Pattern Recognition Framework Based on the Best Rank-(R_1, R_2, …, R_K) Tensor Approximation

B. Cyganek
AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland

ABSTRACT: The paper presents a framework for pattern recognition in digital images based on the best rank-(R_1, R_2, …, R_K) decomposition of the prototype pattern tensors. The tensors are obtained from the patterns defining the classes. In the case of a class with only a single prototype, its pattern tensor is constructed from geometrically deformed versions of that pattern. Pattern recognition is accomplished by testing a distance of the features obtained by projecting the test patterns into the best rank tensor subspace. The method was tested on the number of image groups and showed high accuracy and fast response time. In the paper the software implementation is also discussed.

1 INTRODUCTION

The newest information technologies lead to generation and processing of enormous amounts of data. The associated problems of data processing and information retrieval require not only the newest computer technologies but also the most efficient new algorithms for data representation and recognition of patterns [Duda_2001][Tadeusiewicz_2010]. In this paper we address the problem of pattern recognition in multi-dimensional visual signals based on the best rank-(R_1, R_2, …, R_K) tensor approximation. Its application to pattern recognition in medical images, as well as details of software implementation, is also discussed.

One of the well known methods of data dimensionality reduction and subspace pattern recognition rely on the Principal Component Analysis (PCA) [Duda_2001]. However, in the case of multi-dimensional data, such as in video, hyper-spectral imaging, or ensemble of patterns, qualitatively better results can be obtained with the tensor based approach, as discussed in literature [Chen_2009] [Lathauwer_2000][Muti_2007][Savas_2007][Wang _2004][Wang_2008]. In the tensor approach, rather than vectorizing multi-dimensional data, data is set into a multi-dimensional cube of values in which each different feature of data corresponds to a separate dimension of the representing tensor. However, huge data repositories frequently are characteristic of redundant information. In this respect, the real benefit of the tensor approach is that the dominating dimensions can be extracted which convey the most essential information content. For this purpose the three fundamental tensor decomposition methods can be named:

- The Higher-Order Singular Value Decomposition (HOSVD) [Lathauwer_1997].
- The best rank-1 [Lathauwer_2000] [Wang_2004].
- The best rank-(R_1, R_2, …, R_K) approximations [Lathauwer_2000][Kolda_2008][Wang_2008].

HOSVD can be used to build the orthogonal space for pattern recognition [Savas_2007]. HOSVD with its variants is discussed in [Cyganek_2013]. Its version called truncated HOSVD, results in excessive errors and therefore can be treated only as a coarse approximation or it can serve as an initialization method for the best rank decompositions. In this respect improved results can be obtained with the best rank-1 decomposition [Wang_2004]. However, as was presented by Wang and Ahuja, the rank-(R_1, R_2, …, R_K) approximation can lead to the superior results in respect to dimensionality reduction, reconstruction error, as well as to pattern recognition in the multi-dimensional signals [Wang_2008].

In this paper a method of pattern recognition in digital images is discussed. It is based on the best rank-(R_1, R_2, …, R_K) decomposition of the prototype pattern tensors which are obtained from the patterns defining a class. In the case of a single prototype, a prototype pattern tensor is proposed to be constructed from the geometrically deformed versions of the available pattern. Object recognition is accomplished by comparing distances of the features obtained by projecting the test patterns into the best rank tensor subspaces of different pattern classes. The method was tested on the number of image groups and showed high accuracy and fast response time. In this paper we focus on object recognition in radiograph images. However, the tensor processing methods are computationally very demanding. In this paper we address also this problem. An object-oriented software framework which allows efficient tensor representation and decompositions is presented and discussed. This frameworks is based on our previous realizations [Cyganek_2013][Cyganek_2010].

2 PATTERN RECOGNITION WITH THE BEST RANK TENSOR DECOMPOSITION

2.1 *Pattern representation with tensors*

In this section the basics of the tensor algebra are presented. However, a more in-depth treatment can be found in other publications such as [Lathauwer_1997][Cichocki_2009][Kolda_2008] [Cyganek_2013].

First definition concerns tensor representation in the form of vectors and matrices. The *j-mode vector* (*a fiber*) of the *K-th* order tensor $\mathcal{T} \in \Re^{N_1 \times N_2 \times \ldots N_K}$ is a vector obtained from its elements by varying only one index n_j while keeping all other indices fixed. If from the tensor $\mathcal{T}$ a following matrix

$$\mathbf{T}_{(j)} \in \Re^{N_j \times \left(N_1 N_2 \ldots N_{j-1} N_{j+1} \ldots N_K\right)} \tag{1}$$

is formed, then columns of $\mathbf{T}_{(j)}$ are *j-mode* vectors of $\mathcal{T}$. Also, $\mathbf{T}_{(j)}$ is a matrix representation of the tensor $\mathcal{T}$. The *j*-th index becomes a row index of $\mathbf{T}_{(j)}$, while its column index is a product of all other *K*-1 indices of $\mathcal{T}$. However, a place in memory where an element of the tensor is stored depends on an assumed permutation order of these *K*-1 indices. From the all possible (*K*-1)! order schemes only two are commonly used – the so called forward and backward cycling modes [Lathauwer_1997][Cichocki_2008]. For example, for a 2-mode flattening of a 4D tensor (*K*=4), we obtain the following orderings of the other indices 3-4-1, and 1-4-3, for the forward and backward cycle modes, respectively [Cyganek_2013].

The second important concept is a *p*-mode product of a tensor $\mathcal{T} \in \Re^{N_1 \times N_2 \times \ldots N_K}$ and a matrix $\mathbf{M} \in \Re^{Q \times N_p}$. A result of this operation is the tensor $\mathcal{S} \in \Re^{N_1 \times N_2 \times \ldots N_{p-1} \times Q \times N_{p+1} \times \ldots N_K}$ whose elements are obtained as follows

$$\mathcal{S}_{n_1 n_2 \ldots n_{p-1} q n_{p+1} \ldots n_K} = \left(\mathcal{T} \times_p \mathbf{M}\right)_{n_1 n_2 \ldots n_{p-1} q n_{p+1} \ldots n_K} = \sum_{n_p=1}^{N_p} t_{n_1 n_2 \ldots n_{p-1} n_p n_{p+1} \ldots n_K} m_{q n_p}. \tag{2}$$

It can be shown that the *p-mode* product can be equivalently expressed in terms of the flattened matrices $\mathbf{T}_{(p)}$ and $\mathbf{S}_{(p)}$. If the following is fulfilled

$$\mathcal{S} = \mathcal{T} \times_p \mathbf{M}, \tag{3}$$

then it holds that

$$\mathbf{S}_{(p)} = \mathbf{M}\,\mathbf{T}_{(p)}. \tag{4}$$

2.2 *Best rank tensor decomposition*

In the presented work we exploit the best rank-(R_1, R_2, …, R_K) decomposition of pattern tensors. This decomposition can be defined as follows [Lathauwer_2000][Cyganek_2013]:

Given a tensor $\mathcal{T} \in \Re^{N_1 \times N_2 \times \ldots \times N_K}$ compute an approximating tensor $\tilde{\mathcal{T}}$ having $rank_1\left(\tilde{\mathcal{T}}\right) = R_1$, $rank_2\left(\tilde{\mathcal{T}}\right) = R_2$, …, $rank_K\left(\tilde{\mathcal{T}}\right) = R_K$, which is as close as possible to the input tensor $\mathcal{T}$ (see Figure 1).

The aforementioned tensor decomposition can be stated as minimization of the following least-squares cost function

$$E\left(\tilde{\mathcal{T}}\right) = \left\|\tilde{\mathcal{T}} - \mathcal{T}\right\|_F^2, \tag{5}$$

with the Frobenius norm. It can be shown that the approximated tensor $\tilde{\mathcal{T}}$ conveys as much of the "energy", in the sense of the squared entries of a tensor, as the original tensor $\mathcal{T}$, given the rank constraints. A value of *E* is called the reconstruction error.

It can be also easily observed that the assumed rank conditions mean that the approximation tensor $\tilde{\mathcal{T}}$ can be decomposed as follows

$$\tilde{\mathcal{T}} = \mathcal{Z} \times_1 \mathbf{S}_1 \times_2 \mathbf{S}_2 \ldots \times_K \mathbf{S}_K, \tag{6}$$

where $\times_p$ denotes a *p*-mode product of a tensor with a matrix [Lathauwer_2000][Cyganek_2013].

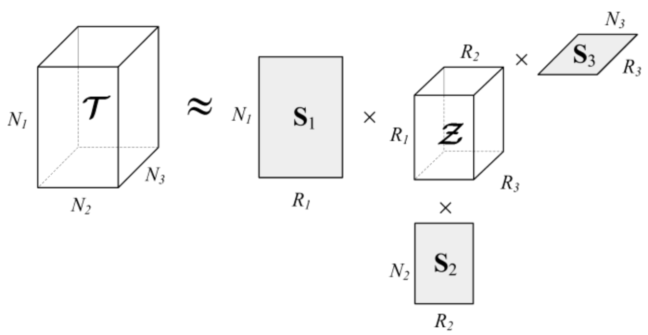

Figure 1: Visualization of the best rank-(R_1, R_2, R_3) decomposition of a 3D tensor.

In the above, each of the matrices $\mathbf{S}_1 \in \Re^{N_1 \times R_1}$, $\mathbf{S}_2 \in \Re^{N_2 \times R_2}$, …, and $\mathbf{S}_K \in \Re^{N_K \times R_K}$ has orthonormal columns (each time, number of columns for $\mathbf{S}_i$ is given by R_i). The core tensor $\mathcal{Z} \in \Re^{R_1 \times R_2 \times \ldots \times R_K}$ has dimensions R_1, R_2, …, R_K. It can be computed from the original tensor $\mathcal{T}$ as follows

$$\mathcal{Z} = \mathcal{T} \times_1 \mathbf{S}_1^T \times_2 \mathbf{S}_2^T \ldots \times_K \mathbf{S}_K^T . \tag{7}$$

Summarizing, to find the best rank-(R_1, R_2, …, R_K) approximation of $\mathcal{T}$ it is sufficient to determine only a set of $\mathbf{S}_i$ in (6), and then $\mathcal{Z}$ is computed from equation (7).

Further analysis is constrained exclusively to 3D tensors. Figure 1 shows visualization of the best rank-(R_1, R_2, R_3) decomposition of the 3D tensors. It can be easily observed that this decomposition can lead to a significant data reduction which can be assessed as follows:

$$C = \frac{R_1 R_2 R_3 + N_1 R_1 + N_2 R_2 + N_3 R_3}{N_1 N_2 N_3} . \tag{8}$$

By proper choice of the ranks R_1, R_2, and R_3 a trade off can be achieved between the compression ratio C in (8) with respect to the approximation error expressed in equation (5). This influences also pattern recognition accuracy, as will be discussed.

2.3 *Pattern classification with the best rank tensor decomposition*

A best rank space described in the previous section can be used to generate specific features of an image $\mathbf{X}$ which can be then used for pattern recognition [Wang_2008]. The features are obtained by projecting the image $\mathbf{X}$ of dimensions $N_1 \times N_2$ into the space spanned by the two matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ in accordance with (7). However, at first $\mathbf{X}$ needs to be represented in an equivalent tensor form $\mathcal{X}$ which is of dimensions $N_1 \times N_2 \times 1$. Then, the feature tensor $\mathcal{F}$ of dimensions $R_1 \times R_2 \times 1$ is obtained, as follows

$$\mathcal{F}_{\mathbf{X}} = \mathcal{X} \times_1 \mathbf{S}_1^T \times_2 \mathbf{S}_2^T . \tag{9}$$

Tensor $\mathcal{T}$ contains training patterns. Depending on the problem and available training patterns (images), two modes of operation are possible. These are as follows:

1. A set of prototype patterns $\mathbf{P}_i$ of the same object is available. These are used to form the input tensor $\mathcal{T}$.
2. If only one prototype $\mathbf{P}$ is available, its different appearances $\mathbf{P}_i$ can be generated by geometrical warping of the available pattern. This process is visualized in Figure 2.

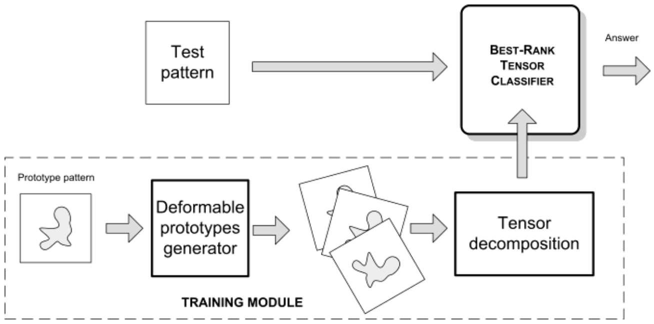

Figure 2: Visualization of the process of the 3D pattern tensor generation by geometrical warping of the prototype pattern.

In either of the above cases, the patterns form a 3D tensor which after the best-rank decomposition spans the space representing that class.

The next step consists of computation of the reference feature out of the set of prototype patterns $\mathbf{P}_i$ from the tensor $\mathcal{T}$. This is computed as follows

$$\mathcal{F}_i = \mathcal{P}_i \times_1 \mathbf{S}_1^T \times_2 \mathbf{S}_2^T , \tag{10}$$

where $\mathcal{P}_i$ denotes a tensor representation of the pattern $\mathbf{P}_i$. However, in our proposition the pattern $\mathbf{P}_i$ is chosen from the set of available patterns which shows the optimal Frobenius norm, that is

$$i = \arg\min_{\mathbf{P}_i} \left\| \mathcal{P}_i \times_1 \mathbf{S}_1^T \times_2 \mathbf{S}_2^T \right\|_F . \tag{11}$$

The above process of building the prototype pattern tensor $\mathcal{T}$, its decomposition and computation of the reference features is repeated for each of the available classes c. Since the process is independent, this training stage can be easily parallelized.

Finally, the classifier returns a class c for which the following is minimized

$$c = \arg\min \left\| \mathcal{F}_{\mathbf{X}} - \mathcal{F}_i^{(c)} \right\|_F . \tag{12}$$

As alluded to previously, the training parameters are the chosen rank values of R_1, R_2, and R_3 in (6). These are usually determined experimentally.

3 COMPUTATION OF THE BEST RANK TENSOR DECOMPOSITION

3.1 *Alternating Least-Squares Method*

Computation of the best rank-(R_1, R_2, …, R_K) decomposition of tensors, given by equations (6) and (7), can be obtained with help of the Alternating Least-Squares (ALS) method, as proposed by Lathauwer *et al.* [Lathauwer_2008]. In each step of this method only one of the matrices $\mathbf{S}_k$ is optimized, whereas other are kept fixed [Chen_2009]. The main concept of this approach is to express the quadratic expression in the components of the unknown matrix $\mathbf{S}_k$ with orthogonal columns with other matrices kept fixed. That is, the following problem is solved

$$\max_{\mathbf{S}_i}\left\{\Psi\left(\mathbf{S}_i\right)\right\} = \max_{\mathbf{S}_i}\left\|\mathcal{T} \times_1 \mathbf{S}_1^T \times_2 \mathbf{S}_2^T \ldots \times_K \mathbf{S}_K^T\right\|^2 . \qquad (13)$$

Columns of $\mathbf{S}_i$ can be obtained finding the orthonormal basis of the dominating subspace of the column space of the approximating matrix $\hat{\mathbf{S}}_i$. As already mentioned, in each step only one matrix $\mathbf{S}_i$ is computed, while other are kept fixed. Such procedure - called the Higher-Order Orthogonal Iteration (HOOI) - is repeated until the stopping condition is fulfilled or a maximal number of iterations is reached [Lathauwer_2000][Cyganek_2013].

3.2 *Software Framework*

The above HOOI procedure has been implemented in our software framework, details of which are described in [Cyganek_2013]. The implementation utilizes C++ classes with basic data types defined as template parameters. Thanks to this, the platform is highly flexible. For instance, the time and memory can be saved by using the fixed point representation of data instead of the floating point. In the presented experiments the 12.12 fixed point representation showed to be sufficient (each data is stored on 3 bytes instead of 8, needed in the case of the floating point representation).

The class hierarchy of our framework is shown in Figure 8. The *Best_Rank_R_DecompFor* is the main class for the best-rank tensor decomposition. However, it is derived from the *TensorAlgebraFor* class which implements all basic operations on tensors, such as the *p*-mode multiplications. Tensors, in turn, are represented by objects of the class *TFlatTensorFor* which conveys tensors in the flattened form. The *Best_Rank_R_DecompFor* class is augmented with the *S_Matrix_Initializer* hierarchy. Its main role is to define the way of initial setup of the values of the S_i matrices for the HOOI process. In our case these were initialized with randomly generated values of uniform distribution. More details on implementation, as well as code details can be found in [DeRecLib_2013] [Cyganek_2013] [Cyganek_2009].

4 EXPERIMENTAL RESULTS

Experiments were carried out with many different groups of images, such as road signs pictograms, handwritten digits, as well as medical images. The latter are reported in this section.

Figure 3 depicts a maxillary radiograph (left), as well as the implant pattern (right).

Figure 3: An example of dental implant recognition in a maxillary radiograph image.

The task is to identify implants in the radiograph images. These are detected as highly contrast areas which, after registration, are fed to the tensor classifier described in the previous sections. Since only one example of the prototype image is usually available, its different appearances are generated by image warping, as described. In these experiments a pattern was rotated in the range of ±12°. Additionally, a Gaussian noise was also added to increase robustness of the method. Examples of deformed versions of the prototype image of an implant are depicted in Figure 4. These form a 3D tensor which, after the best-rank decomposition, is used in recognition process as already discussed.

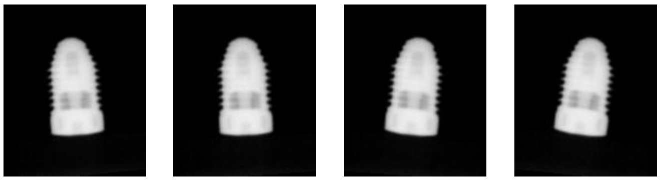

Figure 4: Examples of deformed versions of the prototype image of an implant. These are formed into a 3D tensor which after the best-rank approximation is used in object recognition.

Figure 5 depicts a dental implant found in the maxillary radiograph image.

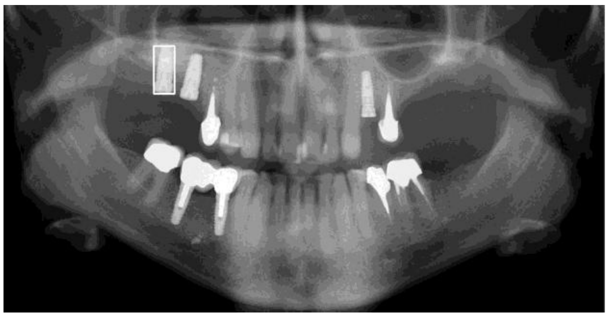

Figure 5: An example of dental implant recognition in a maxillary radiograph image.

Figure 6 shows a plot of the reconstruction error *E*, expressed in equation (5), in respect to the compression ratio *C*, given in equation in (8).

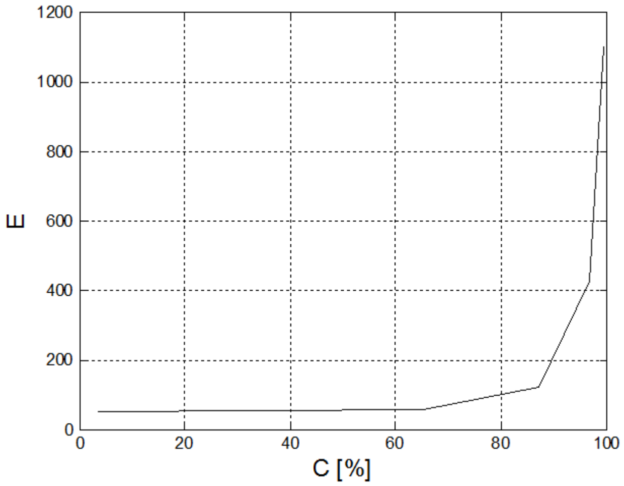

Figure 6: Reconstruction error E in respect to the compression ratio C of the input patterns.

On the other hand, Figure 7 depicts a plot showing measured accuracy of the pattern recognition, also in respect to the compression ratio C. As visible, the accuracy easily reached a level of 95-96% for highly reduced ranks.

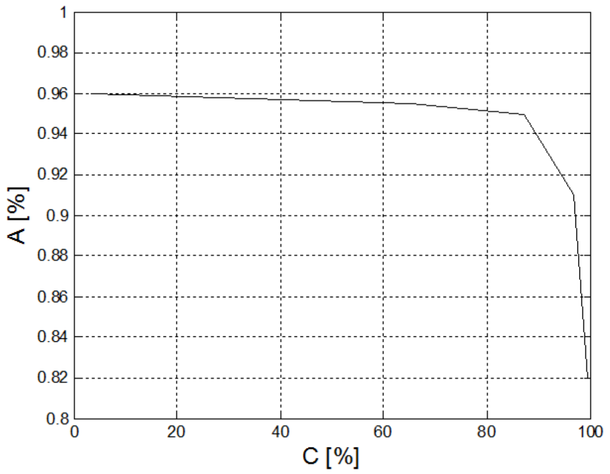

Figure 7: Accuracy A of pattern recognition in respect to the compression ratio C of the input patterns.

In the presented experiments, the input images are of size 56×56 pixels. These are fed from the detection module which returns registered areas of high contrast [Cyganek_2010]. Thus, considering the assumed geometrical deformations, size of the tensor $\mathcal{T}$ is 56×56×13. Then many different rank settings were tested to measure their influence on accuracy. Lowering the rank, results in lower memory requirements and faster computations. However, this comes at the cost of a higher reconstruction error E and lower accuracy A. In our experiments, the high level of accuracy was kept up to the rank values of $R_1=R_2=10$ and $R_3=6$. This corresponds to the compression ratio $C=0.05$ (95%). However, a real compression ratio expressed in data storage reduction is lower since the decomposed matrices $\mathbf{S}_i$ and the core tensor $\mathcal{Z}$ require storage of their integer and fractional parts. In our experiments, mostly to reduce memory requirements, the fixed-point representation was used with only 24 bits (3 bytes) per data. Nevertheless, for monochrome image recognition only $\mathbf{S}_1$ and $\mathbf{S}_2$ are necessary. Similar high accuracy of recognition was observed for other groups of images. Thus, the method is highly effective.

Recognition is very fast since it requires computation of only (9) and (12). The reference features are computed off-line in accordance with (10) and (11).

5 CONCLUSIONS

The paper presents a framework for pattern recognition in digital images with help of the best rank-(R_1, R_2, …, R_K) decomposition of the prototype pattern tensors. The tensors are formed from the patterns defining a class. In the case of a single prototype, a tensor is proposed to be constructed from its geometrically deformed versions. Thus, the deformed model is created. Pattern recognition is accomplished by testing a distance of the features obtained by projection of the patterns into the best rank tensor subspace. The method was tested on the number of image groups and showed high accuracy and fast response time. In the presented experiments with implant recognition in maxillary radiograph images, the reached accuracy is 96%. The presented object-oriented software framework allows computations in the fixed-point data format which greatly limits memory consumption. It was also shown that the training process can be easily parallelized. The software for tensor decomposition is available from the webpage of the book by Cyganek [Cyganek_2013] [DeRecLib_2013].

6 ACKNOWLEDGMENT

The financial support from the Polish National Science Centre NCN in the year 2013, contract no. DEC-2011/01/B/ST6/01994, is greatly acknowledged.

7 REFERENCES

Chen, J. Saad, Y. 2009. On the Tensor SVD and the Optimal Low Rank Orthogonal Approximation of Tensors. *SIAM Journal on Matrix Analysis and Applications* 30(4): 1709-1734.

Cichocki, A., Zdunek, R., Amari, S. 2008. Nonnegative Matrix and Tensor Factorization. *IEEE Signal Processing Magazine* 25(1): 142-145.

Cichocki, A., Zdunek, R., Phan, A.H., Amari, S-I. 2009. *Nonnegative Matrix and Tensor Factorizations. Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Wiley.

Cyganek, B. 2013. *Object Detection and Recognition in Digital Images: Theory and Practice*, Wiley.

Cyganek, B. 2010. Software Framework for Efficient Tensor Representation and Decompositions for Pattern Recognition in Computer Vision. *Advances in Soft Computing Series Image Processing & Communications Challenges II*, Springer-Verlag 84: 185–192.

Cyganek, B., Malisz, P. 2010. Dental Implant Examination Based on the Log-Polar Matching of the Maxillary Radiograph Images in the Anisotropic Scale Space. *IEEE Engineering in Medicine and Biology Conference*, EMBC 2010, Buenos Aires, Argentina: 3093-3096.

Cyganek, B., Siebert, J.P. 2009. *An Introduction to 3D Computer Vision Techniques and Algorithms*, Wiley.

DeRecLib, 2013. http://www.wiley.com/go/cyganekobject

Duda, R.O., Hart, P.E., Stork, D.G.. 2001. *Pattern Classification*. Wiley.

Kolda, T.G., Bader, B.W. 2008. Tensor Decompositions and Applications. SIAM Review.

Lathauwer de L. 1997. Signal Processing Based on Multilinear Algebra. PhD dissertation, Katholieke Universiteit Leuven.

Lathauwer de L, Moor de, B, Vandewalle J. 2000. On the Best Rank-1 and Rank-(R_1, R_2, …, R_N) Approximation of Higher-Order Tensors. SIAM Journal Matrix Analysis and Applications 21(4): 1324-1342.

Muti D, Bourennane S. 2007. Survey on tensor signal algebraic filtering. *Signal Processing* 87: 237-249.

Savas B, Eldén L. 2007. Handwritten digit classification using higher order singular value decomposition. *Pattern Recognition* 40(3): 993-1003.

Tadeusiewicz, R. 2010. Place and Role of Intelligent Systems in Computer Science. *Computer Methods in Materials Science* 10(4): 193-206.

Wang H, Ahuja N: 2004. Compact Representation of Multidimensional Data Using Tensro Rank-One Decomposition. *Proceedings of the 17th International Conference on Pattern Recognition*, Vol. 1: 44-47.

Wang H, Ahuja N. 2008. A Tensor Approximation Approach to Dimensionality Reduction. *International Journal of Computer Vision* 76(3): 217-229.

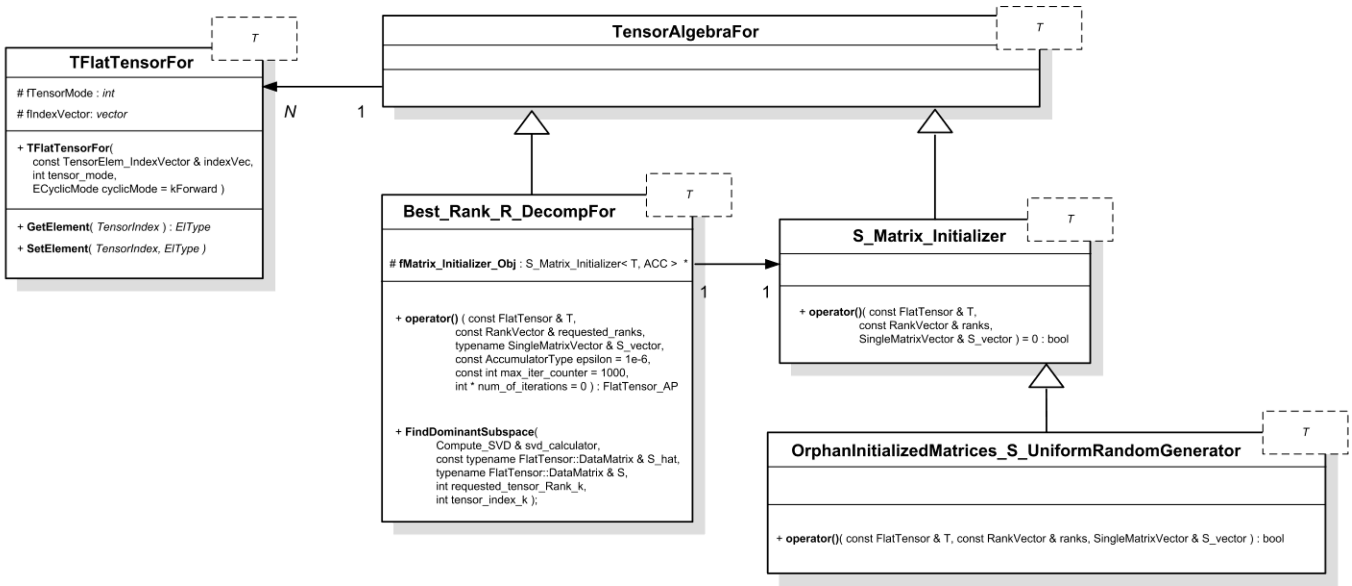

Figure 8. Hierarchy of the classes for tensor best-rank decompositions. The *Best_Rank_R_DecompFor* class performs the best rank-(R_1, R_2, …, R_K) decomposition. The second hierarchy derived from the base *S_Matrix_Initializer* is responsible for different initialization schemes of the **S** matrices in the HOOI algorithm.