Using Independent Component Analysis to process near infrared hyperspectral images for detecting powder food adulteration

Puneet Mishra\textsuperscript{1}, Christophe B.Y. Cordella\textsuperscript{2}, Douglas N. Rutledge\textsuperscript{2}, Pilar Barreiro\textsuperscript{3}, Jean Michel Roger\textsuperscript{4}, Belén Diezma\textsuperscript{3}

\textsuperscript{1} Vision Lab, Campus Drie Eiken, University of Antwerp, Belgium
\textsuperscript{2} UMR Ingénierie Procédés Aliments, AgroParisTech, Inra, Université Paris-Saclay, 91300 Massy, France
\textsuperscript{3} LPF-Tagralia, Universidad Politécnica de Madrid, Spain
\textsuperscript{4} Irstea, UMR ITAP, Montpellier, France
Blind Source Separation

Different pure sources combined to make one mixed signal
Theory of ICA

• Each observed sensor signal is assumed to be weighted sum of pure source signals.

• Weighting coefficients are proportional to concentrations of pure compounds:

\[ x_1 = a_{11} \times s_1 + a_{12} \times s_2 \]
\[ x_2 = a_{21} \times s_1 + a_{22} \times s_2 \]

• In matrix notation:

\[ X = A \times S \quad \ldots \ (1) \]
Procedure

• ICA calculates a demixing matrix, $W$

• $W$ approximates $A^{-1}$, the inverse mixing matrix

• Pure component signals are recovered from measured mixed signals:

$$S = W \times X$$
Algorithm used

• Joint Approximation Diagonalization of Eigenmetrices (JADE) algorithm

• Based on fourth order moment (Kurtosis)

• Gaussian distributions possess zero excess kurtosis

• JADE seeks rotation of mixed vectors to estimate source vectors with high kurtosis values.
Deciding on number of ICs: Random ICAbyBlocks

Perform ICA by splitting data into two blocks and comparing correlations of ICs extracted from the two data blocks.
Peanut allergy

- Peanut allergy is a potentially life-threatening condition.

- European Directive 2003/89/EC makes the labeling of all ingredients mandatory, especially food allergens used in the recipes of packaged foods.

- Ubiquitous nature of peanut in food industry makes dietary avoidance difficult, a risk still persists.
Sample preparation

- Sample mixtures of peanut in wheat flour 0.05 % and 0.01 % by weight
- Size of the ground nuts 500-1000μm (EU-Institute for Reference Materials and Measurements)
- Particle size of wheat flour 100-212μm
- **Aluminum platform** for sample representation
Camera specification

- Line-scan push-broom camera: HySpex (SWIR 320m-e)
- Spectral range 1000–2500 nm
- Spectral sampling every 6 nm and 256 spectral bands.
- Pixel size 408 × 261 μm
Acquisition

Hyperspectral camera setup
Spectral profiles of pure samples

(a). Mean spectra

(b). Spectra after treatment (SNV)
Random ICA by blocks

- Two blocks
- Performed up to 20 ICs
- With 10 repetitions
- 7 ICs had high correlations

Red: High correlation  Blue: Low correlation
IC signals

- IC1 related to non-chemical variation
- At every peak in pure spectra, IC1 is tending to zero
- Bigger particle size of peanut could be a reason
IC2 and IC7 related to starch

(a). Starch (2100 nm)  

(b). OH stretching Starch (1580 nm)

• Wheat flour has higher starch
IC3 and IC6 related to moisture

(a). OH stretching (1940 nm)
(b). OH stretching (1450 nm)

- Wheat flour has higher moisture
IC4 and IC5 related to fatty acid

(a). Amide function (2030 nm)

(b). Fatty acid and overtone (1734, 1395, 1200 nm)

• Peanut has higher fatty acid and amide
Synthetic Unmixing Signal

Calculate the difference between the sum of peanut ICs and the sum of wheat flour ICs

\[
\text{Signal (S)} = IC1 + IC4 + IC5 - (IC2 + IC3 + IC6 + IC7)
\]
Unmixing of HSI

\[ X = \text{Unfolded hyperspectral image} \]

\[ S = \text{Synthetic unmixing signal} \]

\[ S^T = \text{Transpose of synthetic signal} \]

\[ A = \text{Proportions values} \]
A hyperspectral image with known position of peanut was simulated.

Synthetic signal was tested for classification.

High contrast was obtained for pixels representing peanuts.
Image segmentation

• **Connected Component labeling** to detect pixels with enhanced contrast.

• Classification map generated with manual threshold.

• **Regionprop** function in Matlab was used to extract spatial locations
Results

Before processing (0.05 %)
Proportions images and extracted features image for 0.05% peanut traces in wheat flour (146 × 464 pixels).
Before processing (0.01 %)
After processing

Proportions images and extracted features image for 0.01% peanut traces in wheat flour (146 × 464 pixels).
Conclusions

• Detection of peanut traces was possible down to 0.01 %.

• ICA provided an easy understanding of underlying source signals

• Source signals can be easily used for classification and regression analysis.

• HSI with ICA can be used for quantitative prediction of the chemical constituents, with simultaneous representation of their spatial distribution.
Thank you

Peanuts free zone!