



Image elementary manifold and its application in image analysis

Chao Cai, Lingge Li and Changwen Zheng

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

November 7, 2018

Image elementary manifold and its application in image analysis

Chao Cai¹, Lingge Li¹ and Changwen Zheng²

¹ State Key Laboratory for Multispectral Information Processing Technologies
School of Automation, Huazhong University of Science and Technology
Wuhan, Hubei, 430074, P. R. China

² Science & Technology on Integrated Information System Laboratory
Beijing, 10080, P.R.China
caichao@hust.edu.cn

Abstract. Image basis function plays a key role in image information analysis. Due to the complex geometric structure in image, a better image basis or frame often have a very large family with a large number of basis functions lying in a lower dimensionality manifold, such as 2D Gabor functions and Contourlets used in image texture analysis, the corresponding image transform and analysis will be very time consuming. In this article, we propose a novel image representation method called “image elementary manifold”, here, an image elementary manifold can represent all the basis functions lying in the same manifold. A fast elementary manifold based image decomposition and reconstruction algorithm are given. Comparing to traditional image representation methods, elementary manifold based image analysis reduce time consumption, discovers the latent intrinsic structure of images more efficiently and provides the possibility of empirical prediction. Finally, many experiments show the feasibility of image elementary manifold in image analysis.

Keywords: image basis, manifold, image elementary manifold.

1 Introduction

In linear algebra, basis is a set of linearly independent vectors that can represent every vector in a given vector space via linear combination. In mathematical approximation theory, it requires only the linear combination of few vectors to reconstruct signal accurately. For example, the basis functions of Fourier transformation is Sine function and Cosine function, the input signal is limited to $L^1(R)$ [1], then it is extended to $L^2(R)$ [2]. In 1946, Dennis Gabor [3] proposed that the signal to be analyzed is first multiplied by a window function for Time-Frequency Localization. A Gabor system may be a basis for $L^2(R)$, which is generated from a single $L^2(R)$ -function through phase space translations and modulations. Wavelet transform [4, 5] inherits and develops the localization performance of Gabor transform, overcomes its shortcoming of lack of discrete orthogonal basis. Do and Vetterli [6,7] constructed a double filter

bank structure which results in a flexible multiresolution, local, and multidirectional image expansion using contour segments, named Contourlet Transform. The contourlet expansion is shown to achieve the optimal approximation rate for piecewise C^2 smooth images with C^2 smooth contours. For this class of functions, the decay rate of approximation error of contourlet transform is far better than wavelets and Fourier basis [5, 8].

As for images, there are still some problems in subspace decomposition and reconstruction. Traditional image decomposition method requires convolving the image with all the basis functions, but the rotation and dilation of a mother function often generate a large number of basis functions, this leads to a large amount of calculation. For example, Gabor Transform and Contourlet Transform use multidirectional basis functions to achieve better approximation, but the large time consuming limits their application. Actually, there are potential low dimensional manifolds exist in basis functions, generating through shifting, scaling, rotating and so on. Further more, one can decompose images into different sub-manifold space. In this article, we classify the basis functions based on their embedding manifold. Therefore, the computational power and the empirical predictive power of image analysis will be greatly increased.

2 Image elementary manifold

2.1 Manifold embedded property of image basis functions

A manifold is an abstract mathematical space that near each point resembles Euclidean space. Many image sets vary due to a small number degrees of freedom and the set of these images lie in or near to some low-dimensional manifolds embedded in a high dimensional (e.g. equal to the number of image pixels) image spaces [9]. Manifold learning algorithms can infer global structures from locally computed geometric properties (such as distances, angles, and symmetry). Existing manifold learning algorithm includes distance-preserving methods (such as ISomap [10], MVU [11]), angle-preserving methods (such as conformal eigenmaps [12]) and proximity-preserving methods (such as LLE [13]) [14]. Wu Hui, Kilian Q, Saul, et al. prompt a data-driven method for semi-supervised multioutput regression on image manifolds [15].

Manifold embedded property could be found in many kinds of basis functions. Shifting, rotating or scaling always exists in basis functions, and each kind of these basis images should lie in a low-dimensional manifold. Simply looking at the Gabor basis functions in Fig.1, all basis functions in one scale are generated by rotating the zero direction image (the top image), so that these images lie in a one-dimensional manifold. These basis functions could be packaged together as a manifold, and this manifold will be treated as an elementary function in image decomposition and reconstruction. Therefore, the manifold spans the space which is spanned by Gabor basis functions.

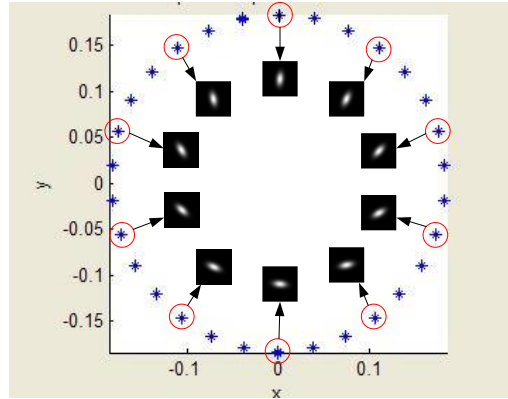


Fig.1. Gabor basis functions and the one-dimensional manifold learned by MVU algorithm. Each asterisk stands for an elementary image which is generated by rotating a zero direction image. The picture that arrow points to is the corresponding basis function.

2.2 Image elementary manifold

As mentioned above, a series of basis functions that lie in a manifold could be packaged together for follow-up computation. A package model will be introduced in the following paragraphs.

Suppose that a manifold S_p of the intrinsic dimension d in the vector input space is a set of some sample points and their deformations, such as shifting, rotating. It could be formulated as follows:

$$S_p = \{x | \exists \alpha, x = s(P, \alpha)\} \quad (1)$$

Where s the deformation function, P is the sample points, $\alpha \in R^d$ is deformation parameter.

A core task of manifold learning is to find the low-dimensional structure (potential embedded manifold) of high-dimensional observed data, that is to find deformation parameter α and deformation function s . Many manifold learning algorithms can give a low dimensional global coordinates from high dimensional input data. If we take these low dimensional variables which have been learned by manifold learning algorithm as the deformation parameters, a deformation mapping between high-dimensional data and low-dimensional variables can be obtained.

In this paper, we use Maximum Variance Unfolding (MVU) algorithm, which attempts to unfold curved manifolds while preserving local geometry. After obtaining the intrinsic variables of image transformation by manifold learning algorithm, the deformation mapping between high-dimensional space and low-dimensional space is obtained by higher-order Taylor expansion, and then the nonlinear manifold is approximated by polynomial.

$$s(P, \alpha) = s(P, 0) + \alpha \left. \frac{\partial s(P, \alpha)}{\partial \alpha} \right|_{\alpha=0} + 0.5\alpha^2 \frac{\partial^2 s(P, \alpha)}{\partial \alpha^2} + O(\alpha^3) \approx P_0 + \alpha T' + 0.5\alpha^2 T'' \quad (2)$$

Here T' is tangent vector, T'' is the second-order partial derivative. As shown in Fig.2: $P_0 + \alpha T'$ is the first-order Taylor expansion to approximate the point P on the manifold surface, $P_0 + \alpha T' + 0.5\alpha^2 T''$ is the second-order Taylor expansion to approximate the point P on the manifold surface.

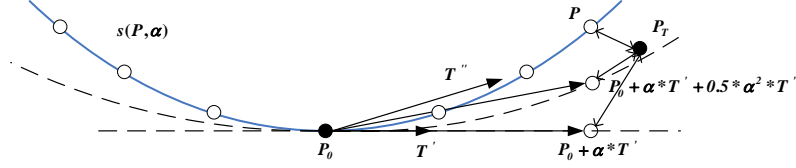


Fig.2. Manifold distance computation based on Taylor-expansion approximation

The approximated manifold s is called an elementary manifold of basis functions. P means that a collection of similar basis functions. All these functions are generated from a central function and modeled by a multivariable polynomial. The elementary manifold s will be an elementary cell in image decomposition and reconstruction.

In many cases, the analytical expression of s is unknown, so, the partial differentiation in equation (2) will be represented by difference in calculation, that is:

$$\frac{\partial s(P, \alpha)}{\partial \alpha} \approx \frac{\Delta s(P, \alpha)}{\Delta \alpha} \quad (3)$$

Traditional transform method convolves the input image with all basis functions, which needs a huge consumption of calculation. Based on elementary manifold, we just compute correlation between input image and manifold instead of a series of basis functions. The correlation details based on manifold will be shown in next sections.

2.3 The decomposition and reconstruction of images based on image elementary manifold

Fig.3 illustrates the distance between manifolds. If S_p and S_E are all represented by polynomials, the problem will be transformed into solving the distance between two polynomial surfaces.

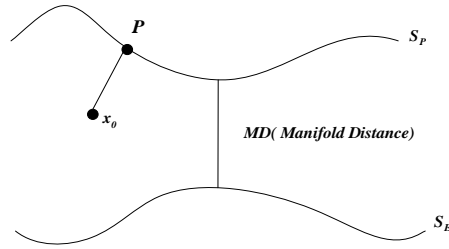


Fig.3. Manifold distance diagram. The manifold distance (MD) between corresponding manifolds S_p and S_E . The correlation between the point x_0 and the manifold S_p .

Based on Taylor expression, manifold could be modeled by a multivariable polynomial whose dimension (D) and order (O) are settable values. Therefore, the correlation between the point and manifold is defined as:

$$r = \max_{\alpha} \text{cov}(x_0, s(P, \alpha)) \quad (4)$$

The Formula (4) constitutes a new polynomial. So, the problem of calculating the correlation between point x_0 and manifold S_p is converted to solving the maximum of polynomial with fewer variables. Solving the maximum of polynomial directly provides the possibility of empirical prediction.

From the point of linear algebra, we can decompose image $I(x, y)$ into linear combination of image basis if the image is in the space spanned by the basis, which is the synthesis equation

$$I(x, y) = \sum_i a_i \tilde{\omega}_i(x, y) \quad (5)$$

Where $\tilde{\omega}_i(x, y)$ is called the dual function of the basis function $\omega_i(x, y)$, a_i is the decomposition coefficient. According to the dual property of basis functions, we have the analysis equation

$$a_i = \langle I(x, y), \omega_i(x, y) \rangle \quad (6)$$

The process of calculating coefficient a_i according to Formula (6) can be considered as decomposition process. And the weighted summation of basis functions according to Formula (5) can be seen as reconstruction process. When the number of basis functions is large, calculating coefficients requires a huge consumption.

In this paper, we introduce the image elementary manifold. Basis functions are classified into different manifolds. Therefore, it only needs to compute the manifold distance for image decomposition and reconstruction. Formula (5) could be converted to Formula (7):

$$I(x, y) = \sum_i r_i \tilde{s}_i \quad (7)$$

Where \tilde{s}_i refers to the dual elementary manifold, r_i is the corresponding decomposition coefficient which is the maximum of the polynomial in Formula (4).

In this work, the concept of inner product of elementary manifolds and the input image is introduced. It replaces the inner product of basis functions and the input image. According to the dual property of basis functions, we can also infer the equation:

$$r_i = \langle I(x, y), s_i \rangle \quad (8)$$

Similarly, the same kind of dual basis functions could be trained as a manifold in image reconstruction. According to Formula (7), using dual elementary manifolds and corresponding coefficients can reconstruct the input image.

2.4 Computation analysis of elementary manifold decomposition

Using elementary manifolds can significantly reduce the calculation of the decomposition and reconstruction. Suppose there are N basis functions, and there are t pixels in each basis image. According to Formula (6), conventional method convolves the

input image (K pixels) with all basis functions. This means that $N*t*K$ times multiplication and $N*t*K$ times addition need to be calculated.

For elementary manifold method, suppose each Taylor-expansion polynomial includes P terms, there are t polynomials. The value of P is determined by the order of the polynomial O and the number of free variables D . According to Formula (4) and (8), the decomposition process needs to calculate $P*t*K$ times multiplication and $P*t*K$ times addition to construct a new polynomial that represents the correlation between point and manifold.

Generally, the number of polynomial terms is far less than the number of basis functions ($P \ll N, P*t*K \ll N*t*K$). Therefore, elementary manifold method will show its advantage. For example, $t=13*13=169$, $P=20$, $N=90$, $K=128*128$, in conventional method, $N*t*K=249,200,640$ times multiplication and 249200640 times addition need to be calculate, while in our method, it only needs to calculate $P*t*K=55,377,920$ times multiplication and 55377920 times addition.

3 Some applications about elementary manifold

In this section, we verify the feasibility of the image elementary manifold, and show applications in image decomposition, image reconstruction and edge detection through a series of experiments. In those experiments, we first test a simple image elementary manifold composed of 13×13 pixels stepped edge images with different directions and displacements (as shown in the Fig.4), these basis functions span a piecewise constant space. Then, other classic basis functions like Gabor functions and Log-Gabor functions are used in elementary manifold method. We use the MVU manifold learning algorithm to get the manifold knowledge.

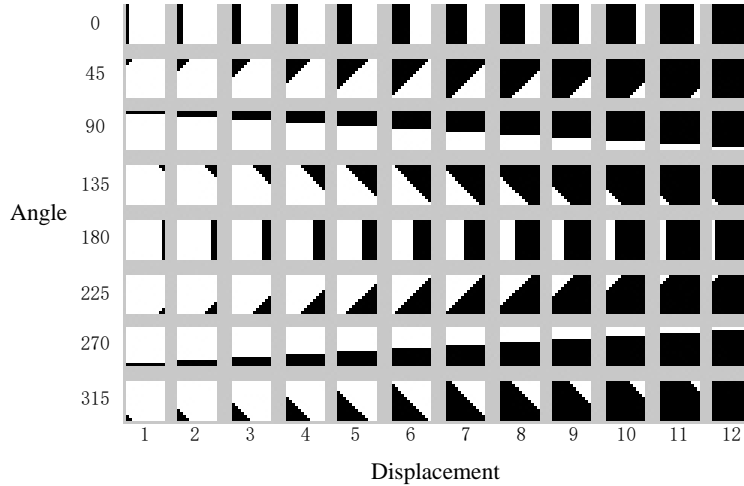


Fig.4. Some 13×13 pixels stepped edge elementary images

3.1 Image Decomposition

By calculating the correlation between each small image block of the input image and the elementary manifold, namely the correlation coefficient, a correlation coefficient matrix can be generated, which is the image decomposition process. Some classic basis functions also can be used in this method, such as Gabor basis functions. As mentioned in Section 2.4, we now present several experiment results with elementary manifold method and compare it with the performance of Gabor transform directly. The decomposition coefficient image is shown in Fig.5 (b), stepped edge images were treated as a two-dimensional manifold for manifold learning. The transform effectively shows the fact image edges. In Fig.5(c), Gabor basis functions are considered lying in two manifolds (real part and imaginary part). The experimental results proved the feasibility of elementary manifold method.



Fig.5. Image decomposition. (a) Input image. (b) Decomposition image with stepped edge elementary manifold. (c) Decomposition image with Gabor basis function manifold.

In this experiment, we choose $t=13*13=169$, $P=20$, $N=90$, the time consuming of our method is far less than the conventional two dimensional convolution method. With the unified data types (“double”), our approach takes 5 seconds while traditional Gabor transform takes 24 seconds for 512×512 pixels image in the same hardware platform.

As mentioned in Section 2, the computation complexity is closely related to the number of polynomial terms. Therefore, we took some experiments to discover the relationship between the terms and the processing time. A 512×512 pixels image is decomposed, with $t=13*13=169$, $N=90$. The result is shown in Table 1. Time consumption is proportional to the number of polynomial terms.

Table.1. The processing time according to increasing polynomial terms. The relation of the number of free variables (D) and orders (O) to polynomial terms (M) and time consumption.

(D, O)	(2,2)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)	(7,3)
P	6	10	20	35	56	84	120
Time consumption(ms)	3033	3997	4939	5889	8403	10747	14472

3.2 Image Reconstruction

According to Formula (7), calculating the linear combination of dual elementary manifolds can reconstruct the input image. In this experiment, we use two different elementary image sets to reconstruct the “Lena” image (256×256 pixels). From Fig.6(a), we find that the result reconstructed by the stepped edge elementary image set (13×13 pixels) mentioned before describes the geometric structure information of the input image correctly, while the hair, facial features and other details are not so accurate. The straight lines in the figure are very clear, although curves are not smooth. As shown in Fig.6(c), it is reconstructed by the stepped edge elementary image set (3×3 pixels), which can almost recover the input image. Fig. 6(b) shows an example of the Log-Gabor transform on the “Lena” image. The reconstruction image is smoother, which mainly results from the different characteristics of basis functions.



Fig.6. Reconstruction image by elementary images with different sizes. (a) Original image. (b) The image is reconstructed by Log-Gabor transform. (c) The image is reconstructed by the stepped edge elementary images (13×13 pixels). (d) The image is reconstructed by the stepped edge elementary images (3×3 pixels).

3.3 Image Edge Detection

Image elementary manifold can be used for image edge detection. For example, in Fig.4, each elementary image is expected to represent an “edge”. The input image is first removed the mean image, and then decomposed into correlation coefficient image, which reflects the intensity of stepped edge. Therefore, the edges of the input image would be highlighted in the transformed image. Since the linear equations of learning samples are known, we can obtain the gradient direction and amplitude of each point of transformed image. We follow on do non-maximum suppression for gradient amplitudes. Then we use the double threshold algorithm to detect and connection edges. Finally, the clear edges would be obtained.

In this experiment, the elementary images are used to detect the edges of a cameraman image, of size 256×256 pixels. The size of each elementary image is 13×13 pixels. The experimental results are shown in Fig.7. It is shown that the edge features of the cameraman and the camera in the image are detected very well, and the lines are continuous and smooth. Due to the limited size of the elementary image, many detailed edges cannot be accurately described.

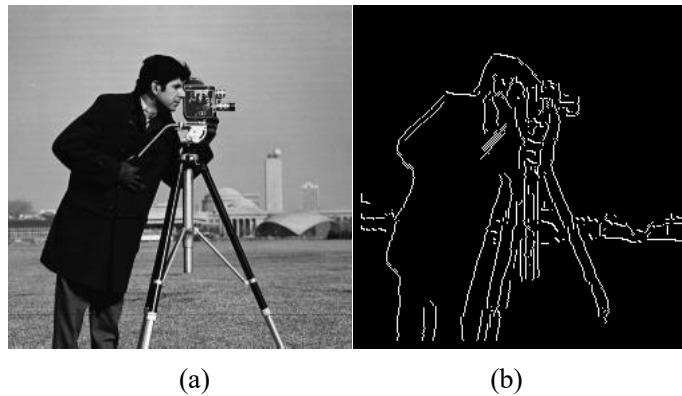


Fig.7. Edge detection images. (a) Input image. (b) Edge image.

4 Conclusion

In this paper we proposed a new elementary manifold method in combined with image multi-scale geometric analysis. It is shown that this method can improve the computing performance of image basis function analysis. In this paper, we calculate the correlation coefficients between elementary manifold and the input image in spatial domain, whereas in traditional method, correlation coefficients image is constructed by using all basis functions convolve with the input image. In order to improve the computing speed, the traditional method of digital image processing deals with images via the transformation from the image to the frequency domain processing. The proposed algorithm could show great advantages when the number of basis functions is considerably larger than the number of Taylor expansion terms. Furthermore, the

elementary manifold spans the space which is spanned by the basis functions that lying in the manifold. Experiments indicate the potential of the elementary manifold method in image processing applications. The disadvantages of this method include two aspects. For one thing, despite the manifold learning algorithm is varied, but there are still unstable situation when training basis functions; for another, the increase of the Taylor expansion order will lead to a sharply increase of computing cost, and if the order of the polynomial is too low will result in fitting error. For future work, we will try to transform the elementary manifold to the frequency domain processing to improve the algorithm efficiency.

References

1. A. Papoulis. The Fourier Intergral and its Applications. McGraw-Hill, New York, NY, second edition, 1987.
2. H. Dym and H. P. McKean. Fourier Serirs and Integral. Academic Press, New York, 1972.
3. D. Gabor, Theory of communication, J. Inst. Elect. Eng. (London), vol. 93, no. 111, pp. 429-457, 1946.
4. Lee D T L, Yamamoto A. Wavelet analysis: Theory and applications [J]. Hewlett Packard journal, 1994, 45: 44-44.
5. S. Mallat, A Wavelet Tour of Signal Processing, 2nd Ed. New York: Academic, 1999.
6. Do M N, Vetterli M. Contourlets: a directional multiresolution image representation[C]//Image Processing. 2002. Proceedings. 2002 International Conference on. IEEE, 2002, 1: I-357-I-360 vol. 1.
7. Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation [J]. Image Processing, IEEE Transactions on, 2005, 14(12): 2091-2106.
8. D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2435–2476, Oct. 1998.
9. Pless R, Souvenir R. A survey of manifold learning for images [J]. IPSJ Transactions on Computer Vision and Applications, 2009, 1(0): 83-94.
10. Tenenbaum JB, de Silva V, Langford JC., "A global geometric framework for nonlinear dimensionality reduction", Science. 290(12), 2319-2323(2000).
11. K. Q. Weinberger and L. K. Saul, "Unsupervised Learning of Image Manifolds by Semidefinite Programming," International Journal of Computer Vision, vol.70, no.1, pp. 77-90, 2006.
12. F. Sha and L.K. Saul, "Analysis and Extension of Spectral Methods for Nonlinear Dimensionality Reduction," Proc. 22nd Int'l Conf. Machine Learning (ICML 2005), ACM Press, 2005, pp. 784–791.
13. S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science, vol.290, no.5500, pp. 2323-2326, 2000.
14. Zhang J, Huang H, Wang J. Manifold learning for visualizing and analyzing high-dimensional data[J]. IEEE Intelligent Systems, 2010: 54-61.
15. Wu, Hui, Spurlock, Scott, Souvenir, Richard. Semi-supervised multi-output image manifold regression [C], IEEE International Conference on Image Processing, vol.70, 2017:2413-2417.