MANIFOLD TRAVERSAL USING DENSITY RIDGES

Jonas N. Myhre, Michael Kampffmeyer, & Robert Jenssen *
Department of Physics and Technology
University of Tromsø– The Arctic University of Norway
{jonas.n.myhre,michael.c.kampffmeyer,robert.jenssen}@uit.no

ABSTRACT

In this work we present two examples of how a manifold learning model can represent the complexity of shape variation in images. Manifold learning techniques for image manifolds can be used to model data in sparse manifold regions. Additionally, they can be used as generative models as they can often better represent or learn structure in the data. We propose a method of estimating the underlying manifold using the ridges of a kernel density estimate as well as tangent space operations that allows interpolation between images along the manifold and offers a novel approach to analyzing the image manifold.

1 INTRODUCTION

The manifold hypothesis, where high-dimensional data is assumed to be concentrated on or near a smooth surface of much lower dimension, is a key concept in learning how to represent data. In this work we present two clear examples of real data sets that are concentrated around low-dimensional manifolds. We apply recent work on estimating and working with low-dimensional manifolds by (Ozertem & Erdogmus, 2011) and (Dollár et al., 2007) on these examples as a proof-of-concept. We describe a complete workflow on the manifold, illustrated by traversal along the manifold.

2 METHOD

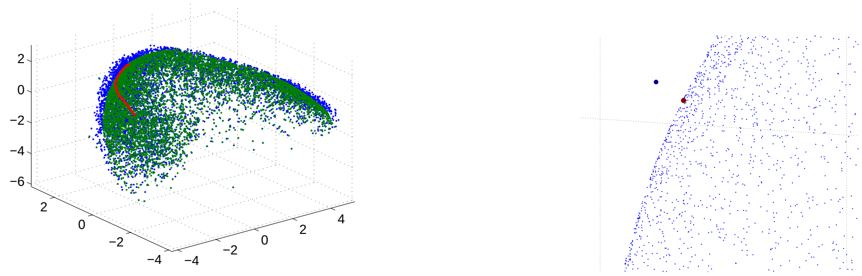
We use principal component analysis to reduce the dimension of the space the manifold is embedded in. As long as the embedding space is of higher dimension than the manifold a linear method causes no harm. We assume that the dimension of the low-dimensional manifold is known. The mapping of a high dimensional image feature vector x to its low dimensional representation z becomes $z = U^T x$, where U is the PCA projection matrix.

Once the ambient dimension is reduced, the manifold can be estimated by a range of techniques. In this work we use the density ridge framework, (Ozertem & Erdogmus, 2011), by solving differential equations as described in (Shaker et al., 2014). The density ridges can be shown theoretically to be close to the underlying manifold bounded by Hausdorff distance (Genovese et al., 2014). The ridge can be found by following the Hessian eigenvector flow $\dot{x}= Q_\perp(x)Q_\perp^T(x)g(x)$, where $Q_\perp(x)$ is the orthogonal subspace of eigenvectors of the Hessian.

Given an estimate of the underlying manifold, we can use the work presented in Section 4 in Dollár et al. (2007), based on the local tangent spaces of the manifold, to work with the manifold. Due to space limitations we choose to focus on geodesic distances. To approximate geodesics on the estimated manifold a scheme alternating between keeping points projected onto the ridge and gradient descent to shorten distances between points on the manifold.

Finally, to approximately reconstruct the image $\widetilde{x}$ from the low dimensional representation z, two approaches were tested. The first approach utilizes the orthogonal PCA projection matrix to reverse the linear transformation via $x = Uz$. Alternatively, a non-linear mapping was used, by training a small 2-layer MLP (60 hidden units) for reconstruction on parts of the data. The advantage of using the inverse PCA is that no additional training is required. Eventhough the MLP requires enough

*http://site.uit.no/ml/



(a) The ridge for the MNIST dataset (b) Projection from noisy points onto ridge

Figure 1: Density ridge estimation results for the MNIST dataset

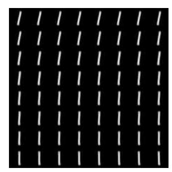


Figure 2: The interpolation results for the MNIST dataset using MLP reconstruction

training data to learn the inverse mapping, our results illustrate that its reconstruction is generally better.

3 Experimental results

We evaluate our method on two real-world datasets, namely the ones in the MNIST dataset (LeCun et al., 1998) and the Frey Faces dataset [1]. These datasets are both known to have a low dimensional manifold structure in a high dimensional image space.

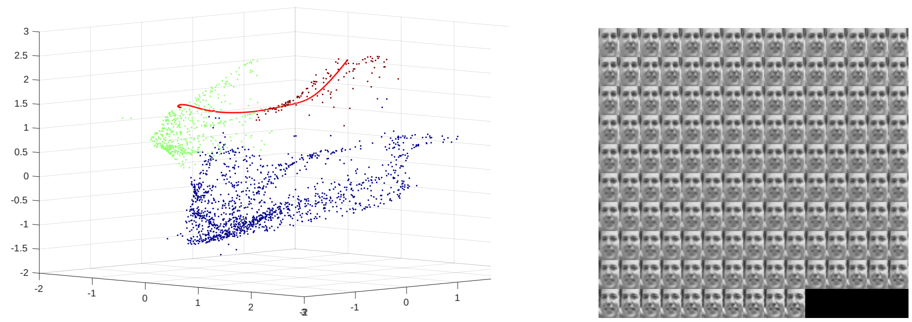
3.1 MNIST

Figure 1a shows the result of the dimensionality reduction for the MNIST dataset using PCA (in blue), as well as the manifold estimation using density ridge estimation (in green). Figure 1b shows the effect of the density ridge estimation, where the data point in blue is projected onto the estimated manifold in red. A clear 3-dimensional manifold structure can be observed. Interpolation along the geodesic path shown by the red curve in Figure 1a, yields the results in Figure 2. The first image (top left) and the last image (bottom right) are images from the MNIST dataset, which were projected onto the estimated manifold. All images inbetween are interpolated images along the geodesic path and are not in the original dataset. It can be seen that during the interpolation the angle in which the one leans changes smoothly from leaning to the right to being straight.

3.2 Frey Faces

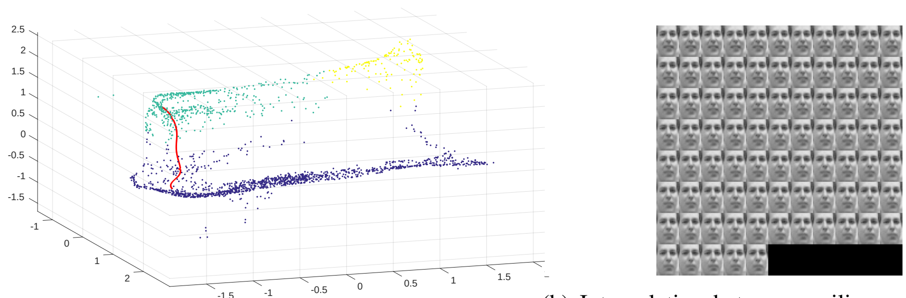
Figure 3 shows the result of the dimensionality reduction for the Frey Face dataset. As for the MNIST dataset, we can see a low-dimensional manifold structure. The colors indicate the three different modes that were found. Closer investigation showed that the blue points mainly correspond to frowning faces, whereas the green and the red points correspond to smiling faces. Figure 3b illustrates the results for the interpolation between two smiling faces along the red line in Figure 3a. Again, the first image (top left) is the original start image, and the bottom right is the original end image, whereas the images inbetween are interpolated. Results show that the head slowly turns from

[1] Obtained with kind permission from Brandon Frey, University of Toronto.



(a) The ridge for the Frey Faces dataset (b) Interpolation between smiling images

Figure 3: Results for interpolating between smiling images



(a) The ridge for the Frey Faces dataset (b) Interpolation between smiling and frowning images

Figure 4: Results for interpolating between smiling and frowning images

looking to the right to looking to the left. Note that all images appear to be realistic images. Figure 4b illustrates the results for the interpolation between the smiling and the frowning face. The path along the ridge can be seen in Figure 4a. Again, the first and last image correspond to the original images, whereas the images inbetween correspond to the interpolation along the geodesic path. The results show a smooth interpolation from frowning to smiling.

To end the experiments we include that the mean square errors between the input images and the MLP reconstructions shows that it is in fact useful. MSE MLP: 0.0094 and 366.14 and MSE PCA: $0, 06$ and 2.587×10^4 for MNIST and Frey faces respectively.

4 Discussion and Conclusion

In this work we propose that low dimensional manifold approaches can model complex shape variation. A limitation of the method is the fact that it requires the data to lie near a smooth low dimensional manifold. However, we have visual examples where this assumption is valid for real-world datasets. In both datasets a smooth interpolation was achieved and reconstruction produced realistic images.

Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation. This work was partially funded by the Norwegian Research Council grant no. 239844 on developing the *Next Generation Learning Machines*.

References

Piotr Dollár, Vincent Rabaud, and Serge Belongie. Non-isometric manifold learning: Analysis and an algorithm. In *Proceedings of the 24th international conference on Machine learning*, pp. 241–

248. ACM, 2007.

Christopher R Genovese, Marco Perone-Pacifico, Isabella Verdinelli, Larry Wasserman, et al. Nonparametric ridge estimation. *The Annals of Statistics*, 42(4):1511–1545, 2014.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Umut Ozertem and Deniz Erdogmus. Locally defined principal curves and surfaces. *The Journal of Machine Learning Research*, 12:1249–1286, 2011.

Matineh Shaker, Jonas N Myhre, M Devrim Kaba, and Deniz Erdogmus. Invertible nonlinear cluster unwrapping. In *Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on*, pp. 1–6. IEEE, 2014.