issues impact global monitoring, *GIS world*, 5(10): pp.52-55

Fuchs, C., Gülch, E. & Förstner, W. 1998: OEEPE Survey on 3DCity Models. OEEPE Official Publication No. 35.

Brenner, C. 2000: Towards Fully Automatic Generation of City Models. *International Archives of Photogrammetry & Remote Sensing*, vol. 33, part B3, pp. 85-92, Amsterdam.

Fischler, M. and Bolles, R., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM 24*(6), pp. 381–395.

Faugeras, O. and Luong, Q.-T., 2001. *The Geometry of Multiple Images*. MIT Press, Cambridge, USA.

Förstner, W. 1999: 3D-City Models: Automatic and Semiautomatic Acquisition Methods.

D. Fritsch and R. Spiller (eds.), Photogrammetric Week `99, *Wichmann Verlag*, pp. 291–303.

Fotogrametria y Teledeteccion. Baillard, C. and Zisserman, A., 1999. Automatic reconstruction of piecewise planar models from multiple views. In: Proc. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 559.565.

Gr¨un, A. andWang, X., 2001. News from CyberCity Modeler. In: *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, Balkema Publishers, Lisse, The Netherlands, pp. 93–101.

Gr¨un, A. and Dan, H., 1997. TOBAGO . a topology builder for the automated generation of building models. In: *A. Gr¨un, E. Baltsavias and O. Henricsson (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkh¨auser, Basel, pp. 149.160.

G¨ulch, E., M¨uller, H. and L¨abe, T., 1999. Integration of Automatic Processes into Semi-Automatic Building Extraction. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. (32) 3-2W5, pp. 177–186.

Habib, A., and M. Morgan, 2002. Epipolar image resampling from push-broom imagery: investigation and preliminary implementation, *Korean Electronics and Telecommunications Research Institute (ETRI)*, Daejeon, Korea, 107 p.

Habib, A., and R. Al-Ruzouq, 2004. Line-based modified iterated Hough transform for automatic registration of multisource imagery, *Photogrammetric Record*, 19(105):pp. 5–21.

Habib, A., M. Morgan, and Y. Lee, 2001a. Integrating data from terrestrial mobile mapping systems and aerial imagery for change detection purposes, *Proceedings of the Third Mobile Mapping Symposium*, 3-5 January, Cairo, Egypt, unpaginated CD-ROM.

Habib, A., Y., Lee, and M. Morgan, 2001b. Surface matching and change detection using modified iterative Hough transform for robust parameter estimation, *Photogrammetric Record*, 17 (98):pp. 303-315.

Kraus, K. & Pfeifer, N. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry & Remote Sensing* 53, pp. 193-203.

Kulschewski, K., 1999. Recognition of Buildings Using a Dynamic Bayesian Network. In: *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, SMATI ’99, M¨unchen*, pp. 121–132.

Li, D., S. Haigang, and X. Ping, 2002. Automatic change detection of geo-spatial data from imagery. *Mapping and Remote Sensing, commission II, IC WG II/ IV*:pp. 245-252.

Lillesand, T., and R. Kiefer. 2000. Remote sensing and image interpretation, *4thed. John Wiley and Sons*, New York, 724 p.

Mayer, H., 1999. Automatic Object Extraction from Aerial Imagery – A Survey Focusing on Buildings. *Computer Vision and Image Understanding 74(2)*, pp. 138–149.

P. Meer, B. Georgescu, 2001. Edge detection with embedded confidence. *IEEE Trans. Pattern Anal. Machine Intell.*, 23, 1351-1365, December.

Rottensteiner, F., 2003. Automatic Generation of High Quality Building Models from Lidar Data. *IEEE Computer Graphics and Applications* 23(6), pp. 42–50.

Scharstein, D. and Szeliski, R., 2002. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision* 47(1), pp. 7–42.

Shufelt, J., 1999. Performance Evaluation and Analysis of Monocular Building Extraction from Aerial Imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(4), pp. 311–326.

Singh, A., 1989. Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing*, 10(6):pp. 989-1003.

Schickler, W. & Thorpe, A. 2001: Surface estimation based on LIDAR. *Proceedings of the ASPRS Annual Conference* St. Louis, Missouri, April 2001.

Townshend, J., C. Justice, C. Gurny, and J. McManus, 1992. The impact of misregistration on change detection, *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):pp. 1054-1060.

Weidner, U. and F¨orstner, W., 1995. Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal of Photogrammetry and Remote Sensing* 50, pp. 38–49.

Y. Boykov, O. Veksler, and R. Zabih. 2001. *Fast approximate energy minimization via graph cuts*. IEEE TPAMI, 23(11):1222–1239,