Lattice Computing: applications

Manuel Graña

ENGINE project, Wroclaw University of Technology (WrUT), Poland
Computational Intelligence Group, Basque Country University (UPV/EHU), Spain

TFML 2015,
Bedlewo, Poland, February 18th, 2015

Summary of the talk

- Introduce Lattice Computing paradigm
- Focus on Lattice Autoassociative Memories
- Applications
  - Hyperspectral image unmixing
  - Face recognition
  - MRI classification
  - fMRI processing
  - Multivariate Mathematical Morphology
    - Hyperspectral image
    - brain networks on resting state fMRI

Contents

Contents

Lattice Computing

Definition

Lattice Computing is the class of algorithms built in the realm of Lattice Theory.

- Define computations in the ring of the real valued spaces endowed with some (inf, sup) operators $(\mathbb{R}^n, \vee, \wedge, +)$,
- or use lattice theory to produce generalizations or fusions of conventional approaches.

Contents

Mathematical Morphology

- Classical application of lattice theory to signal and image processing
- Filtering and detection
  - Erosion and dilation operators buit on infimum and supremum operators
    - non-linear convolution-like processes with structural elements
  - Filters: Opening and closing
  - Degmentation: morphological gradient and watershed
  - Detection: top-hat, hit-and-miss

Formal Concept Analysis

- Application of lattice theory to semantic analysis
- Ontology induction from data
  - intensional (attributes) and extensional (instances) representation of concepts
  - building the lattice induced by the partial order of concepts

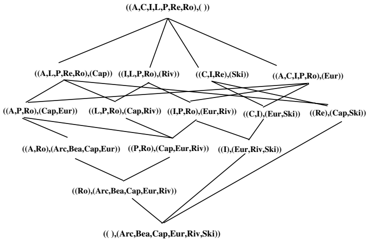


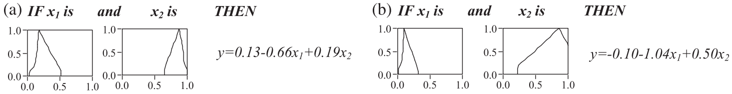
Fig. 1. Concept Lattice of the *European Cities* context.

Lattice Associative Memories

- Builiding learning algorithms with morphological operators
- Associative Memories
  - Store and recall patterns
  - Dual memories from infimum and supremum operators
  - Nice properties:
    - Infinite storage capacity of real valued patterns
    - Robustness against specific erosive/dilative noise
    - Not-nice: sensitivity to general additive noise

Kaburlasos' Lattice Interval Numbers

- A new general data type: Intervals Numbers (IN)
  - Conventional data types can be mapped into IN
  - The lattice valuation function allows to define error measures
  - Variations of conventional learning algoritms
    - Generalization of Fuzzy-ART
    - Lattice Self Organizing Map



Contents

LAAM definitions

- LAAMs are auto-associative neural networks
  - neuron functional activations built on morphological (lattice) operations.
- LAAMs present interesting properties such as perfect recall, unlimited storage and one-step convergence.
- Proposed by Ritter et al.[1] [2]
- We found applications besides image storage and retrieval

[1] G. X. Ritter, P. Sussner, and J. L. Diaz-de Leon. Morphological associative memories. Neural Networks, IEEE Transactions on, 9(2):281–293, 1998.

[2] G. X. Ritter, J. L. Diaz-de Leon, and P. Sussner. Morphological bidirectional associative memories. Neural Networks, 12(6):851–867, 1999.

LAAM definitions

- Input/output pattern pairs

$$(X, Y) = \left\{ \left(\mathbf{x}^{\xi}, \mathbf{y}^{\xi} \right); \xi = 1, .., k \right\}$$

- Linear heteroassociative neural network

$$W = \sum_{\xi} \mathbf{y}^{\xi} \cdot \left(\mathbf{x}^{\xi} \right)'.$$

- Erosive and dilative LAMs, respectively

$$W_{XY} = \bigwedge_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right] \text{ and } M_{XY} = \bigvee_{\xi=1}^{k} \left[\mathbf{y}^{\xi} \times \left(-\mathbf{x}^{\xi} \right)' \right],$$

where $\times$ is any of the $\boxed{\vee}$ or $\boxed{\wedge}$ operators,

LAAM definitions

- Operator $\boxed{\vee}$ denotes the max matrix product

$$C = A \boxed{\vee} B = [c_{ij}] \Leftrightarrow c_{ij} = \bigvee_{k=1..n} \left\{ a_{ik} + b_{kj} \right\},$$

- Operator $\boxed{\wedge}$ denotes the min matrix product

$$C = A \boxed{\wedge} B = [c_{ij}] \Leftrightarrow c_{ij} = \bigwedge_{k=1..n} \left\{ a_{ik} + b_{kj} \right\}.$$

LAAM definitions and properties

Definition

When $X = Y$ then W_{XX} and M_{XX} are called Lattice Auto-Associative Memories (LAAMs).

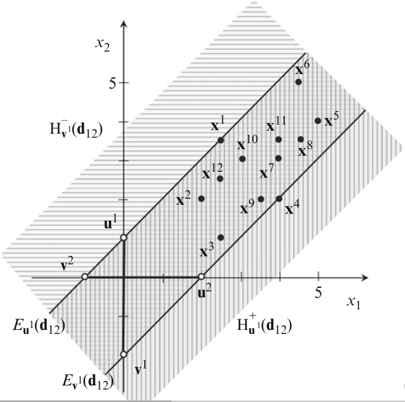
- Perfect recall for an unlimited number of real-valued stored patterns

$$W_{XX} \boxed{\vee} X = X = M_{XX} \boxed{\wedge} X$$

- Convergence in one step for any input pattern. i.e. reaching a fixed point in one step
  - if $W_{XX} \boxed{\vee} \mathbf{z} = \mathbf{v}$ then $W_{XX} \boxed{\vee} \mathbf{v} = \mathbf{v}$
  - if $M_{XX} \boxed{\vee} \mathbf{z} = \mathbf{u}$ then $M_{XX} \boxed{\wedge} \mathbf{u} = \mathbf{u}$.

Fixed points of M_{XX} and W_{XX}[a]

[a] G.X.Ritter, G.Urcid, "Lattice algebra approach to endmember determination in hyperspectral imagery," in P. Hawkes (Ed.), Advances in imaging and electron physics, Vol. 160, 113–169. Elsevier, Burlington, MA (2010)



Contents

Linear Mixing Model

- Linear Mixing Model (LMM):

$$\mathbf{x} = \sum_{i=1}^{M} a_i \mathbf{e}_i + \mathbf{w} = \mathbf{E}\mathbf{a} + \mathbf{w}, \tag{1}$$

  - $\mathbf{x}$ is the d-dimension input vector,
  - $\mathbf{E}$ is the $d \times M$ matrix of d-dimension *endmembers* $\mathbf{e}_i, i = 1, .., M$,
    - defining a convex region covering the measured data.
    - endmembers are affine independent
  - $\mathbf{a}$ is the M-dimension abundance vector, and
    - non-negative $a_i \geq 0, i = 1, .., M$,
    - fully additive to 1: $\sum_{i=1}^{M} a_i = 1$.
  - $\mathbf{w}$ is the d-dimension additive observation noise vector.

Endmember induction Algorithm

Definition

Endmember Induction algorithms (EIA): extracting a set of endmembers E from the data X

- Types of EIA
  - Geometric: searching for simplex covering
  - Algebraic (PCA, ICA, NNMF)
  - Lattice computing: equivalence between lattice independence and affine independence

Ritter's EIA

Algorithm 2 Endmember Threshold Selection Algorithm (ETSA) based on [27,28]

(1) Given a set of vectors $X = \left\{ \mathbf{x}^1, ..., \mathbf{x}^k \right\} \subset \mathbb{R}^n$ compute the min and max auto-associative memories W_{XX} M_{XX} from the data. Their column vector sets W and M will be the candidate endmembers.

(2) Register W and M relative to the data set adding the maximum and minimum values of the data dimensions (bands in the hyperspectral image). Obtain $\overline{W}$ and $\overline{M}$ as follows:
   (a) Compute $u_i = \bigvee_{\xi=1}^{n} x_i^{\xi}$ and $v_i = \bigwedge_{\xi=1}^{n} x_i^{\xi}$.
   (b) Compute $\overline{\mathbf{m}}^i = \mathbf{m}^i + v_i$ and $\overline{\mathbf{w}}^i = \mathbf{w}^i + u_i$

(3) Remove lattice dependent vectors from the joint set $\overline{W} \bigcup \overline{M}$.

(4) Compute the standard deviation along each dimension of the candidate endmember vectors, denoted by the vector $\overrightarrow{\sigma} = \{\sigma_1, \ldots, \sigma_n\}$.

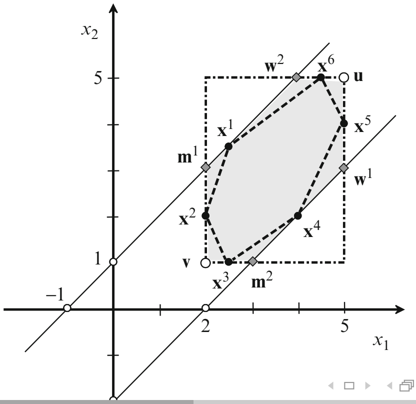
(5) Assume the first vector in the set $\mathbf{v}_1 \in \overline{W} \bigcup \overline{M}$ as the first endmember, $E = \{\mathbf{v}_1\}$

(6) Iterate for the remaining vectors $\mathbf{v} \in \overline{W} \bigcup \overline{M}$
   (a) If $\|\mathbf{v} - \mathbf{e}\| < \gamma \overrightarrow{\sigma}$ for any $\mathbf{e} \in E$ then discard $\mathbf{v}$ otherwise include $\mathbf{v}$ in E

Figure : A specification of Ritter's EIA

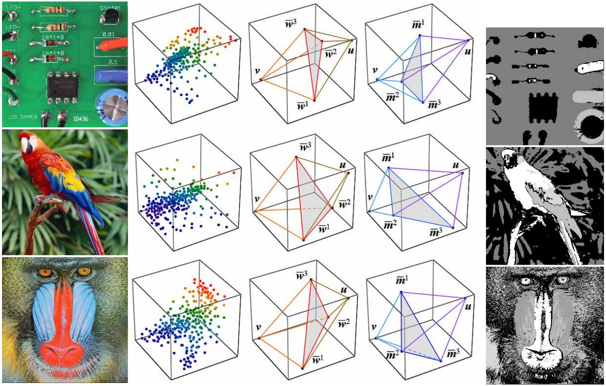
Convex Polytope from Ritter's EIA[a]

[a] G.X.Ritter,G.Urcid,"Lattice algebra approach to endmember determination in hyperspectral imagery," in P. Hawkes (Ed.), Advances in imaging and electron physics, Vol. 160, 113–169. Elsevier, Burlington, MA (2010)



Ritter's EIA endmembers in RGB images[a]

[a] G. Urcid, JC Valdiviezo-N, GX Ritter, Lattice algebra approach to color image segmentation, JMIV 42 (2-3), 150-162 (2012)



Graña's EIA[a]

[a]M. Graña, I. Villaverde, J.O. Maldonado, C. Hernandez, Two lattice computing approaches for the unsupervised segmentation of hyperspectral images, Neurocomputing 72:2111–2120 (2009)

Algorithm 3 Endmember Induction Heuristic Algorithm (EIHA)

(1) Shift the data sample to zero mean $\{\mathbf{f}^c(i) = \mathbf{f}(i) - \overrightarrow{\mu}; i = 1, .., n\}$.

(2) Initialize the set of vertices $E = \{\mathbf{e}_1\}$ with a randomly picked sample. Initialize the set of lattice independent binary signatures $X = \{\mathbf{x}_1\} = \{(e_k^1 > 0; k = 1, .., d)\}$

(3) Construct the AMM's based on the lattice independent binary signatures: M_{XX} and W_{XX}.

(4) For each pixel $\mathbf{f}^c(i)$
   - (a) compute the noise corrections sign vectors$\mathbf{f}^+(i) = (\mathbf{f}^c(i) + \alpha\overrightarrow{\sigma} > \mathbf{0})$ and $\mathbf{f}^-(i) = (\mathbf{f}^c(i) - \alpha\overrightarrow{\sigma} > \mathbf{0})$
   - (b) compute $y^+ = M_{XX} \boxtimes \mathbf{f}^+(i)$
   - (c) compute $y^- = W_{XX} \boxtimes \mathbf{f}^-(i)$
   - (d) if $y^+ \notin X$ or $y^- \notin X$ then $\mathbf{f}^c(i)$ is a new vertex to be added to E, execute once 3 with the new E and resume the exploration of the data sample.
   - (e) if $y^+ \in X$ and $\mathbf{f}^c(i) > \mathbf{e}_{y^+}$ the pixel spectral signature is more extreme than the stored vertex, then substitute $\mathbf{e}_{y^+}$ with $\mathbf{f}^c(i)$.
   - (f) if $y^- \in X$ and $\mathbf{f}^c(i) < \mathbf{e}_{y^-}$ the new data point is more extreme than the stored vertex, then substitute $\mathbf{e}_{y^-}$ with $\mathbf{f}^c(i)$.

(5) The final set of endmembers is the set of original data vectors $\mathbf{f}(i)$ corresponding to the sign vectors selected as members of E.

Hyperspectral images

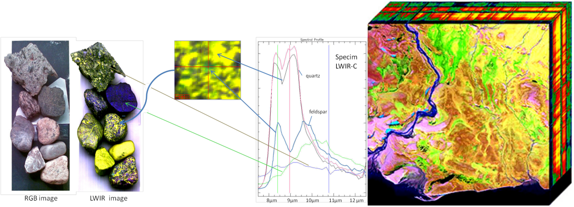
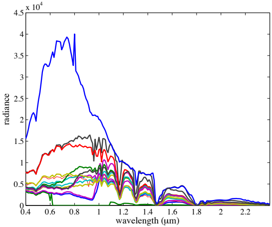


Figure : Hyperspectral imaging, source: wikipedia

Hyperspectral image unmixing

(a)



(b)

Figure : (a) patch of washington dc image, (c) EIHA endmembers

Hyperspectral image unmixing

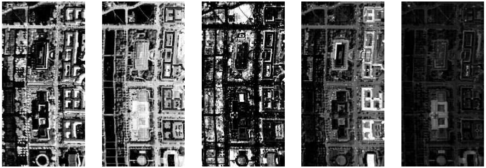


Figure : LSU estimated abundances from Washington DC patch

Lattice Independent Component Analysis (LICA)

- A non-linear version of Independent Component Analysis
  - Statistical Independence $->$ Lattice independence
  - Endmembers $==$ Lattice Independent sources
  - Abundance computation $==$ feature extraction

LICA

Algorithm 1 LICA feature extraction .

1. Given training data matrix

$$X_{TR} = \{\mathbf{x}_j; j = 1, \ldots, m\} \in \mathbb{R}^{N \times m}$$

   and testing data matrix

$$X_{TE} = \{\mathbf{x}_j; j = 1, \ldots, m/3\} \in \mathbb{R}^{N \times m/3}$$

2. Apply on X_{TR} an EIA to induce the set of k endmembers

$$E = \{\mathbf{e}_j; j = 1, \ldots, k\}$$

3. Unmix train and test data: $A_{TR} = E^{\#} X_{TR}^T$ and $A_{TE} = E^{\#} X_{TE}^T$.

Application examples

- Focus on recent works in our reasearch group
- LICA applications
  - Face recognition: feature extraction
  - DWI data classification Alzheimer's Disease
- Multivariate Mathematical Morphology
  - resting state fMRI processing
  - hyperspectral image spectral-spatial classification

Contents

Face recognition

- 1st Experiment comparing LICA with PCA, ICA, LDA[3]
- Classification by Extreme Learning Machines, Random Forest and SVM
- Four umbalanced face databases from the FERET database

[3] Ion Marques, Manuel Graña, Face recognition with Lattice Independent Component Analysis and Extreme Learning Machines. Soft Computing,16(9):1525-1537 (2012)

Face data processing pipeline

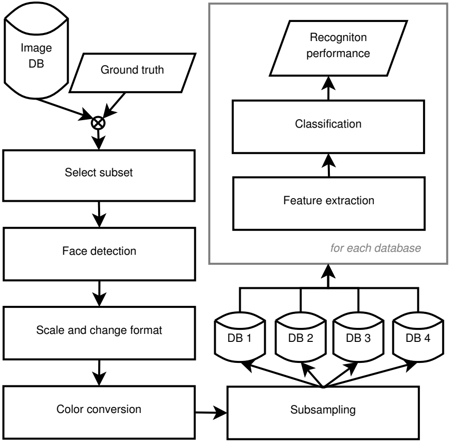
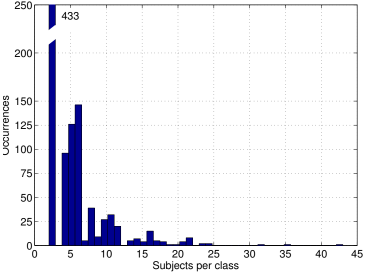


Figure :

Face sample

Figure : subject sample

Databases



	DB 1	DB 2	DB 3	DB 4
Number of samples	5169	3249	832	347
Number of classes	994	635	265	79
Mean (samples per class)	4.3924	3.1396	5.2835	5.2002
Standard deviation (samples per class)	5.8560	3.4498	4.9904	4.5012
Median (samples per class)	2	2	4	4
Mode (samples per class)	2	2	2	2

Figure : features of the face databases in the experiment

Face detection

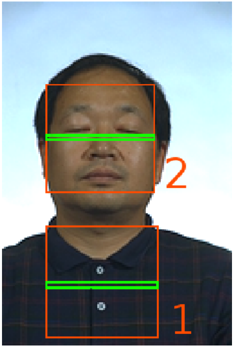
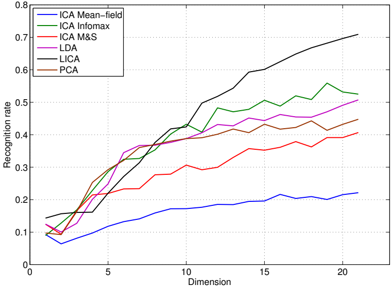


Figure : Face detection candidates by Viola's algorithm, source: SciLab, SIVP toolbox

Face bases

Figure : Rows: Instances of 5 basis from ICA Infomax, ICA Molguey & Schuster, LICA, PCA

Face recognition results



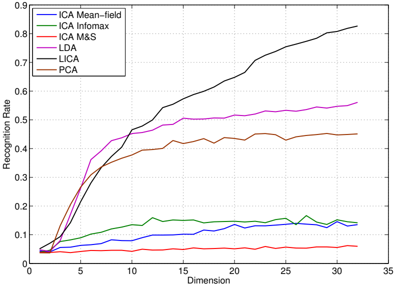


Figure : face recognition results on databases of increasing size

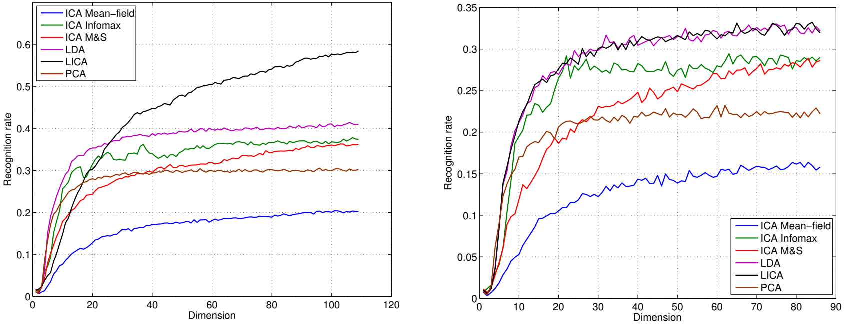


Figure : face recognition results cont.

Fusion of features

- The 2nd experiment performs the fusion of features obtained by LICA and linear algorithms[4]
- Classification by ELM
- Four different databases tested
- Conclusion: LICA-fusion enhances the linear features

[4] Ion Marques, Manuel Graña Fusion of lattice independent and linear features improving face identification. Neurocomputing 114:80–85 (2013)

Fusion pipeline

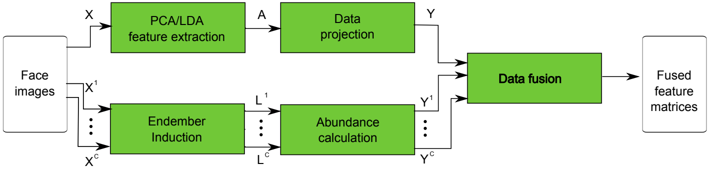


Figure : Pipeline of LICA and linear feature fusion

Feature fusion

- Dataset matrix X:
  $X = \{\mathbf{x}_i^c; i = 1, \ldots, n; c \in \{1, 2, \ldots, C\}\} \in \mathbb{R}^{n \times N}$,
- Dataset class restricted c:
  $X^c = \left\{\mathbf{x}_j^c \in X; j = 1, \ldots, M\right\} \in \mathbb{R}^{M \times N}$,
  - class restricted abundance matrix: $A^c = (E^c)^{\#} X^{cT}$,
- Data features obtained by linear algorithm $Y = \Phi X^T$

Feature fusion (cont.)

- Class restricted abundance coefficients $A^c = \{\mathbf{a}_i^c; i = 1, \ldots, M\} \in \mathbb{R}^{M_c \times M}$
- Linear feature matrix $Y = \{\mathbf{y}_i^c; i = 1, \ldots, n; c \in \{1, 2, \ldots, C\}\} \in \mathbb{R}^{d \times n}$
- Fused i-th feature vector $\mathbf{z}_i^c \in \mathbb{R}^d$ of a face of class c is

$$\mathbf{z}_i^c = \mathbf{a}_{j(i)}^c \| \left[y_{i,M_c+1}^c, \ldots, y_{i,d}^c \right], \tag{2}$$

Face databases

Table : Summary characteristics of the experimental databases.

Name	Number of images	Number of subjects	Variations
AT&T Database of Faces	400	40	Pose, expression, light*
MUCT Face Database	3755	276	Pose, expression, light
PICS (Stirling)	312	36	Pose, expression
Yale Face Database	165	15	Expression, light, glasses

Face feature fusion results

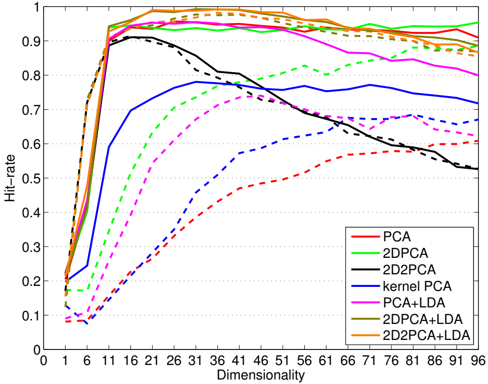


Figure : Recognition rate using ELM classifier for the AT&T database. Dotted lines correspond to standard feature extraction methods. Solid lines correspond to feature fusion approach.

Face feature fusion results (cont.)

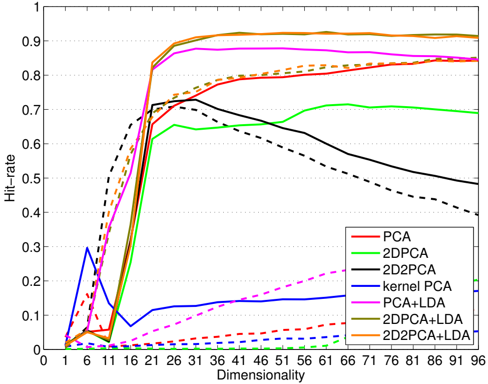


Figure : Recognition rate using ELM classifier for the MUCT database. Dotted lines correspond to standard feature extraction methods. Solid lines correspond to feature fusion approach.

Face feature fusion results (cont.)

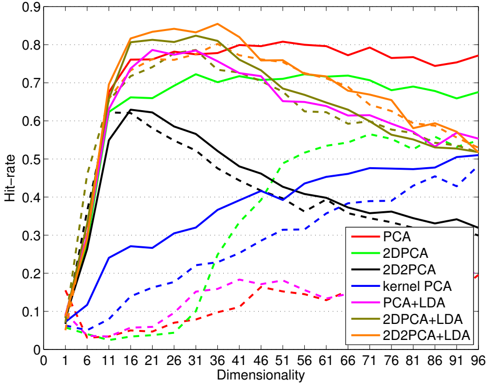


Figure : Recognition rate using ELM classifier for the PICS database. Dotted lines correspond to standard feature extraction methods. Solid lines correspond to feature fusion approach.

Face feature fusion results (cont.)

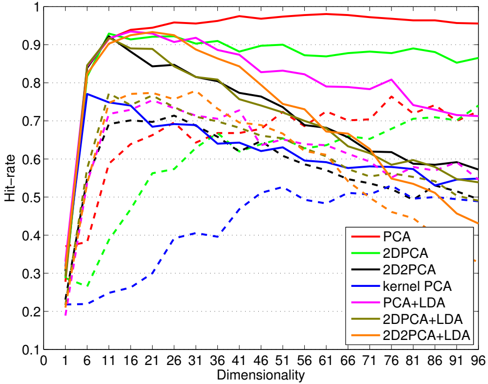


Figure : Recognition rate using ELM classifier for the Yalefaces database. Dotted lines correspond to standard feature extraction methods. Solid lines correspond feature fusion.

Contents

Diffusion MRI data classification

- Discrimination of Alzheimer's disease (AD) patients from diffusion MRI data [5]
- Database collected by collaborating clinicians at Hospital Santiago, Vitoria
- Classification by SVM, RVM, 1-NN
- LICA residuals are used for feature selection
  - localization of voxel sites for classification with clinical significance
  - classification performance

[5] M. Termenon, M. Graña, A. Besga, J. Echeveste, A. Gonzalez-Pinto, Lattice Independent Component Analysis feature selection on Diffusion Weighted Imaging for Alzheimer's Disease Classification, Neurocomputing 114:132–141 (2013)

DWI, DTI and FA, MD

- Diffusion Weighted Imaging (DWI) measures the diffusion of water molecules inside the brain along several directions
  - *in vivo* information about the integrity of the White Matter (WM) fibers.
- Diffusion Tensor Imaging (DTI) is the diffusion covariance tensor at each voxel.
- Scalar diffusion measures computed from DTI are
  - Fractional Anisotropy (FA) privileged diffusion direction
  - Mean Diffusivity (MD), magnitude of the diffusion process
- DTI studies about WM abnormalities in AD have found differences between AD patients and controls

Preprocessing pipeline

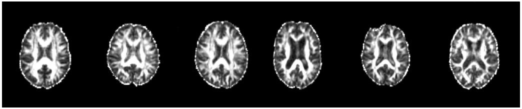
Algorithm 1 T1 and DWI data processing pipeline to obtain spatially normalized FA data.

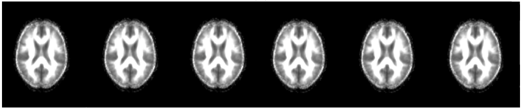
1. Convert DICOM to nifti
2. Skull stripping T1-weighted volumes
3. Affine registration of T1-weighted skull stripped volumes to template MNI152.
4. Correct DWI scans.
5. Obtain skull stripped brain masks for each DWI corrected scans.
6. Apply diffusion tensor analysis computing DTI and FA.
7. Rigid registration 6DoF of FA data to T1-weighted normalized volumes, from Step3.

LICA for feature detection in FA

- Linear Mixing Model $X = AS + \epsilon$,
- S extracted by an EIA from the set of FA volumes
- Abundance estimation by LSU $\hat{A} = XS^{\#}$, or FCLSU
- Residual error $R = (X - \hat{A}S)^2$.
- $P(i,j,k)$ Pearson's correlation of $R(i,j,k)$ with the categorical variable (AD=1, HC=0)
  - Feature sites $|P(i,j,k)| > P_\alpha$
    - where P_α is the α-percentile of the e.p.d. of $P(i,j,k)$

LICA for feature detection in FA


(a)


(b)

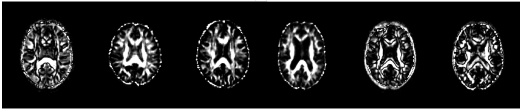

(c)

Figure : (a) original FA data, (b) reconstruction from FCLSU estimated abundances, (c) residual R

Feature localization

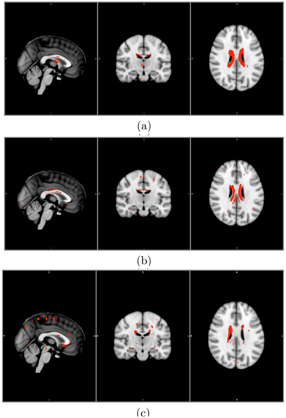


Figure : Feature localization in the brain (a) LICA residual, (b) bare FA data, (c) VBM

Feature localization

- LICA residuals produce feature localization that correspond to biomarkers in the limbic system in agreement with the medical literature,
  - hippocampus,
  - amygdala , and
  - the brainstem.

DWI Classification results

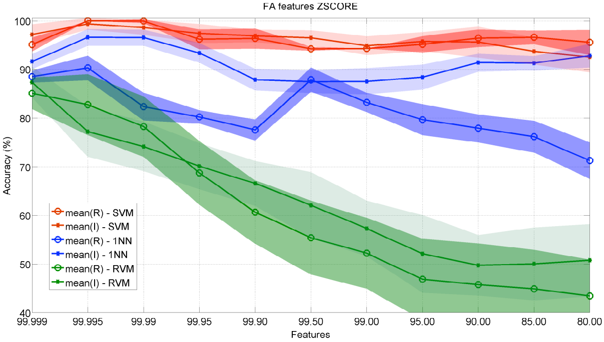


Figure : LICA residual R vs. bare FA, accuracy results for decreasing P_α increasing number of features

DWI Classification results

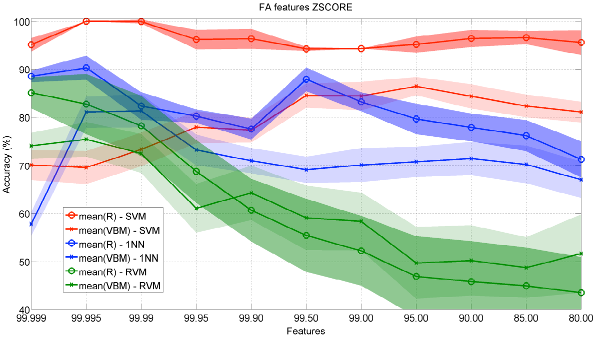


Figure : LICA residual R vs. VBM, accuracy results for decreasing P_α increasing number of features

Contents

Multivariate Mathematival Morphology

Morphological operations are mappings between complete lattices, denoted $\mathbb{L}$ or $\mathbb{M}$,

erosion is a mapping $\varepsilon : \mathbb{L} \rightarrow \mathbb{M}$ conmuting with the infimum operation, $\varepsilon\left(\bigwedge Y\right) = \bigwedge_{y \in Y} \varepsilon\left(y\right)$; $\forall Y \subseteq \mathbb{L}$

dilation is a mapping $\delta : \mathbb{L} \rightarrow \mathbb{M}$ conmuting with the supremum operation, $\delta\left(\bigvee Y\right) = \bigvee_{y \in Y} \delta\left(y\right)$.

high dimensional vectors have no natural total order

Multivariate ordering

Definition

A *h*-ordering is defined by a surjective map of the original partially ordered set onto a complete lattice $h : X \to \mathbb{L}$,

- The order in $\mathbb{L}$ induces a total order in X,

$$r \leq_h r' \Leftrightarrow h(r) \leq h\left(r'\right) \tag{3}$$

Definition

Supervised h-ordering the mapping is built by supervised classification

- satisfying $h(b) = \perp$, $\forall b \in B$, and $h(f) = \top$, $\forall f \in F$,
- for background and foreground $B, F \subset X$, $B \cap F = \emptyset$,
  - $\perp$ and $\top$ are the bottom and top elements of $\mathbb{L}$

Supervised erosion and dilation

Definition

The supervised *h*-erosion by structural object S is

$$\varepsilon_{h,S}(I)(p) = I(q) \text{ s.t. } I(q) = \bigwedge_h \{I(s); s \in S_p\}$$

Definition

The supervised *h*-dilation by structural object S is

$$\delta_{h,S}(I)(p) = I(q) \text{ s.t. } I(q) = \bigvee_h \{I(s); s \in S_p\}$$

where $\bigwedge_h$ and $\bigvee_h$ are the infimum and supremum defined by the reduced ordering $\leq_h$

LAAM h-function

Definition

Given $\mathbf{c} \in \mathbb{R}^n$ and $X = \{\mathbf{x}_i\}_{i=1}^{K}$, $\mathbf{x}_i \in \mathbb{R}^n$; the LAAM based h_X-function is

$$h_X(\mathbf{c}) = \zeta\left(\mathbf{x}^{\#}, \mathbf{c}\right), \tag{4}$$

- $\mathbf{x}^{\#} \in \mathbb{R}^n$ is a LAAM recall result

$$\mathbf{x}^{\#} = M_{xx} \boxed{\wedge} \mathbf{c}$$

or

$$\mathbf{x}^{\#} = W_{xx} \boxed{\vee} \mathbf{c}$$

- $\zeta(\mathbf{a}, \mathbf{b})$ is the Chebyshev distance $\zeta(\mathbf{a}, \mathbf{b}) = \bigvee_i |a_i - b_i|$.

One-side ordering

Definition

one-side LAAM-supervised ordering:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \mathbf{x} \leq_X \mathbf{y} \Longleftrightarrow h_X(\mathbf{x}) \leq h_X(\mathbf{y}). \tag{5}$$

- $h_X : \mathbb{R}^n \to \mathbb{L}_X$, where $\mathbb{L}_X = \left(\mathbb{R}_0^+, <\right)$, $\perp_X = 0$
- the Background set B s.t. $h_X(\mathbf{b}) = \perp_X = 0$
  - is the set of fixed points of the LAAM $B = \mathcal{F}(X)$

B/F ordering

Definition

The relative background/foreground supervised LAAM *h*-function:

$$h_r(\mathbf{c}) = h_F(\mathbf{c}) - h_B(\mathbf{c}), \tag{6}$$

Given training sets B and F

Definition

relative LAAM-supervised ordering denoted $\leq_r$:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,\ \mathbf{x} \leq_r \mathbf{y} \iff h_r(\mathbf{x}) \leq h_r(\mathbf{y}) \tag{7}$$

B/F ordering

- $h_r(\mathbf{c}) : \mathbb{R}^n \to \mathbb{L}_{B/F}$ where $\mathbb{L}_{B/F} = (\mathbb{R}, <)$,
  - $h_r(\mathbf{b}) > 0$; $\mathbf{b} \in \mathcal{F}(B)$
  - $h_r(\mathbf{f}) < 0$; $\mathbf{f} \in \mathcal{F}(F)$
  - no bottom or top elements
  - $h_r(\mathbf{c}) = 0$; decision boundary $\mathbf{c} \in C_r$

Contents

Resting state fMRI

- Resting state fMRI data has been used to study brain functional connectivity
  - correlation of low frequency oscillations in diverse areas of the brain reveal functional resting networks.
  - connections discovered provide a brain fingerprint,
    - default-mode network: Doing nothing network
- Not imposing constraints on subject cognitive abilities.
  - Example: in the study of brain maturation there is no single cognitive task which is appropriate across the aging population.

Schizophrenia

Schizophrenia is a severe psychiatric disease that is characterized by delusions and hallucinations, loss of emotion and disrupted thinking.

Functional disconnection between brain regions is suspected to cause these symptoms, because of known aberrant effects on gray and white matter in brain regions that overlap with the default mode network.

Resting state fMRI studies have indicated aberrant default mode functional connectivity in schizophrenic patients.

Auditory Hallucinations

Goal of our work[6] is to find differences in connectivity betwee patients with and without auditory hallucinations.

Classification provides the detection power value of the connectivity features extracted

Features have biological meaning and anatomical correspondence

[6] D. Chyzhyk, M Graña, D Öngür, AK Shinn, Discrimination of Schizophrenia Auditory Hallucinators by Machine Learning of resting-state Funcitonal MRI, International Journal of Neural Systems (online first)

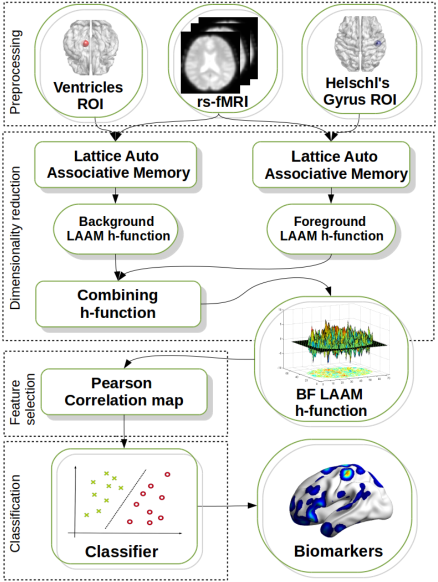
Experiment goal

experiments provide a proof of concept of LAAM multivariate morphology approach for functional connectivity

- finding discriminating features of **healthy** control subjects, **schizophrenia** patients **with** and **without** auditory hallucinations.

comparison with activity based features: regional homogeneity (ReHo), fractional amplitude of low frequency fluctuations (fALFF)

Experiment layout



Experiment layout

1. Compute both h_X and $h_{B/F}$ maps related to the key seed region
2. Select most salient voxel sites to extract features
3. Crossvalidation classification experiments == detection power
4. Localization of features
    1. network effect related to the auditory hallucinations.

Materials

- 68 men and women, ages 18-65 years, divided in three groups:
  - (i) SZAH: 26 schizophrenia patients with a history of AH,
  - (ii) SZNAH: 14 schizophrenia patients without a history of AH, and
  - (iii) HC: 28 healthy control subjects.

Preprocessing

- fMRI: 240 BOLD volumes and one anatomical T1-weighted per subject
  - skull extraction
  - manually AC-PC transformed.
  - The functional images coregistered to the T1-weighted anatomical image.
  - slice timing,
  - head motion correction
  - smoothing (FWHM=4mm)
  - spatial normalization to (MNI) template
  - temporal filtering (0.01-0.08 Hz)
  - linear trend removing
  - All the subjects have less than 1mm maximum displacement and less than 1º of angular motion.

Seed regions for connectivity

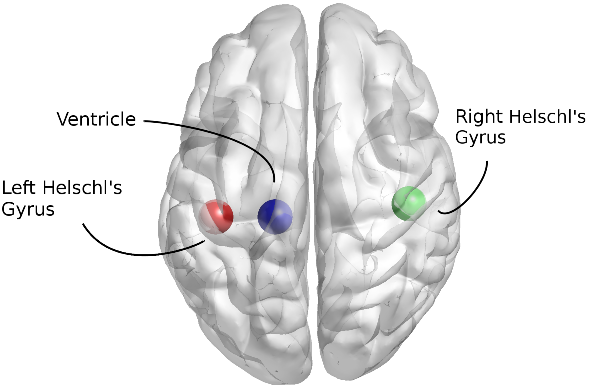


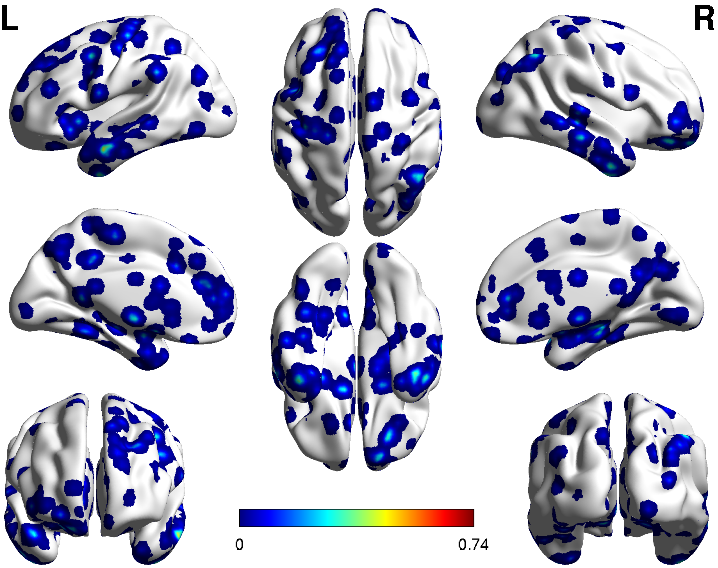
Figure : The ROIs used for lattice auto-associative memory (LAAM) based connectivity analysis. Left and right Heschl's gyrii foregrounds, ventricle background.

Classification results

Table : Average **accuracy** of cross-validation results, feature vector size *per* columns.

	Measure	Feat.Map.	**HG**	500	1000	5000	10000
SZAH vs. SZNAH	Func. Conn.	OS-LAAM	L	**97.5**	**97.5**	**97.5**	**92.5**
			R	**92.5**	**92.5**	**95**	**95.2**
		BF-LAAM	L	**100**	**97.5**	**95**	**90**
			R	**100**	**100**	**100**	**100**
	Local Act.	ReHo	-	**100**	**100**	**100**	**100**
		ALFF	-	85	87.5	**92.5**	**92.5**
		fALFF	-	**97.5**	**100**	**100**	**97.5**

Localizations from the B/F *h*-function LHG discriminating Hallucinators



Localizations of the features

Table 4. Cortical brain regions of the feature voxel sites corresponding to **BF-LAAM** *h***-map** feature vector of size 1000, for the classification of **SZAH vs. SZNAH**. Regions highlighted in bold typeface have also been reported in[81]. CS = Cluster size. H=Brain Hemisphere, L=Left, R=Right.

BF-LAAM LHG			Coordinates			BF-LAAM RHG			Coordinates		
Region	H	CS	x	y	z	Region	H	CS	x	y	z
Middle Frontal Gyrus	L	10	-36	33	45	Frontal Pole	L/R	7/11	-15/6	60/69	-18/-9
Inferior Frontal Gyrus	L	10	-60	18	-3	Superior Frontal Gyrus	L	5	-3	36	48
Middle Temporal Gyrus	L	33	-57	-6	-27	Superior Temporal Gyrus	L	5	-72	-24	6
Temporal Pole	R	11	30	27	-36	Postcentral Gyrus	L	7	-45	-33	63
Precentral Gyrus	L	10	-21	-24	63	Juxtapositional Lobule Cortex		8	0	9	66
Parahippocampal Gyrus	R	15	27	-24	-15	Angular Gyrus	R	5	51	-54	48
						Cingulate Gyrus	R	5	6	-33	18
						Lateral Occipital Cortex	R	6	42	-63	45
						Occipital Fusiform Gyrus	L	5	-36	-75	-9
						Parahippocampal Gyrus	L	21	-21	-12	-21

Some comments on the results

- Functional connectivity and local activity have similar discrimination power
  - Localizations do not overlap == different biological interpretations
- LAAM-based functional connectivity achieved similar detection power for LHG and RHG
  - conventional approaches did not report detections from RHG connectivity
- Findings fit into postulatd models for hallucinations

Contents

Hyperspectral image spectral-spatial classification

- Independent SVM spectral classification per pixel
- Multivariate mathematical morphology provide the spatial information
    - Watershed regions from morphological gradient
        - assume homogeneous class inside each region
    - Spatial correction of SVM results

Hyperspectral image and baseline SVM classification

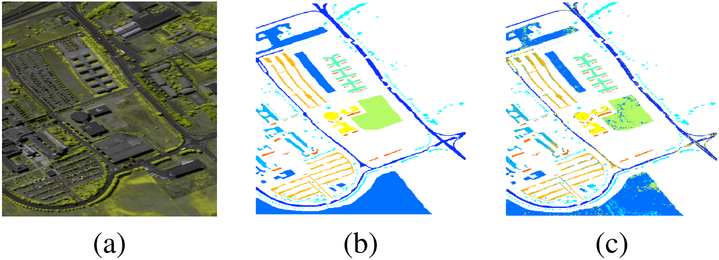


Figure : (a) Pavia image, (b) ground truth, (c) pixelwise SVM classification

Supervised morphological gradient

Definition

The h-supervised morphological gradient:

$$g_{h,S}(I) = h\left(\delta_{h,S}(I)\right) - h\left(\varepsilon_{h,S}(I)\right),$$

where $\varepsilon_{h,S}(I)$ and $\delta_{h,S}(I)$ are the h-supervised erosion and dilation

Unsupervised selection of LAAM training data

- An EIA induces a set of endmembers $E = \{\mathbf{e}_i\}_{i=1}^{p}$. Compute $D = [d_{i,j}]_{i,j=1}^{p}$, where $d_{ij} = |\mathbf{e}_i, \mathbf{e}_j|$
- One-side h-supervised ordering
  - $X = \{\mathbf{e}_{k^*} \in E\}$ such that $k^* = \arg\min_k \left\{\frac{1}{p-1}\sum_{i \neq k} d_{ik}\right\}_{i=1}^{p}$.
- Background/Foreground h-supervised orderings
  - $F = \{\mathbf{e}_{i^*} \in E\}$ and $B = \{\mathbf{e}_{j^*} \in E\}$ such that $(i^*, j^*) = \arg\max_{i,j}\{(d_{ij})\}$

Endmembers

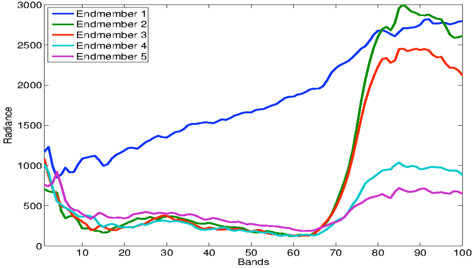


Figure : Endmembers found in the hyperspectral image

Morphological gradient results

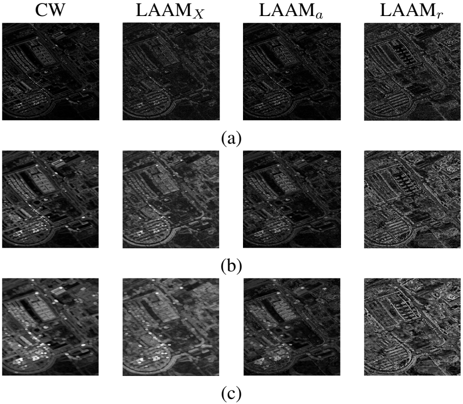
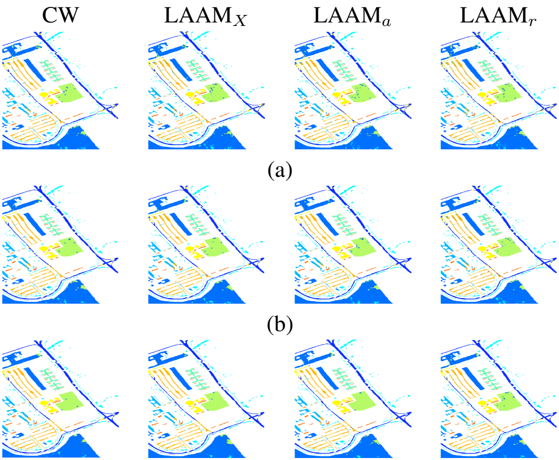


Figure : Morphological gradients with increasing structural element size

Classification results



(a)

(b)

(c)

Classification results

Method		OA	AA	κ
Pixel-wise SVM		88.97	91.60	0.8565
SVM + NWHED	CW	93.41	94.39	0.9135
	LAAM_X	93.65	94.72	0.9167
	LAAM_r	92.61	93.84	0.9034
SVM+WHED	CW	95.46	95.86	0.9403
	LAAM_X	95.27	96.11	0.9378
	LAAM_r	94.91	95.71	0.9332

Table : Classification results of the Pavia University hyperspectral image: OA, AA, and Kappa (κ) values. Morphological structural element disc shaped of radius $r = 5$.

Class specific sensitivities

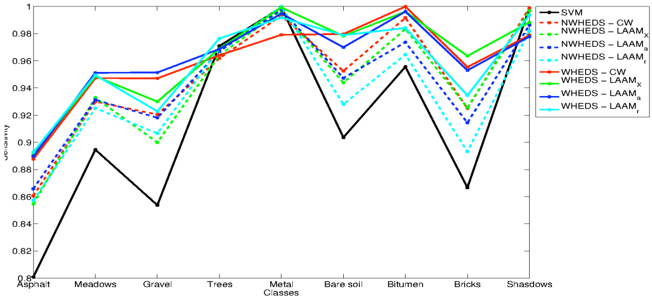


Figure : Sensitivity per Class, structural element of radius 3

Concluding remarks

- Lattice Computing proposes a new paradigm for the definition of Intelligent Systems algorithmms
    - does not involve statistical techniques, is model-free
        - relies mostly on lattice operators and lattice theory
- I have concentrated on the LAAMs stream of research
- Increasing range of practical applications with competitive results

Future work avenues

- Sparse bayesian hyperspectral unmixing based on Ritter's EIA
- Multi-class Supervised Multivariate Mathematica Morphology
- LICA fMRI group analysis for detection

People involved in this research

- Darya Chyzhyk
- Maite Termenon
- Ion Marques
- Miguel A Veganzones
- With the collaboration of Prof. G Ritter (U. Florida), AK Shinn (U. Harvar), A Besga (University hospital Alava)