HOSVD Based Image Processing Techniques

András Rövid, Imre J. Rudas, Szabolcs Sergyán
Obuda University
John von Neumann Faculty of Informatics
Bécsi út 96/b, 1034 Budapest
Hungary
rovid.andras@nik.uni-obuda.hu, rudas@uni-obuda.hu
sergyan.szabolcs@nik.uni-obuda.hu

László Szeidl
Széchenyi István University
System Theory Laboratory
Egyetem tér 1., 9026 Győr
Hungary
szeidl@sze.hu

Abstract: In the framework of the paper an improved method for image resolution enhancement is introduced. The enhancement is performed through another representation domain, where the image function is expressed with the help of polylinear functions on higher order singular value decomposition (HOSVD) basis. The paper gives a detailed description on how to determine the polylinear functions corresponding to an image and how to process them in order to obtain a higher resolution image. Furthermore, the proposed approach and the Fourier-based approach will be compared from the point of view of their effectiveness.

Key–Words: Image resolution, HOSVD, approximation, numerical reconstruction

1 Introduction

Nowadays the importance of image processing and machine vision-based applications is increasing significantly. In order to perform the desired task more accurately and reliably, the related algorithms should be developed. An important factor regarding the processing is the applied data representation domain, in which the processing is performed. Representing an image in other domain may give new possibilities regarding its processing. In many cases these representations are related to expressing the image intensity function as a combination of simpler functions (components) having useful predefined properties. Let us mention some applications and some concepts of the related methods and algorithms from the field of image processing, where the applied domain plays a significant role.

The image resolution enhancement, filtering [1], image compression [2], etc. can be performed much more effectively when working in another image representation domain [3]. In the frequency domain for instance, the image compression is much more efficient than in the spatial one. On the other hand, to represent the image in frequency domain without meaningful quality decline, relatively large number of trigonometric components is needed.

Another well known application is the resolution enhancement of an image, investigated in the framework of this paper in more details, focusing on its realization and effectiveness in the proposed domain.

Numerical reconstruction or recovering of a continuous intensity surface from discrete image data samples is considered, for example, when the image is rescaled or remapped from one pixel grid to another one. In case of image enlargement the color components or in case of grayscale images the intensity of missing pixels should be estimated. One common way for estimating the color values of such pixels is interpolating the discrete source image. There are several issues which affect the perceived quality of the interpolated images: sharpness of edges, freedom from artifacts and reconstruction of high frequency details.

There are numerous methods approximating the image intensity function based on the color and location of known image points, such as the bilinear, bicubic or spline interpolation all working in the spatial domain of the input image.

A new data representation domain is introduced in the paper, connected to the higher order singular value decomposition (HOSVD), which is a generalized form of the well known singular value decomposition (SVD). As shown in the upcoming sections, any n-variable smooth function can be expressed with the help of a system of orthonormal one-variable smooth functions on HOSVD basis. The main aim of the paper is to numerically reconstruct these specially determined one-variable functions using the HOSVD and to show that how this approach can give support for certain image processing tasks and problems.

The paper is organized as follows: Section 2 gives a closer view on how to express a multidimensional function using polylinear functions on HOSVD basis, and how to reconstruct these polylinear functions,
Section 3 shows how this representation can be applied in image processing for resolution enhancement, while in section 4 the proposed method is compared to the well known Fourier transformation from the approximation error point of view. Section 5 shows the experimental results and finally, conclusions are reported.

2 Theoretical background

The approximation methods of mathematics are widely used in theory and practice for several problems. If we consider an \(n\)-variable smooth function

\[
f(x), \quad x = (x_1, \ldots, x_N)^T, \quad x_n \in [a_n, b_n], \quad 1 \leq n \leq N,
\]

then we can approximate the function \(f(x)\) with a series

\[
f(x) = \sum_{k_1=1}^{I_1} \cdots \sum_{k_N=1}^{I_N} \alpha_{k_1, \ldots, k_N} p_{k_1, k_2}(x_1) \cdots p_{k_N, k_N}(x_N).
\] (1)

where the system of orthonormal functions \(p_{k_1, k_2}(x_n)\) can be chosen in classical way by orthonormal polynomials or trigonometric functions in separate variables and the numbers of functions \(I_n\) playing role in (1) are large enough. With the help of Higher Order Singular Value Decomposition (HOSVD) a new approximation method was developed in [7], [5] in which a specially determined system of orthonormal functions can be used depending on function \(f(x)\), instead of some other systems of orthonormal polynomials or trigonometric functions.

Assume that the function \(f(x)\) can be given with some functions \(\tilde{w}_{n,i}(x_n), x_n \in [a_n, b_n]\) in the form

\[
f(x) = \sum_{k_1=1}^{I_1} \cdots \sum_{k_N=1}^{I_N} \alpha_{k_1, \ldots, k_N} \tilde{w}_{1,k_1}(x_1) \cdots \tilde{w}_{N,k_N}(x_N).
\] (2)

Denote by \(A \otimes_n \mathbf{U}\) the \(n\)-mode tensor-matrix product, \(A \otimes_n \mathbf{U} = \{a_{k_1, \ldots, k_N} \tilde{w}_{1,k_1}(x_1) \cdots \tilde{w}_{N,k_N}(x_N)\}\), \(1 \leq k_n \leq I_n, 1 \leq n \leq N\) and let us use the following notations (see [4]).

- \(A \otimes_n \mathbf{U}\): the \(n\)-mode tensor-matrix product,
- \(A \otimes_n \mathbf{U} \otimes_n \mathbf{U}\): the multiple product as \(A \otimes_n \mathbf{U} \otimes_n \mathbf{U}\),
- \(A \otimes_n \mathbf{U} \otimes_n \mathbf{U}\): the \(n\)-mode tensor-matrix product is defined by the following way. Let \(U\) be an \(K_n \times M_n\)-matrix, then \(A \otimes_n \mathbf{U}\) is a \(M_1 \times \ldots \times M_{n-1} \times K_n \times M_{n+1} \times \ldots \times M_N\)-tensor for which the relation

\[
(A \otimes_n \mathbf{U})(m_1, \ldots, m_{n-1}, k_n, m_{n+1}, \ldots, m_N) \overset{def}{=} \sum_{1 \leq m_n \leq M_n} a_{m_1, \ldots, m_n, \ldots, m_N} U_{k_n, m_n}
\]

holds. Detailed discussion of tensor notations and operations is given in [4]. We also note that we use the sign \(\otimes\) instead of some other systems of orthonormal polynomials or trigonometric functions.

\[
f(x) = A \otimes_n \mathbf{U} \cdot \tilde{w}_n(x_n),
\] (3)

where \(\tilde{w}_n(x_n) = (\tilde{w}_{n,1}(x_n), \ldots, \tilde{w}_{n,I_n}(x_n))^T, 1 \leq n \leq N\). Based on HOSVD it was proved in [6] that under mild conditions the (3) can be represented in the form

\[
f(x) = D \otimes_n \mathbf{U} \cdot \tilde{w}_n(x_n),
\] (4)

where

- \(D \in \mathbb{R}^{r_1 \times \ldots \times r_N}\) is a special (so called core) tensor with the properties:

1. \(r_n = \text{rank}_n(A)\) is the \(n\)-mode rank of the tensor \(A\), i.e. rank of the linear space spanned by the \(n\)-mode vectors of \(A\):

   \[
   \langle (a_{1, \ldots, i_n-1, i_n+1, \ldots, i_N, \ldots, a_{1, \ldots, i_n-1, i_n, i_n+1, \ldots, i_N})^T : 1 \leq i_n \leq I_n, 1 \leq n \leq N
   \]

2. all-orthogonality of tensor \(D\): two subtensors \(D_{i_n=\alpha}\) and \(D_{i_n=\beta}\) (the \(n\)-th indices \(i_n = \alpha\) and \(i_n = \beta\) of the elements of the tensor \(D\) keeping fix) orthogonal for all possible values of \(\alpha, \beta\) : \(\langle D_{i_n=\alpha}, D_{i_n=\beta} \rangle = 0\) when \(\alpha \neq \beta\). Here the scalar product \(\langle D_{i_n=\alpha}, D_{i_n=\beta} \rangle\) denotes the sum of products of the appropriate elements of subtensors \(D_{i_n=\alpha}\) and \(D_{i_n=\beta}\).

3. ordering: \(\|D_{i_n=1}\| \geq \|D_{i_n=2}\| \geq \cdots \geq \|D_{i_n=r_n}\| > 0\) for all possible values of \(n\) \((\|D_{i_n=\alpha}\| = \langle D_{i_n=\alpha}, D_{i_n=\alpha} \rangle\) denotes the Kronecker-norm of the tensor \(D_{i_n=\alpha}\).

- Components \(w_{n,i}(x_n)\) of the vector valued functions

   \[
w_{n}(x_n) = (w_{n,1}(x_n), \ldots, w_{n,r_n}(x_n))^T, 1 \leq n \leq N,
   \]

are orthonormal in \(L_2\)-sense on the interval \([a_n, b_n]\), i.e.

\[
\forall n : \int_{a_n}^{b_n} w_{n,i}(x_n) w_{n,j}(x_n) dx = \delta_{i,j},
\]

\[1 \leq i, j \leq r_n,\]
where $\delta_{i,j}$ is a Kronecker-function ($\delta_{i,j} = 1$, if $i = j$ and $\delta_{i,j} = 0$, if $i \neq j$).

The form (4) was called in [6] HOSVD canonical form of the function (2).

Let us decompose the intervals $[a_n, b_n], n = 1..N$ into $M_n$ number of disjunct subintervals $\Delta_{n,m_n}, 1 \leq m_n \leq M_n$ as follows:

$$\xi_{n,0} = a_n < \xi_{n,1} < \ldots < \xi_{n,M_n} = b_n,$$

$$\Delta_{n,m_n} = [\xi_{n,m_n}, \xi_{n,m_n-1}].$$

Assume that the functions $w_{n,k_n}(x_n), x_n \in [a_n, b_n], 1 \leq n \leq N$ in the equation (2) are piecewise continuously differentiable and assume also that we can observe the values of the function $f(x)$ in the points

$$y_{i_1,\ldots,i_N} = (x_{1,i_1},\ldots,x_{N,i_N}), 1 \leq i_n \leq M_n. \quad (5)$$

where

$$x_{n,m_n} \in \Delta_{n,m_n}, 1 \leq m_n \leq M_n, 1 \leq n \leq N$$

Based on the HOSVD a new method was developed in [6] for numerical reconstruction of the canonical form of the function $f(x)$ using the values $f(y_{i_1,\ldots,i_N}), 1 \leq i_n \leq M_n, 1 \leq i_n \leq N$. We discretize function $f(x)$ for all grid points as:

$$b_{m_1,\ldots,m_N} = f(y_{m_1,\ldots,m_N}).$$

Then we construct $N$ dimensional tensor $B = (b_{m_1,\ldots,m_N})$ from the values $b_{m_1,\ldots,m_N}$. Obviously, the size of this tensor is $M_1 \times \ldots \times M_N$. Further, we discretize vector valued functions $w_{n}(x_{n})$ over the discretization points $x_{n,m_n}$ and construct matrices $W_n$ from the discretized values as:

$$w_n = \begin{pmatrix}
w_{n,1}(x_{n,1}) & w_{n,2}(x_{n,1}) & \cdots & w_{n,r_n}(x_{n,1}) \\
w_{n,1}(x_{n,2}) & w_{n,2}(x_{n,2}) & \cdots & w_{n,r_n}(x_{n,2}) \\
\vdots & \vdots & \ddots & \vdots \\
w_{n,1}(x_{n,M_n}) & w_{n,2}(x_{n,M_n}) & \cdots & w_{n,r_n}(x_{n,M_n})
\end{pmatrix} \quad (6)$$

Then tensor $B$ can simply be given by (4) and (6) as

$$B = D \otimes_{n=1}^{N} W_n. \quad (7)$$

For further details see [5].

3 Resolution Enhancement on HOSVD basis

Let $f(x), x = (x_1, x_2, x_3)^T$ represent the image function, where $x_1$ and $x_2$ correspond to the vertical and horizontal coordinates of the pixel, respectively. $x_3$ is related to the color components of the pixel, i.e. the red, green and blue color components in case of RGB image. Function $f(x)$ can be approximated (based on notes discussed in the previous section) in the following way:

$$f(x) = \sum_{k_1=1}^{I_1} \sum_{k_2=1}^{I_2} \sum_{k_3=1}^{I_3} a_{k_1,k_2,k_3} \bar{w}_{i_1,k_1}(x_1) \cdot \bar{w}_{i_2,k_2}(x_2) \cdot \bar{w}_{i_3,k_3}(x_3). \quad (8)$$

The red, green and blue color components of pixels can be stored in a $m \times n \times 3$ tensor, where $n$ and $m$ correspond to the width and height of the image, respectively. Let $B$ denote this tensor. The first step is to reconstruct the functions $\bar{w}_{n,k_n}, 1 \leq n \leq 3, 1 \leq k_n \leq I_n$ based on the HOSVD of tensor $B$ as follows:

$$B = D \otimes_{n=1}^{3} U(n) \quad (9)$$

where $D$ is the so called core tensor. Vectors corresponding to the columns of matrices $U(n), 1 \leq n \leq 3$ as described in the previous section are representing the discretized form of functions $\bar{w}_{n,k_n}(x_{n})$ corresponding to the appropriate dimension $n, 1 \leq n \leq 3$.

Our goal is to demonstrate the effectiveness of image scaling in the HOSVD based domain.

Let $s \in \{1,2,\ldots\}$ denote the number of pixels having to be injected between each neighbouring pixel pair in horizontal and vertical directions. First, let us consider the first column $U(1)^{(1)}$ of matrix $U^{(1)}$.

Based on the previous sections, it can be seen, that the value $\bar{w}_{1,1}(1)$ corresponds to the 1st element of $U^{(1)}$, $\bar{w}_{1,1}(2)$ to the 2nd element, ..., $\bar{w}_{1,1}(M_n)$ to the $M_n$th element of $U^{(1)}$. To scale the image in the HOSVD-based domain, the $U^{(1)}, i = 1.2$ matrices should be updated, depending on $s$, as follows: The number of columns remains the same, the number of lines will be extended according to the factor $s$. Let us denote the such obtained matrix as $V^{(1)}$. For example let consider the column $U^{(1)}_{1}$ of $U^{(1)}$. The elements of $V^{(1)}_{1}$ are determined as follows: $V^{(1)}_{1}(1) := U^{(1)}_{1}(1)$, $V^{(1)}_{1}(s+2) := U^{(1)}_{1}(2), V^{(1)}_{1}(2s+3) := U^{(1)}_{1}(3), \ldots$, $V^{(1)}_{1}((M_n-1)s+M_n) := U^{(1)}_{1}(M_n)$. The missing elements of $V^{(1)}_{1}$ can be determined by interpolation.

In the paper the cubic spline interpolation was applied. The remaining columns should be processed similarly. After every matrix element has been determined the enlarged image can be obtained using the equation (9).

4 Fourier vs. HOSVD

The proposed approach introduced in the previous sections uses orthonormal functions $\bar{w}_{n,i}(x_{n}), x_{n} \in
\([a_n, b_n]\) (see Section 2) for approximating an \(n\)-variable smooth function. We saw how these functions \(\tilde{w}_{n,i}(x_n)\) can numerically be reconstructed and what properties they have. Comparing the proposed approach to the Fourier transformation, similarities can be observed in their behaviour. As it is well known, the Fourier Transformation is connected to trigonometric functions, while in case of HOSVD approach the functions \(\tilde{w}_{n,i}(x_n)\) are considered, which are specific to the approximated \(n\)-variable function. In both cases the functions are forming an orthonormal basis. Since in case of HOSVD the functions are specific ones, much fewer number of components is needed then in case of Fourier based approach to achieve the same approximation accuracy.

Let us mention some common, widely used applications of both approaches.

In case of Fourier based smoothing, some of higher frequencies from the frequency domain are dismissed, resulting a smoothed image (low pass filter).

In case of HOSVD similar effect can be observed when dismissing polylinear functions corresponding to smaller singular values for certain dimensions. The same concept can be used also for data compression.

In the opposite case, i.e. when maintaining only the functions corresponding to the smaller singular values, will result an edge detector. In case of Fourier approach detecting edges in an image is equivalent to dismissing the smaller frequency components (high pass filter).

The examples show that in case of HOSVD much smaller number of basis functions is enough than in case of Fourier-based approach to represent the image without significant information loss. When dismissing high frequency components, there is a frequency threshold depending on the concrete image, below which by dismissing further frequencies some kind of waves can be observed in the image as noise. In case of the proposed approach the ratio of maintained and dismissed components may be significantly smaller in order to achieve the same quality than in case when trigonometric functions are applied. Vectors corresponding to the columns of matrices \(\mathbf{U}^{(n)}\), \(1 \leq n \leq 3\) as described in the previous section contain the control points of functions \(\tilde{w}_{n,k_n}(x_n)\) corresponding to the appropriate dimension \(n, 1 \leq n \leq 3\). It means that there will be as many one-variable functions for a dimension as many columns there are in the orthonormal matrix corresponding to that dimension. The number of these functions can be further decreased by dismissing some columns from the orthonormal matrices obtained by HOSVD. Let \(C_n, 0 \leq C_n \leq I_n, n = 1..N\) stand for the number of dismissed columns in \(n\)th dimension. Dismissing some of the columns is equivalent to dismissing some of the frequencies in case of Fourier approach.

\[
f(x) = \sum_{k_1=1}^{C_1} \sum_{k_2=1}^{C_2} \sum_{k_3=1}^{C_3} \alpha_{k_1,k_2,k_3} \tilde{w}_{1,k_1}(x_1) \cdot \tilde{w}_{2,k_2}(x_2) \cdot \tilde{w}_{3,k_3}(x_3).
\]

The below examples clearly show that the proposed approach has good compression capabilities, which further extends its applicability also in the field of image processing.

5 Examples

Part-1 (HOSVD vs. Fourier)

In this section some approximations can be observed, performed by the proposed and by the Fourier-based approach. As the number of the used components decreases, the observable differences in quality become more significant. In the examples below in both the HOSVD-based and Fourier-based cases the same number of components have been used in order to show how the form of determined functions influences the quality. The differences in quality of below images are much more observable in color images (see the electronic version of the paper).

Part-2 (Resolution Enhancement)

The pictures are illustrating the effectiveness of the image zooming using the proposed approach. The results obtained by the HOSVD are compared to the results obtained by the bilinear and bicubic image interpolation methods.

6 Conclusion

In the present paper a new image representation domain and reconstruction technique has been introduced. The results show that how the efficiency of the certain tasks depends on the applied domain. Image rescaling has been performed using the proposed technique and has been compared to other well known image interpolation methods. Using this technique the resulted image maintains the edges more accurately then the other well-known image interpolation methods. Furthermore, some properties of the proposed representation domain have been compared to the corresponding properties of the Fourier-based approximation. The results show that in the proposed domain some tasks can be performed more efficiently than in other domains.
Figure 1: Original image (24bit RGB)

Figure 2: HOSVD-based approximation using 7500 components composed from polylinear functions on HOSVD basis

Figure 3: Fourier-based approximation using 7500 components composed from trigonometric functions

Figure 4: HOSVD-based approximation using 2700 components composed from polylinear functions on HOSVD basis

Figure 5: Fourier-based approximation using 2700 components composed from trigonometric functions

Figure 6: The original image.
Acknowledgements: The research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and in part by the project TAMOP 421B at the Széchenyi University of Győr.

References:


