

Body-wide Anatomy Recognition in PET/CT Images

Huiqian Wang^{a,b}, Jayaram K. Udupa^b, Dewey Odhner^b, Yubing Tong^b, Liming Zhao^{b,c}, Drew A. Torigian^b

^aCollege of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China; ^bMedical Image Processing Group Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, United States; ^cLaboratory of Materials and Metallurgy, College of Materials Science and Engineering, Chongqing University, Chongqing 400030, P. R. China.

ABSTRACT

With the rapid growth of positron emission tomography/computed tomography (PET/CT)-based medical applications, body-wide anatomy recognition on whole-body PET/CT images becomes crucial for quantifying body-wide disease burden. This, however, is a challenging problem and seldom studied due to unclear anatomy reference frame and low spatial resolution of PET images as well as low contrast and spatial resolution of the associated low-dose CT images. We previously developed an automatic anatomy recognition (AAR) system [15] whose applicability was demonstrated on diagnostic computed tomography (CT) and magnetic resonance (MR) images in different body regions on 35 objects. The aim of the present work is to investigate strategies for adapting the previous AAR system to low-dose CT and PET images toward automated body-wide disease quantification. Our adaptation of the previous AAR methodology to PET/CT images in this paper focuses on 16 objects in three body regions – thorax, abdomen, and pelvis – and consists of the following steps: collecting whole-body PET/CT images from existing patient image databases, delineating all objects in these images, modifying the previous hierarchical models built from diagnostic CT images to account for differences in appearance in low-dose CT and PET images, automatically locating objects in these images following object hierarchy, and evaluating performance. Our preliminary evaluations indicate that the performance of the AAR approach on low-dose CT images achieves object localization accuracy within about 2 voxels, which is comparable to the accuracies achieved on diagnostic contrast-enhanced CT images. Object recognition on low-dose CT images from PET/CT examinations without requiring diagnostic contrast-enhanced CT seems feasible.

Keywords: image segmentation, whole body, PET/CT, object recognition, anatomy recognition, fuzzy models

1. INTRODUCTION

Whole-body PET/CT scanning provides co-registered functional and anatomic images of patients in one single procedure. It has recently become a standard method of imaging for the management of oncology patients. With the rapid growth of PET/CT-based medical applications, quantification of body-wide disease burden in whole-body PET/CT images becomes crucial. At present, this quantification is done mostly manually by sampling regions of interest (ROIs) in different parts of the body on CT images and estimating PET Standardized Uptake Values (SUVs) in the ROIs. This approach is inefficient, subjective, suffers from intra- and inter-operator variability, introduces sampling bias in estimating the disease, and may not offer a full body-wide impression of the disease. Automated quantification of disease burden is therefore needed for full and effective use of the information provided in PET/CT images in order to optimize disease management and patient care. A critical first step toward this goal is automatic body-wide segmentation of anatomy in whole-body PET/CT images. If this becomes feasible, then the pathological regions can be subsequently anatomically localized and segmented, and their extent and involvement in various anatomic objects can be quantified automatically.

Anatomy segmentation however is very challenging on PET/CT images due to an unclear anatomy reference frame, the low spatial resolution of PET images, and the low contrast and spatial resolution of the unenhanced low-dose CT images. Many image segmentation methods have been investigated and applied to PET/CT clinical applications, mainly for segmenting lesions, including thresholding, gradient-based, region growing, clustering, deformable model, and graph-based approaches [1-3]. However, most published papers focus only on pathological region recognition and not anatomy recognition, exceptions being Refs [4-8]. The latter however have confined to mostly one object in one body region and did not consider a whole body region or the whole body. Although many papers have been published on segmenting

multiple objects in diagnostic quality contrast-enhanced CT images [9-14], they have all focused upon a specific body region, and have not demonstrated their applicability body-wide or even body-region-wide in diagnostic CT or PET/CT images.

We previously developed an automatic anatomy recognition (AAR) system [15] whose applicability was demonstrated on diagnostic CT and MR images in three different body regions on 35 objects. It consists of two related processes: *recognition* and *delineation* [15]. Recognition is a high-level process of determining the whereabouts of the object in the image. Delineation is a low-level process of determining the precise spatial extent of the object in the image. With acceptable recognition, good delineation results can be obtained. In the previous paper [15], we demonstrated that the AAR methodology is effective in different body regions on a variety of objects in diagnostic CT and MR images.

The aim of the present work is to investigate strategies for adapting the previous AAR system to PET/CT images. We focus only on the recognition task and investigate ways to achieve best localization of objects throughout the body in whole-body PET/CT images.

2. WHOLE-BODY AAR

In the AAR approach, we build a fuzzy anatomy model of the body region with all its major objects arranged in a hierarchy to subsequently hierarchically recognize objects in any given image. Subsequent to recognition, model-driven iterative relative fuzzy connectedness (IRFC) algorithms are used to delineate the objects in the image, also following the hierarchical order. In the previous paper [15], we demonstrated that good recognition can guarantee good delineation accuracy in diagnostic CT and MR images. In this paper, we focus only on the recognition task and investigate it closely to achieve good object localization on PET/CT images. In the rest of this description, when we refer to *ICT*, it is understood that the reference is to low-dose unenhanced CT images of PET/CT acquisitions. Diagnostic contrast-enhanced CT will be referred to as *dCT* throughout. Our adaptation of the original AAR approach to PET/CT images in this paper consists of the following two steps: constructing whole-body fuzzy models following AAR principles, and recognizing objects throughout the body by using the whole-body fuzzy anatomy models. We will highlight mainly the differences from the previous AAR approach; please refer to [15] for the original methodology.

2.1 Constructing whole-body fuzzy anatomy models

For a body region B of subject group G , the Fuzzy Anatomy Model, $FAM(B, G)$, is a quintuple: $FAM(B, G) = (H, M, \rho, \lambda, \eta)$. H here is a hierarchy of objects in B , represented as a tree. This tree structure offers an order to specify an offspring object in relation to its parent object. M is a set of fuzzy models, one model for each of the L objects in B , $M = \{FM(O_k) : k = 1, \dots, L\}$. ρ describes the parent-to-offspring relationship in H over G . λ is a set of scale factor ranges, one range for each object, estimated over G . η denotes a set of measurements pertaining to the object assembly in B including intensity properties that are used in delineation.

In our case, B is the “whole body” which for the purpose of this paper is defined as the body torso extending from 15 mm above the lung apices to the inferior aspect of the ischial tuberosities in the pelvis. This region covers the Thorax,

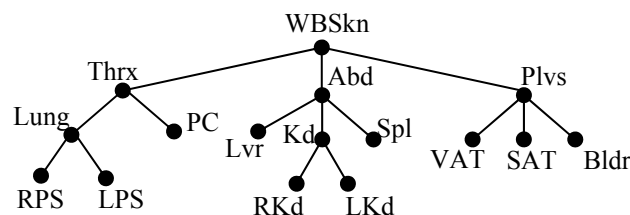


Figure 1. Anatomic hierarchy for whole-body WB. WBSkn: the outer skin boundary enclosing the body torso region; Thrx: the union of thoracic objects; PC: Pericardial region; Lung: Lungs; LPS: Left Lung; RPS: Right lung; Abd: the union of abdominal objects; Lvr: Liver; Kd: Kidney; RKd: Right kidney; LKd: Left kidney; Spl: Spleen; Plvs: the union of pelvic objects; VAT: Visceral Adipose Tissue; SAT: Subcutaneous Adipose Tissue; Bldr: Bladder.

Abdomen, and Pelvis, as per our definition of these body regions. We built $FAM(B, G)$ for whole body by using the anatomical hierarchy depicted in Figure 1. The hierarchies were simplified for thorax and abdomen from those used in [15] in view of the smaller number of objects employed in the present work. We also added a new body region, namely the pelvis. In Figure 1, Thrx, Abd, and Plvs represent composite objects denoting the union of the considered objects in the Thorax, Abdomen, and Pelvis, respectively.

The fuzzy model of each object encodes the variations in the object's form over G and is independent of image intensity and modality. As a result, M can be used in different modalities and so also ρ and λ . The only component that needs change when using $FAM(B, G)$ in different modalities for a given B and G is η , which can be updated with minimal retraining. The fuzzy model $FM(O_k)$ of each object O_k is built from training binary images as described in [15]. See [15] also for how ρ , λ , and η are estimated from given training binary and gray images.

2.2 Recognizing objects

The recognition process proceeds hierarchically following the assumed object hierarchy. The goal of this process is to find an optimal pose for the fuzzy model of every object in the given test image. In this paper, we adopted the thresholded optimal search method of [15] for recognition. In the training process, an optimal threshold interval is determined automatically for each object. This learned optimal interval is used to threshold the given test image and search for the optimal pose for the model within a certain region in the pose space (as determined by the parent-to-offspring relationship ρ) by minimizing the sum of the volume of false positive and false negative regions between the model and the thresholded binary image. The hierarchical order determines the initial pose for the optimal search.

Recognition in different kinds of images: We tested whole body AAR on I CT images. The optimal threshold intervals are determined for every object for I CT images. The respective optimal threshold is then used in the corresponding image in the recognition stage.

3. EXPERIMENTAL RESULTS

3.1 Data sets

This retrospective study was conducted following approval from the Institutional Review Board at the Hospital of the University of Pennsylvania along with a Health Insurance Portability and Accountability Act (HIPAA) waiver. All image data sets were selected from the existing patient image database of our health system. We employed 20 whole-body [18F]-fluorodeoxyglucose (FDG) PET/CT images for testing our methods (16 for patients with hands over their heads and 4 for patients with hands by their sides while imaging). The image details are as follows. I CT: male subjects of age 31-71 years, voxel size 1.2 mm x 1.2 mm x 4 mm. For both body regions and all organs, we follow the definitions developed in [15]. We considered the following objects: Whole-body skin (WBSkn), Thoracic skin (TSkn), the thoracic organs together as a composite object (Thrx), Pericardial region (PC), Left Pleural Space (LPS), Right Pleural Space (RPS), both Pleural Spaces together as a composite object (Lung), Abdominal skin boundary (ASkn), the abdominal organs together as a composite object (Abd), Liver (Lvr), Right kidney (RKd), Left kidney (LKd), both kidneys together as a composite object (Kd), Spleen (Spl), Pelvic skin (PSkn), the Pelvic organs together as a composite object (Plvs), Visceral Adipose Tissue in the pelvis (VAT), Subcutaneous Adipose Tissue in the pelvis (SAT), and Bladder (Bldr). All objects were delineated in all 3D image sets following strictly their definition using manual painting/ tracing/ editing, iterative live wire, live wire, and thresholding methods.

3.2 Validation of Recognition

In AAR, good recognition should accurately determine the whereabouts of the fuzzy model in the image and thus provide enough accurate seed points for IRFC to yield good delineation. For this consideration, the performance of recognition on PET/CT was evaluated by two metrics: position error and scale error. For a target object, position error indicates the distance between the centers of the fuzzy model at recognition and the known true object.

$$\text{Position error} = \sqrt[3]{(x_{truth} - x_{result})^2 + (y_{truth} - y_{result})^2 + (z_{truth} - z_{result})^2} \quad (1)$$

where $(x_{truth}, y_{truth}, z_{truth})$ is the location of the geometric center of true object, and $(x_{result}, y_{result}, z_{result})$ is the location of the geometric center of the model at recognition.

Size error is expressed as a ratio of the estimated object size (scale) to the true size.

$$\text{Scale error} = \frac{S_{result}}{S_{truth}} \quad (2)$$

where S_{result} and S_{truth} are the scale of the model at recognition and scale of the true object, respectively. Value 1 indicates perfect estimation of the unknown object size.

3.3 Evaluation strategies

For evaluation, we conduct two strategies: (1) Models are built from 8 *ICT* images and then deployed on 8 test images and the process is repeated 3 times. (2) Same as (1) except by using the leave-one-out strategy repeated 16 times.

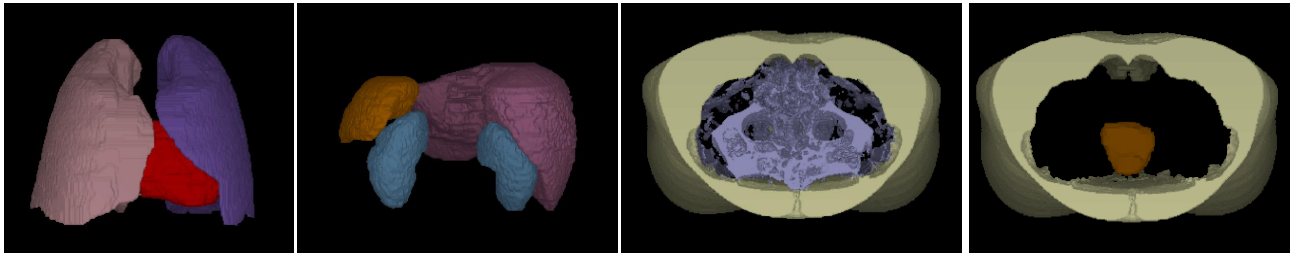


Figure 2. Objects from one training set for each body region are displayed via surface rendering. L to R: Thorax {RPS, PC, LPS}. Abdomen {Spl, LKd, RKd, Lvr}. Pelvis {SAT, VAT}. Pelvis {SAT, Bldr}.



Figure 3. Volume renditions of fuzzy models of objects in different combinations for the three body regions. L to R: Top row: Thorax. 1st picture: {RPS, PC, LPS}. Middle row: Abdomen. 3rd picture: {Lvr, LKd, RKd, Spl}. Bottom row: Abdomen and Pelvis. 1st picture: {Spl}; 2nd picture: {SAT, Bldr}; 3rd picture: {SAT, VAT}; 4th picture: {Bldr}.

Model building: In Figure 2, the objects defined in the image of one of the subjects are displayed for each body region. We have examined all data sets in this manner to ensure the continuousness of the manual segmentations in 3D space. Figure 3 displays fuzzy models of objects in different combinations for the three body regions. The models are built from 8 whole-body *ICT* images which are used in the multifold strategy of evaluation.

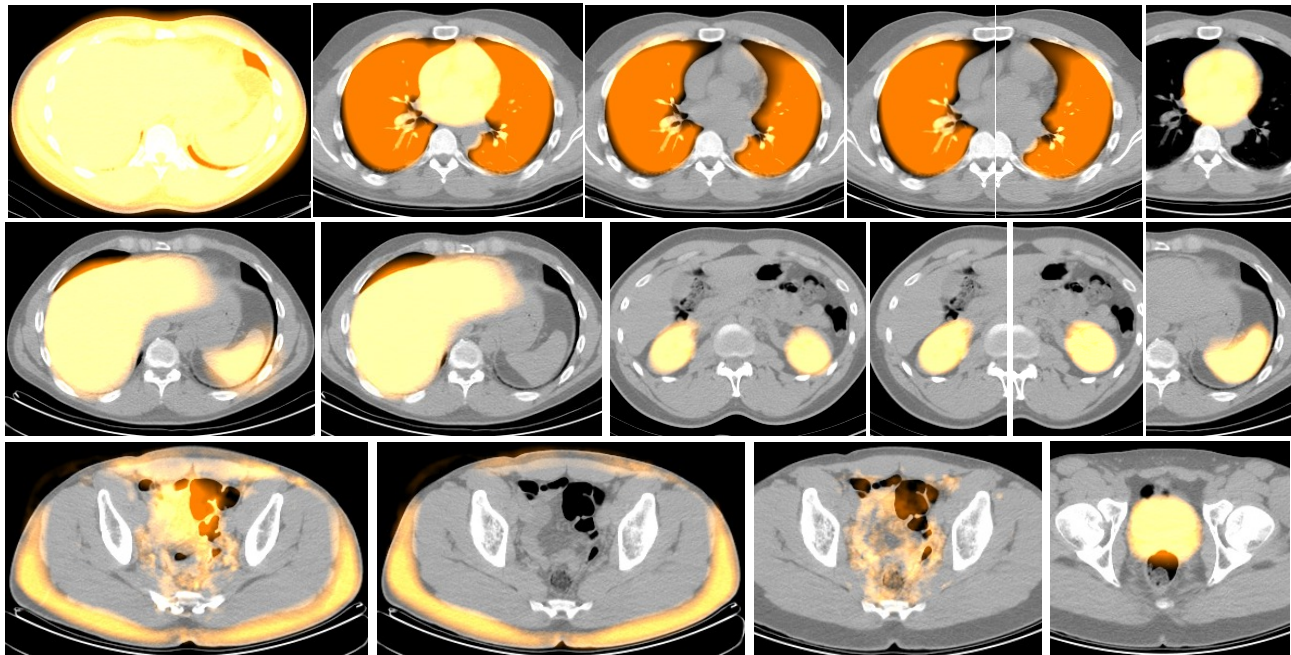


Figure 4. Sample recognition results for Thorax, Abdomen and Plvis on *ICT* images. The models are shown overlaid on the test image slices. Top to Bottom, L to R: WBSkn, Thrx, Lung, RPS, LPS, PC, Abd, Lvr, Kd, RKd, LKd, Spl, Plvs, VAT, SAT, Bldr.

Recognition on *ICT* images using models from *ICT*: Recognition results for *ICT* images are summarized in Figure 4 and Tables 1-2. Figure 4 shows the recognition result of each object for the multifold strategy, in which the recognition result is overlaid on the test image slices. The estimated location error and scale error for the multifold and leave-one-out strategies are shown in Tables 1-2, respectively.

Table 1. Recognition results (mean, standard deviation) for whole-body *ICT* images using multifold strategy.

	WBSkn	Thrx	Lung	RPS	LPS	PC	Abd	Lvr	Kd	RKd	LKd	Spl	Plvs	VAT	SAT	Bldr	Mean
Location error (mm)	4.4	6.8	5.8	7.2	6.1	12.7	11.5	15.1	11.4	8.3	12.9	22.3	12.7	15.3	13.4	9.9	11.0
	2.4	5.5	4.6	4.4	3.2	8.9	4.5	7.5	10.4	9.3	15.2	24.2	7.6	8.0	7.2	5.7	8.0
Size error	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.1	1.0	1.1	1.1	1.0	1.0	1.0	1.0	1.1	1.0
	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.0	0.1	0.0	0.1	0.1

Table 2. Recognition results (mean, standard deviation) for whole-body *ICT* images using leave-one-out strategy.

	WBSkn	Thrx	Lung	RPS	LPS	PC	Abd	Lvr	Kd	RKd	LKd	Spl	Plvs	VAT	SAT	Bldr	Mean
Location error (mm)	4.7	6.1	5.7	5.9	5.2	12.0	11.5	11.0	10.2	5.7	13.1	16.6	13.6	14.2	14.3	8.0	9.9
	2.4	5.5	3.8	3.0	2.2	8.8	3.9	3.8	8.1	4.5	13.6	19.2	7.4	6.5	7.2	4.8	6.5
Size error	1.00	0.98	0.99	1.00	0.99	1.06	1.06	1.08	1.02	1.07	1.07	0.98	1.00	0.99	1.00	0.98	1.02
	0.01	0.02	0.02	0.03	0.04	0.06	0.08	0.07	0.05	0.10	0.09	0.17	0.02	0.09	0.02	0.15	0.06

Overall, for multifold strategy the mean position error is within 11.0 mm (~ 2 voxels) and the scale error is always close to 1. The leave-one-out strategy shows better performance since it has a larger number of training data sets for model building which is favorable for recognition. Generally speaking, AAR shows consistent and excellent recognition performance on whole-body *ICT* images. For Spl, the relatively lower performance might be caused by the small number of training data sets in the multifold strategy. Note that the location error is lower (~ 3 voxels) for leave-one-out strategy using more training data sets.

3.4 Computational considerations

The computational times are estimated on a Dell platform with the following specifications: 4-core Intel Xeon 3.6 GHz CPU with 8 GB RAM and running the Linux-jb18 3.7.10-1.16 operating system. Mean computational time per object for the different steps are as follows: optimal threshold training: 1-5 secs; model building: 1-3 secs; object recognition: 7-32 secs.

3.5 Comparison with other methods

Published works that studied anatomy recognition on *ICT* images are few [4-8]. Among these, Refs [6, 7] did not present segmentation results quantitatively. As the involved objects in these works are different from those considered by us, it is hard for us to draw conclusions on performance. Other methods [9-14] which demonstrated multi-object localization in *dCT* images have similar (or poorer) performance compared to our results on *ICT* images.

4. CONCLUSIONS

In this paper, we investigated strategies for adopting the AAR approach of [15] to whole-body PET/CT images. Despite the low spatial and contrast resolution of *ICT* images from PET/CT, AAR shows good performance on *ICT*, yielding overall object localization within 1-3 voxels. The greatest strength of the approach is the union of two independent and cooperative parts: encoding deep anatomic information in model building and exploitation of this information in hierarchical recognition where optimal threshold intervals are estimated with minimal retraining for different imaging modalities. However, there are some limitations appearing in the experiments. For some objects, like VAT and SAT, the results are not as good as others. This might be caused by the large variations of these objects from subject to subject which is difficult to encode into a model. For Bldr, the different situations of the bladder being full, part full, or empty may also cause difficulty in model building and recognition.

Acknowledgements

The training at University of Pennsylvania of Ms. Huiqian Wang was partly supported by China Scholarship Council and the Fundamental Research Funds for the Central Universities CDJXS11171159. This work was partly supported by an NIH grant HL105212.

REFERENCES

- [1] C. Greco, K. Rosenzweig, G. L. Cascini *et al.*, "Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC)," *Lung Cancer*, 57(2), 125-134 (2007).
- [2] E. Day, J. Betler, D. Parida *et al.*, "A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients," *Medical Physics*, 36(10), 4349-4358 (2009).
- [3] U. Bagci, J. K. Udupa, N. Mendhiratta *et al.*, "Joint Segmentation of Anatomical and Functional Images: Applications in Quantification of Lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT Images," *Medical Image Analysis*, 17(8), 929-945 (2013).
- [4] M. Baiker, J. Milles, J. Dijkstra *et al.*, "Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data," *Medical image analysis*, 14(6), 723-737 (2010).
- [5] A. M. Ali, and A. A. Farag, [Automatic lung segmentation of volumetric low-dose CT scans using graph cuts] Springer, (2008).
- [6] H. Lamecker, S. Zachow, A. Wittmers *et al.*, "Automatic segmentation of mandibles in low-dose CT-data," *International Journal of Computer Assisted Radiology and Surgery*, 1, 393 (2006).

- [7] C.-Y. Hsu, C.-Y. Liu, and C.-M. Chen, "Automatic segmentation of liver PET images," *Computerized Medical Imaging and Graphics*, 32(7), 601-610 (2008).
- [8] Y. Zhan, X. S. Zhou, Z. Peng *et al.*, [Active scheduling of organ detection and segmentation in whole-body medical images] Springer, (2008).
- [9] C. Lu, Y. Zheng, N. Birkbeck *et al.*, [Precise segmentation of multiple organs in CT volumes using learning-based approach and information theory] Springer, (2012).
- [10] M. G. Linguraru, J. A. Pura, V. Pamulapati *et al.*, "Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT," *Medical Image Analysis*, 16(4), 904-914 (2012).
- [11] T. Okada, K. Yokota, M. Hori *et al.*, [Construction of hierarchical multi-organ statistical atlases and their application to multi-organ segmentation from CT images] Springer, (2008).
- [12] C. Chu, M. Oda, T. Kitasaka *et al.*, "Multi-organ segmentation from 3D abdominal CT images using patient-specific weighted-probabilistic atlas." *Proc. SPIE* 8669, 86693Y-86693Y-7.
- [13] A. Criminisi, D. Robertson, E. Konukoglu *et al.*, "Regression forests for efficient anatomy detection and localization in computed tomography scans," *Medical image analysis*, 17(8), 1293-1303 (2013).
- [14] X. Zhou, S. Wang, H. Chen *et al.*, "Automatic localization of solid organs on 3D CT images by a collaborative majority voting decision based on ensemble learning," *Computerized Medical Imaging and Graphics*, 36(4), 304-313 (2012).
- [15] J. K. Udupa, D. Odhner, L. Zhao *et al.*, "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," *Medical Image Analysis*, 18(5), 752-771 (2014).