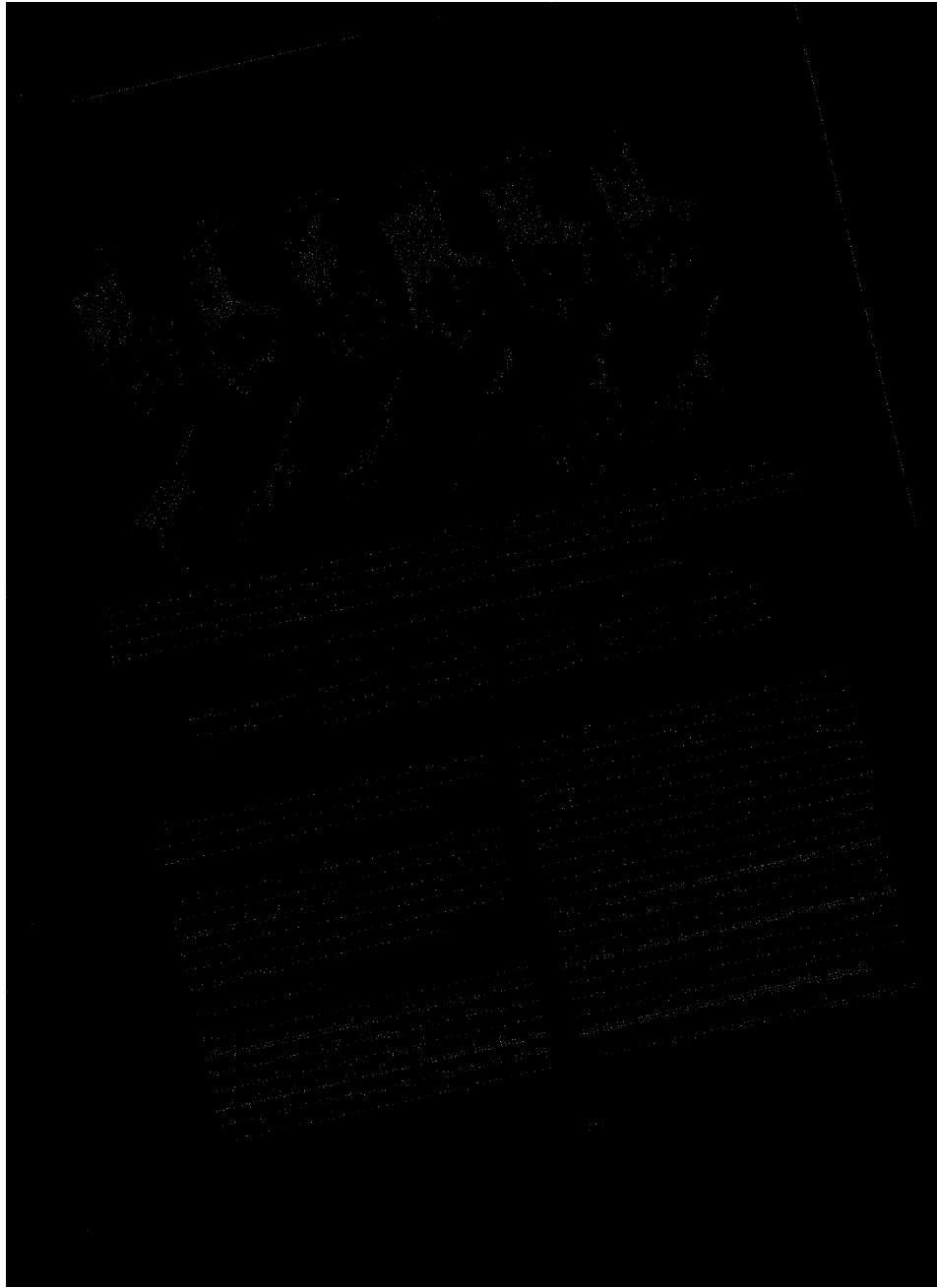


Distinction Task 2.3: Document analysis and recognition

1. Task 1

The negative image of doc.jpg with candidate points only



- We have obtained a new negative image containing the candidate points selected based on the strategy "The centres of connected components are considered candidate points" when applied to the original negative image.

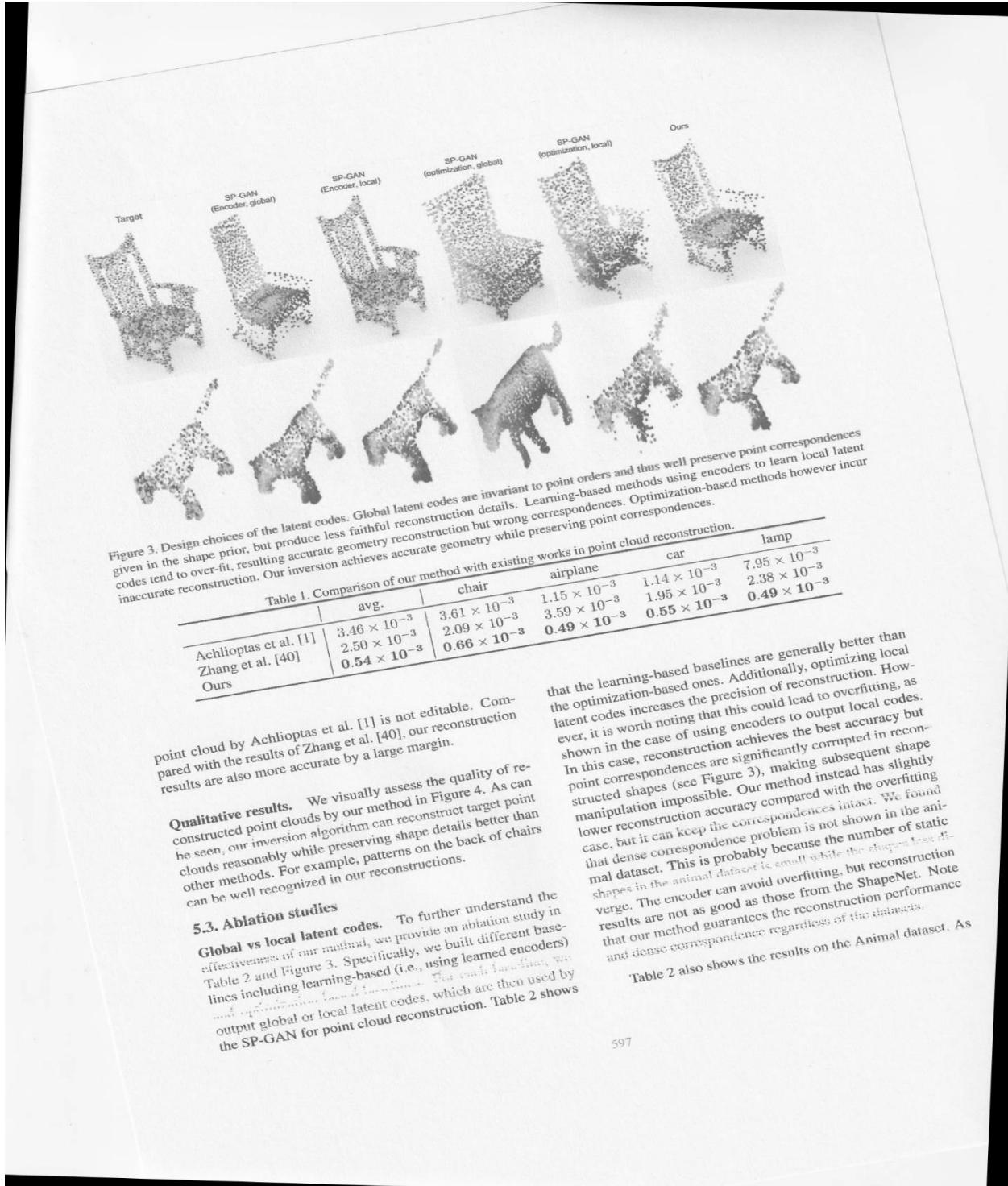
2. Task 2

- a. Candidate points Strategy - All foreground pixels in the image are considered candidate points.

Strategy: a, Threshold: 10

Estimated skew angle: -2 degrees.

Skew estimation took 69.20 seconds.



Strategy: a, Threshold: 15

Estimated skew angle: -2 degrees.

Skew estimation took 73 seconds.

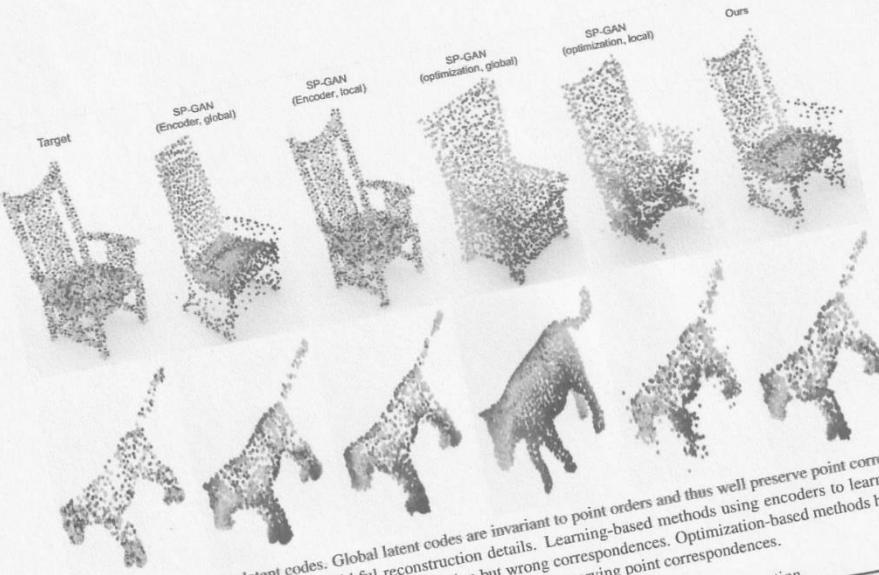


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based (i.e., using learned encoders). For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the slugs have diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Strategy: a, Threshold: 20

Estimated skew angle: -2 degrees.

Skew estimation took 76.39 seconds.

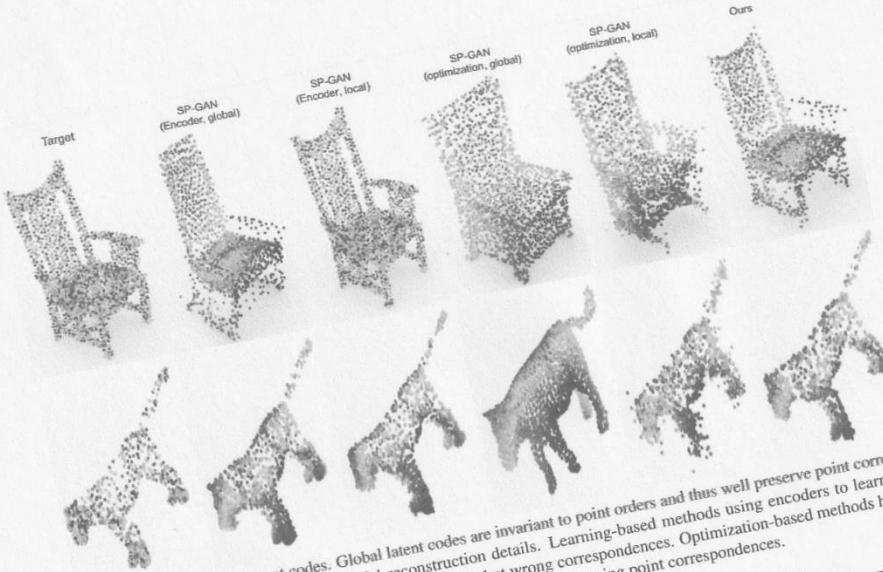


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based (i.e., using learned encoders). The main baseline we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes less diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

- b. **Candidate points Strategy** - The centres of connected components are considered candidate points. The centre of a connected component has x- and y-coordinate as the means of x- and y-coordinates of all pixels in that connected component.

Strategy: b, Threshold: 10

Estimated skew angle: -11 degrees.

Skew estimation took 70.22 seconds.

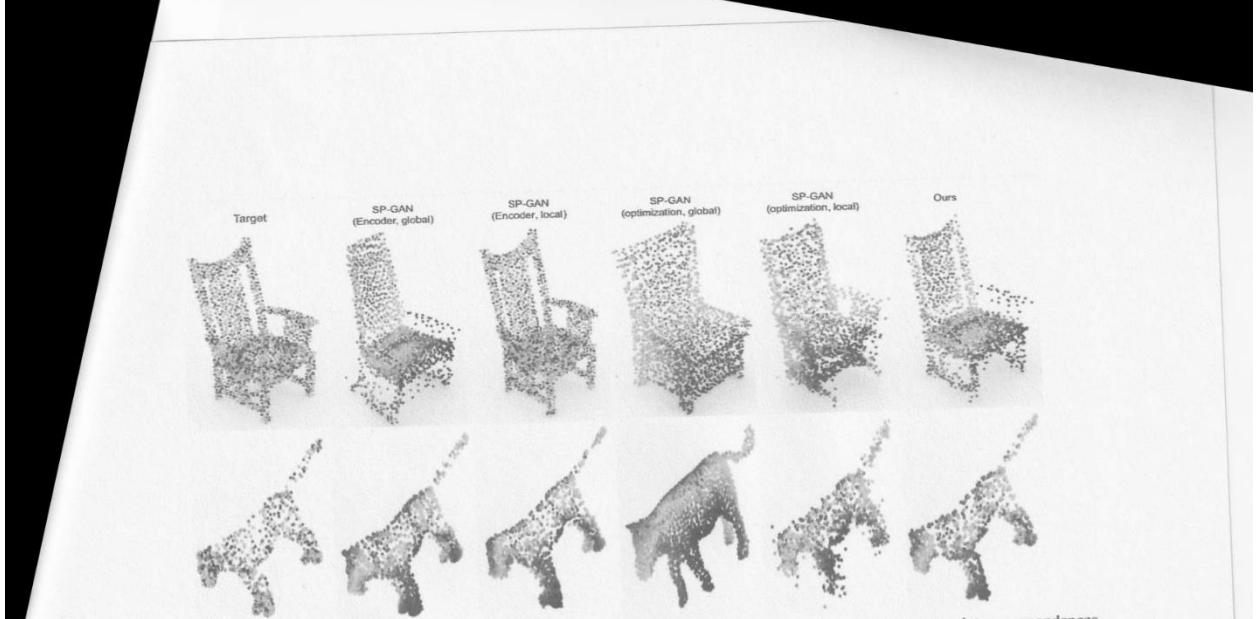


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based ones. For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes less diverse. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Strategy: b, Threshold: 15
Estimated skew angle: -13 degrees.
Skew estimation took 70.25 seconds.

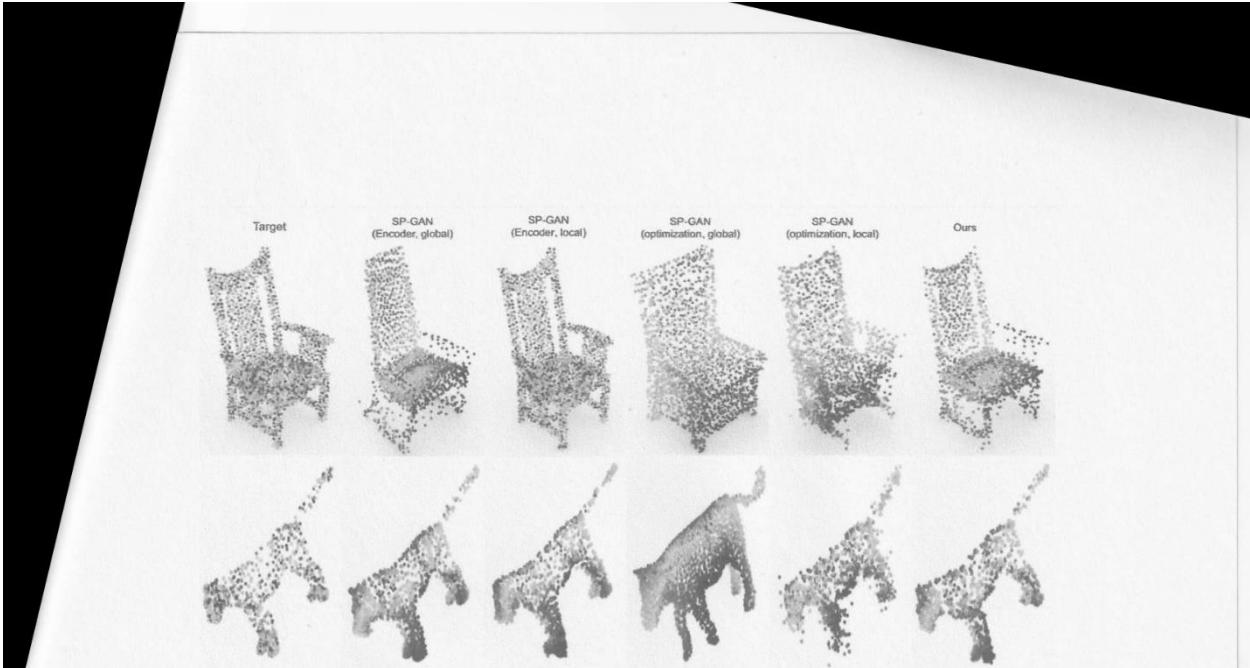


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based ones. For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes less diverse. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Strategy: b, Threshold: 20
Estimated skew angle: -13 degrees.
Skew estimation took 64.61 seconds.

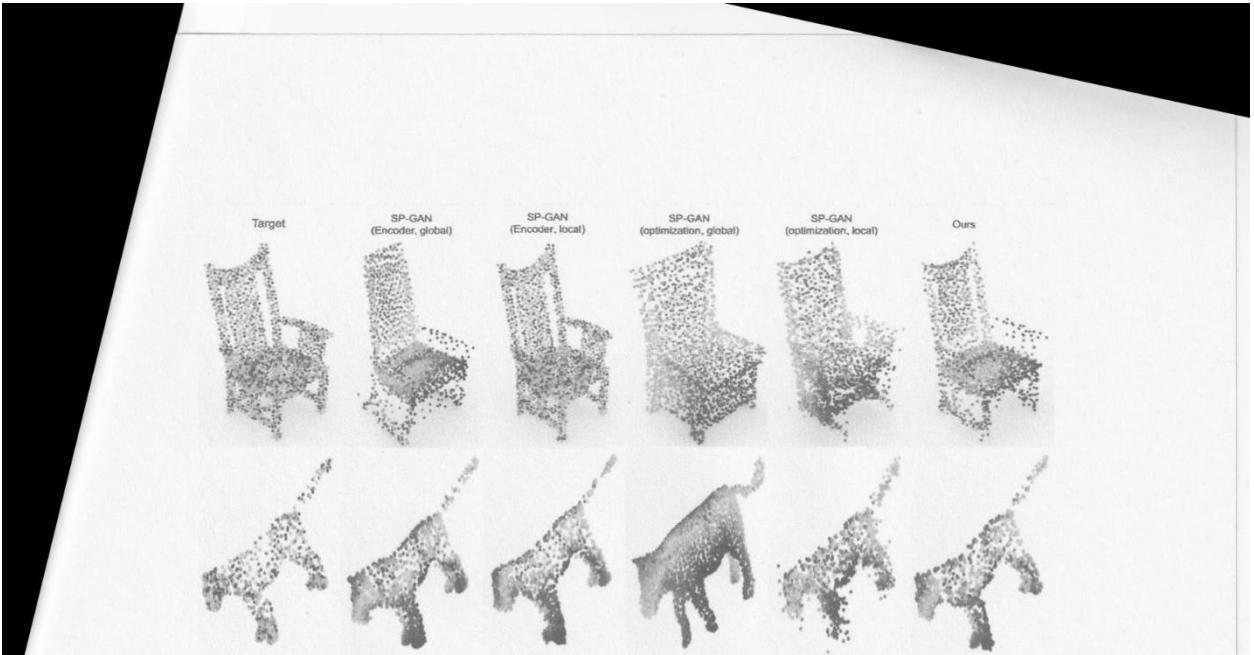


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based ones. For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes themselves diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

- c. Candidate points Strategy - The point which has maximum y-coordinate in each connected component is chosen as a candidate point.

Strategy: c, Threshold: 10

Estimated skew angle: -11 degrees.

Skew estimation took 68.99 seconds.

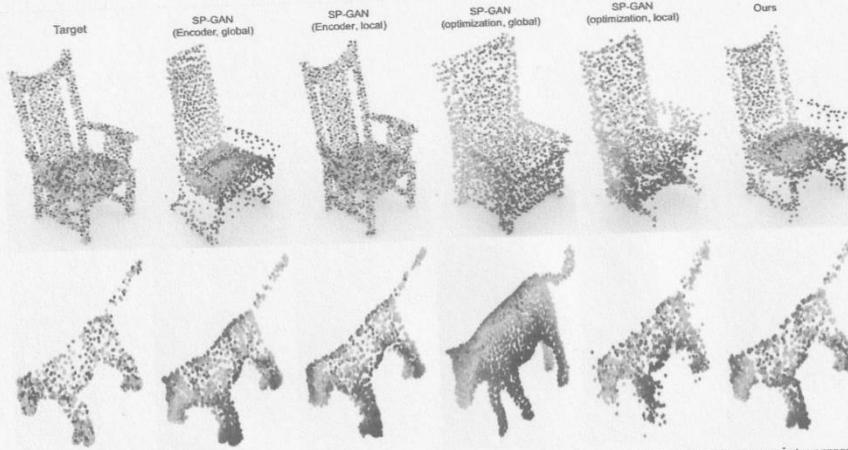


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders to output global or local latent codes, which are then used by SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the slugs tree diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Strategy: c, Threshold: 15
Estimated skew angle: -13 degrees.
Skew estimation took 69.11 seconds.

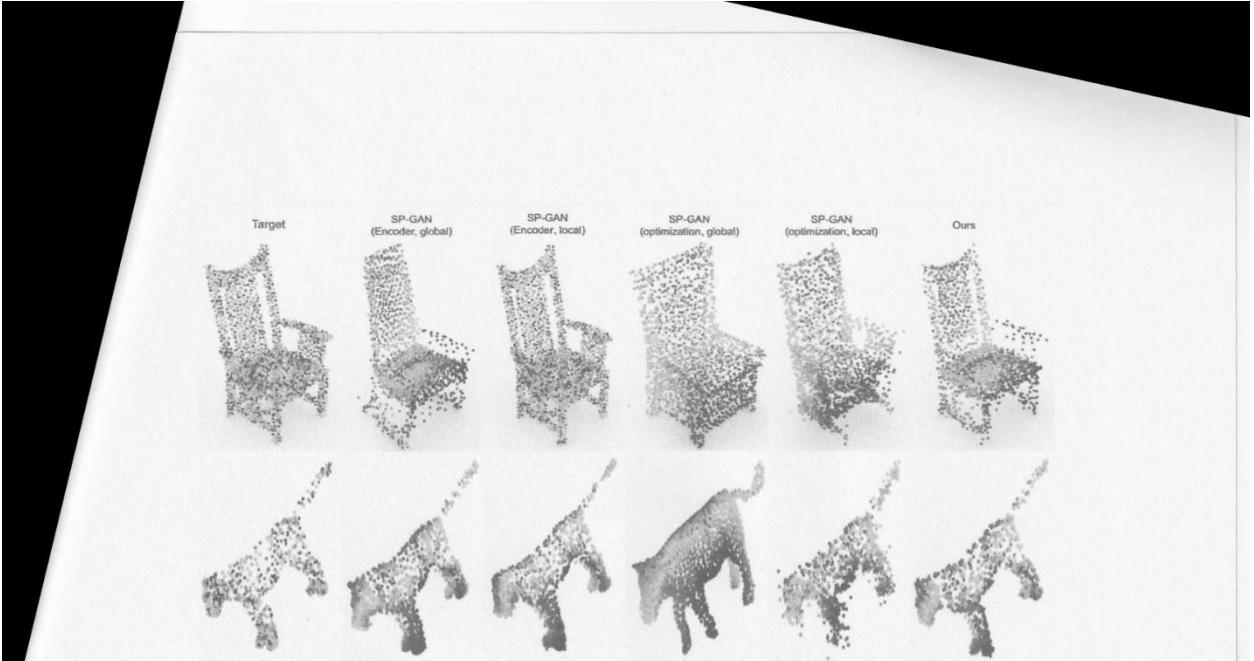


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based ones. The main baseline will output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes less diverse. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Strategy: c, Threshold: 20
Estimated skew angle: -13 degrees.
Skew estimation took 68.91 seconds.

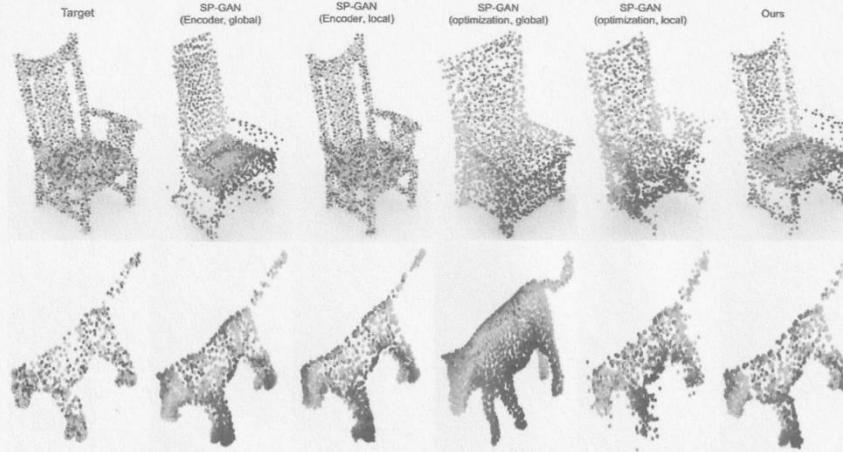


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based baselines. For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes here diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

Deskewed Image 1

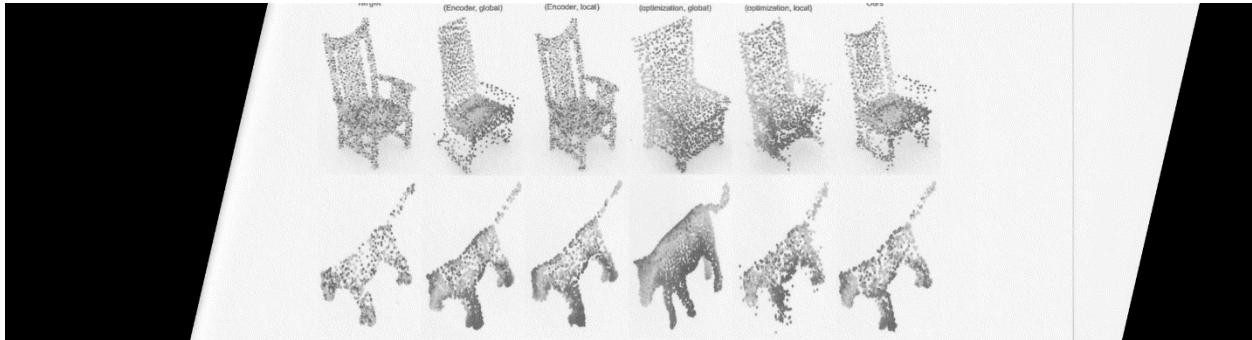


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

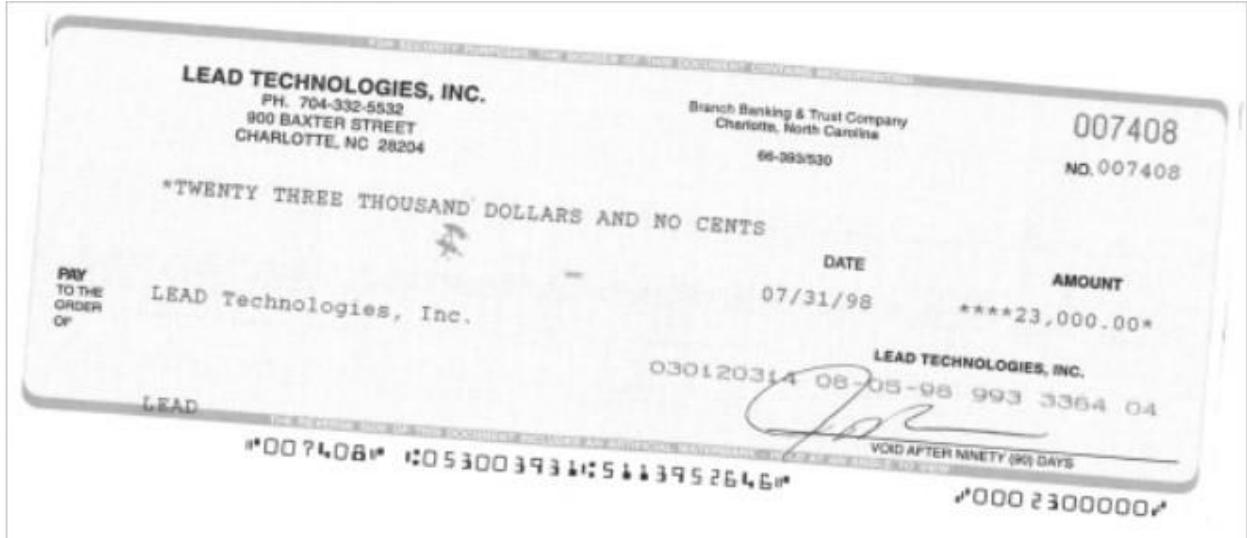
Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based (i.e., using learned encoders). Then each baseline will output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes have diverse. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

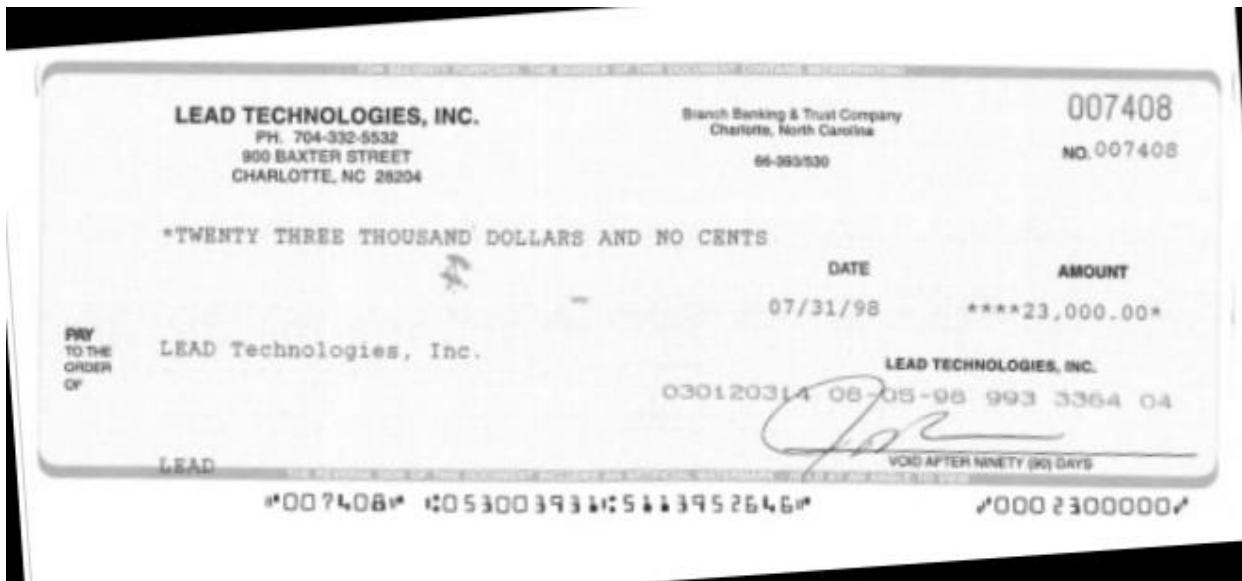
Table 2 also shows the results on the Animal dataset. As

- For the "doc_1" Image, the candidate point strategy used was "The point which has maximum y-coordinate in each connected component" with a density threshold of "15".
- This strategy resulted in the most precise deskewed image, featuring a skew angle of 77 degrees, and with a computational time of 70.72 seconds.

Skewed Image 2



Deskewed Image 2

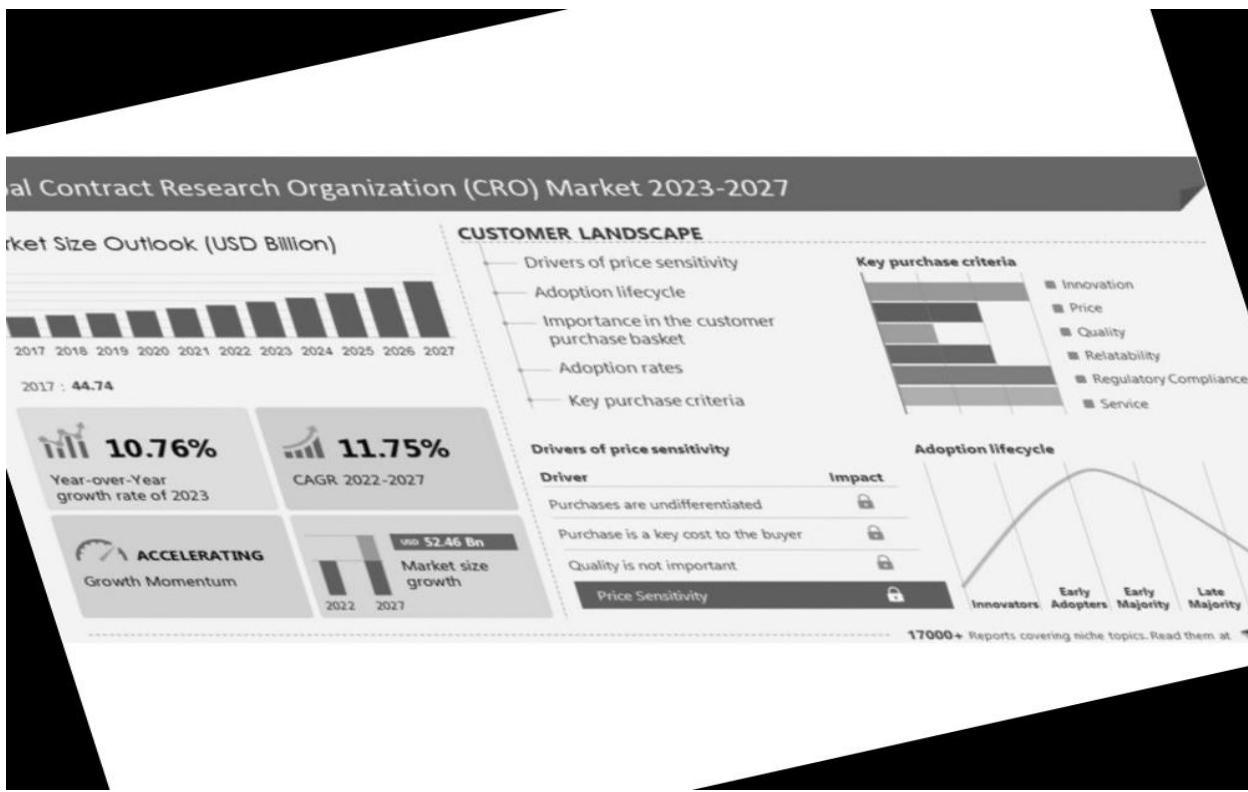


- For the "doc_2" Image, the candidate point strategy used was "The point which has maximum y-coordinate in each connected component" with a density threshold of "10".
- This strategy resulted in the most precise deskewed image, featuring a skew angle of 4 degrees, and with a computational time of 0.294 seconds.

Skewed Image 3



Deskewed Image 3



- For the "doc_3" Image, the candidate point strategy used was "The point which has maximum y-coordinate in each connected component" with a density threshold of "10".
- This strategy resulted in the most precise deskewed image, featuring a skew angle of 15 degrees, and with a computational time of 1.76 seconds.

Skewed Image 4

403038

YONG TAT HARDWARE TRADING

(JM0292487-D)

NO 4,JALAN PERJIRANAN 10,

TAMAN AIR BIRU,

81700 PASIR GUDANG,

JOHOR.

TEL : 07-2614733 FAX : 07-2614733

YONGTAT99@YAHOO.COM.MY

GST ID : 001570349056

TAX INVOICE

Doc No. : CS00035863 Date: 13/03/2018

Cashier : USER Time: 12:49:00

Salesperson : Ref.:

Item	Qty	S/Price	(GST) S/Price	(GST) Amount	Tax
2687	1.00 PCS	48.00	50.88	50.88	SR
10" WING COMPASS	1	50.88			
5736	1.00 SET	47.17	50.00	50.00	SR
1/8"-3/4" 12PCS HOLLOW PUNCH SET	1	50.00	50.00	50.00	
Total Qty:	2			100.88	

Total Sales (Excluding GST) : 95.17

Discount : 0.00

Total GST : 5.71

Rounding : 0.02

Total Sales (Inclusive of GST) : 100.90

CASH : 101.00

Change : 0.10

GST SUMMARY

Tax Code	%	Amt (RM)	Tax (RM)
SR	6	95.17	5.71
Total :		95.17	5.71

GOODS SOLD ARE NOT RETURNABLE, THANK YOU

Deskewed Image 4

A03078

YONG TAT HARDWARE TRADING (JM0292487-D)

NO 4, JALAN PERJIRANAN 10,
TAMAN AIR BIRU,
81700 PASIR GUDANG,
JOHOR.

TEL : 07-2614733 FAX : 07-2614733
YONGTAT99@YAHOO.COM.MY
GST ID : 001570349056

TAX INVOICE

Doc No. : CS00035863 Date: 13/03/2018

Cashier : USER Time: 12:49:00

Salesperson : Ref.:

Item	Qty	S/Price	(GST) S/Price	(GST) Amount	Tax
2587	1.00 PCS	48.00	50.88	50.88	SR
10" WING COMPASS	1.00	6.00			
5736	1.00 SET	47.17	50.00	50.00	SR
1/8"-3/4" 12PCS HOLLOW PUNCH SET	1.00	6.00			
Total Qty:	2			100.88	

Total Sales (Excluding GST) : 95.17

Discount : 0.00

Total GST : 5.71

Rounding : 0.02

Total Sales (Inclusive of GST) : 100.90

CASH : 101.00

Change : 0.10

GST SUMMARY

Tax Code	%	Amt (RM)	Tax (RM)
SR	8	95.17	5.71
Total :		95.17	5.71

GOODS SOLD ARE NOT RETURNABLE, THANK YOU

- For the "doc_4" Image, the candidate point strategy used was "The point which has maximum y-coordinate in each connected component" with a density threshold of "10".
- This strategy resulted in the most precise deskewed image, featuring a skew angle of -2 degrees, and with a computational time of 1.74 seconds.

IMPORTED

OPICI

ITALIAN SELECTION



FINE MARSALA WINE

SWEET

MARSALA

ITALIA PARTICOLARE

DENOMINAZIONE DI ORIGINE CONTROLLATA

ITALIAN DESSERT WINE

ALC. 17% BY VOL. - CONTAINS SULFITES - CONT. 750 ML.

PRODUCED AND BOTTLED BY FRATELLI FICI MARSALA (ITALY) - I.C.R.F. TP/2932
COD. ACCISA: IT00TPA00042L

PRODUCT OF ITALY

IMPORTED BY OPICI WINES GLEN ROCK N. J.

Deskewed Image 5



- For the "doc_5" Image, the candidate point strategy used was "The point which has maximum y-coordinate in each connected component" with a density threshold of "10".
- This strategy resulted in the most precise deskewed image, featuring a skew angle of 5 degrees, and with a computational time of 1.21 seconds.
- Overall, the Hough transform accurately worked in every sample image, we applied a lower density threshold of 10 to documents with a lower volume of text and applied a higher density threshold value with images with a larger volume of text.

4. Task 4

Text detected for Skewed doc.jpg.

```
doc = cv.imread('E:\T2 2023\SIT789 CV\Assignments\Task 2.2C\doc.jpg', 0)
text = pytesseract.image_to_string(doc)

print(text)
```

```
sponden® 28
¢ point corres
ders jeatt jocal jatent
ods poweve® yncut
```

Text detected for the deskewed version of the doc.jpg image.

Figure 3. Design choices of the latent codes. Global latent codes are inv

codes tend to over-fit, resulting accurate geometry
inaccurate reconstruction. Our inversion achieves accurate
Table 1. Comparison of our method with existin

given in the shape prior, but produce less faithful reconstruction details. Learning-t
econstruction but wrong correspondences, Optimization-b:
geometry while preserving point correspondences.

t to pr snd thus well preserve point @
sed methods using encoders to learn local latent
ased methods however incur
g works in point cloud reconstruction.

avg. chair airplane lamp
Achlioptas et al. [1] 3.46 × 10 ⁻³ 3.61 × 10 ⁻³ 3 ge x to-? 7.95 × 10 ⁻³
Zhang et al. [48] 2.50 × 10 ⁻³ 209×10 ⁻³ 3.59x 10 ⁻³ 1.95 × 10 ⁻²) 1 238% 10
Ours 0.54 × 10 ⁻⁵ 0.66 × 10 ⁻³ 9.49x10 ⁻² 0.55x10 ⁻³ 0.49x fo

point cloud by Achlioptas et al. [1] is not editable. Com-
pared with the results of Zhang et al. [48], our reconstruction
results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of re-
constructed point clouds by our method in Figure 4. As can
be si C et pc
clouds reasonably while preserving shape details better than
other methods. For example, patterns on the back of chairs
can be well recognized in our reconstructions.

```

| avg. | chair airplane lamp
Achlioptas et al. [1] | 3.46 x 10^-3 | 3.61 x 10^-3 ge x to -? 7.95 x 10^-3
Zhang et al. [40] | 2.50 x 10^-3 | 2.09x10^-3 3.59x 10^7 1.95 x 10^-2) 1 238% 10
Ours | 0.54 x 10^-5 | 0.66 x 10^-3 9.49x10^-2 0.55x10^-3 0.49x 10^-3

```

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen C et pc clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

our inversion algorithm can reconstruct point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes.

output global or local latent codes

AN for point cloud reconstruction.

Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting

We found that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting

case, but it can keep the correspondences more accurate and

- When Text recognition (OCR) is performed on a **skewed document image like 'doc.jpg'** using Tesseract-OCR, it was only able to **detect very little text** or fails to recognize any text altogether.
- Contrarily, performing Text recognition (OCR) on the **deskewed document image** yielded significantly **more favourable results**, detecting a **substantially higher number of texts** within the document.
- Deskewing the document **rectifies the skew angle and enhances image quality**, leading to a notable improvement in OCR performance. This allows the OCR engine to accurately recognize a greater number of text elements, resulting in a more favourable and precise outcome.

- Ultimately, after saving the result of the OCR's **text-to-PDF function**, a noticeable distinction emerges: the PDF version of the **skewed document remains uneditable**, lacking any text recognition, while the PDF version of the **deskewed document is editable**.
- In the editable PDF, certain portions of the **text have been highlighted** to demonstrate the **successful recognition** and conversion of the text content.
- Similar results were obtained when performing Text recognition (OCR) on different sets of deskewed documents which can be seen in the pdf below.

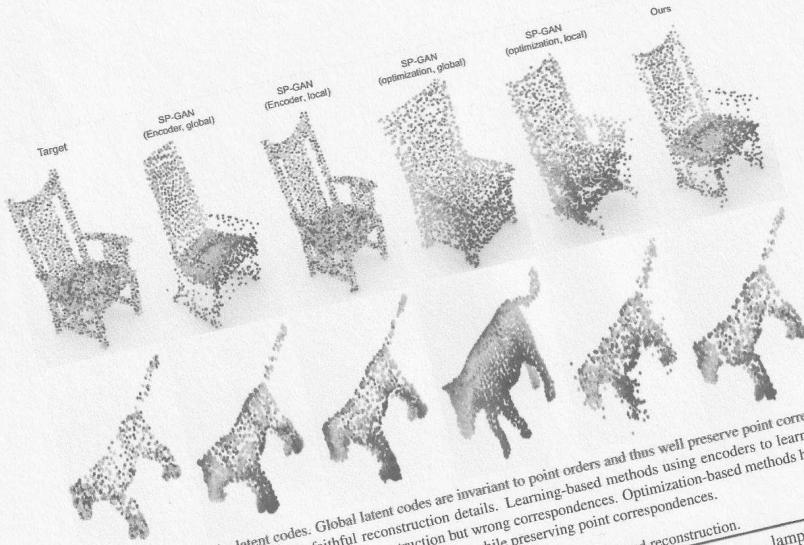


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we build different baselines including learning-based (i.e., using learned encoders) and optimization-based (i.e., using learned encoders). For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes have diverse. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

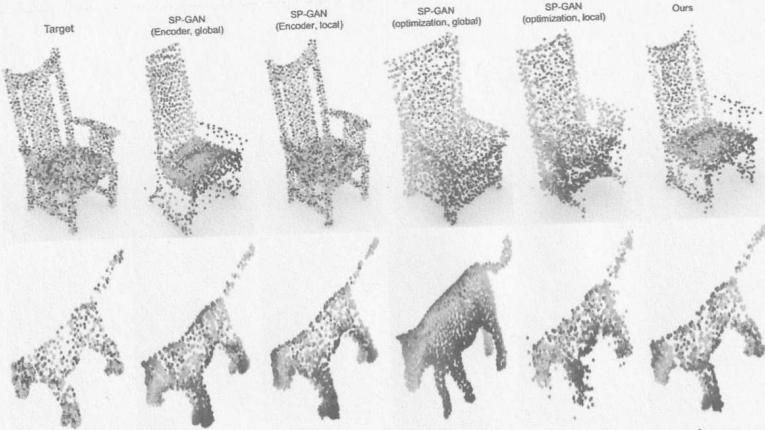


Figure 3. Design choices of the latent codes. Global latent codes are invariant to point orders and thus well preserve point correspondences given in the shape prior, but produce less faithful reconstruction details. Learning-based methods using encoders to learn local latent codes tend to over-fit, resulting accurate geometry reconstruction but wrong correspondences. Optimization-based methods however incur inaccurate reconstruction. Our inversion achieves accurate geometry while preserving point correspondences.

Table 1. Comparison of our method with existing works in point cloud reconstruction.

	avg.	chair	airplane	car	lamp
Achlioptas et al. [1]	3.46×10^{-3}	3.61×10^{-3}	1.15×10^{-3}	1.14×10^{-3}	7.95×10^{-3}
Zhang et al. [40]	2.50×10^{-3}	2.09×10^{-3}	3.59×10^{-3}	1.95×10^{-3}	2.38×10^{-3}
Ours	0.54×10^{-3}	0.66×10^{-3}	0.49×10^{-3}	0.55×10^{-3}	0.49×10^{-3}

point cloud by Achlioptas et al. [1] is not editable. Compared with the results of Zhang et al. [40], our reconstruction results are also more accurate by a large margin.

Qualitative results. We visually assess the quality of reconstructed point clouds by our method in Figure 4. As can be seen, our inversion algorithm can reconstruct target point clouds reasonably while preserving shape details better than other methods. For example, patterns on the back of chairs can be well recognized in our reconstructions.

5.3. Ablation studies

Global vs local latent codes. To further understand the effectiveness of our method, we provide an ablation study in Table 2 and Figure 3. Specifically, we built different baselines including learning-based (i.e., using learned encoders) and optimization-based ones. For each baseline, we output global or local latent codes, which are then used by the SP-GAN for point cloud reconstruction. Table 2 shows

that the learning-based baselines are generally better than the optimization-based ones. Additionally, optimizing local latent codes increases the precision of reconstruction. However, it is worth noting that this could lead to overfitting, as shown in the case of using encoders to output local codes. In this case, reconstruction achieves the best accuracy but point correspondences are significantly corrupted in reconstructed shapes (see Figure 3), making subsequent shape manipulation impossible. Our method instead has slightly lower reconstruction accuracy compared with the overfitting case, but it can keep the correspondences intact. We found that dense correspondence problem is not shown in the animal dataset. This is probably because the number of static shapes in the animal dataset is small while the shapes less diverge. The encoder can avoid overfitting, but reconstruction results are not as good as those from the ShapeNet. Note that our method guarantees the reconstruction performance and dense correspondence regardless of the datasets.

Table 2 also shows the results on the Animal dataset. As

A03078

YONG TAT HARDWARE TRADING
 (JM0292487-D)

NO 4, JALAN PERJIRANAN 10,
 TAMAN AIR BIRU,
 81700 PASIR GUDANG,
 JOHOR.

TEL : 07-2614733 FAX : 07-2614733

YONGTAT99@YAHOO.COM.MY

GST ID : 001570349056

TAX INVOICE

Doc No. :	CS00035863	Date:	13/03/2018
Cashier :	USER	Time:	12:49:00
Salesperson :		Ref.:	

Item	Qty	S/Price	(GST) S/Price	(GST) Amount	Tax
2587	1.00 PCS	48.00	50.88	50.88	SR
10" WING COMPASS		68.00			
5736	1.00 SET	47.17	50.00	50.00	SR
1/8"-3/4" 12PCS HOLLOW PUNCH SET		68.00			
Total Qty:	2			100.88	

Total Sales (Excluding GST) : 95.17

Discount : 0.00

Total GST : 5.71

Rounding : 0.02

Total Sales (Inclusive of GST) : 100.90

CASH : 101.00

Change : 0.10

GST SUMMARY

Tax Code	%	Amt (RM)	Tax (RM)
SR	6	95.17	5.71
Total :		95.17	5.71

GOODS SOLD ARE NOT RETURNABLE, THANK YOU