Supplementary Material: Machine Vision Guided 3D Medical Image Compression for Efficient Transmission and Accurate Segmentation in the Clouds

Zihao Liu[1], Xiaowei Xu[2], Tao Liu[1], Qi Liu[1], Yanzhi Wang[3], Yiyu Shi[2], Wujie Wen[1], Meiping Huang[4], Haiyun Yuan[4], Jian Zhuang[4]
[1] Flordia International University, [2] University of Notre Dame, [3] Northeastern University, [4] Guangdong General Hospital
{zliu021,tliu023,qliu020,wwen}@fiu.edu, {xxu8,yshi4}@nd.edu,yanz.wang@northeastern.edu

In this supplementary material, we introduce four mapping functions and perform performance comparison in Section 1. In Section 2, we present more results with compression rate of 40. In Section 3, we then provide the image by image segmentation results to further demonstrate that the improvement by our method is statistically significant.

1 Mapping Function Exploration

We explored four types of mapping functions–reciprocal mapping (RM) (depicted as 'ours' in our paper), linear mapping (LM), discrete mapping (DM) and power mapping (PM), to search the best mapping between statistic index (SI) and quantization step (QS). Then we compared segmentation accuracy and compression rates using these mapping methods in Section 1.4.

1.1 Linear Mapping

We first assumed the relationship between SI and QS obeyed a linear mapping (LM) function:

$$Q_n = a \cdot \delta_n + b, \quad s.t. \quad Q_{min} \leq Q_n \leq Q_{max}, \tag{1}$$

where Q_n is the quantization step at subband n, Q_{min} and Q_{max} are the smallest and largest QS, and a and b are the fitting parameters.

1.2 Discrete Mapping

We further developed a seemingly more intuitive discrete mapping (DM) method defined as:

$$Q_n = \begin{cases} Q_{min} & \delta_n > T_1 \\ 2 \cdot Q_{min} & T_2 < \delta_n \leq T_1 \\ 4 \cdot Q_{min} & T_3 < \delta_n \leq T_2 \\ Q_{max} & \delta_n \leq T_3 \end{cases} \tag{2}$$

where T_1, T_2 and T_3 indicate the average intervals between the largest and smallest δ_n.

Table 1: Segmentation results of four candidate mapping methods (LM, DM, PM and RM (depicted as "ours" in our paper)) using DenseVoxNet and HVSMR2016 dataset.

		LM	**DM**	**PM**	**RM (Ours)**
Myocardium	**Dice**	0.771±0.0438	0.821±0.043	0.834±0.0366	**0.834±0.0386**
	Hausdorff	36.823±10.079	32.372±8.551	31.012±7.911	**31±7.940**
	ASD	0.822±0.846	0.669±0.761	**0.651±0.724**	0.652±0.671
Blood Pool	**Dice**	0.903±0.026	0.911±0.027	0.913±0.026	**0.914±0.024**
	Hausdorff	43.007±11.358	41.792±9.796	41.014±9.769	**40.93±9.52**
	ASD	0.664±0.514	0.576±0.439	0.556±0.453	**0.556±0.432**
Compression Rate		**∼40x**	∼35x	∼22x	∼30x

1.3 Power Mapping

We adopted reciprocal mapping function in our paper (as Ours) to get QS for each subband from corresponding SIs. We also considered another non-linear mapping function, power mapping (PM),

$$Q_n = a \cdot \delta_n^b, \tag{3}$$

where a and b are fitting parameters. This mapping methods introduced more non-linearity between Q_n and δ_n.

1.4 Results Comparison

We compared segmentation accuracy and compression rate of these three mapping methods with the method in our paper. The experiment setup is the same as that in our paper for fair comparison. As shown in Table 3, we can observe that non-linear mapping methods (DM, PM and RM (ours)) can always achieve significant higher segmentation accuracy than linear mapping (LM) at all three segmentation measurement metrics, which demonstrates the advantage of the non-linear relationship between SI and QS. For the four mapping methods, it can be noticed that usually high accuracy corresponds to low compression rate except for RM (Ours). *This clearly indicates that our adopted RM is the best option to balance the accuracy and compression rate, e.g. offering almost the best accuracy at a relatively high compression rate.*

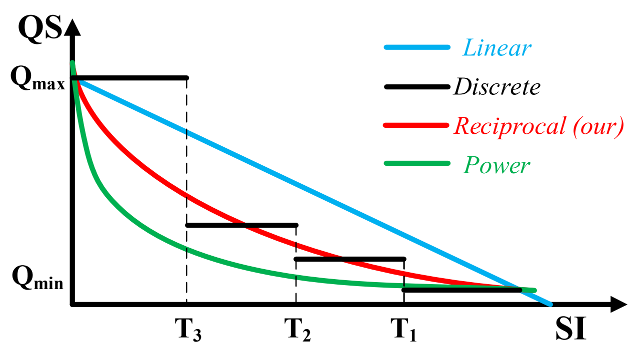

Figure 1: Mapping method comparison for result illustration.

The above phenomenon can be further explained with the illustration in Figure 1. The quantization steps in LM are always larger than non-linear mapping methods (PM and RM) at different SIs, thus resulting with a higher compression rate but a lower segmentation accuracy due to the more significant distortions of important features to DNN. Compared with RM (ours), PM has a steeper slope and a smaller QSs, thus decreasing compression rate (22× v.s. 30×). However, as we can see from Table 3, the segmentation accuracy is not improved w.r.t. RM. This is because the additional information saved by PM has almost no contributions to DNN segmentation, i.e. redundant data. Although DM is similar with RM in terms of the magnitude of quantization steps (QS),

Table 2: Segmentation results of our method and JPEG-2000 using DenseVoxNet and HVSMR2016 dataset at different compression rates (30$\times$ and 40$\times$).

		Ours	**JPEG2000**	**Ours**	**JPEG2000**
Myocardium	**Dice**	**0.834$\pm$0.0386**	0.816$\pm$0.042	0.808$\pm$0.043	0.802$\pm$0.039
	Hausdorff	**31$\pm$7.940**	33.513$\pm$7.566	34.112$\pm$8.551	34.506$\pm$7.962
	ASD	**0.652$\pm$0.671**	0.722$\pm$0.746	0.767$\pm$0.791	0.802$\pm$0.824
Blood Pool	**Dice**	**0.914$\pm$0.024**	0.912$\pm$0.024	0.910$\pm$0.023	0.909$\pm$0.025
	Hausdorff	**40.93$\pm$9.52**	41.031$\pm$9.648	41.192$\pm$9.706	41.448$\pm$9.671
	ASD	**0.556$\pm$0.432**	0.564$\pm$0.414	0.576$\pm$0.439	0.582$\pm$0.453
Compression Rate		$\sim$30x	$\sim$30x	$\sim$40x	$\sim$40x

it only exploits QS with coarse-grained intervals. Therefore, DM can achieve a higher compression rate but lower segmentation accuracy due to the non-optimized (coarse-grained) quantization steps.

Therefore, we finally select RM for our compression framework which can achieve a high compression rate with the best segmentation accuracy.

2 Results with Different Compression Rates

The comparison of our method and JPEG-2000 with different compression rates is shown in Table 2. At 30$\times$ compression rate, our method has remarkable segmentation accuracy improvement compared with JPEG-2000. However, at 40$\times$ compression rate, the segmentation accuracy is improved. One reasonable explanation is that too aggressive compression will cause serious information loss resulting with low performance. When the compression rate is 30x, the necessary information for segmentation is still retained. However, when it comes to 40x, this information is seriously dropped which causes a large accuracy degradation. This phenomenon is common in image compression, and compression methods can only drop unnecessary information within their capacities for particular domains. Otherwise, performance loss will always happen.

3 Image by Image Segmentation Results Comparison

To better illustrate that our improvement is statistically significant, we further report the image by image segmentation results of the HVSMAR dataset by using our method and JPEG-2000 under two different DNN structures–DenseVoxNet and 3D-DSN. Table 3 and 4 clearly show that our method wins in almost every case w.r.t. JPEG2000 for Dice, Hausdorff and ASD characterized from various Myocardoum and Blood Pool images under the same compression rate for both networks. Moreover, compared with the original uncompressed images, our method always offers the similar or even better results due to filtering the unnecessary image features that can mislead the segmentation.

Table 3: Image by image segmentation results of original (uncompressed), our method, JPEG-2000 using DenseVoxNet and HVSMR2016 dataset. The compression rate of our method and JPEG-2000 is 30. The images compressed by ours can be segmented with almost the same accuracy, or sometimes even better than original ones, significantly beating those compressed by JPEG-2000.

	Myocardoum								
	Dice			**Hausdorff**			**ASD**		
	Original	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**
Img. 1	0.876	**0.876**	0.850	19.647	**19.197**	22.782	0.344	**0.327**	0.380
Img. 2	0.786	**0.774**	0.755	37.229	**37.815**	39.434	1.426	**1.453**	1.53
Img. 3	0.849	**0.855**	0.848	32.419	**31.516**	31.718	1.371	**1.3**	1.526
Img. 4	0.851	**0.846**	0.839	27.258	**28.083**	29	0.122	**0.089**	0.083
Img. 5	0.831	**0.818**	0.791	37.842	**38.615**	41.109	0.099	**0.088**	0.088
	Blood Pool								
	Dice			**Hausdorff**			**ASD**		
	Original	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**
Img. 1	0.898	**0.9**	0.897	25.278	**24.434**	24.515	1.316	**1.247**	1.226
Img. 2	0.9	**0.9**	0.898	45.200	**44.878**	44.922	0.323	**0.311**	0.341
Img. 3	0.953	**0.954**	0.953	41.219	**41.677**	41.158	0.552	**0.532**	0.553
Img. 4	0.895	**0.897**	0.895	49.669	**48.363**	48.518	0.683	**0.62**	0.649
Img. 5	0.927	**0.919**	0.911	43.806	**45.31**	46.776	0.127	**0.11**	0.139

Table 4: Image by image segmentation results of original (uncompressed), our method, JPEG-2000 using 3D-DSN and HVSMR 2016 dataset. The compression rate of our method and JPEG-2000 is 30. The images compressed by ours can be segmented with almost the same accuracy, or sometimes even better than original ones, significantly beating those compressed by JPEG-2000.

	Myocardoum								
	Dice			**Hausdorff**			**ASD**		
	Original	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**
Img. 1	0.829	**0.829**	0.801	20.273	**20.149**	22.045	0.446	**0.476**	0.607
Img. 2	0.688	**0.692**	0.703	39.712	**38.652**	38.026	0.401	**0.386**	0.39
Img. 3	0.799	**0.802**	0.781	35.567	**34.929**	35.185	0.451	**0.5**	0.526
Img. 4	0.834	**0.839**	0.836	25.199	**24.880**	25.160	0.093	**0.091**	0.088
Img. 5	0.773	**0.769**	0.748	40.976	**40.902**	44.744	0.157	**0.171**	0.164
	Blood Pool								
	Dice			**Hausdorff**			**ASD**		
	Original	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**	**Original**	**Ours**	**JPEG-2000**
Img. 1	0.903	**0.905**	0.896	23.324	**22.782**	25.239	0.568	**0.559**	0.57
Img. 2	0.873	**0.866**	0.863	45.837	**46.551**	46.755	0.229	**0.235**	0.239
Img. 3	0.949	**0.950**	0.949	40.410	**40.386**	40.841	0.221	**0.179**	0.197
Img. 4	0.915	**0.913**	0.910	35.931	**36.905**	38.730	0.88	**0.076**	0.079
Img. 5	0.906	**0.900**	0.883	47.074	**47.403**	51.254	0.068	**0.066**	0.067