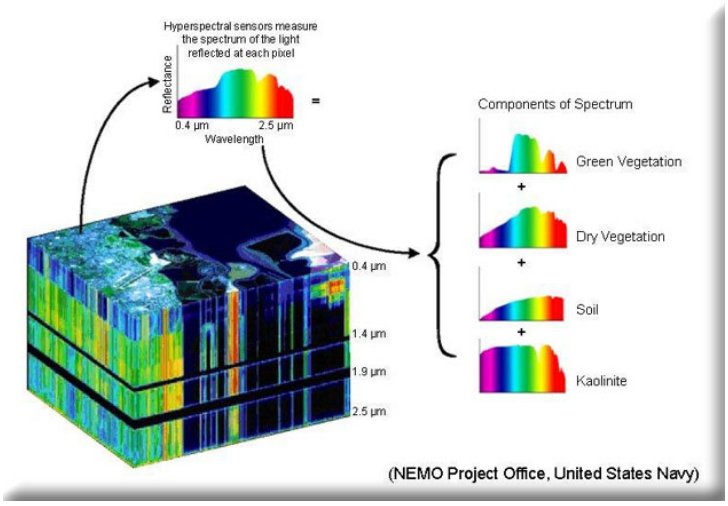
A Compressive Sensing and Unmixing Scheme for Hyperspectral Data Processing

Authors: Chengbo Li, Ting Sun, Kevin Kelly, Yin Zhang
Publication: IEEE Trans. On Image Processing, March, 2012
Speaker: Woongbi Lee

Short summary: In this paper, a low complexity scheme is proposed for hyperspectral data compression and reconstruction. The data reconstruction minimizes the total variation of the abundance fractions subject to a preprocessed fidelity equation with a significantly reduced size, and other side constraints.

I. INTRODUCTION

Hyperspectral imaging is a technique to identify and quantify distinct material substances(재료물질) from observed spectral data. It employs hyperspectral sensors to collect information as a set of images. Each image represents a range of the electromagnetic spectrum, which is known as spectral bands containing the visible, near-infrared, and shortwave infrared spectral bands. Hyperspectral imaging has a wide range of applications such as terrain classification, mineral detection and exploration, pharmaceutical counterfeiting, environmental monitoring, and military surveillance.



<출처: Google Image>

Hyperspectral imaging is typically low resolution and is a mixture of several different material substances, termed endmembers (pure signature), each possessing a characteristic hyperspectral signature. Hyperspectral unmixing is to decompose each pixel spectrum to identify and quantify the relative abundance of each endmember. The representative endmembers for a given scene are known a priori and their signatures can be obtained from a spectral library or are unknown but the hyperspectral data is fully accessible.

Hyperspectral data cubes have huge volume so that it is difficult to directly process and analyze them in real time. But, the hyperspectral data are highly compressible with two-fold compressibility: 1) each spatial image is compressible, and 2) the entire cube, when treated as a matrix, is of low rank.

In this paper, data are acquired by means of compressive sensing (CS), similar to extension of the single pixel camera. Data reconstruction and unmixing are combined into a single step of much lower complexity. They propose a compressive sensing and unmixing (CSU) scheme which formulates an unmixing model based on total variation (TV) minimization, develops an efficient algorithm to solve it, and provides experimental and numerical evidence to validate the scheme.

II. System Model

A. Notations

n_e	number of significant endmembers
$w_i^T \in \mathbb{R}^{n_b}$, for $i = 1,...,n_e$	spectral signature of an endmember
$n_b \geq n_e$	number of spectral bands
$x_i \in \mathbb{R}^{n_b}$	hyperspectral data vector at the i^{th} pixel
$h_i^T \in \mathbb{R}^{n_e}$	abundance fractions of the endmembers for any $i \in \{1,...,n_p\}$
n_p	number of pixels
$X = \left[x_1,...,x_{n_p} \right]^T \in \mathbb{R}^{n_p \times n_b}$	a matrix representing the hyperspectral cube

$W = \left[w_1, ..., w_{n_e} \right]^T \in \mathbb{R}^{n_e \times n_b}$	mixing matrix containing the endmember spectral signatures
$H = \left[h_1, ..., h_{n_p} \right]^T \in \mathbb{R}^{n_p \times n_e}$	a matrix holding the respective abundance fractions
$\mathbf{1}_s$	column vector of all ones with length s
$A \in \mathbb{R}^{m \times n_p}$	measurement matrix
$F \in \mathbb{R}^{m \times n_b}$	observation matrix
m	number of samples for each spectral band

B. Problem Formulation

The hyperspectral vector x_i at the i-th pixel can be regarded as a linear combination of the endmember spectral signatures, and the weights are gathered in a nonnegative abundance vector h_i.

$$X = HW,\ H\mathbf{1}_{n_e} = \mathbf{1}_{n_p},\ H \geq 0 \tag{1}$$

where $X = \left[x_1, ..., x_{n_p} \right]^T \in \mathbb{R}^{n_p \times n_b}$, $H = \left[h_1, ..., h_{n_p} \right]^T \in \mathbb{R}^{n_p \times n_e}$, and $W = \left[w_1, ..., w_{n_e} \right]^T \in \mathbb{R}^{n_e \times n_b}$.

$$\begin{bmatrix} - & x_1 & - \\ - & x_2 & - \\ & \vdots & \\ - & x_{n_p} & - \end{bmatrix} = \begin{bmatrix} - & h_1 & - \\ - & h_2 & - \\ & \vdots & \\ - & h_{n_p} & - \end{bmatrix} \begin{bmatrix} - & w_1 & - \\ - & w_2 & - \\ & \vdots & \\ - & w_{n_e} & - \end{bmatrix}$$

Since each column of X represents a 2D image corresponding to a particular spectral band, we can collect the compressed hyperspectral data $F \in \mathbb{R}^{m \times n_b}$ by randomly sampling all the columns of X using the same measurement matrix $A \in \mathbb{R}^{m \times n_p}$, where $m < n_p$ is the number of samples for each column.

$$AX = F \tag{2}$$

Combining (1) and (2), we obtain

$$AHW = F,\ H\mathbf{1}_{n_e} = \mathbf{1}_{n_p},\ H \geq 0 \tag{3}$$

Assuming the endmember spectral signatures in W are known, we aim to find the abundance distributions, H in (3), given the measurement matrix A and the compressed hyperspectral data F.

Instead of l_1 minimization of compressive sensing, TV (Total Variation) regularization is generally more advantageous on image problems since it can better preserve edges or boundaries in images. TV regularization puts emphasis on sparsity in the gradient map of the image and is suitable when the gradient of the underlying image is sparse. With assumption that the gradient of each image composed by abundance fractions for each endmember is mostly and approximately piecewise constant, we propose to recover the abundance matrix H by solving the following unmixing model:

$$\min_{H\in\mathbb{R}^{n_p\times n_e}} \sum_{j=1}^{n_e} TV\left(He_j\right) \text{ s.t. } AHW = F, H\mathbf{1}_{n_e} = \mathbf{1}_{n_p}, H \geq 0$$

where e_j is the j-th standard unit vector in $\mathbb{R}^{n_e}$.

$$TV\left(He_j\right) \triangleq \sum_{i=1}^{n_p} \left\| D_i\left(He_j\right)\right\|$$

$\|\bullet\|$ is the 2-norm in $\mathbb{R}^2$, and $D_i \in \mathbb{R}^{2\times n_p}$ denotes the discrete gradient operator at the *i*-th pixel.

III. SVD PREPROCESSING

In eq. (3), $AHW = F$, the size is $m\times n_b$, where $m << n_p, m << n_b$. In this section, we propose a preprocessing procedure based on singular value decomposition (SVD) of the observation matrix F to decrease the size of eq. (3) from $m\times n_b$ to $m\times n_e$. Typically, $n_e << n_b$.

Let $A\in\mathbb{R}^{m\times n_p}$ and $W\in\mathbb{R}^{n_e\times n_b}$ be full-rank, and $F\in\mathbb{R}^{m\times n_b}$ be rank- n_e with $n_e < \min\left\{n_b, n_p, m\right\}$. Let $F = U_e\sum_e V_e^T$ be the economy-size SVD of F where $\sum_e \in \mathbb{R}^{n_e\times n_e}$ is diagonal and positive definite, $U_e \in \mathbb{R}^{m\times n_e}$ and $V_e \in \mathbb{R}^{n_b\times n_e}$ both have orthonormal columns. Assume that $\text{rank}\left(WV_e\right) = n_e$, then the two linear systems below for $H\in\mathbb{R}^{n_p\times n_e}$ have the same solution set;

$$AHW = F \Leftrightarrow AHWV_e = U_e\sum_e \tag{4}$$

IV. ALGORITHMS

$$\min_{H} \sum_{j=1}^{n_e} TV\left(He_j\right) \text{ s.t. } AHWV_e = U_e \sum_e \tag{5}$$

To separate the discrete gradient operator from the non-differentiable TV term, splitting variables $v_{ij} = D_i\left(He_j\right)$ for $i = 1,...,n_p$ and $j = 1,...,n_e$. Then (5) becomes

$$\min_{H, v_{ij}} \sum_{i,j} \left\| v_{ij} \right\| \text{ s.t. } D_i\left(He_j\right) = v_{ij}, AHWV_e = U_e \sum_e \tag{6}$$

The augmented Lagrangian function for (6) can be written as

$$\begin{aligned} L_A\left(H, v_{ij}\right) \triangleq \sum_{i,j} \left\{ \left\| v_{ij} \right\| - \lambda_{ij}^T \left(D_i\left(He_j\right) - v_{ij} \right) + \frac{\alpha}{2} \left\| D_i\left(He_j\right) - v_{ij} \right\|_2^2 \right\} - \\ \left\langle \Pi, AHW - F \right\rangle + \frac{\beta}{2} \left\| AHW - F \right\|_F^2 - \nu^T \left(H\mathbf{1}_{n_e} - \mathbf{1}_{n_p} \right) + \frac{\gamma}{2} \left\| H\mathbf{1}_{n_e} - \mathbf{1}_{n_p} \right\|_2^2 \end{aligned} \tag{7}$$

where λ_{ij}, Π, ν are multipliers of appropriate sizes, and $\alpha, \beta, \gamma > 0$ are penalty parameters.

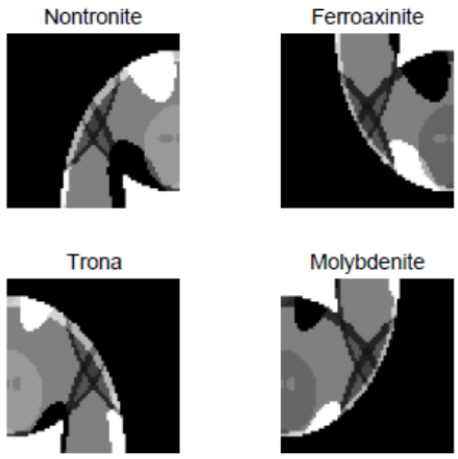
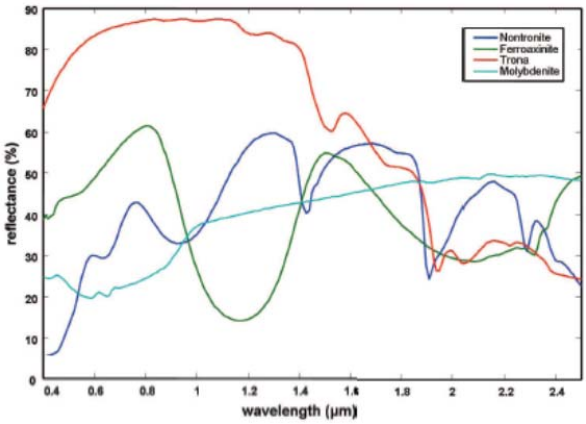
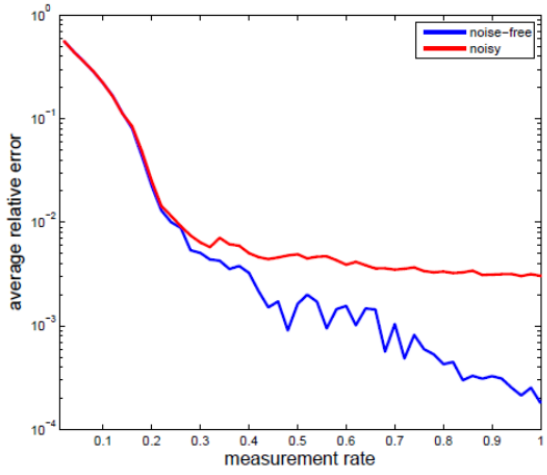
V. NUMERICAL RESULTS: SYNTHETIC DATA

In the experiments, we use randomized Walsh-Hadamard matrices as measurement matrices, *A*, considering that they permit fast transformation and easy hardware implementation. A Walsh-Hadamard matrix is randomized by choosing *m* random row from it and applying a random permutation to its columns.

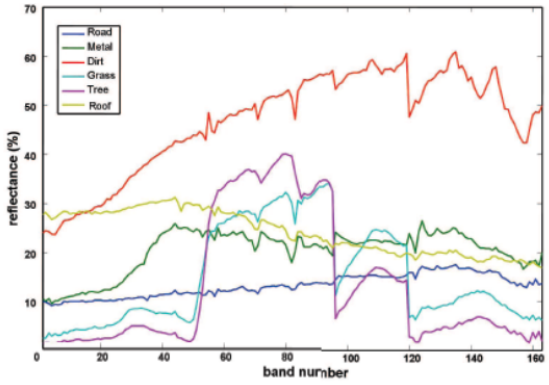
Walsh-Hadamard (WH) Matrix (Matlab – hadamard)
The *N*-by-*N* WH matrices $\mathbf{H}_N$ are defined by $\mathbf{H}_1 = 1$ $\mathbf{H}_N = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{bmatrix}$

A. Test Results on Synthetic Data

We selected $n_e = 4$ endmembers from the ASTER Spectral Library: nontronite, ferroaxinite, trona, and molybdenite. $n_b = 211$ spectral bands were selected in the range of 0.4 to 2.5 μm. The distributions of abundance fractions (물질별 점유비율) corresponding to 4 endmembers were given with a spatial resolution of $n_p = 64 \times 64$.

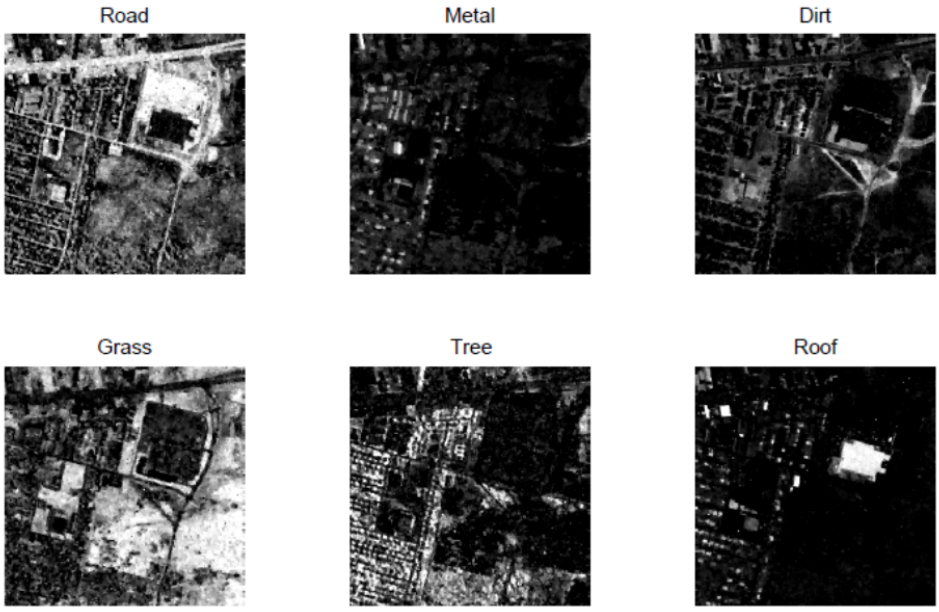
Synthetic abundance distributions ($H \in \mathbb{R}^{n_p \times n_e}$)	Endmember spectral signatures
	
Recoverability for noisy and noise-free cases	
	- Relative errors in computed abundance fractions versus measurement rate of compressed data on 100 distinct testing points with or without additive noise. - When measurement rate is greater than 20% in both noisy and noise-free cases, the CSU scheme attains relative error less than 1%.

In the second test, from the publicly available HYDIC Urban hyperspectral data, $n_b = 163$ bands in a range from 0.4 to 2.5 μm, $n_p = 307 \times 307$, $n_e = 6$ significant endmembers: road, metal, dirt, grass, tree, and roof.

"Urban" image and endmember selection	Spectral signatures with water absorption bands abandoned
	

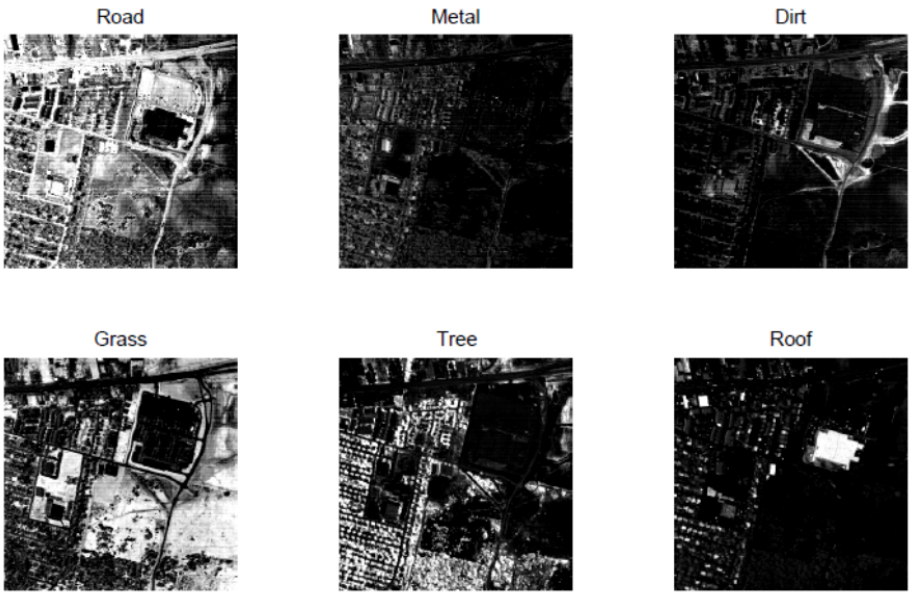
Unmixing results from 25% measurements are given below.

Computed abundance: solution obtained from 25% of measurements



Estimated abundance: least squares solution

from directly solving $AHW = F$ for H with 100% data

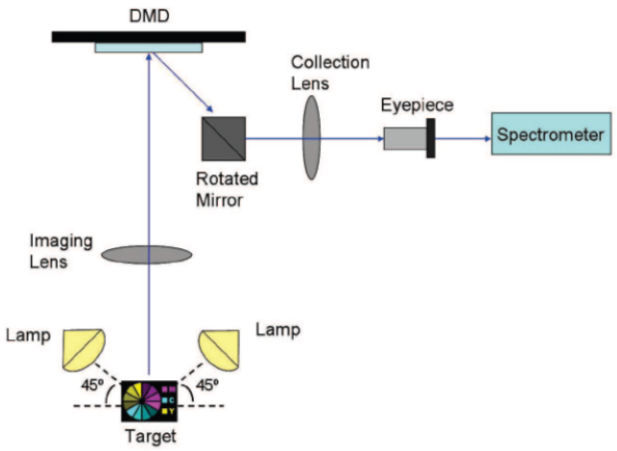
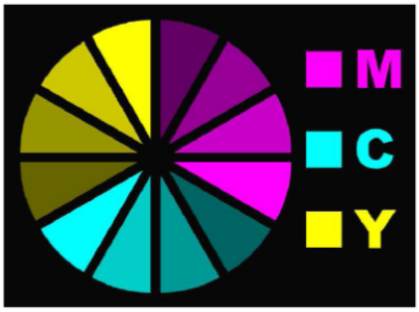
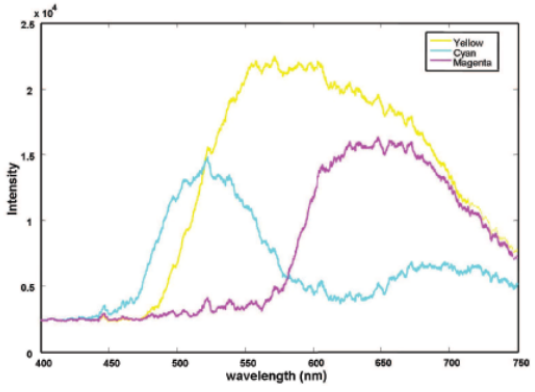


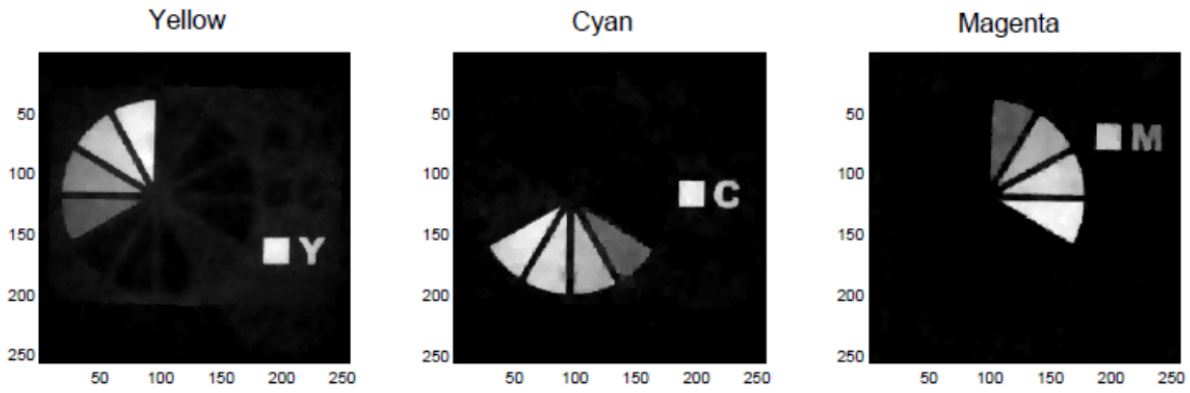
VI. Experimental Results: Hardware-Measured Data

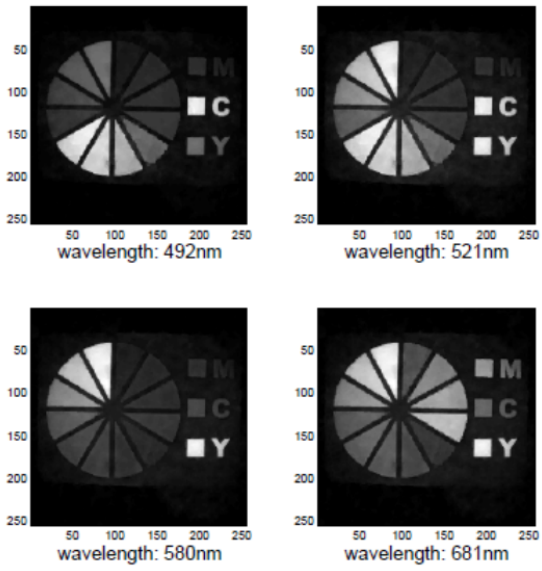
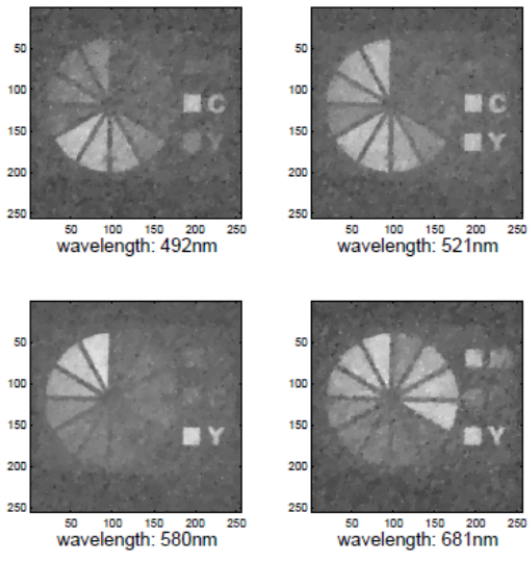
A. Hardware Implementation

A compressive sensing hyperspectral imaging system is implemented based on a digital micro-mirror device (DMD). This system incorporates a micro-mirror array driven by pseudo-random patterns and one spectrometer.

$n_b = 175$ in the range of 0.4 to 0.75 μm, $n_p = 256 \times 256$, $n_e = 3$: yellow, cyan, and magenta

Single-pixel camera schematic for hyperspectral data acquisition	
	
Target image “Color wheel”	Measured spectral signatures of the three endmembers
	

Estimated abundance (H): CS unmixing solution from 10% measurements


Four slices computed by the proposed approach ($X = HW$)	Four slices computed slice-by-slice by TV minimization
	

VII. Conclusion

This paper proposes a compressive sensing and unmixing (CSU) scheme for hyperspectral data processing that does not require forming or storing any full-size data cube. The CSU scheme consists of three major steps: 1) data acquisition by compressive sensing; 2) data preprocessing

by SCD; and (3) data unmixing by solving a compressed unmixing model with total-variation regularization on abundance fraction distributions.

VIII. DISCUSSION

After meeting, please write discussion in the meeting and update your presentation file.

Appendix

Reference

[1]
[2]
[3]

Dry and Noncontact EEG Sensors for Mobile Brain-Computer interface.

Yu Mike Chi et al. (Gert Cauwenberghs*)

IEEE Transactions on Neural Systems and Rehabilitation Engineering (2012)

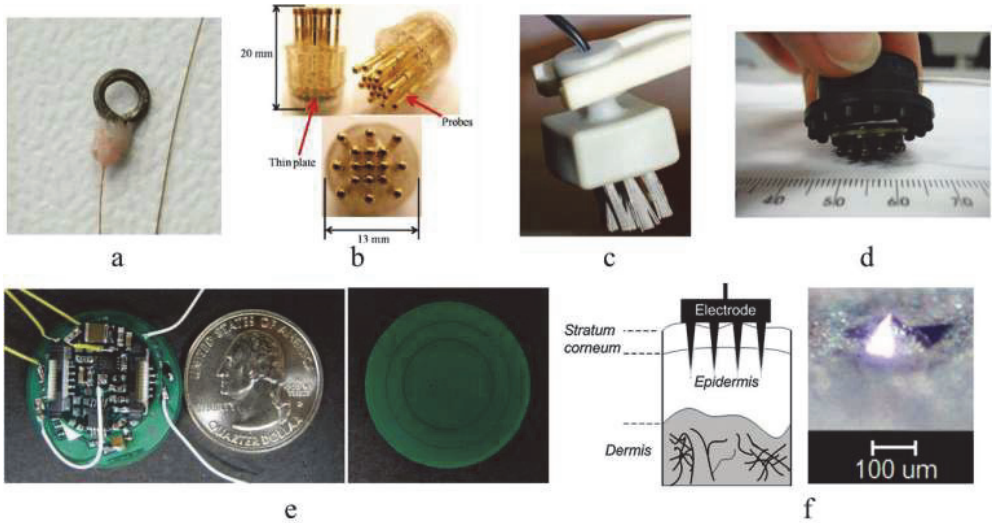
Presenter : SeungChan Lee

GIST, Dept. of Information and Communication, INFONET Lab.

Gwangju Institute of Science and Technology

Background

- Various EEG electrodes

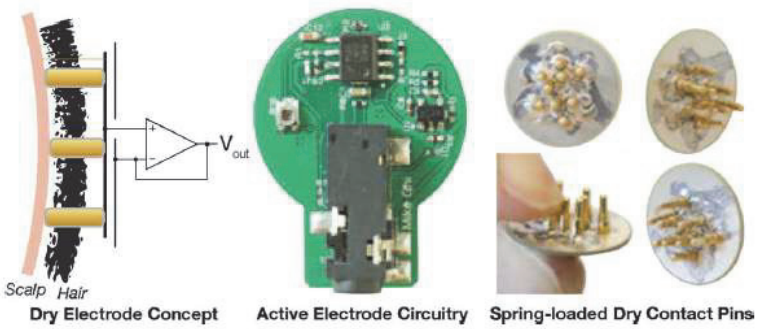


(a) a miniature passive ring electrode (b) a spring-loaded dry electrode (c) a bristle-type dry electrode (d) the Quasar hybrid EEG biosensor (e) a non-contact-type active dry EEG sensor (f) Diagram of a micro-tip electrode and the pyramidal shape of a micro-tip

Introduction

- Introduction
  - In EEG- based BCI systems, conventional BCI systems need extensive preparations such as scalp abrasion, conductive gels for good signal quality. Moreover, multiple wired electrodes are difficult to escape from laboratory scale experiments.
  - To overcome these problems, extensive research produced a variety of dry electrodes.
  - In this paper, they introduced dry and non-contact electrodes and evaluate their performance with SSVEP paradigms.
- Contents
  - Introduction of their dry and non-contact electrodes
  - Offline sensor benchmark with SSVEP paradigm
  - Online decoding test with mobile application

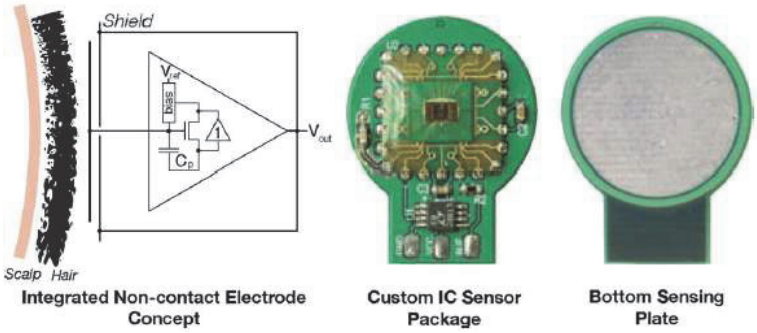
Dry electrodes



Dry Electrode Concept | Active Electrode Circuitry | Spring-loaded Dry Contact Pins

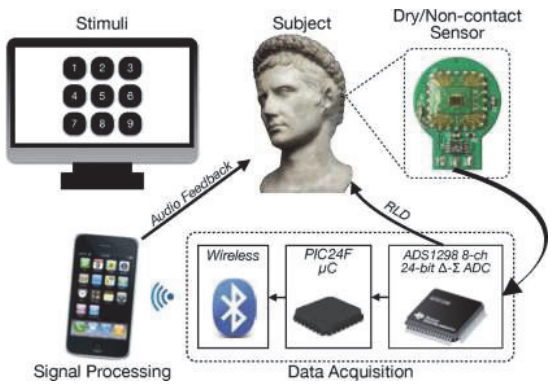
- Structure
  - Lower plate : a set of spring-loaded pins, a male snap connector
  - Upper PCB : active electrode circuitry (CMOS-input opamp, LMP7702)
  - Unity gain buffer (gain=1) with shielded cable
  - No discomfort, injury hazard

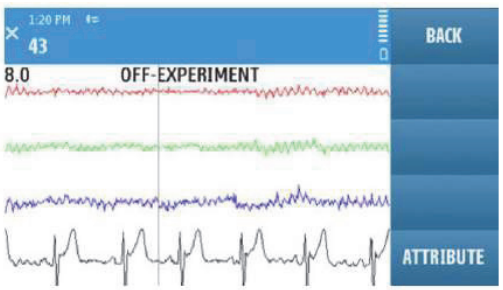
Non-contact electrodes



- Structure
  - The electrodes operates via capacitive coupling on top of hair.
  - Based on a custom VLSI integrated analog front-end circuit

System design and mobile application





- Data acqusition
  - 24bit delta-sigma ADCs(TI ADS1298), PIC24F low-power microcontroller, onboard Bluetooth module, two AAA batteries (10 hours working time)
- Mobile signal processing
  - Nokia N97 cellular phone (640x360 pixel 3.5 inch touchscreen LCD)
  - Canonical correlation analysis(CCA) : band-pass filter and correlation calculation

CCA

- CCA is a multivariable statistical method used when there are two sets of data, which may have some underlying correlation.
- It finds a pair of linear combinations, for two sets, such that the correlation between the two canonical variables is maximized.
- Consider two multidimensional random variables X, Y and their linear combinations $x = X^T W_x$ and $y = X^T W_y$ respectively.
- CCA finds the weight vectors, Wx and Wy, which maximize the correlation between x and y, by solving the following problem:

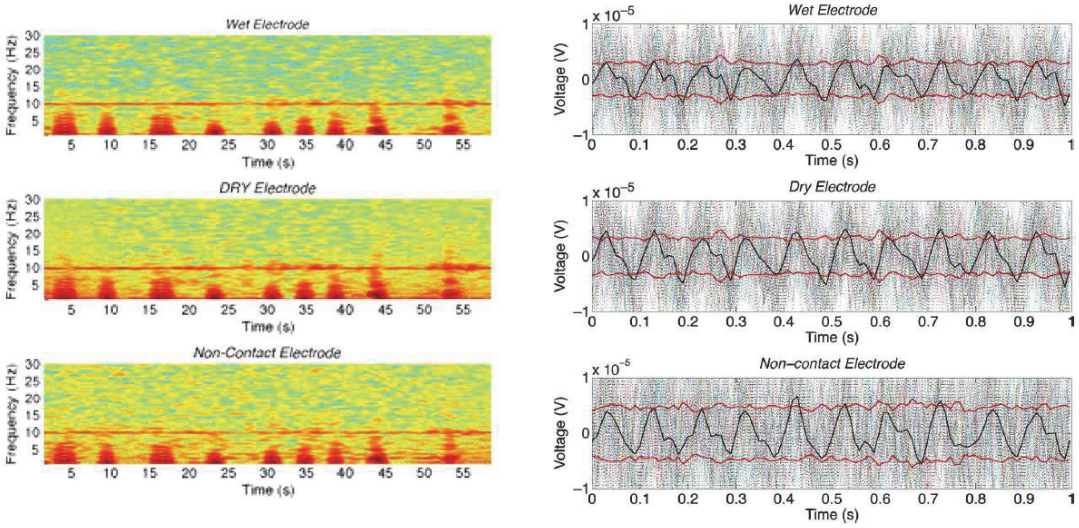
$$\max_{W_x, W_y} \rho(x, y) = \frac{E[x^T y]}{\sqrt{E[x^T x]E[y^T y]}} = \frac{E[W_x^T X Y^T W_y]}{\sqrt{E[W_x^T X X^T W_x]E[W_y^T Y Y^T W_y]}}.$$

- The maximum of ρ with respect to Wx and Wy is the maximum canonical correlation. Projections onto Wx and Wy, i.e. x and y, are called canonical variants.

Offline sensor benchmark

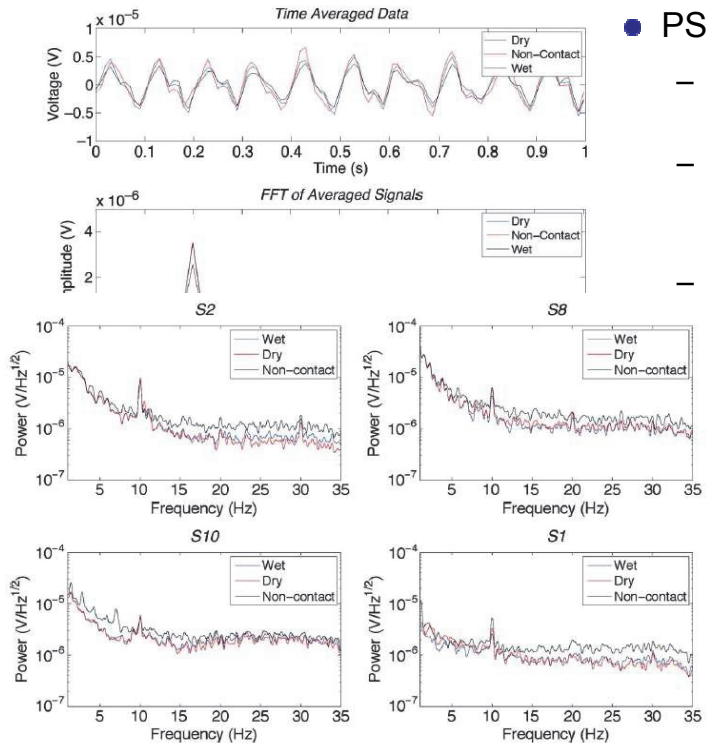
- Test setting
  - Comparison electrodes : wet Ag/AgCl electrodes, proposed dry electrodes, proposed non-contact electrodes
  - Three sensors array are attached in a triad over the occipital region as closely together as possible.
  - 10 subjects
  - Each subject gaze at a single SSVEP target stimulus(10Hz) displayed on a CRT monitor for a 1-min duration.
  - Each subject repeated this task three times, and the best dataset was used for analysis.

Offline sensor benchmark



- Result plots
  - Left spectrograms are one of the 60s trials shown 10Hz SSVEP stimulus.
  - Right graphs show detailed signals with the average in black, the standard deviation in red with the raw signals.

Offline sensor benchmark



- PSD
  - In the four subjects shown the 10Hz stimulus is clearly visible.
  - PSD from the wet electrode almost perfectly matches that from the dry electrode.
  - The PSD of the noncontact electrode's signals also shows the 10Hz stimulus. But, there is greater amount of broadband noise due to their high coupling impedance.

Offline sensor benchmark

TABLE I
SIGNAL CORRELATION BETWEEN DIFFERENT ELECTRODES

Subject	SSVEP Amplitude (μV)			Sensor Correlation			SNR (dB)		
	Wet	Dry	NC	Wet vs. Dry	Wet vs. NC	Dry vs. NC	Wet	Dry	NC
1	1.1	1.7	2.2	0.88	0.85	0.74	-15.16	-10.97	-10.36
2	3.7	3.7	3.2	0.98	0.88	0.85	-6.49	-7.00	-8.46
3	1.9	2.0	2.1	0.90	0.78	0.70	-11.71	-12.24	-12.96
4	2.2	2.2	2.4	0.97	0.80	0.78	-7.69	-8.09	-8.22
5	1.1	1.1	1.0	0.97	0.96	0.94	-12.24	-11.87	-13.05
6	1.6	1.2	1.4	0.75	0.71	0.55	-6.61	-10.49	-9.67
7	1.6	1.1	1.8	0.91	0.86	0.88	-14.33	-13.61	-10.72
8	2.5	3.4	3.5	0.93	0.73	0.70	-6.85	-5.17	-7.47
9	1.4	0.8	0.8	0.89	0.85	0.85	-13.08	-17.42	-17.64
10	1.4	1.8	1.4	0.95	0.57	0.59	-15.60	-13.29	-18.21
Mean	**1.8**	**1.9**	**2.0**	**0.91**	**0.80**	**0.76**	**-10.98**	**-11.01**	**-11.68**
STD	0.8	1.0	0.9	0.07	0.11	0.13	3.71	3.56	3.78

$$\mathrm{SNR} = 10\ \log_{10} \frac{\bar{X}(10\ \mathrm{Hz})^2_{\mathrm{rms}}}{var(x) - \bar{X}(10\ \mathrm{Hz})^2_{\mathrm{rms}}}.$$

- Correlation and SNR
  - Over half the subjects has a correlation of greater than 0.9 between the wet and dry electrodes.
  - Correlation values of the wet versus noncontact electrode were lower. But, half the subject had correlation values of above 0.8.

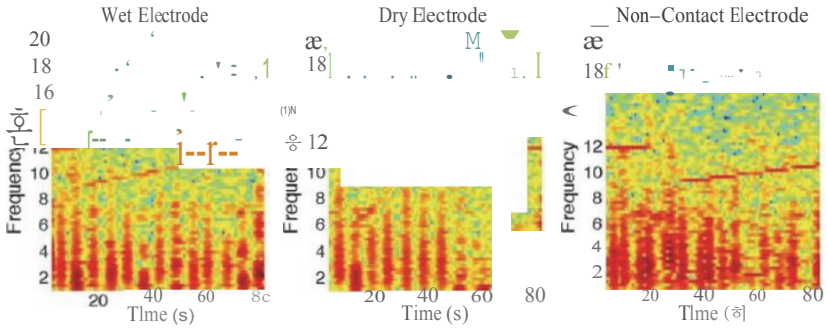
Online decoding test

- Test setting
  - Subjects 1 and 2 were recalled to perform an SSVEP phone dialing task using the mobile signal processing platform.
  - Procedure
    - 4s sliding window with 1s increments
    - Two consecutive decisions are constructed as a successful input and trigger an audio feedback to notify the subject
    - Noncontact electrodes
      - 6s sliding window with four consecutive decisions due to degraded SNR
  - Predetermined 12 digit sequence
  - Signal decoding performed using CCA analysis

Online decoding test

TABLE III
RESUL: FROM ONLINE BCI TESTS

		Accuracy			Detection Time (1.)			ITR (bitvmin)		
		Wet	Dry	NC	Wet	Dry	NC	Wel	Dry	NC
Subject 1	Trial 1	0.83	0.92	1.00	6.2	5.7	10.3	23.0	28.1	19.3
	Trial 2	0.83	0.83	1.00	5.9	5.8	9.7	23.9	22.6	20.5
	Trial 3	0.83	1.00	1.00	6.4	5.6	*9A*	20.5	*3-'A*	21.0
Subjecl 2	Trial 1	0.83	0.83	0.50	6.2	5.9	12.8	23.0	23.9	4.0
	Trial 2	0.83	0.92	0.75	5.9	6.3	9.7	23.9	27.3	11.9
	Trial 3	0.92	0.83	0.75	5.7	6.3	11.0	29.2	22.6	10A
	Mean	0.85	0.89	0.83	6.04	5.92	10.49	23.9	26.5	14.5
	STO	0.03	0.07	0.20	0.26	0.31	1.29	2.90	4.52	6.85



Online decoding test

- Discussion of online test
  - In subject 1's spectrograms for the three different electrodes, the different SSVEP frequencies are clearly visible.
  - The wet and dry electrodes were could both be successfully used for BCI.
  - The dry electrode trials achieved superior performance to the wet electrode trials because the wet electrodes was tested last
  - Noncontact electrodes
    - Subject 1achieve 100% accuracy with noncontact electrodes because of longer detection window. But they achieve lower ITR (19 bits/min).
    - Subject 2 had difficulty with utilizing the noncontact electrodes due to thicker hair.

Conclusion

- Quantitative benchmarking show that dry and noncontact electrodes are capable of resolving SSVEP-type signals.
  - The dry electrode only shows a slight amount of signal degradation.
  - The noncontact electrodes show more signal degradation and susceptibility to movement artifacts.
- However, the online test demonstrate that both electrodes can be successfully utilized in BCI applications.
- The signal quality of noncontact electrodes is possible to still resolved with careful circuit design.

INFONET Seminar Application Group 05/25/2013

Exemplar-Based Processing for Speech Recognition

Tara N. Bhuvana Ramabhadra.
IEEE SIGNAL PROCESSING MAGAZINE

Presenter Pavel Ni

Gwangju Institute of Science and Technology

Introduction

Automatic Speech Recognition is the translation of spoken words in to text.
(Voice dial, Apple Siri, Google One Voice, Samsung S voice)

Recognition and classification of speech requires modeling of speech production and uncertainty in it. Vocal tract complexity noise corruption, and vocal tract variations amongst different people arise uncertainty.

The goal of modeling is to establish a generalization from the set of observed data such that accurate classification can be made about unseen data i.e. speaker independent speech recognition.

Introduction

Construction of the model leads to two categories of approaches for modeling the observed data:

- Global-data model uses all available training data to build a model before the test sample seen
  - allow for a generalization of the observed data only if distribution estimated by the model provides a reasonable description of unseen data.
  - with limited training data unable of representing the fine detail in distribution of the data. Therefore require large amount of training data.

- exemplar based modeling since the model is build from a few relevant training examples for each test sample.
  - Building an instance of the model based only on the relevant and informative exemplars. Doesn't need large data however for each query builds local model.

Speech recognition problem

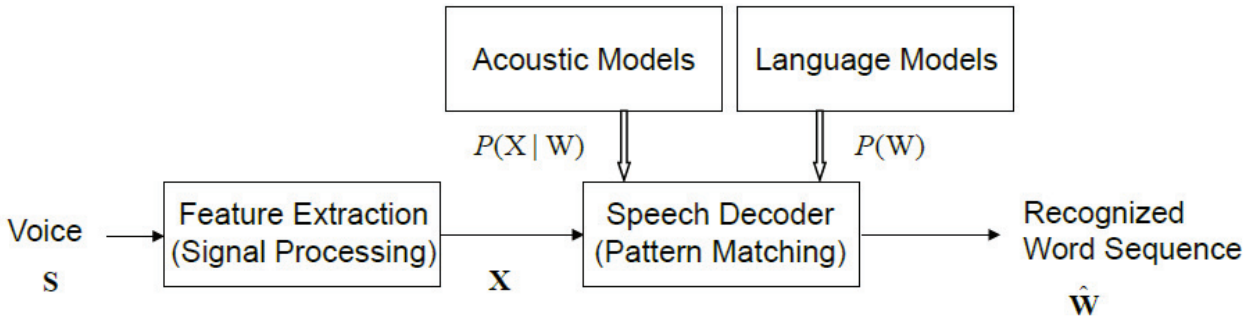


Figure 1. Block diagram of speech recognition system

Where X is a set of observations, W sequence of words.

Acoustic model: physics of sound speech, models of vocal tract
Language model: syntax and semantics

$$\hat{W} = \arg\max_{W} P(W \mid X) = \arg\max_{W} \frac{P(X \mid W)P(W)}{P(X)}$$

In speech recognition we need to find sequence of words $\hat{W}$. The common solution is to find the word sequence that maximizes the posterior probability $P(W \mid X)$

System overview

Features - such as power, pitch, and vocal tract configuration fr om the speech signal.
Frames – fixed length features

3 stages of recognition:
- Exemplar selection
- Instance modeling
- Frame decoding

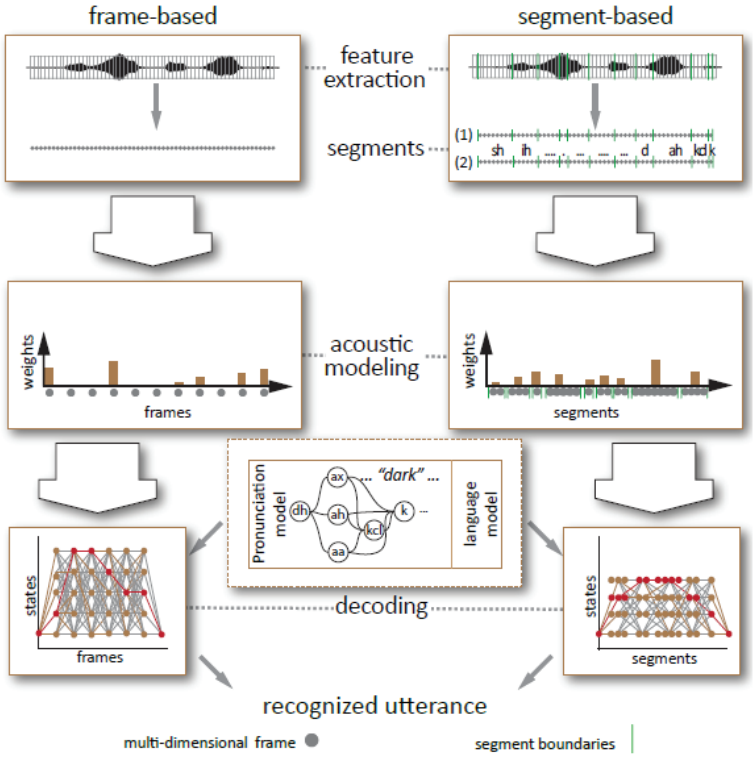


Figure 2. Frame and segment based recognition

System overview

- Exemplar selection

Identifies instances from the training data which are more relevant to test inst ances. Training set selectively downsampled to simplify the search and then appropriate exemplars are taken from this reduced set.

- Instance modeling

Set of training instances which most relevant for test instance is used to mode l the test instance. Weight distribution using k-NN (near neighbors) or sparse r epresentation (SR or compressive sensing)

- Frame decoding

Recognizing an entire sentence.

k-Nearest Neighbors

k-NN algorithm classifies test point based on *k* closest neighbors in the training set.
Exemplar selection: exemplars for individual frames, fixed length sequences of multiple frames

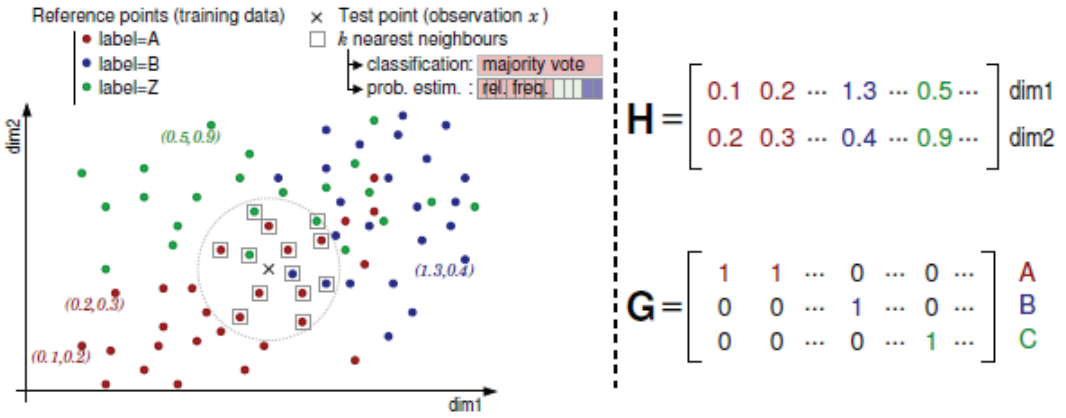


Figure 3. An example of k-NN in a 2-D feature space. In the left panel the distance between an observed feature vector *x* and a set of exemplars are shown, and in the right panel the mathematical description of the exemplars and the exemplar-label association, H and G respectively, is visualized. The columns of H and G correspond to the exemplars in the left panel.

k-Nearest Neighbors

In order to extract useful information for speech recognition or classification exemplars are associated with label. (phone classes and word labels)
Matrix G is a binary matrix that associates each exemplar in H with class labels.

Classification made by maximum likelihood.

Sparse Representation classification

Test vector y, dictionary H with exemplars h_i from training set.
Where H is a $m \times N$ matrix where m is the dimension of each feature vector x and N is the total number of all training examples from all class es. (m<<N)

Then vector y can be written as linear combination of all training examples.

$$y = H\beta$$

β should be sparse with non-zero elements for the elements in H which belong to the same class.

Results: Sparse representation reconstruction features was tested on TIMIT database. Phonetic Error Rate 18.6%. 0.8% improved over a state-of-art GMM/HMM

Thank you