Material Decomposition Using Ensemble Learning for Energy-Resolved Computed Tomography

Yanye Lu*, Markus Kowarschik, Qiushi Ren, Rebecca Fahrig, Joachim Hornegger and Andreas Maier

***Abstract*—Material decomposition facilitates the differentiation of different materials in X-ray imaging. As an alternative to the previous empirical material decomposition methods, we performed material decomposition using ensemble learning methods in this work. Three representative ensemble methods with two decision trees as the base learning algorithms were implemented to perform material decomposition in both simulation study and experimental study. The results were quantitatively evaluated for comparison study. The performance of the base learning algorithms was improved by using appropriate ensemble methods. The results indicate that it is feasible and promising to perform material decomposition using ensemble learning, which is valuable to be further investigated.**

***Index Terms*—Material decomposition, machine learning, ensemble learning, energy-resolved computed tomography (CT).**

I. Introduction

Energy-resolved computed tomography (CT) has been developed to facilitate detecting spectral information from the energy-dependent attenuation properties of objects, potentially increasing the measuring accuracy of the objects. Such technology can be implemented using either energy-resolved detectors with one polychromatic X-ray spectrum or various X-ray spectra with conventional detectors. Nowadays energy-resolved CT attracts more and more research interests, not only because it facilitates quantitative measurements, but also due to its advantage of allowing material decomposition. Material decomposition, which has great potential in medical applications, can decompose the materials that have the same range of attenuation gray values in conventional CT images by analyzing the different attenuation behaviors across various energy bins.

Material decomposition in energy-resolved CT can be performed either before CT reconstruction in projection domain or after CT reconstructions in image domain. A traditional way to perform material decomposition is so called the basis material decomposition (BMD) method [1], which is actually based on the mixture rule. The BMD method proposes that X-ray attenuation coefficients can be represented as the superposition of basis functions, relying on explicitly modeling the energy-resolved imaging system. However, it is difficult to measure polychromatic spectrum information, which limits the accuracy of material composition. The spectral information of projections is typically affected by the spectral distortions caused by X-ray detectors. Furthermore, energy-resolved detection introduces more bias such as noise to the projections. Therefore, empirical calibrations such as spectral weighting or spectrum calibration are employed to get the attenuation behavior of the basis material to facilitate the BMD method. Nevertheless, deriving an explicit transmission model in realistic scans is still challenging due to the nonlinearity and nonparametric models. In addition, the approximation is generally vulnerable to the nonlinearity of the measure system. As an alternative, supervised machine learning can be utilized to model the relationship between spectral measurements and the relevant measurement system. Several related works have been reported in the literature [2][3]. Moreover, we also proposed a learning-based material decomposition pipeline [4][5], which employs machine learning algorithms to generate the material-specific models for material decomposition tasks. The machine learning algorithms are crucial to the pipeline, largely conditioning the success of the material decomposition endeavor.

Ensemble learning with decision trees has been demonstrated great potential in modeling relationship between input observations and output values [6]. Based on the previous studies, the aim of this work is to make an effort to investigate the feasibility and performance of using ensemble learning for material decomposition. Three representative ensemble methods with two decision trees as the base learning algorithms were investigated using the learning-based material decomposition pipeline to perform material decomposition in both simulation study and experimental study. The results were quantitatively evaluated to compare to a BMD-based method for comparison study.

II. Methods

A. Material Decomposition Model Estimation

The aim of material decomposition task is to build material-specific decomposition models to decompose the effective material-specific projection $\hat{p_j}$ from the energy-resolved observations $q_1, \cdots, q_B$. For this goal, a target function of the material decomposition model is defined as follow:

$$\hat{p_j} = D(\mathbf{q}), \tag{1}$$

where $D(\mathbf{q})$ is an as-yet-unknown general function estimator to model the material decomposition approach with the obser-

*Y. Lu is with the Pattern Recognition Lab, Department of Computer Science, University of Erlangen-Nuremberg, 91058 Erlangen, Germany and with the Advanced Therapies, Siemens Healthineers, 91301 Forchheim, Germany. e-mail: (yanye.lu@fau.de).

M. Kowarschik and R. Fahrig are with the Advanced Therapies, Siemens Healthineers, 91301 Forchheim, Germany.

Q. Ren is with the Department of Biomedical Engineering, Peking University, 100871 Beijing, China.

J. Hornegger and A. Maier are with the Pattern Recognition Lab, Department of Computer Science, University of Erlangen-Nuremberg, 91058 Erlangen, Germany.

vations $\mathbf{q} = (q_1, \ldots, q_B)$. A hypothesis model $H(\mathbf{x})$ is defined to approximate $D(\mathbf{q})$, where $\mathbf{x}$ are vectors that numerically represent the observations $\mathbf{q}$. The model constructions rely on various machine learning algorithms, where the hypothesis model $H(\mathbf{x})$ is optimized to find the optimal choice $\overset{*}{H}(\mathbf{x})$ that minimizes the expected value of a loss function $\mathcal{L}$ on the training samples:

$$\overset{*}{H}(\mathbf{x}) = \arg\min_{H(\mathbf{x})} \frac{1}{m} \sum_{i=1}^{m} \mathcal{L}(\hat{\mathbf{y}}, \mathbf{y}^{(\mathbf{i})}). \tag{2}$$

Once the optimal hypothesis model $\overset{*}{H}(\mathbf{x})$ was determined, we can obtain an estimated material decomposition model $\hat{D}(\mathbf{q})$.

B. Leaning-Based Material Decomposition Pipeline

In this work, a learning-based material decomposition pipeline was employed to build material-specific decomposition models for material decomposition tasks. We need the instance-label pairs that consist of energy-resolved projections and their relevant material-specific projections respectively for generating material-specific decomposition models in the learning phase. First the instance-label pairs are split into two categories of training set and test set, then these datasets are processed by a feature extractor to yield feature vectors. These feature vectors from the training sets are applied by machine learning algorithms to learn decomposition models for recognizing the input unit with target output unit. Once the models have been learned, the performance is evaluated by the accuracy of the model on predicting outcome on the test set. In the material decomposition phase, the input energy-resolved projections also need to be processed by the same feature extractor with same parameters that used in the learning phase.

In order to extract information from the spectral measurements, a raw-pixel-value feature extractor was built to extract feature information from energy-resolved projections, extracting the pixel values per-pixel from each energy bin of the input energy-resolved projections to yield the feature vectors. The total number of the features is the number of energy bins, and the total number of the instances is the pixel amount of the input projections.

C. Ensemble Learning with Decision Trees

In contrast to ordinary leaning methods that employ one learner, ensemble learning methods train multiple learners and combine them to learn models. In this work, we investigated three representative ensemble methods of Boosting [7], Bagging [8] and Stacking [9]. The Boosting combines base learners in an iterative way, gradient descent with the shrinkage rate of 0.3 was involved to minimize the residual error for optimizing the performance. The Bagging employs base learning algorithms in parallel using bootstrap sampling, then combines them by averaging their outputs, the iteration was set to 20. The Stacking trains a group of base learners as the first-level learner, whose outputs are regarded as input features of a second-level learner, for which a multi-response linear regression [10] was employed. Random Tree [11] and REPTree [12] were used as the base learning algorithms in this work. Both set 1 as the minimum number of instances per leaf, 0.001 as the minimum variance proportion of train variance for split, as well as the maximum depth of the tree was set to be unlimited.

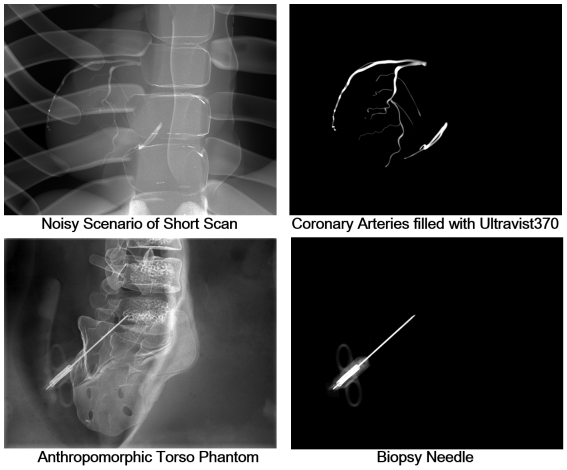

Figure 1. Example images showing the simulation scenarios (top row) and the real scan scenarios (bottom row). The right images are the corresponding material decomposition objects in this study.

D. Experimental Setup

In simulation study, we generated a virtue CT system to simulate the geometry of a research Artis zeego C-arm angiography system (Siemens Healthineers, Forchheim, Germany). The peak voltage was set to 90 kV and the time current product to 2.5 mAs. A flat panel photon-counting energy-resolved detector of 620×480 pixels with a pixel size of 0.4×0.4 mm was simulated for detecting three energy bins (10-40 keV, 40-70 keV and 70-100 keV) with a cross-talk of 3 keV. The source-to-patient distance was 750 mm while the source-to-detector distance was 1200 mm. 200-degree short scans with an angular increment of 1.5-degree were performed to acquire the energy-resolved projections of a modified XCAT phantom using an append buffer based rendering procedure [13]. The energy-dependent X-ray absorption coefficients were obtained from the NIST [14] database. The field of view (FOV) was centered around the heart to focus on the costal arch and the coronary arteries that were filled with Ultravist370. The ground truth of the material decomposition object was simulated as material-specific projections as well. Both noiseless projections and noisy projections with Poisson noise were created respectively. Furthermore, we applied a joint bilateral filter (JBF) [15] as the pre-processing of noise reduction to the noisy data to yield the JBF-denoised data.

In experimental study, we scanned an anthropomorphic torso phantom (Sawbones Europe AB, Malmo, Sweden) containing a SAWBONES spine with a biopsy needle inserted at different peak voltage setting of 40 kVp (0.9 mAs), 70 kVp (0.8 mAs), and 125 kVp (0.2mAs) using an angiography CT system [16]. The flat panel detector was operated with 4x4 binning that allows an effective resolution of 620×480 pixels with a pixel size of 0.616×0.616 mm. The source-to-isocenter

Scenario	Noise Level	Reference Method		Base Learner		Boosting		Bagging		Stacking	
		R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD
Short Scan	Noisy	**0.46±0.12**	0.36±0.19	0.31±0.13	0.30±0.13	0.32±0.13	0.30±0.13	0.44±0.15	**0.39±0.15**	0.31±0.13	0.28±0.12
	JBF-Denoised	0.73±0.13	0.42±0.15	0.71±0.09	0.65±0.10	0.74±0.08	0.69±0.10	**0.82±0.07**	**0.72±0.10**	0.73±0.08	0.61±0.08
	Noiseless	0.85±0.05	0.61±0.14	0.83±0.06	0.79±0.09	0.86±0.06	0.83±0.08	**0.91±0.04**	**0.86±0.08**	0.85±0.04	0.75±0.04
Torso	Noisy	0.56±0.22	0.35±0.19	0.73±0.08	0.67±0.10	0.78±0.05	0.73±0.09	**0.79±0.07**	**0.75±0.10**	0.78±0.08	0.74±0.10
	JBF-Denoised	0.58±0.24	0.36±0.20	0.78±0.08	0.72±0.10	0.83±0.05	0.80±0.09	**0.84±0.04**	**0.80±0.08**	0.83±0.08	0.80±0.10

(a) Random Tree as base learning algorithm

Scenario	Noise Level	Reference Method		Base Learner		Boosting		Bagging		Stacking	
		R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD	R±SD	SSIM±SD
Short Scan	Noisy	0.46±0.12	0.36±0.19	0.47±0.16	0.41±0.15	**0.52±0.16**	**0.44±0.16**	0.51±0.14	0.39±0.12	0.48±0.16	0.41±0.15
	JBF-Denoised	0.73±0.13	0.42±0.15	0.76±0.10	0.67±0.16	**0.83±0.07**	**0.72±0.12**	0.79±0.07	0.57±0.15	0.77±0.08	0.68±0.12
	Noiseless	0.85±0.05	0.61±0.14	0.84±0.09	0.80±0.14	**0.91±0.05**	**0.85±0.10**	0.86±0.03	0.70±0.04	0.86±0.07	0.82±0.10
Torso	Noisy	0.56±0.22	0.35±0.19	0.80±0.08	0.74±0.10	**0.83±0.05**	**0.78±0.09**	0.80±0.07	0.75±0.10	0.79±0.08	0.74±0.10
	JBF-Denoised	0.58±0.24	0.36±0.20	0.84±0.08	0.81±0.10	**0.88±0.05**	**0.84±0.10**	0.85±0.07	0.81±0.10	0.84±0.08	0.80±0.10

(b) REPTree as base learning algorithm

Table I

MEAN R± STANDARD DEVIATION AND MEAN SSIM ± STANDARD DEVIATION ACROSS ALL QUANTITATIVE MEASUREMENTS OF THE MATERIAL DECOMPOSITION RESULTS WITH (A) THE RANDOM TREE AND (B) THE REPTREE.

distance was 700 mm while the source-to-detector distance was 1200 mm. Rotation of 199 degrees short scans with an average angular increment of 1 degree were performed. In order to obtain the ground truth projections of the needle, we performed a fourth scan at 125 kVp (0.2mAs) after careful removal of the needle. Subtraction of the two 125 kVp scans yielded a material-specific projections of the needle for the labeling and the performance evaluation. Similar to the simulation study, the experimental data was pre-processed by the JBF as well.

E. Performance Evaluation

The experiments were performed in accordance with the learning-based material decomposition pipeline. In the learning phase, the mixed feature vectors that concatenate feature vectors from different noise levels were used to learn generic material decomposition models. For each experiment, the original datasets were split into training set and test set using the hold-out validation with a sampling interval of 1 (corresponding hold-out rates: 49.62%). The performance of each method was quantified by the Pearson's correlation coefficient (R) and the structural similarity (SSIM) index [17]. The mean value and standard deviation of the quantitative results from the test sets were used to evaluate the overall performance of the material decomposition models. Furthermore, we used a BMD-based empirical material decomposition estimator with polynomial fitting [18] as the reference method for comparison studies. The machine learning algorithms were implemented by the Waikato Environment for Knowledge Analysis (Weka) [19], and the other methods were implemented in the Java-based framework CONRAD [20].

III. RESULTS

Figure 2 and 3 demonstrate the decomposed material-specific projections of the object materials in the simulation study and experimental study respectively. The contrast of the images was enhanced by histogram stretching with 0.3% saturated pixels in the image and normalized to [0, 1]. It can be seen that, the performance of the base learning algorithms was improved using the Boosting and the Bagging, but deteriorated

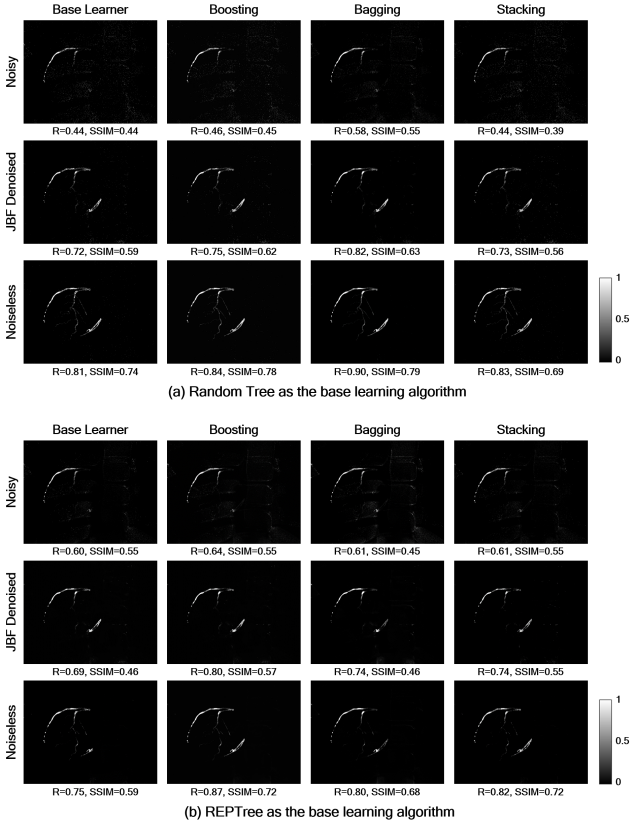

Figure 2. Decomposed material-specific projections of the coronary arteries filled with Ultravist370 in the **simulation** study using the **generic** material decomposition models with (a) the Random Tree and (b) the REPTree. The central projections are presented.

using the Stacking. Table I shows the performance on R and SSIM (average± standard deviation) across all quantitative measurements from the experiments. It can be noticed that, there are two combinations, the Random Tree with the Bagging and the REPTree with the Boosting, that demonstrate superior performance comparing to the other methods.

IV. CONCLUSION

In this work, three representative ensemble methods of the Boosting, the Bagging and the Stacking with two decision trees

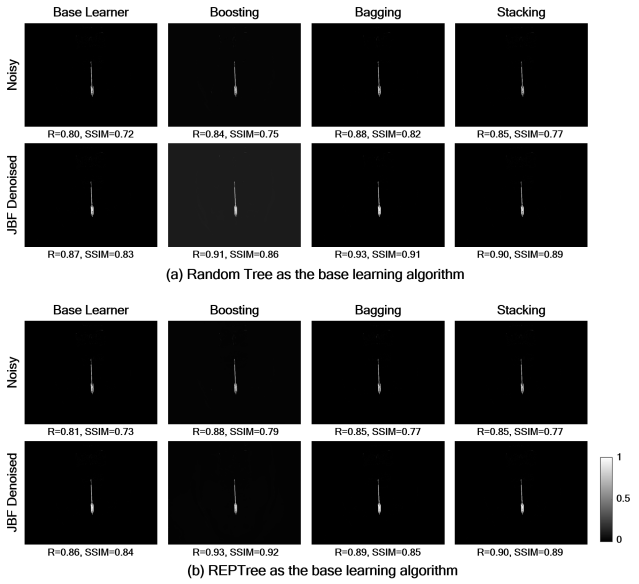

Figure 3. Decomposed material-specific projections of the biopsy needle in the **experimental** study using the **generic** material decomposition models with (a) the Random Tree and (b) the REPTree. The central projections are presented.

were investigated to perform material decomposition in both simulation study and experimental study. The performance of the base learning algorithms was improved by using appropriate ensemble methods, such as the Random Tree with the Bagging and the REPTree with the Boosting. The results indicate that it is feasible and promising to perform material decomposition using ensemble learning, which is valuable to be further investigated.

V. Acknowledgment

The authors gratefully acknowledge the funding supports from the Research Training Group 1773 Heterogeneous Image Systems by the German Research Foundation (DFG), the German Academic Exchange Service (DAAD), and the Siemens Healthineers Advanced Therapies respectively.

Disclaimer: The concepts and information presented in this paper are based on research and are not commercially available.

References

[1] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in X-ray computerised tomography," *Physics in Medicine and Biology*, vol. 21, no. 5, p. 733, 1976.

[2] K. C. Zimmerman and T. G. Schmidt, "Experimental comparison of empirical material decomposition methods for spectral CT," *Physics in Medicine and Biology*, vol. 60, no. 8, pp. 3175–3191, 2015.

[3] M. Touch, D. P. Clark, W. Barber, and C. T. Badea, "A neural network-based method for spectral distortion correction in photon counting X-ray CT," *Physics in Medicine and Biology*, vol. 61, no. 16, p. 6132, 2016.

[4] Y. Lu, J. Geret, M. Unberath, M. Manhart, Q. Ren, R. Fahrig, J. Hornegger, and A. Maier, "Projection-based material decomposition by machine learning using image-based features for computed tomography," in *Proceedings of the Fully3D 2015*, pp. 448–451, June 2015.

[5] Y. Lu, M. Berger, M. Manhart, J. hwan Choi, M. Hoheisel, M. Kowarschik, R. Fahrig, Q. Ren, J. Hornegger, and A. Maier, "Bridge to real data: Empirical multiple material calibration for learning-based material decomposition," in *Proceedings of the 2016 IEEE International Symposium on Biomedical Imaging*, pp. 457–460, 2016.

[6] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*, pp. 1–15, Springer, 2000.

[7] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis*, vol. 38, no. 4, pp. 367 – 378, 2002. Nonlinear Methods and Data Mining.

[8] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[9] "Stacked generalization," *Neural Networks*, vol. 5, no. 2, pp. 241 – 259, 1992.

[10] K. M. Ting and I. H. Witten, "Issues in stacked generalization," *J. Artif. Intell. Res.(JAIR)*, vol. 10, pp. 271–289, 1999.

[11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[12] J. Quinlan, "Simplifying decision trees," *International Journal of Man-Machine Studies*, vol. 27, no. 3, pp. 221–234, 1987.

[13] A. Maier, H. G. Hofmann, C. Schwemmer, J. Hornegger, A. Keil, and R. Fahrig, "Fast simulation of X-ray projections of spline-based surfaces using an append buffer," *Physics in Medicine and Biology*, vol. 57, pp. 6193–6210, 10 2012.

[14] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients," tech. rep., National Inst. of Standards and Technology-PL, Gaithersburg, MD (United States). Ionizing Radiation Div., 1995.

[15] Y. Lu, M. Manhart, O. Taubmann, T. Zobel, Q. Yang, J. hwan Choi, M. Wu, A. Doerfler, R. Fahrig, Q. Ren, J. Hornegger, and A. Maier, "Projection-based denoising method for photon-counting energy-resolving detectors," in *Bildverarbeitung für die Medizin 2015*, (Lübeck, Germany), pp. 137–142, 2015.

[16] K. Müller, S. Datta, M. Ahmad, J.-H. Choi, T. Moore, L. Pung, C. Niebler, G. E. Gold, A. Maier, and R. Fahrig, "Interventional dual-energy imaging¡ªfeasibility of rapid kv-switching on a c-arm CT system," *Medical Physics*, vol. 43, no. 10, pp. 5537–5546, 2016.

[17] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600–612, April 2004.

[18] C. Maaß, E. Meyer, and M. Kachelrieß, "Exact dual energy material decomposition from inconsistent rays (MDIR)," *Medical Physics*, vol. 38, no. 2, pp. 691–700, 2011.

[19] E. Hall, Mark andFrank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: An update," *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.

[20] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig, "CONRAD-A software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, p. 111914, 2013.