Projection-based Material Decomposition by Machine Learning using Image-based Features for Computed Tomography

Yanye Lu*[§], Jan Geret*, Mathias Unberath*[†], Michael Manhart*, Qiushi Ren[§], Rebecca Fahrig[‡],

Joachim Hornegger* and Andreas Maier*†

*Pattern Recognition Lab, Department of Computer Science,

Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany

[†]Erlangen Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany

[‡]Department of Radiology, Stanford University, CA, USA

[§]Department of Biomedical Engineering, Peking University, Beijing, China

Abstract—In this paper, we extracted appropriate image-based features from energy-resolved computed tomography images and incorporated these features in a machine learning material decomposition process to separate bone and contrast agent in projection domain. We performed various simulation experiments to investigate the performance of this new technique. Comparing the feature based approach to a polynomial method, the linear correlation of the Ultravist images to the ground truth in the noise and motion scenario increased from 0.49 to 0.89, the structural similarity index from 0.06 to 0.88 by using appropriate features, respectively. The results suggest that it is possible to decompose materials by using appropriate image-based features.

Index Terms—Material decomposition, Machine learning, Image features, Projection-based

I. INTRODUCTION

Computed Tomography (CT) is a widely used X-ray imaging technique as it allows the measurement of the human body morphology in diagnostic medical imaging. The output of a CT measurement and data reconstruction is typically a set of images representing the X-ray attenuation characteristics of the scanned patient. However, in practice polychromatic Xray sources are commonly employed for medical CT but most of them are processed as mono-energetic CT measurements by conventional CT detectors, which cannot take advantage of the energy information in the X-ray beam.

Recently, energy-selective X-ray detectors are attracting more research interests in Spectral CT development. Such detectors split the polychromatic X-ray spectrum into several energy bins as well as measure the X-ray attenuation of objects in each energy bins separately. Moreover, they enable material decompositon. Firschig et al. successfully distinguished iodine-based contrast agent from all other materials using an energy-selective detector [1]. Le and Molloi evaluated least squares parameter estimation methods for material decomposition with energy discriminating detectors[2]. Wang et al. separated more than two types of materials simultaneously by employing an energy-resolved photon-counting detector with more than two energy windows [3]. Zimmerman and Gilat-Schmidt investigated the use of artificial neural networks for material decomposition and suggested the preliminary experimental feasibility [4].

As an alternative to a linear model, any method that allows the modelling of non-linear processes can be used to estimate a function using machine learning [5]. On the other side, local features, which are only small segments of the original image, can differ very much from global features because of changing neighborhoods. These features contain various information and could serve as training features to build a classifier. In this study, we extracted appropriate image-based features from energy-resolved computed tomography images and incorporate these features in a machine learning material decomposition process to separate bone and contrast agent in projection domain. We performed various simulation experiments to investigate the performance of this new technique. The results are evaluated using linear correlation and the structural similarity index (SSIM) as quantitative metrics.

II. METHODS

The simulation studies are carried out in order to examine if it is possible to do material decomposition in computed tomography using imaged-based features incorporated in machine learning approaches. All methods are implemented in Javabased framework CONRAD [6] and will be made available as open source software.

A. Data generation

The data for the projection images generation rely on the simulation of x-ray images [7]. An X-ray spectrum with the same half layer values as a commercially available C-arm system is used. As shown in Figure 1, we simulated two different scenarios. The motion case includes one heart cycle and the lung motion from full exhale to full inhale; the nomotion case simulates a 200 degree vertical rotation around the torso (short scan). Both projection sets are centered around the heart to set focus on the costal arch and the coronary arteries, which should be decomposed in the results. Energy-dependent X-ray absorption coefficients for elemental data

and compounds such as bone were obtained from the NIST database[8]. The coronary arteries are filled with an iodinebased contrast medium that has a similar density as the Ultravist370 contrast agent. The projection set was simulated with 620x480 pixels with a pixel size of 0.4x0.4 mm. The source-topatient distance 750 mm while the source-to-detector distance 1200 mm. In the case of 200 degree vertical rotation around the torso, we had an average angular increment of 1.5 per projection. A photon-counting energy-resolving detector with three energy bins was used in this simulation. The energy bins were equally spaced between 10 and 100 keV with a spectral bandwidth of 30 keV. This means channel 1 ranged from 10 to 40 keV, channel 2 from 40 to 70 keV and channel 3 from 70 to 100 keV. The simulated cross-talk among the three energy bins was 3 keV. The peakvoltage was set to 90 kV and the time current product to 2.5 mAs, which is comparable to the dose setting per projection in a clinical 3D scan. For every set of projections noisy and noiseless images were created.



Figure 1. Illustration of the simulated projection images. Top row: The 200 degree vertical rotation projection (left) as well as the heart cycle and lung motion projection (right). Bottom row: two materials that need to be decomposed, Ultravist 370 and bone.

B. Image-based Features extraction

Energy-selective photo-counting detectors yield CT projection data with multiple image channels. Therefore, the pixel intensity values of each image channel can be used as training features to build the classifiers [1]. This results in three features per pixel in this study. The second approach is polynomial combinations of the energy channels as features for learning the classifiers, which proposed by Maass et al. [9]. Nonlinear processes are also included in the feature set. Local image features could serve as training features, therefore we employed the local gray level co-occurrence matrix (GLCM), local histogram, vesselness and the ImageJ Trainable Weka segmentation [10] to extract image features for training classifiers. All features were used for pixel-based analysis of the images in the machine learning material decomposition process. For GLCM and histogram feature extractions we tested patch sizes varing from 3x3 to 9x9, 15x15, 25x25 and 35x35. The vesselness measure is used for identification of vascular structures in angiographic images, which could also be a useful feature to differentiate bone and contrast agent injected into the vessel. The ImageJ Trainable Weka segmentation includes a huge amount of features. Therefore we narrowed the selection of features to only six features: Hessian, Lipschitz filter, Gabor filter, Derivatives filter, Structure filter and Entropy.

C. Classification Methods

Three general classifiers from Weka were used: linear regression, Reduced Error Pruning Tree (REPTree) and Bootstrap Aggregating (Bagging) using REPTrees, were used in this study.

D. Evaluation Methods

We used two basic approaches, linear correlation and SSIM index measurement, to determine the similarity of the images between the ground truth and the results for evaluation. To investigate the performance of the proposed approaches, we compared these approaches with the methods using the energy channels of the CT images and polynomial combinations of the energy channels as features for material decomposition, which have proven to be successful in prior research.

III. RESULTS

We extracted a large number of features for different classifier learning sets, which produced a huge amount of results. In these results, the bagging classifier shows superior performance than the other two classifiers, therefore we only present the results with Bagging. Figure 2, 3, 4, 5 show the prediction results for bone and contrast agent with respective approaches in projection domain. The mean and standard deviation of the linear correlation and SSIM between the prediction results from the different material decomposition approaches and the ground truth were calculated and presented in Table I.

IV. DISCUSSION

We applied these approaches to the raw data of the projections without any arithmetic operations or filtering. The utilization of energy-selective channels and their polynomial combinations as features provided adequate results in distinguishing Ultravist contrast agent and bone from all other materials. However, the new approaches using image-based features for material decomposition achieved varying results depending on the utilized features and classifiers. Although prediction results from the GLCM and vesselness features could separate the contrast agent from the other materials, had problems in decomposing bone from the remaining materials persisted. The material decomposition with histogram features was a challenging task, because the difference between the lowest and the highest gray value was too big. A further investigation on the parameters of the histogram could lead to better results, but even when the histogram was built with the local minimum and maximum value of the used patch the results did not improve significantly. For the trainable



Figure 2. Decomposition results of the noiseless and no-motion scenario. Top row: Predicitons for Ultravist of the short scan image. Bottom row: Predicitons for bone of the short scan image. All predicitions result from the 90 percent bagging classifier and the noiseless images.



Figure 3. Decomposition results of the noise and no-motion scenario. Top row: Predicitons for Ultravist of the short scan image. Bottom row: Predicitons for bone of the short scan image. All predicitions result from the 90 percent bagging classifier and the noise images.

Image	Material		Channel		Polynomial		Weka	
			r	SSIM	r	SSIM	r	SSIM
Heart and breathing motion	Bone	noiseless	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.98 (0.00)	0.98 (0.00)
		noise	0.94 (0.00)	0.94 (0.00)	0.92 (0.00)	0.92 (0.00)	0.98 (0.00)	0.98 (0.00)
	Ultravist	noiseless	0.79 (0.03)	0.32 (0.02)	0.82 (0.02)	0.41 (0.02)	0.89 (0.02)	0.87 (0.03)
		noise	0.49 (0.05)	0.06 (0.01)	0.49 (0.05)	0.06 (0.01)	0.89 (0.02)	0.88 (0.02)
Short scan(no motion)	Bone	noiseless	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.99 (0.00)	0.98 (0.01)	0.97 (0.01)
		noise	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.01)	0.95 (0.02)
	Ultravist	noiseless	0.71 (0.02)	0.21 (0.01)	0.75 (0.03)	0.29 (0.03)	0.82 (0.11)	0.47 (0.12)
		noise	0.35 (0.02)	0.03 (0.00)	0.35 (0.22)	0.03 (0.00)	0.55 (0.10)	0.40 (0.08)

Table I

MEAN AND STANDARD DEVIATION (IN BRACKETS) OF THE LINEAR CORRELATION R AND SSIM FOR THE PREDICTION WITH THE 90 PERCENT BAGGING CLASSIFIER.

Weka segmentation, the decomposition results were very good, which means comparable or in some cases (i.e. the noise and motion scenario) even better than the results from the polynomial combination of the energy selective channels. The reason why the trainable Weka segmentation features worked better than the other proposed features was that a set of selected image features is used. Compared with the classic decomposing methods (i.e. energy channel and polynomial combinations), the trainable Weka segmentation features are robust against noise .

However, the results in this study are only based on simulations. Furthermore, the extraction of the training data for the classifiers is strongly dependent on the ground truth image. If we test this new image-based feature material decomposition approach with real data, the region for the ground truth information must be chosen wisely, otherwise the material decomposition results would vary a lot. We will study these challenges further in the future.

V. CONCLUSION

We investigated the performance, applicability, appropriateness of image-based feature material decomposition at projection domain for computer tomography. Comparing the feature based approach to a polynomial method, the linear correlation of the Ultravist images to the ground truth in the noise and motion scenario increased from 0.49 to 4 0.89, the SSIM from



Figure 4. Decomposition results of the noiseless and motion scenario. Top row: Predicitons for Ultravist of the heart cycle and lung motion image. Bottom row: Predicitons for bone of the heart cycle and lung motion image. All predicitions result from the 90 percent bagging classifier and the noiseless images.



Figure 5. Decomposition results of the noise and motion scenario. Top row: Predicitons for Ultravist of the heart cycle and lung motion image. Bottom row: Predicitons for bone of the heart cycle and lung motion image. All predicitions result from the 90 percent bagging classifier and the noise images.

0.06 to 0.88 by using appropriate features, respectively. The simulations experiments show that an appropriate image-based feature material decomposition process such as the trainable Weka segmentation features could lead to pleasing results and should encourage further investigations.

REFERENCES

- M. Firsching, P. Takoukam Talla, T. Michel, and G. Anton, "Material resolving x-ray imaging using spectrum reconstruction with medipix2," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 591, no. 1, pp. 19–23, 2008.
- [2] H. Q. Le and S. Molloi, "Least squares parameter estimation methods for material decomposition with energy discriminating detectors," *Medical Physics*, vol. 38, no. 1, pp. 245–255, 2011.
- [3] X. Wang, D. Meier, K. Taguchi, D. J. Wagenaar, B. E. Patt, and E. C. Frey, "Material separation in x-ray ct with energy resolved photoncounting detectors," *Medical Physics*, vol. 38, no. 3, pp. 1534–1546, 2011.
- [4] G.-S. T. Zimmerman K C, "Experimental investigation of multi-energy ct material decomposition using artificial neural networks," in *Proceedings* of the third international conference on image formation in x-ray computed tomography, 2014, pp. 208–211.
- [5] E. Hall, Mark and Frank, 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.
- [6] 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, pp. –, 2013.

- [7] 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, no. 19, p. 6193, 2012.
- [8] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients," National Inst. of Standards and Technology-PL, Gaithersburg, MD (United States). Ionizing Radiation Div., Tech. Rep., 1995.
- [9] N. Maass, S. Sawall, M. Knaup, and M. Kachelriess, "Empirical multiple energy calibration (emec) for material-selective ct," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, 2011 IEEE, Oct 2011, pp. 4222–4229.
- [10] I. Arganda-Carreras, A. Cardona, V. Kaynig, and J. Schindelin, "Trainable weka segmentation," *Fiji website*, 2013.