



CHALMERS

Chalmers Publication Library

A Multidimensional Signal Processing Approach for Classification of Microwave Measurements with Application to Stroke Type Diagnosis

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '11), Boston, MA, AUG 30-SEP 03, 2011

Citation for the published paper:

Yousefi Mesri, S. ; Khazaeli Najafabadi, M. ; McKelvey, T. (2011) "A Multidimensional Signal Processing Approach for Classification of Microwave Measurements with Application to Stroke Type Diagnosis". 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '11), Boston, MA, AUG 30-SEP 03, 2011 pp. 6465-6469.

Downloaded from: <http://publications.lib.chalmers.se/publication/146769>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

A Multidimensional Signal Processing Approach for Classification of Microwave Measurements with Application to Stroke Type Diagnosis

Hamed Yousefi Mesri

Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

Email: mesri@student.chalmers.se

Masoud Khazaeli Najafabadi

Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

Email: khazaeli@student.chalmers.se

Tomas McKelvey

Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

Email: mckelvey@chalmers.se

Abstract— A multidimensional signal processing method is described for detection of bleeding stroke based on microwave measurements from an antenna array placed around the head of the patient. The method is data driven and the algorithm uses samples from a healthy control group to calculate the feature used for classification. The feature is derived using a tensor approach and the higher order singular value decomposition is a key component. A leave-one-out validation method is used to evaluate the properties of the method using clinical data.

Index Terms— classification, HOSVD, microwave scattering data, multichannel signal processing, multilinear algebra

I. INTRODUCTION

Electromagnetic waves at microwave frequencies can interact with the human body both through reflection at the interface between the body and the surrounding media but also penetrate and interact with internal body tissues. The propagation of electromagnetic waves is determined by geometry and dielectric properties of the various body tissues. Hence, changes to the dielectric properties inside the body could thus be detected by a change in the electromagnetic properties around the body. Blood has a different dielectric properties compared to brain tissue and hence accumulation of blood in the brain would lead to a change in the electromagnetic properties. By employing a microwave antenna array placed around the head of a patient, electromagnetic measurements in the form of the scattering matrix at several frequencies would thus convey information regarding the electrical properties of the antenna system which includes the electromagnetic properties also inside the brain. Microwave measurement systems thus have the potential ability to be used for non-

invasive medical monitoring and diagnosis for selected purposes. Prototype systems exists for imaging breast cancer using a microwave tomography approach [1] and for monitoring and detection of bleedings in stroke patients [2, 3]. Microwave array measurements are inferior in comparison with Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) regarding resolution. However, the microwave technique has no hazardous radiation and in some diagnostic cases, such as cancer, it has better contrast [4]. Furthermore a microwave measurement system is expected to be both inexpensive and lightweight thereby enabling portable units for use for example in an ambulance.

Stroke has in principle two origins. Either a blood vessel has ruptured in the brain and a blood pool is formed leading to the stroke symptoms or a cerebral vessel has become occluded by a clot locally restricting the blood flow to the brain and ultimately leading to necrosis. The first case is referred to as an intra cerebral hemorrhagic stroke (ICH) while the second situation is known as an ischemic stroke (IS). If an ischemic stroke is diagnosed within 3 hours from onset of symptoms thrombolytic therapy can be initiated with the aim to dissolve the clot and restore the circulation [5]. Today the diagnosis is made by CT or MRI. However, since in most clinics CT and MRI are heavily utilized, in many IS cases the diagnosis is made to late to enable usage of thrombolysis and hence leading to a worse outcome [6,7]. This paper is focused on the application of detecting bleedings within the brain cavity from microwave measurements.

A technique based on multidimensional signal processing techniques and Higher Order Singular Value Decomposition (HOSVD) is developed to derive suitable features from the microwave scattering dataset and investigate the power of a classifier based on the derived features.

A clinical trial has been performed by Medfield Diagnostics AB in collaboration with Sahlgrenska University Hospital, Sweden and the data are used to evaluate the developed technique.

II. CLASSIFICATION ALGORITHM

A. Measurement Data

The microwave measurement device consists of 10 transmitting and receiving antennas and the measurement for each pairs of antennas is done by sweeping the frequency range from 100 MHz to 3 GHz yielding on estimate of the array system microwave scattering data. The data is organized in a 3 dimensional complex-valued tensor in which the 1st and 2nd dimensions correspond to transmitting and receiving antenna index and 3rd dimension corresponds to microwave measurement frequency. Therefore, it is natural to utilize methods which consider all the interactions between different components in the 3D data structure and construct a classifier based on multidimensional signal processing techniques. In the previous classifiers, used on microwave scattering data, this structure has been disregarded[3].

The measured microwave signals are the S-parameters of the microwave transmission and are in general complex numbers.

The reflection and transmission coefficients are collected using 2 antennas in each time slot. The incident wave in the illumination area of the head splits into two main parts after arriving the head. The major part of the wave is reflected by the body surface. The remaining, rather small, part of the signal penetrates inside the brain and scatters. The receiving antennas on both sides will receive the signals which are the sum of all scattered and reflected signals on the transmitting side and only scattered signal on the receiving side. Generally all the transmission coefficients look almost the same but they are totally different from the reflection coefficients. Due to their higher amplitude and less information we have disregarded reflective coefficients and just considered transmission coefficients.

B. Methodology

The basic idea behind using HOSVD for classification of microwave scattering data is to use the knowledge about how the data is organized in order to extract powerful features and check whether this structural information of the data would lead to better performance in the classifier. First we investigate how the microwave data is structured.

Since the measurement device is able to collect just one reflection and one transmission coefficient at each time slot, the measurement should be done with fixed sending and receiving antenna (one channel) for all the frequency range and then by fixing the sending antenna and switching to the next receiving antenna. Therefore, if we ignore the switching time between two consecutive measurements, we can model our measurement setup to a multichannel setup with single input and multiple output structure (SIMO). This idea will lead us to use multichannel setup to restructure the data into its actual 2D format in which each measurement is a 2D matrix with different channels in the first dimension and different wave frequencies in the second dimension.

C. Orthogonal base construction using HOSVD

In order to define the HOSVD first we define the n-mode product of a tensor and a matrix.

Definition C-1 (n-mode product of a tensor and a matrix)

The n-mode product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{C}^{J_n \times I_n}$ denoted by $\mathcal{A} \times_n \mathbf{U}$ is an $(I_1 \times I_2 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N)$ tensor of which the entries are given by

$$(\mathcal{A} \times_n \mathbf{U})_{i_1 i_2 \dots i_{n-1} j_n i_{n+1} \dots i_N} = \sum_{i_n} a_{i_1 i_2 \dots i_{n-1} i_n i_{n+1} \dots i_N} u_{j_n i_n} \quad (1)$$

The Nth order SVD or Higher order SVD is the generalization of tucker3 decomposition [8,9].

Every complex $(I_1 \times I_2 \times \dots \times I_N)$ tensor can be written as the product

$$\begin{aligned} \mathcal{A} &= \mathcal{S} \times_{I_1} \mathbf{U}^{(1)} \times_{I_2} \mathbf{U}^{(2)} \dots \times_{I_N} \mathbf{U}^{(N)} \\ &= \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} s_{i_1 i_2 \dots i_N} \mathbf{u}_{i_1}^{(1)} \circ \mathbf{u}_{i_2}^{(2)} \circ \dots \circ \mathbf{u}_{i_N}^{(N)} \end{aligned} \quad (2)$$

In which,

$\mathbf{U}^{(n)} = [\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \dots \mathbf{u}_{i_n}^{(n)}]$ is a unitary $(I_n \times I_n)$ matrix,

$\mathbf{u}_1 \circ \mathbf{u}_2$ indicates the outer product of two vectors \mathbf{u}_1 and \mathbf{u}_2 ,

\mathcal{S} is a complex $(I_1 \times I_2 \times \dots \times I_N)$ tensor of which the subtensors $\mathcal{S}_{i_n=\alpha}$ obtained by fixing the nth index to α , have the properties of:

all-orthogonality: two subtensors $\mathcal{S}_{i_n=\alpha}$ and $\mathcal{S}_{i_n=\beta}$ are orthogonal for all possible values of n , α and β subject to $\alpha \neq \beta$,

ordering:

$$\|\mathcal{S}_{i_n=1}\| \geq \|\mathcal{S}_{i_n=2}\| \geq \dots \geq \|\mathcal{S}_{i_n=I_n}\| \geq 0 \quad (3)$$

for all possible values of n [10,11,12].

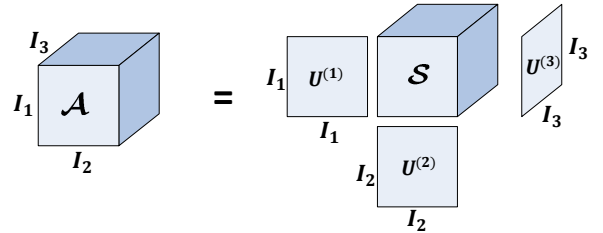


Figure 1: Visualization of Higher order SVD for a third-order tensor

By comparing the matrix SVD and HOSVD, it is obvious that in the higher order case the core matrix \mathbf{S} is replaced by the core tensor \mathcal{S} . Instead of being pseudo-diagonal which means that non zero elements could only occur when the indices $i_1 = i_2 = \dots = i_N$. Tensor \mathcal{S} is in general a full tensor, but it obeys the weaker condition of all-orthogonality instead. The elements of \mathcal{S} are not necessarily positive in general and they can be complex when \mathcal{A} is a complex

valued tensor. The entries of \mathcal{S} show the interaction between different elements of \mathcal{A} as depicted in Figure 1 [10,12,13].

An important property of HOSVD is that the matrices $\mathbf{U}^{(n)}$ are orthogonal. Therefore, by using HOSVD decomposition it is possible to write each n^{th} order tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ as sum of n^{th} order rank-1 tensors as illustrated in Figure 2.

The description of a tensor as the sum of rank-1 tensors can be formulated as

$$\mathcal{A} = \sum_{i=1}^{I_3} \mathbf{A}_i \times_3 \mathbf{u}_i^{(3)} \quad (4)$$

in which

$\mathbf{A}_i = (\mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)})_{:,i}$ for $i=1$ to I_3 are orthogonal bases with respect to 3rd mode and

$$\mathbf{U}^{(3)} = [\mathbf{u}_1^{(3)} \ \mathbf{u}_2^{(3)} \ \dots \ \mathbf{u}_{I_3}^{(3)}] \quad (5)$$

Using this property of HOSVD, it is possible to construct a subspace based on the training data. In the case of a 3D tensor the bases will be 2D matrices[12].

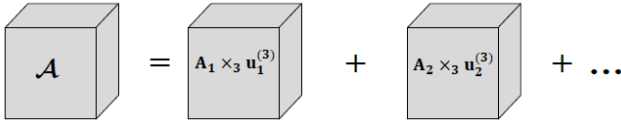


Figure 2: Tensor as a sum of rank-1 tensor

D. Feature extraction using HOSVD

In this section, we describe the feature extraction algorithm used in the HOSVD classifier. As a normal classification problem, there is a need to define a labeled data set as training data for classification. Training data in this project are based on measurements from a healthy control group. First, we stack several training sample matrices together and form a 3D tensor $\mathcal{H} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$ in which I_1 is different channels, I_2 is frequency range and I_3 is different measurements. In the next step, the mean of all healthy samples is calculated ($\bar{\mathbf{H}}$).

$$\bar{\mathbf{H}} = \frac{1}{I_3} \sum_{i=1}^{I_3} \mathbf{H}_{:,i} \quad (6)$$

Then the mean value is subtracted from each training sample in order to calculate deviations from mean (\mathcal{H}_d). These data will have the same structure as the original data. As it was discussed in section C, using HOSVD, we can construct orthonormal bases (\mathbf{A}_i) from training samples (\mathcal{H}_d). Afterwards, for each new measurement (\mathbf{X}), the deviation from mean will be calculated as:

$$\mathbf{X}_d = \mathbf{X} - \bar{\mathbf{H}} \quad (7)$$

Having \mathbf{A}_i , $\bar{\mathbf{H}}$ and \mathbf{X}_d ready we can start to extract features. Figure 3 depicts the whole feature extraction algorithm.

In fact, feature extraction unit considers matrices $\bar{\mathbf{H}}$ and \mathbf{X}_d and a set of orthonormal matrices \mathbf{A}_i as the bases for a

subspace (here we call it deviation space) and extracts features of $\bar{\mathbf{H}}$ and \mathbf{X}_d which characterize their relations with deviation space. In general, in order to classify different matrices in an n -dimensional space, there are n different components which can be used for classification. Our assumption was that information correlated to ICH or IS would change the direction or the norm of the data with respect to deviation space. In order to have better detection of this change, we have extracted a feature which we call it ‘‘HOSVD angle’’ and observe that it can lead to a total separation between healthy and ICH patients.

Before getting to this feature consider these definitions:

Definition D-1 (Inner product of matrices)

The inner product of two matrices $\mathbf{X} \in \mathbb{C}^{m \times n}$ and $\mathbf{Y} \in \mathbb{C}^{m \times n}$ is defined as:

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \sum_{i=1}^m \sum_{j=1}^n x_{ij} y_{ij}^* \quad (8)$$

Definition D-2 (Frobenius norm of a matrix)

For a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$, the Frobenius norm is defined as:

$$\|\mathbf{X}\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n x_{ij} x_{ij}^*} \quad (9)$$

Definition D-3 (Projection of a matrix on a subspace)

The projection of a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ on a subspace \mathcal{A} represented by orthogonal bases $\mathbf{A}_i \in \mathbb{C}^{m \times n}$ is defined as:

$$\mathbf{P}_{\mathcal{A}}(\mathbf{X}) = \sum_{i=1}^r \langle \mathbf{A}_i, \mathbf{X} \rangle \cdot \mathbf{A}_i \quad (10)$$

In which $\langle \mathbf{A}_i, \mathbf{X} \rangle$ is the scalar product \mathbf{A}_i and \mathbf{X} and $r = \text{rank}(\mathcal{A})$.

Definition D-4 (Angle between a matrix and a subspace)

The angle between a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ and a subspace \mathcal{A} is defined as:

$$\theta_{\mathbf{X}, \mathcal{A}} = \cos^{-1} \left(\frac{\langle \mathbf{P}_{\mathcal{A}}(\mathbf{X}), \mathbf{X} \rangle}{\|\mathbf{P}_{\mathcal{A}}(\mathbf{X})\| \cdot \|\mathbf{X}\|} \right) \quad (11)$$

HOSVD angle has been defined as the difference of the angle between new measurement and healthy deviation subspace and the angle between healthy average and healthy deviation subspace.

$$\theta_{\text{HOSVD}} = \theta_{\bar{\mathbf{H}}, \mathcal{H}_d} - \theta_{\mathbf{X}_d, \mathcal{H}_d} \quad (12)$$

In which $\theta_{\mathbf{X}_d, \mathcal{H}_d}$ is corresponding to the angle between new measurement and healthy deviation subspace. Therefore, we have $\theta_{\mathbf{X}_d, \mathcal{H}_d} = \theta_{\mathbf{H}_d, \mathcal{H}_d}$ for new healthy measurement and $\theta_{\mathbf{X}_d, \mathcal{H}_d} = \theta_{\mathbf{B}_d, \mathcal{H}_d}$ for new ICH measurement.

In the feature extraction step both X_d and \bar{H} belong to a high dimensional space and considering their relation is rather complicated.

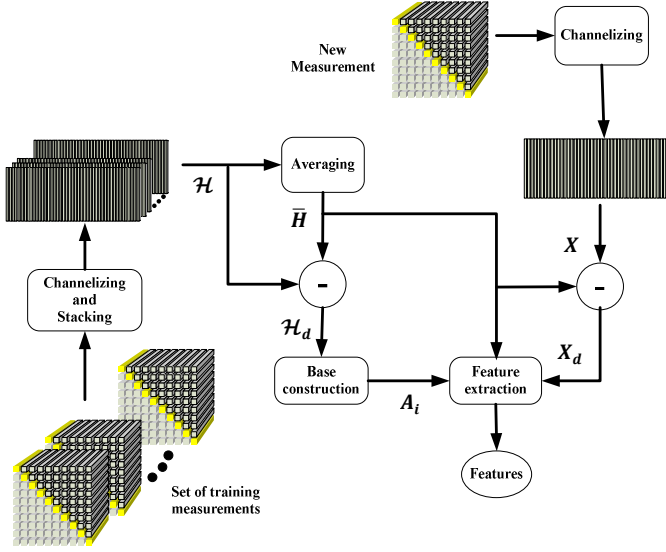


Figure 3: Feature extraction algorithm

E. Leave-one-out approach

Leave-one-out approach is one of the cross validation methods to validate the result of the classifier [14]. In our case we have 25 healthy subjects, 10 ICH patients and 12 IS patients. Leave-one-out approach is done in following steps:

1. Building training samples from all but the first healthy sample.
2. Constructing orthogonal bases from training samples.
3. Classifying all patients (22) and the left out healthy sample.
4. Repeating steps 1 to 3 for the next healthy samples

F. Data preprocessing using HOSVD

As it discussed earlier, according to (2) for a given tensor \mathcal{A} , we can find an exact HOSVD of $rank-(R_1, R_2, \dots, R_N)$ where $R_n = rank_n(\mathcal{A})$. If we disregard parts of the core tensor and corresponding columns in the factor matrices $\mathbf{U}^{(n)}$ and reconstruct the tensor, the rank of reconstructed tensor i.e. $rank-(R_1, R_2, \dots, R_N)$ will be less than the rank of the original tensor $R_n < rank_n(\mathcal{A})$. In this case the reconstructed tensor will be an approximation of \mathcal{A} by truncated HOSVD, which is the main idea of compressing the data using HOSVD. For each tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$, we can approximate the tensor as the sum of rank-1 components as follows:

$$\hat{\mathcal{A}} = \sum_{i_1=1}^{P_1} \sum_{i_2=1}^{P_2} \dots \sum_{i_N=1}^{P_N} s_{i_1 i_2 \dots i_N} \mathbf{u}_{i_1}^{(1)} \circ \mathbf{u}_{i_2}^{(2)} \circ \dots \circ \mathbf{u}_{i_N}^{(N)} \quad (13)$$

Where P_n are the number of components (orthogonal vectors in each factor matrices) and $P_n \leq I_n$. Therefore, the approximation $\hat{\mathcal{A}}$ of a tensor \mathcal{A} [11]. (see Figure 4 for an illustration)

The ordering constraint of the tensor \mathcal{S} imposes that the Frobenius-norm of the horizontal (frontal and respectively vertical) matrices does not increase as the index i_1 (i_2 and respectively i_3) is increased [10,13]. Therefore, it is possible to consider that the strongest components of the data will be constructed by mostly the first ordered values in the core tensor \mathcal{S} . This is the basic idea behind using truncated HOSVD for data preprocessing [10]. In this case, considering different measurements for one patient, it is possible to form the data in a 4D tensor structure and compress the data using truncated HOSVD with respect to the 4th mode.

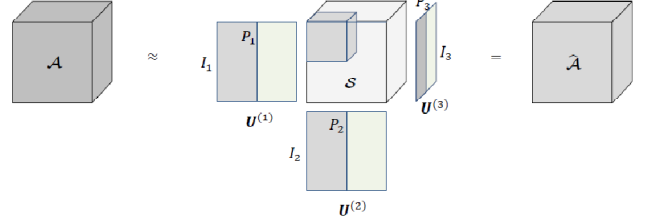


Figure 4: Truncated HOSVD of a 3rd mode tensor

III. EXPERIMENTAL RESULTS FOR CLINICAL DATA

A clinical study has been done at Sahlgrenska University Hospital, Sweden, and the data has been collected for different patients with IS and ICH. The same type of measurements have also been collected on healthy volunteers to form the training set. The data consists of 20 measurements for 10 ICH patients (2 measurements for each patient in two consecutive days), 24 measurements for 12 IS patients and 25 measurements for 25 healthy volunteers.

A. Separation of ICH from Healthy

In order to classify ICH patients, the classifier has been trained by healthy subjects i.e. to determine the bases \mathbf{A}_i for the deviation space \mathcal{H}_d . The results have been plotted in different box plots for each measurement (Figure 5). The deviation in each boxplot is due to changing the training set according to the leave-one-out approach as discussed in section II.E. Results for all healthy subjects have been gathered in one group to be able to visualize the total variation for healthy measurements and compare with the other groups.

HOSVD angle can lead to total separation of ICH from Healthy subjects for the data at hand. It is also observed that the threshold value can be easily set to 0 and negative angles will show the ICH patients while positive angles will show the healthy samples. As it is illustrated in Figure 5.

B. Separation of IS from ICH

The most important application of this classifier is to classify ICH from IS patients. In this section the HOSVD based classifier and the preprocessing method based on truncated HOSVD was used. For each patient all the 6 consecutive measurements have been considered as a 4D tensor and by setting the compression factor to 5 (i.e. the

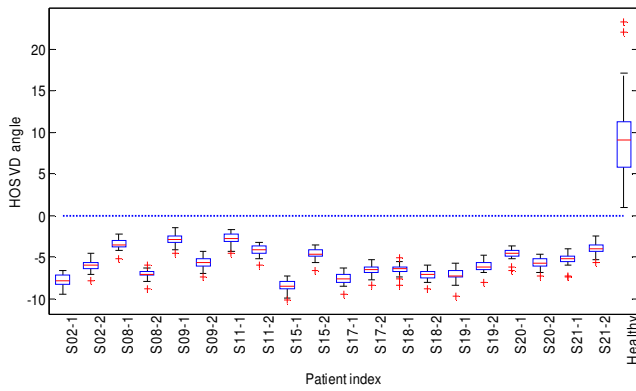


Figure 5: Distribution of HOSVD angle for ICH patients

first 5 slices of the core tensor is used for data reconstruction) the compression was done in frequency (3rd) direction. Then the dataset is reconstructed back as a 4D tensor and the measurements are separated respectively. The same procedure has been done for healthy measurements and the training set is also reconstructed with compression factor 5. The reconstructed data is feed into the same algorithm as depicted in Figure 3. In this section the factor of onset of the stroke which is the time slot between the first symptoms of stroke and measurement has been also considered.

The results show that IS subjects have larger HOSVD angle values compared to ICH as depicted in Figure 6. It can also be concluded that the earlier the measurement is taken the better the classification results are.

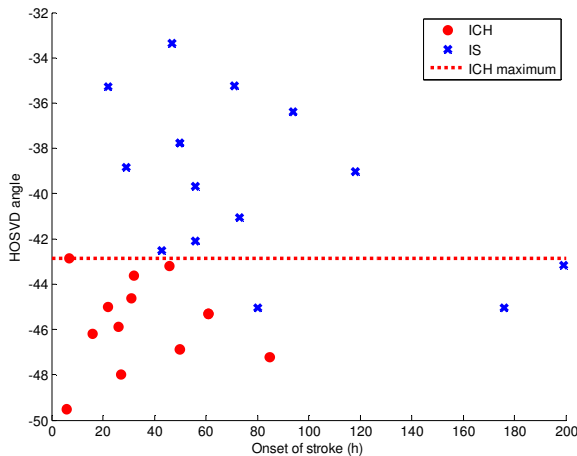


Figure 6: Classification results for separation of IS and ICH vs. onset time of stroke; the classification results show better performance for earlier measurements.

IV. CONCLUSION

In this paper a multidimensional signal processing method based on HOSVD was presented for the discrimination of ICH from IS patients. The basic idea of this method is to consider the actual structure of the data in microwave measurements and construct the classifier based on the orthogonal bases which can be obtained using HOSVD.

The performance of the classifier has been analyzed using the leave-one-out validation approach. The data from clinical trial are examined using the constructed classifier showing a performance of 100% sensitivity and specificity for classification of ICH patient patients from Healthy subjects. The implementation of HOSVD for preprocessing the data has a great impact on the classification of IS from ICH and the correlation between the feature values and the onset of stroke.

V. REFERENCES

- [1] Fhager A. ; Padhi, S. ; Howard, J. (2009). 3D Image Reconstruction in Microwave Tomography Using an Efficient FDTD Model. *IEEE Antennas and Wireless Propagation Letters*. vol. 8 s. 1353 - 1356
- [2] Fhager, A. ; McKelvey, T. ; Persson, M. (2010). Stroke Detection Using a Broad Band Microwave Antenna System. *EUCAP2010, 4th European Conference on Antennas and Propagation, April 12-16, Barcelona, Spain*. s. C13P1-2
- [3] Mohammad Ali Khorshidi, Tomas McKelvey, and Mikael Persson and Hana Dobsicek Trefna, "Classification of Microwave Scattering Data based on a Subspace Distance with Application to Detection of Bleeding Stroke," The Third International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, December 2009.
- [4] Semenov S, Kellam J, Althausen P, Williams T, Abubakar A, Bulyshev A, Sizov Y. "Microwave tomography for functional imaging of extremity soft tissues: feasibility assessment" *Physics in Medicine and Biology* 2007 Sep 21;52(18):5705-19. Epub 2007 Sep 4.
- [5] Adams, H.P. et. al., "Guidelines for Thrombolytic Therapy for Acute Stroke: A Supplement to the Guidelines for the Management of Patients With Acute Ischemic Stroke: A Statement for Healthcare Professionals From a Special Writing Group of the Stroke Council, American Heart Association", *Circulation*, vol. 94(5), pp 1167-1174, 1996.
- [6] Fisher M., Macleod M. and Davis S Donnan G., "Stroke," *The Lancet*, vol. 371, no. 9624, pp. 1612 - 1623, May 2008.
- [7] J. Saver, "Time is brain – quantified," *Stroke*, vol. 37, pp. 263-266, 2006.
- [8] L.R Tucker, "Some mathematical Notes on Three-Mode Factor Analysis," *Psychometrika*, vol. 31, no. 3, pp. 279-311, September 1966.
- [9] L.R Tucker, "The extension of factor analysis to Three-Dimensional Matrices," in *Contributions to mathematical psychology*, H. Gulliksen and N. Frederiksen, Ed. NewYork: Holt, Rinehart and Winston, 1964, pp. 110--127.
- [10] Lieven De Lathauwer, Joos Vandewalle Bart De Moor, "A Multilinear Singular Value Decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, 2000.
- [11] Lars Eldén, Berkant Savas, "Handwritten digit classification using higher order singular value decomposition," *Pattern Recognition*, vol. 40, no. 3, pp. 993-1003, March 2007.
- [12] Lieven De Lathauwer, "Signal Processing Based on Multilinear Algebra," E.E. Department (ESAT), KATHOLIEKE UNIVERSITET LEUVEN, Leuven, PhD Thesis September 1997.
- [13] Tamara G. Kolda and Brett W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455--500, September 2009, doi:10.1137/07070111X.
- [14] Christopher M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag New York, Inc. Secaucus, NJ, USA, 2006