Mineral identification in hyperspectral imaging using Sparse-PCA

Bardia Yousefi
Clemente Ibarra-Castanedo
François Huot
Georges Beaudoin
Xavier P. V. Maldague
Martin Chamberland

Canada Research Chair in Multipolar Infrared Vision – Vision Infrarouge Multipolaire: MiViM
Département de génie électrique et de génie informatique- Université Laval, Quebec, Canada

Chaire de recherche industrielle CRSNG - Agnico Eagle en exploration minérale
Département de géologie et de génie géologique - Université Laval, Quebec, Canada

TELOPS, Inc., 100-2600 St-Jean-Baptiste Ave, Quebec, Qc, G2E 6J5, Canada

OUTLINE

HYPERSPECTRAL IMAGING

PROBLEM STATEMENT

OBJECTIVE OF THE PROJECT

THE STEPS OF PROJECT

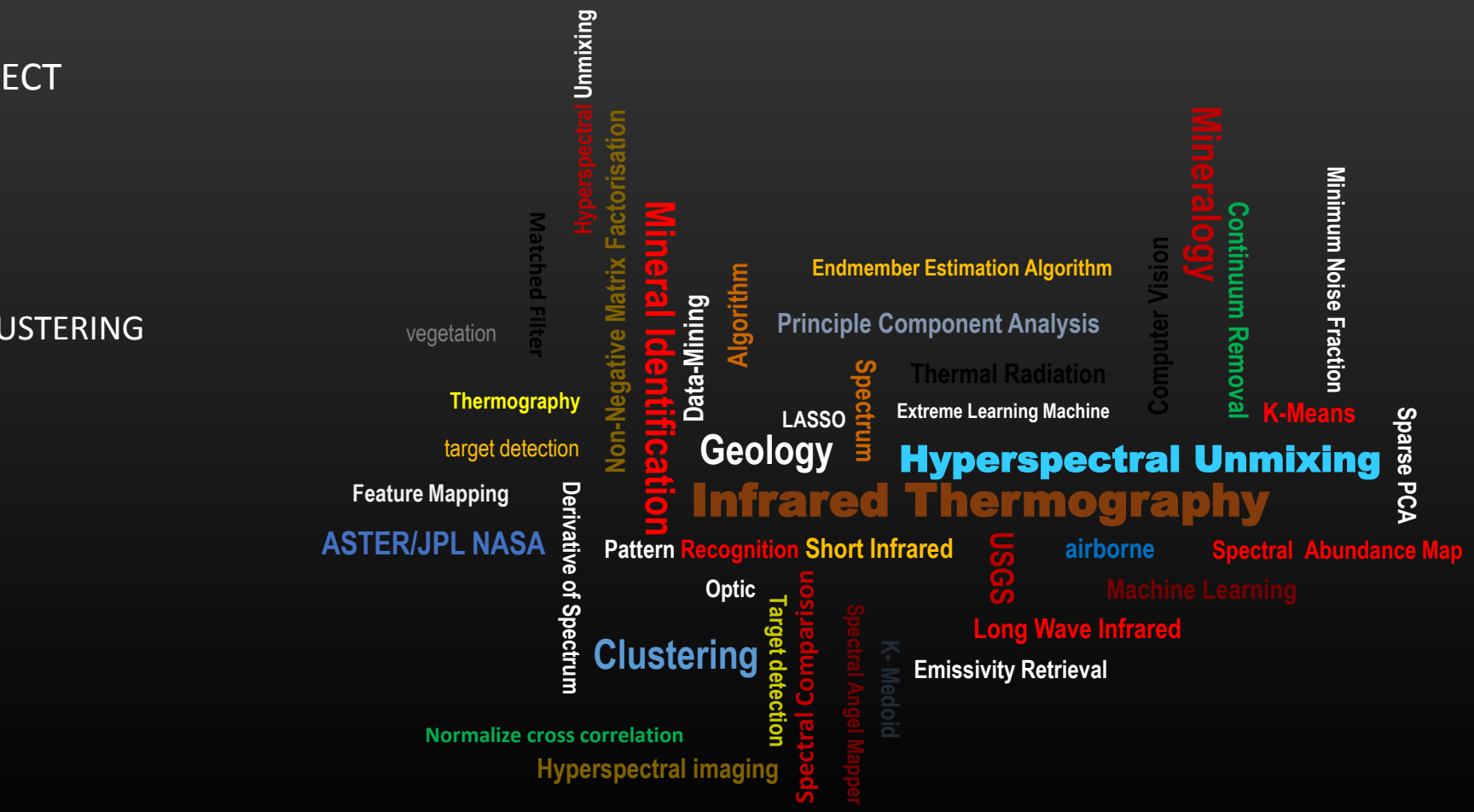
MEASUREMENT

DATA - MINING

TWO ALGORITHMS FOR CLUSTERING

RESULT OF CLUSTERING

CONCLUSION

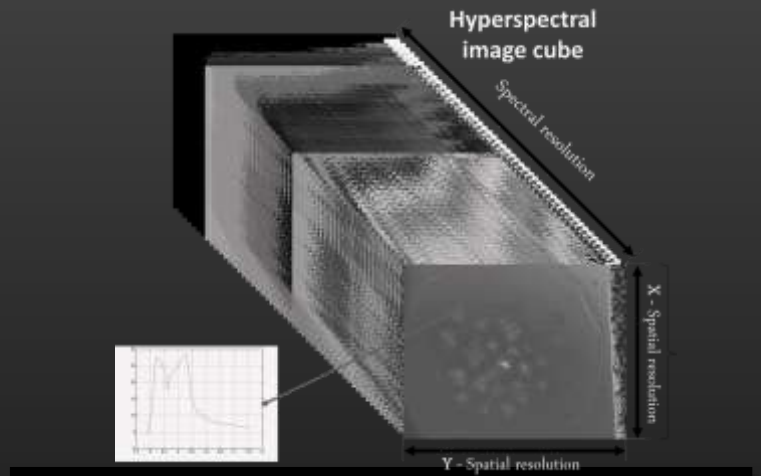


HYPERSPECTRAL IMAGING

Hyperspectral imaging, like other spectral imaging, collects and processes information from across the electromagnetic spectrum. The goal of hyperspectral imaging is to obtain the spectrum for each pixel in the image of a scene, with the purpose of finding objects, identifying materials, or detecting processes.

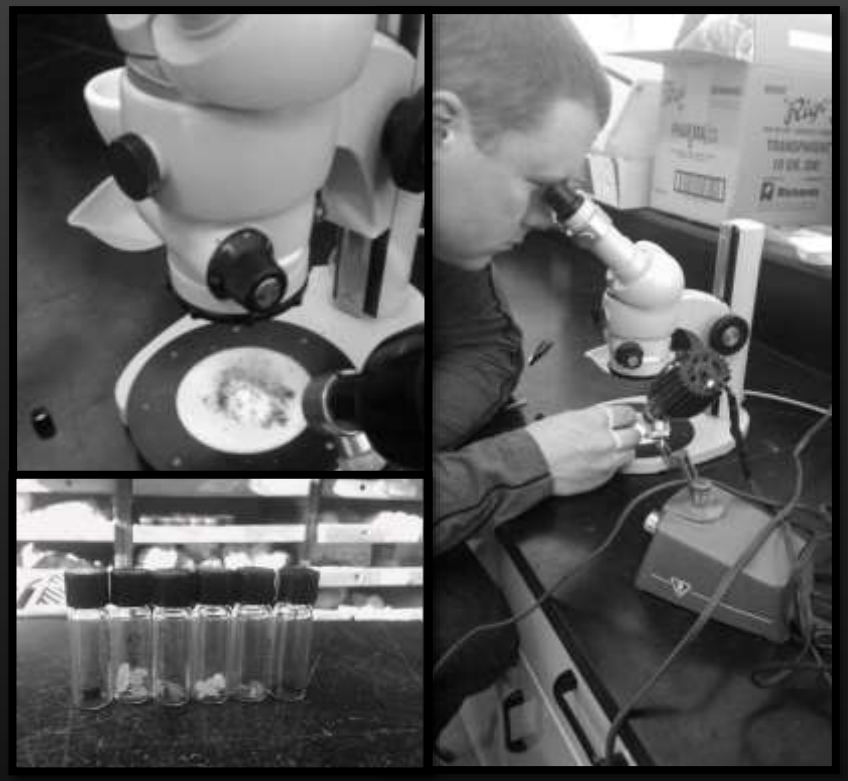
Applications

- vegetation
- urban area
- geology
- target detection
- medical
- etc



northropgrumman.com

PROBLEM STATEMENT



.Mineral identification is challenging in the field of geology and mineralogy.

.It relates to ***geological research*** and is usually conducted by geologists (mineralogy experts).

.It is extensively investigated by hyperspectral **remote *sensing (and airborne) sensing*** and has been the subject of many research studies.

.Manual identification of mineral samples by a ***mineralogy expert*** which is a ***time consuming*** process and provides high level of ***disparity*** due to fatigue or inadequate methods for this specific application.

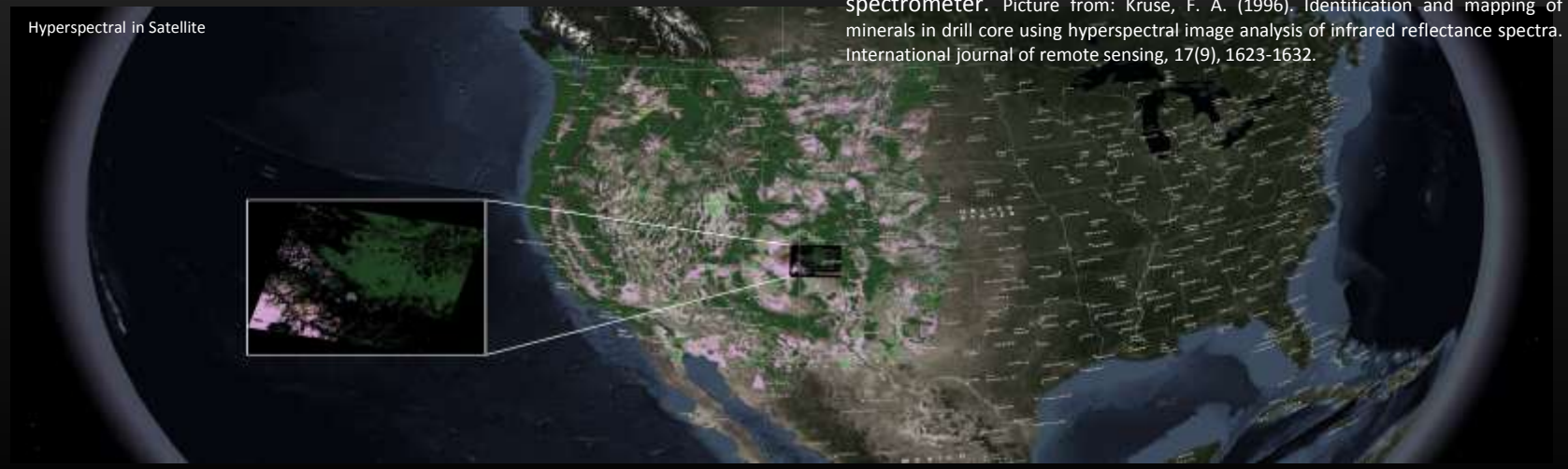
OBJECTIVE OF THE PROJECT

The main objective of this research is to automatic identification of mineral using hyperspectral imageries in laboratory(indoor) situation.

Making automated system for mineral identification

It use(d/s) for long time but most of the researches in this area are for remote sensing, airborne, astronomy applications. (e.g. Only in 2015 the geological based hyperspectral infrared imagery (contributions) papers were more 100>)

Photograph showing measurements being made with the PIMA II field spectrometer. Picture from: Kruse, F. A. (1996). Identification and mapping of minerals in drill core using hyperspectral image analysis of infrared reflectance spectra. International journal of remote sensing, 17(9), 1623-1632.



THE STEPS OF PROJECT

To identify the minerals in hyperspectral images, there are several should be considered from mineral preparation to data-mining. These steps are summarized in two general categories:

Measurement

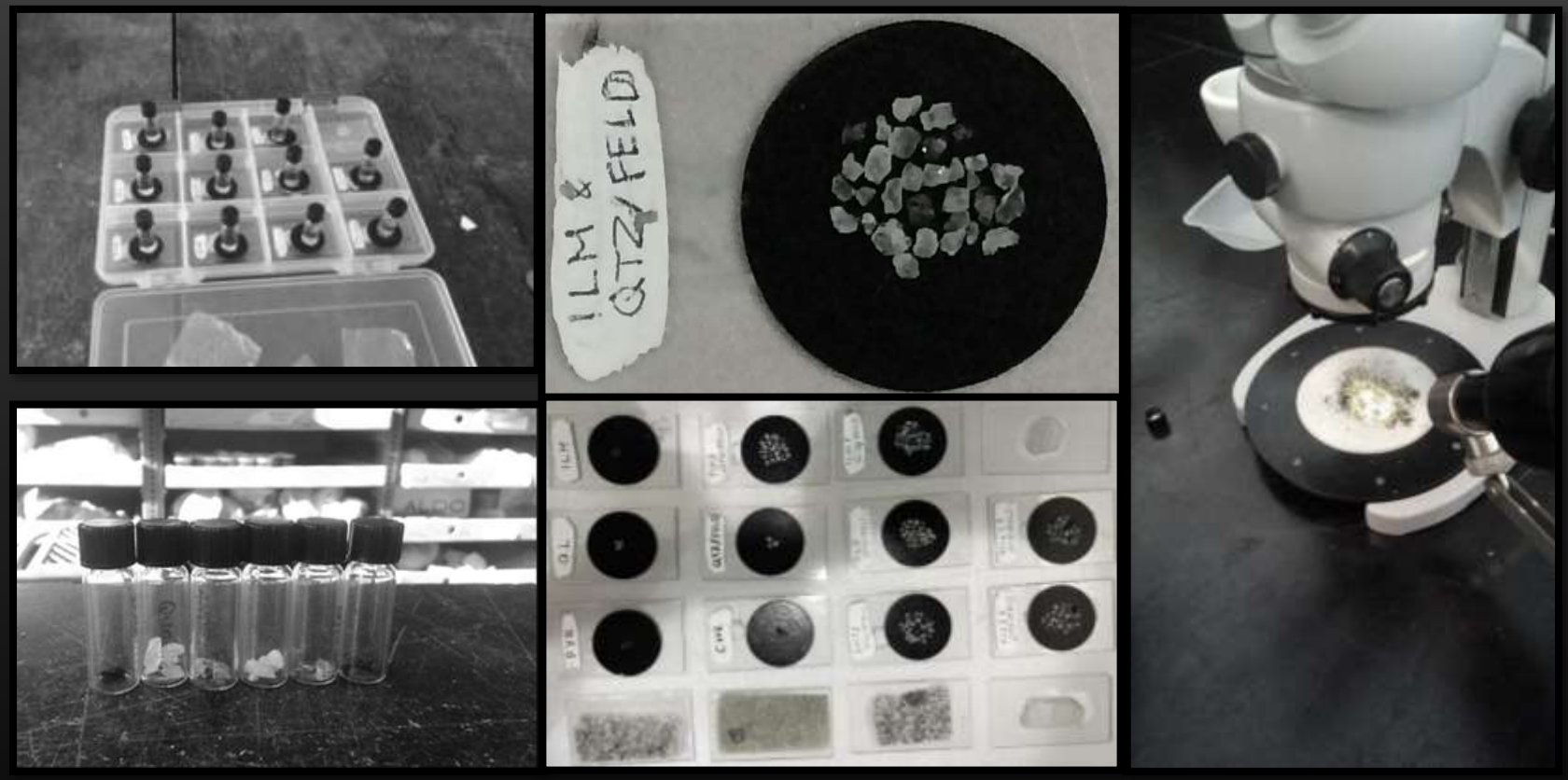
- Mineral preparation
- Conducting the experiments
- Finding the reference datasets

Data-mining

- Data retrieval (Continuum removal)
- Extraction of the spectral difference
- Clustering (without training) / Classification (with training)

MEASUREMENT

Mineral preparation



Illumemite Pyrope Quartz Olivine Chromite Biotite Diopside Epidote Goethite Gahnite Titanite Marcasite Chalcopyrite Sphalerite Pyrite Pentlandite Cassiterite Fluorite Kynite Scheelite Smithsonite Tourmaline Wolframite Ferroaxinite Sillimanite

MEASUREMENT

Conducting the experiments

7.7-11.8 µm- Long Wave Infrared (LWIR)

Imageries are done at Telops Inc. 02 April 2015

LW MACRO LENS

New accessory for “microscope-type” measurements

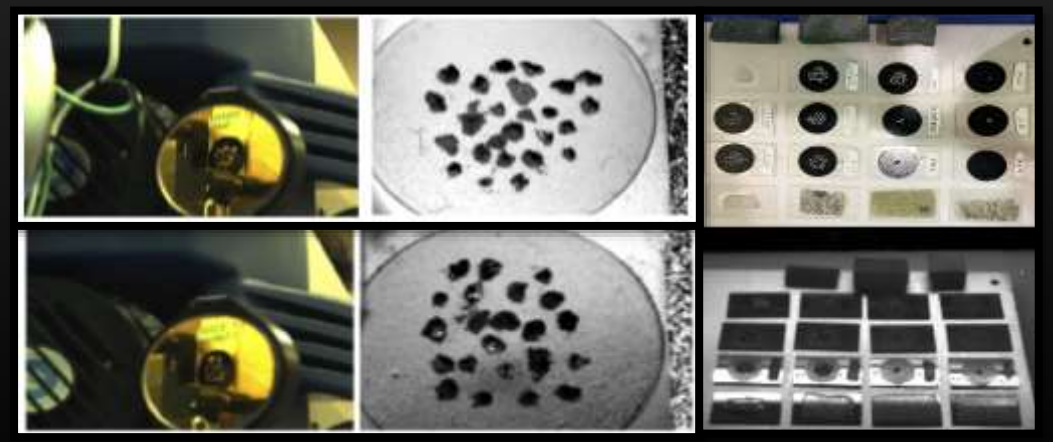
.Provides a pixel footprint of 0.1mm

.Working distance: 30cm (from the lens)

.Easy installation by the user

.Replaces the entrance window similar to the 0.25x telescope

.Full radiometric calibration



After Before

M E A S U R E M E N T

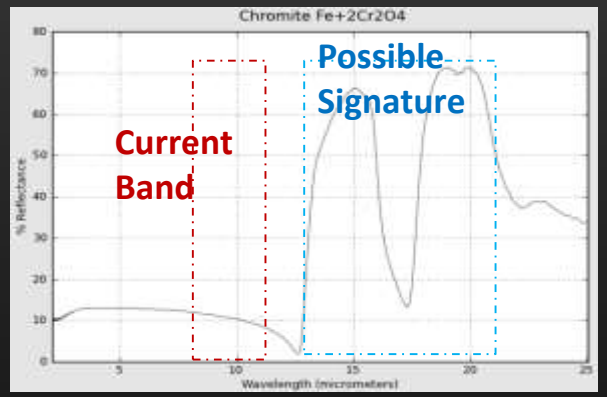
Finding the reference datasets

.ASTER Spectral Library - Version 2.0

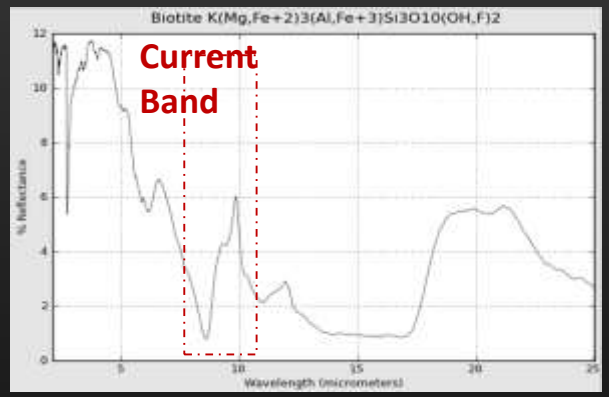
.Calculating the reference spectra from the pure minerals

Minerals do not have signature in LWIR

Illumemite
Chromite
Gahnite Titanite
Marcasite
Chalcopyrite
Sphalerite
Pyrite
Pentlandite
Cassiterite
Fluorite
Wolframite
Ferroaxinite
Sillimanite



Picture is taken from ASTER spectral library



Picture is taken from ASTER spectral library

Minerals have signature in LWIR

Pyrope
Quartz
Olivine
Biotite
Diopside
Epidote
Kynite
Scheelite
Smithsonite
Tourmaline
Goethite

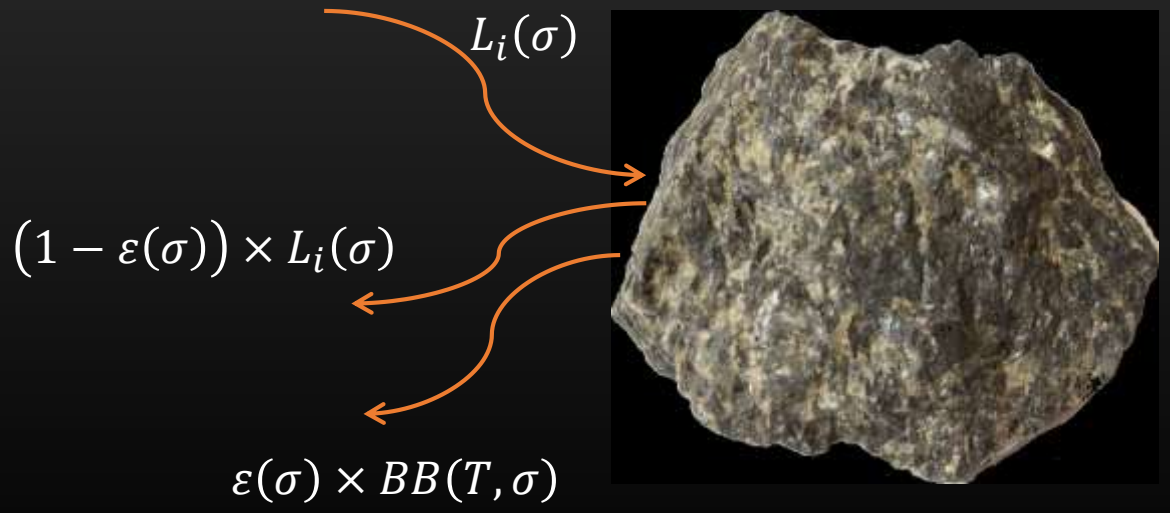
MEASUREMENT PRINCIPLE

Spectral Reflectance from 2 Measurements

$$L_m(\sigma) = \varepsilon(\sigma) \times BB(T,\sigma) + \big(1 - \varepsilon(\sigma)\big) \times L_i(\sigma)$$

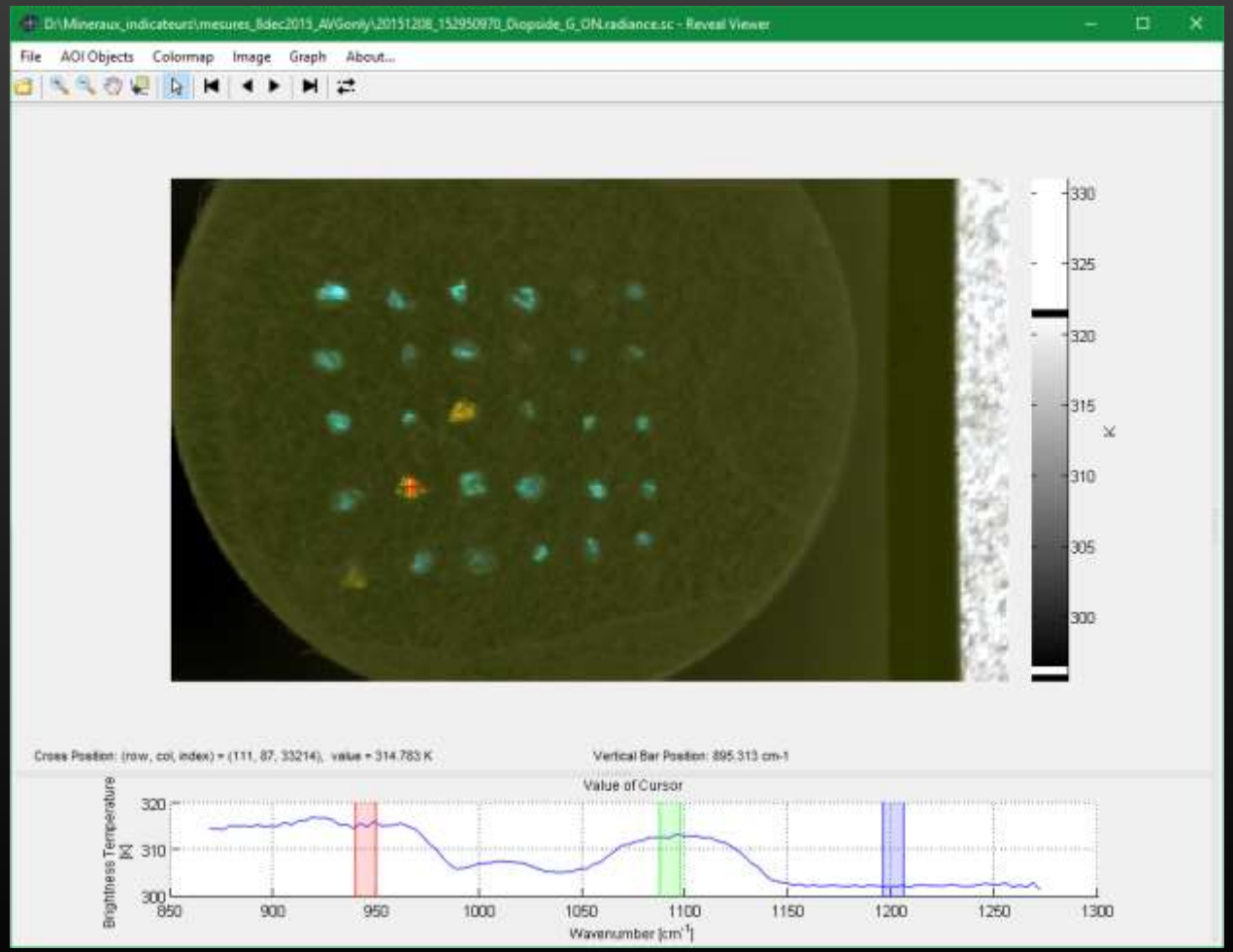
$$L_{m_{ON}}(\sigma) - L_{m_{OFF}}(\sigma) = \big(1 - \varepsilon(\sigma)\big) \times \big(L_{iON}(\sigma) - L_{iOFF}(\sigma)\big)$$

$$\big(1 - \varepsilon(\sigma)\big) = R(\sigma) = \frac{L_{m_{ON}}(\sigma) - L_{m_{OFF}}(\sigma)}{L_{i_{ON}}(\sigma) - L_{i_{OFF}}(\sigma)}$$



MEASUREMENT PRINCIPLE

Spectral Reflectance from 2 Measurements



DATA - MINING

Data retrieval (Continuum removal)

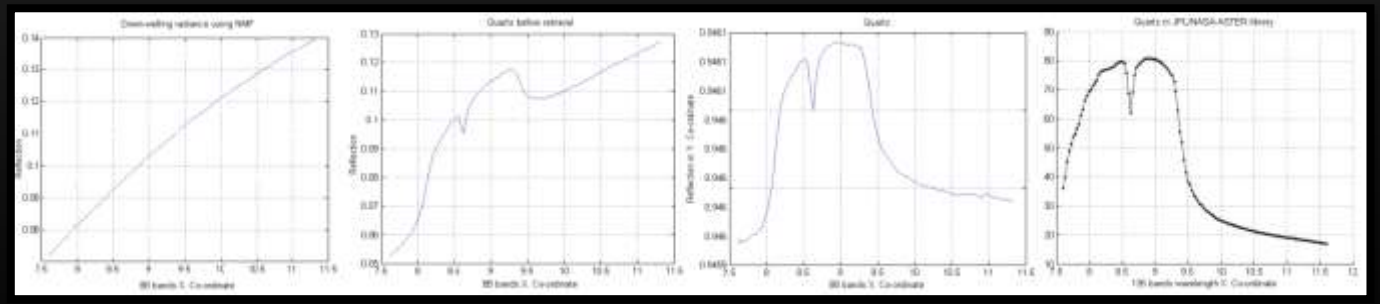
Having the continuum (background reflection) on the spectra make comparison impossible.
For removing this continuum, we require the reflection from the sample (we obtain it from infragold plate in our experiments)

The experiment is conducted for heating source ON and OFF. The continuum removed radiance is calculated by following formula:

$$\dot{R} = \frac{L_{m_{ON}} - L_{m_{OFF}}}{L_{i_{ON}} - L_{i_{OFF}}}$$

Where $L_{m_{ON}}$ and $L_{m_{OFF}}$ are measured radiance while the heating source is ON and OFF, respectively. L_i is the radiation of Infragold while the heating source is ON , $L_{i_{ON}}$, and when it is OFF , $L_{i_{OFF}}$ and $\dot{L}$ represents the continuum removed spectra.
L_i is a single spectrum obtained from the infragold region. For selecting the best representative spectrum for L_i , we used Non-negative Matrix Factorization(NMF).

Infragold plate



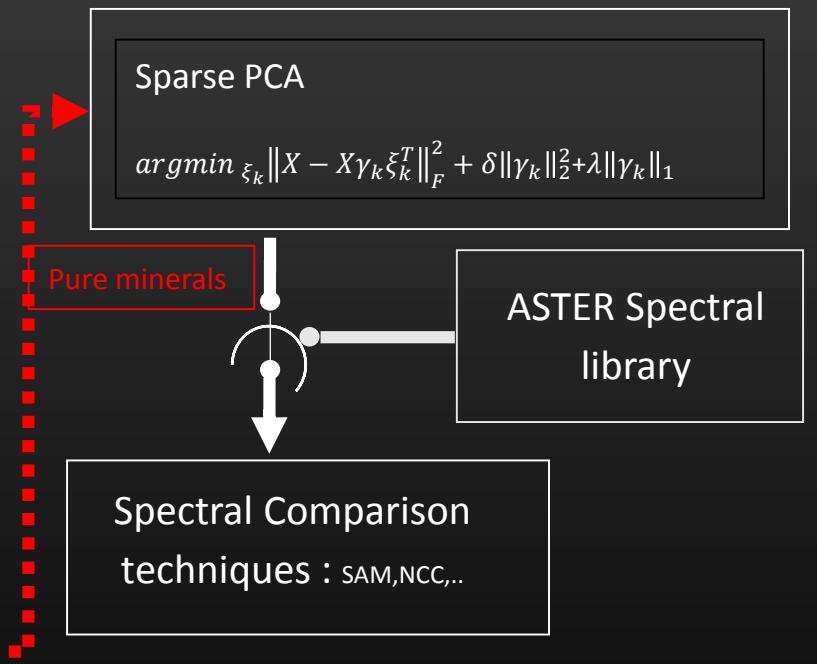
DATA - MINING

Extraction of the spectral difference

Some Spectral comparison techniques :

. Spectral Angle Mapper (SAM)
. Spectral information divergence (SID)
. Normalized cross Correlation (NCC)
. Matched Filter (MF)
. First Derivative (FD) of the spectrum
. Fast Fourier Transform (FFT)
. Orthogonal subspace projection (OSP)
. Adaptive matched subspace detector (AMSD)

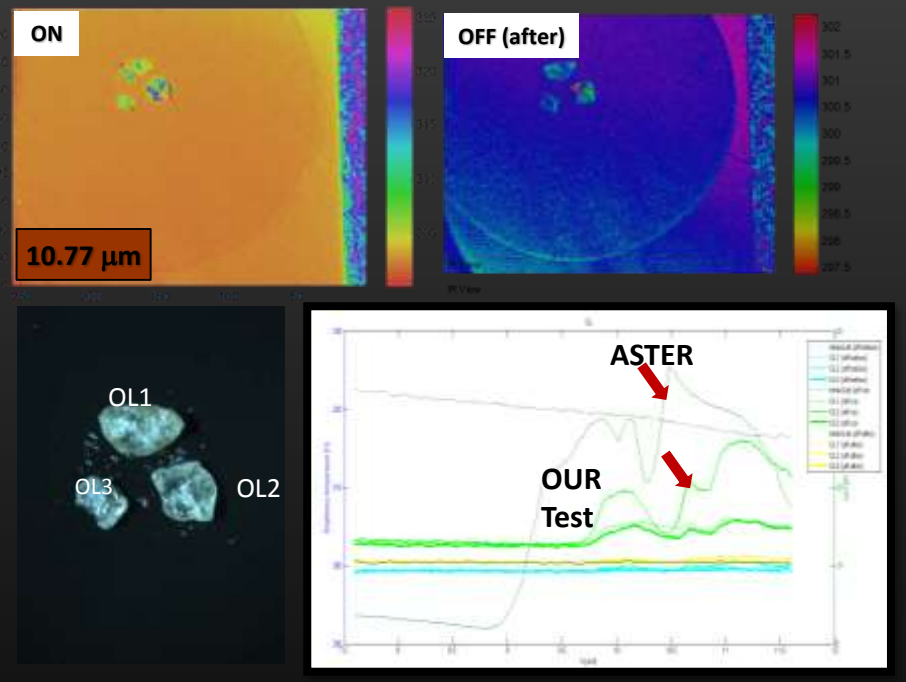
We have tested other way of having reference spectrum applying Sparse Principal Component Analysis (SPCA) to the spectra of pure minerals. But it is noticeable that there is an identical similarity between using SPCA or ASTER spectra as reference.



DATA - MINING

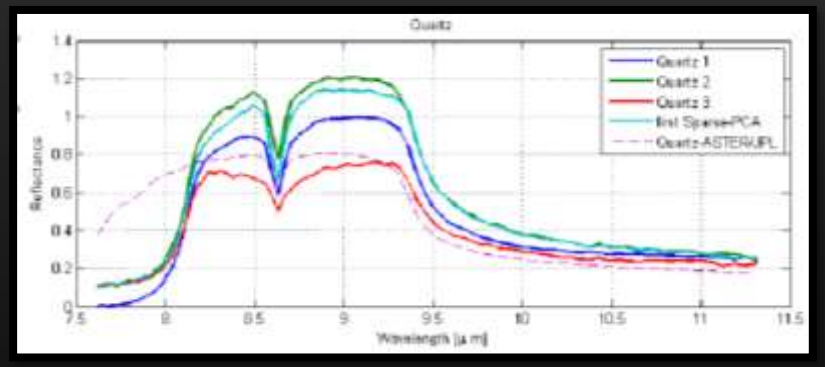
Results of Data retrieval (Continuum removal)

One example of Results of Data retrieval (Continuum removal):



FileName = '20150402_211927875_OL_ON.radiance.sc';

Spectral Comparison Technique	Quartz	Quartz Similarity (%)	Olivine	Olivine Similarity (%)	Pyrope	Pyrope Similarity (%)
SAM	0.483	51.70	0.269	73.02	0.233	76.70
NCC	0.605	60.55	0.933	93.36	0.952	95.20

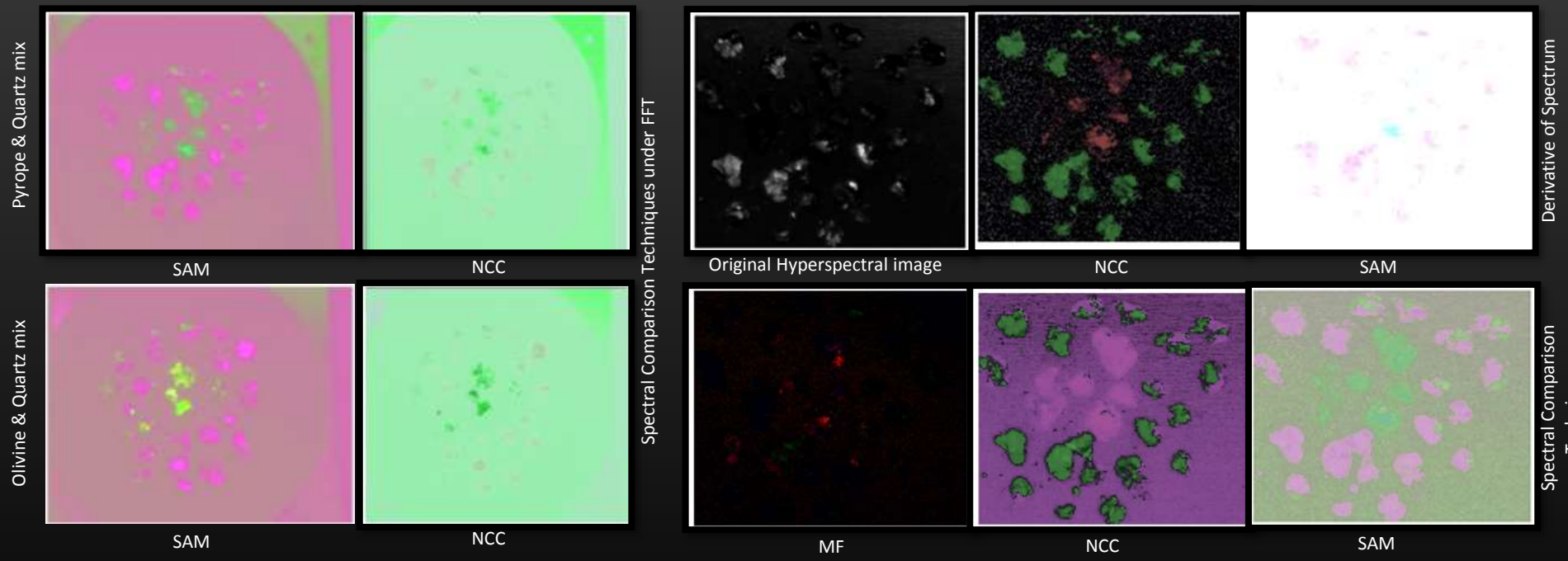


One example of Reference spectrum applying SPCA and compare to ASTER spectrum.

DATA - MINING

Extraction of the spectral difference- False colors

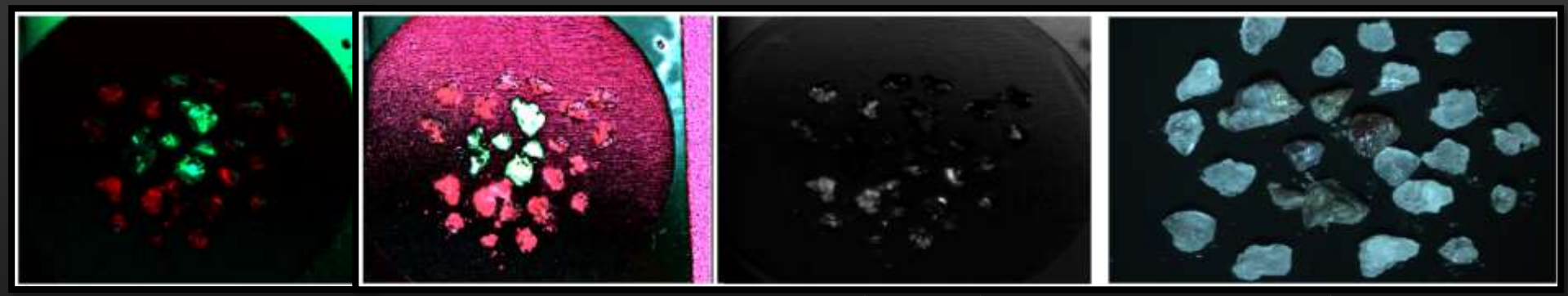
Applying the spectral comparison techniques that can be represented by false colors. It helps to discriminate the mineral contents from eachother.



DATA - MINING

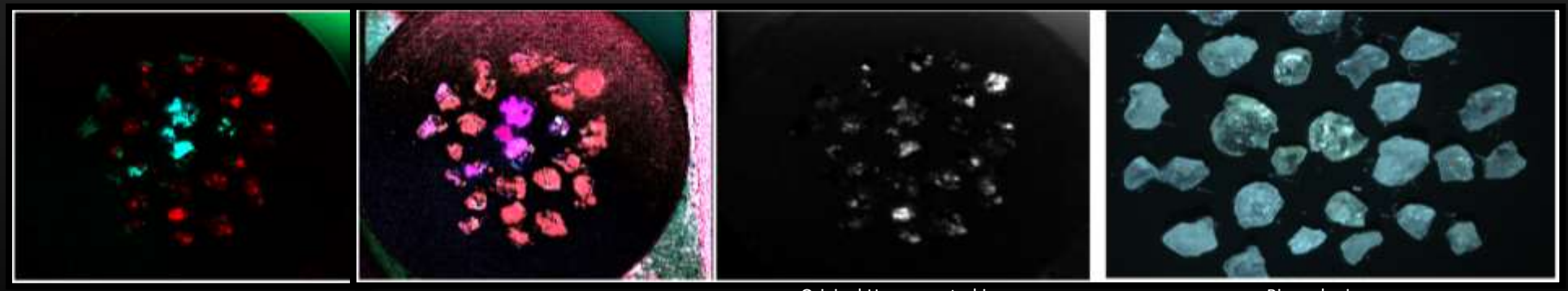
Extraction of the spectral difference- False colors

Orthogonal subspace projection (OSP) & Adaptive matched subspace detector (AMSD) algorithm



OSP | AMSD | Original Hyperspectral image | Binocular image

Pyrope & Quartz mix

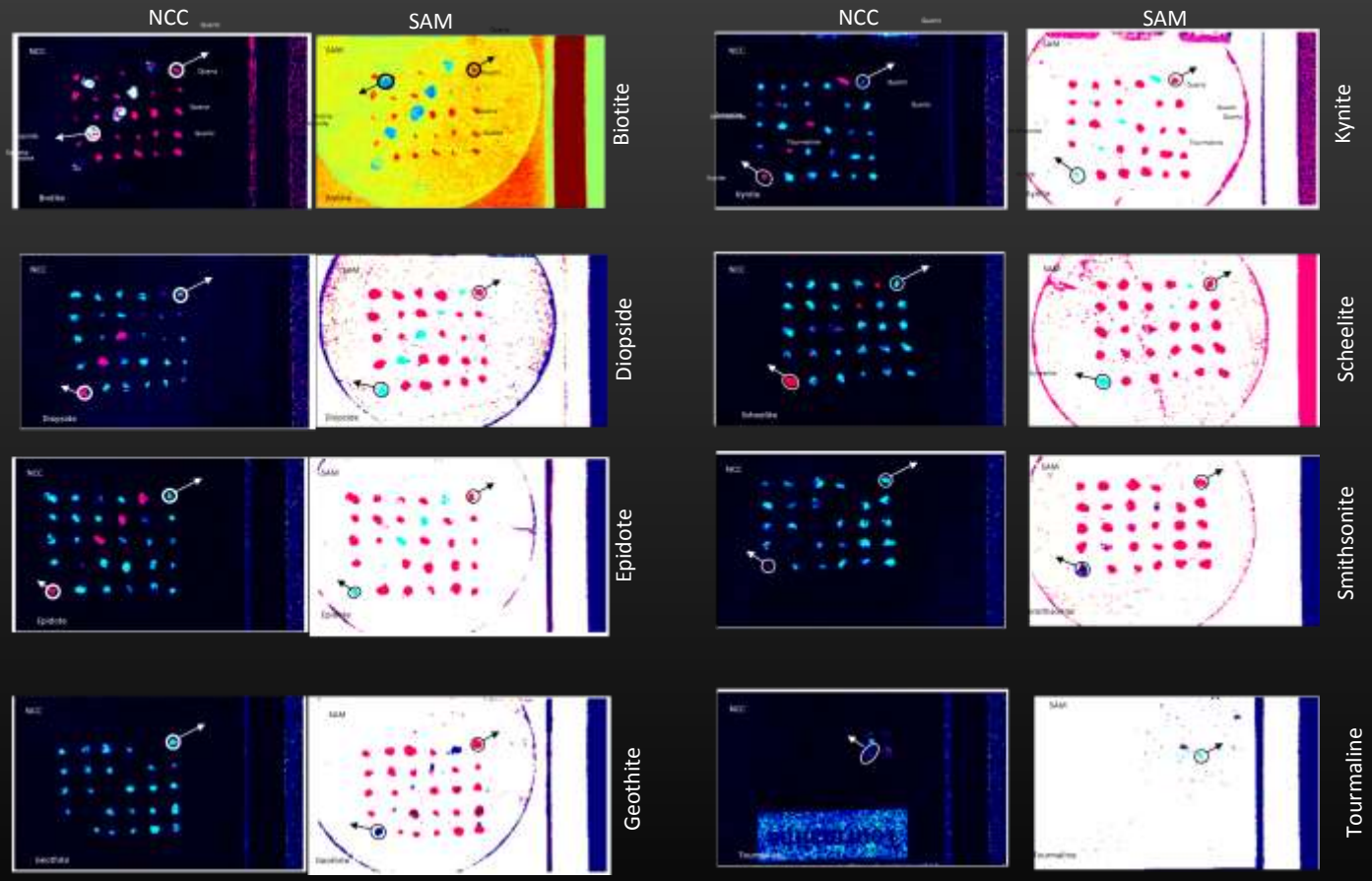


OSP | AMSD | Original Hyperspectral image | Binocular image

Olivine & Quartz mix

DATA - MINING

Extraction of the spectral difference- False colors

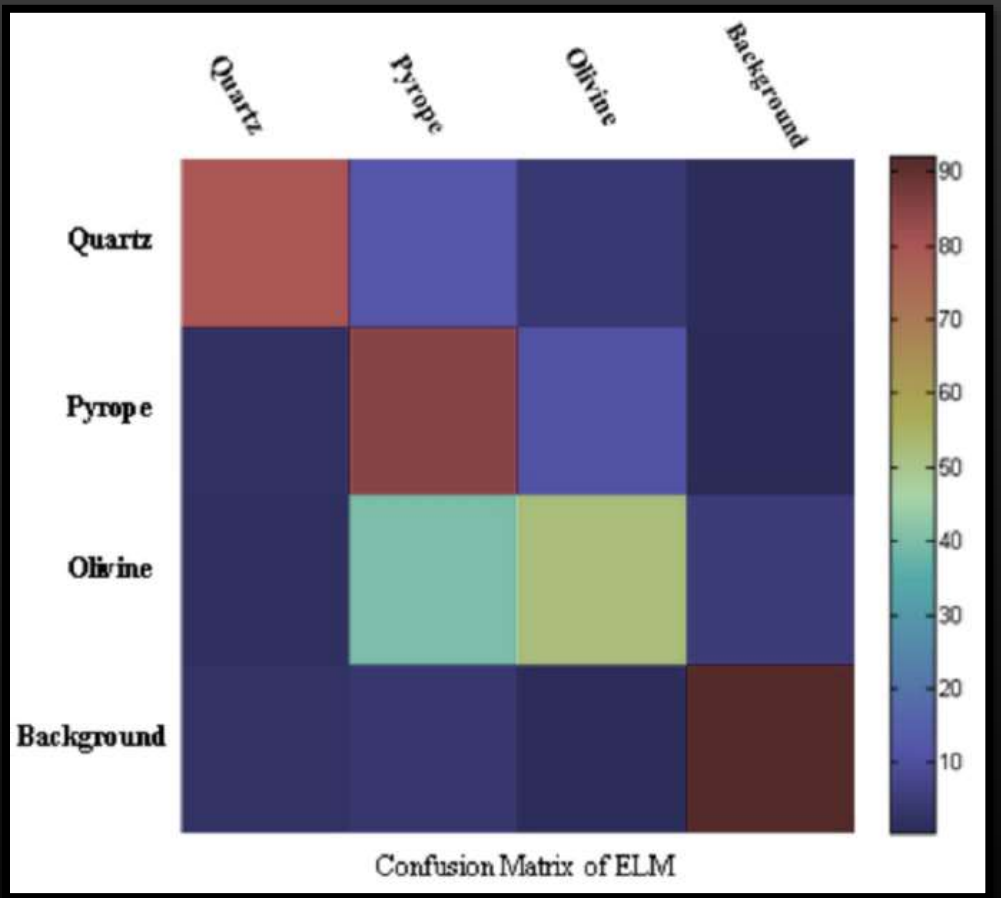


DATA - MINING

Clustering (without training) / Classification (with training)

Clustering or Classification? Which one is more substantial for automatic identification of mienrals. For answering this question, we made a system to classify and compare it with clustering.
Classification is very much dependent on the training stage. There are limited number of minerals exist for training.

The confusion matrix of classification applying Extreme Learning Machine(ELM) is showed. ELM is a single hidden layer neural network.
For training, arround 3000 pixels-spectra have been used having 3 attributes amounts (SAM,SID,NCC). Classification is not a good approach for automatic identification of minerals due to diffeculties in the training of networks. Also it is more expensive than unsupervised approaches because of memory usage and having dependency to the training stage.



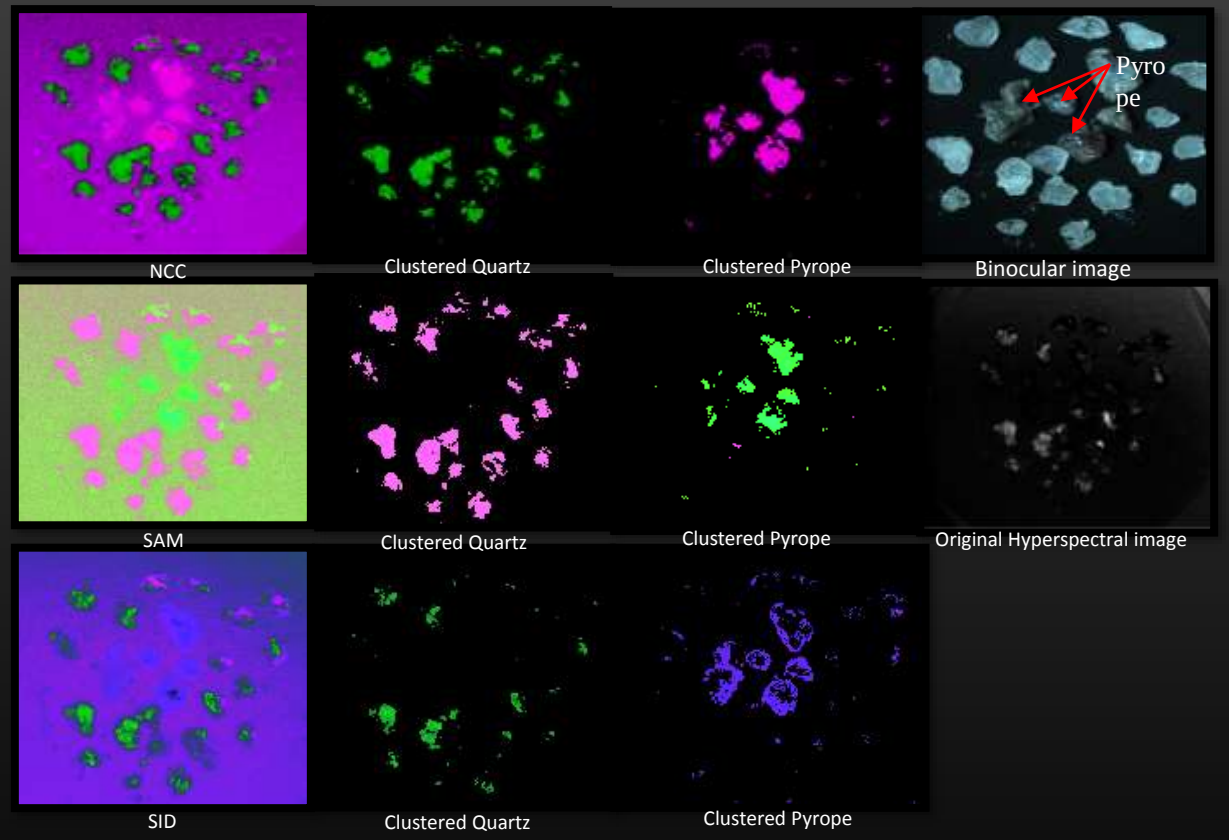
DATA - MINING

Clustering (without training)

The clustering applied for identification of the mineral and the results depicted in the figure.

A color based K-means approach have been used for clustering the false colors.

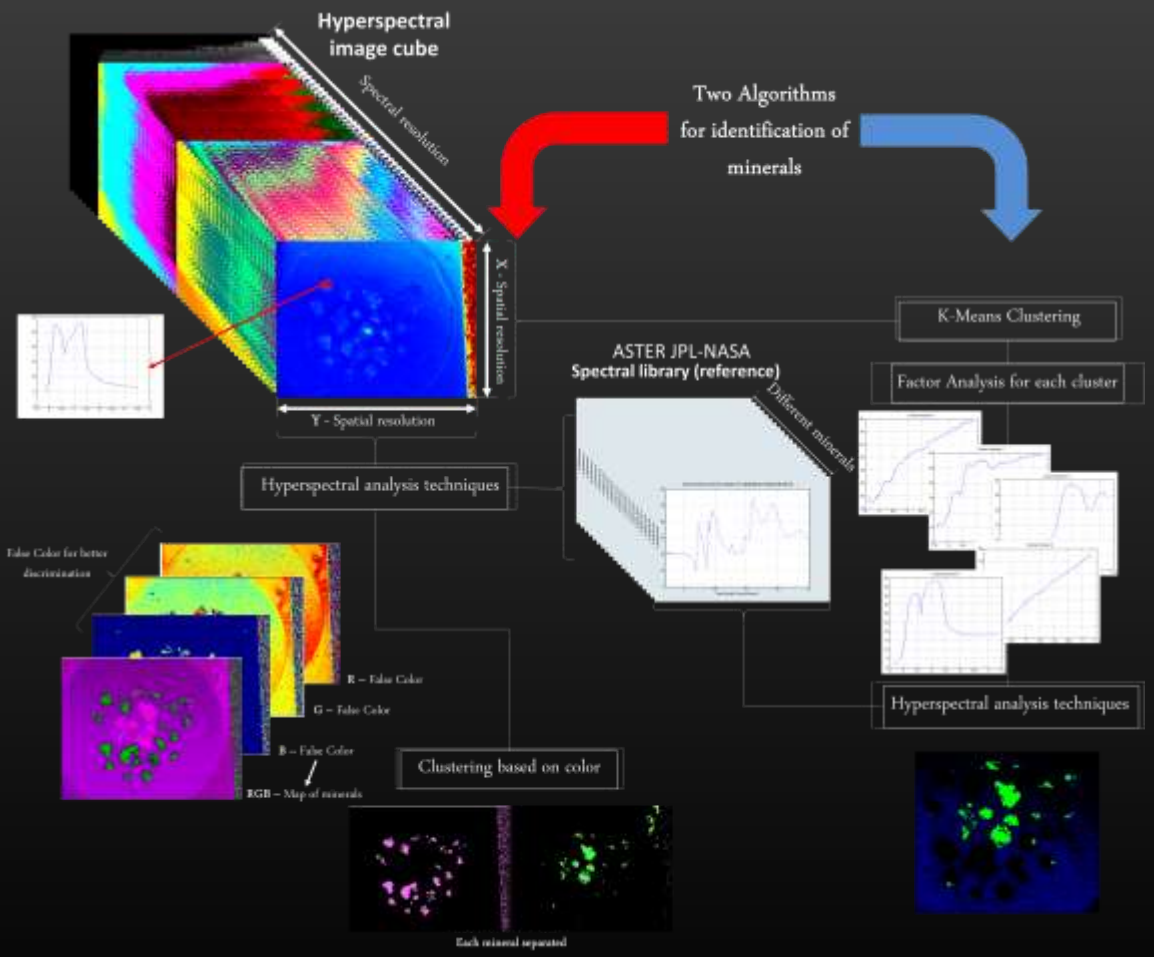
It is can be performed on any data and it does not require training stage (it just needs labeling part).
. It is fast.
. It does not occupied much of the memory (for training) .
. It is easier to use as compared to classification approach.



Pyrope & Quartz Mixture

TWO ALGORITHMS FOR CLUSTERING

We found that the clustering is a solution for idnetification of the minerals. Here there are two ways to identify the minerals:



TWO ALGORITHMS FOR CLUSTERING

There are two algorithms corresponding to two different ways for indetification of the minerals on the false colors.

Algorithm 1: **First applying the spectral comparison techniques then clustering the false colors.**

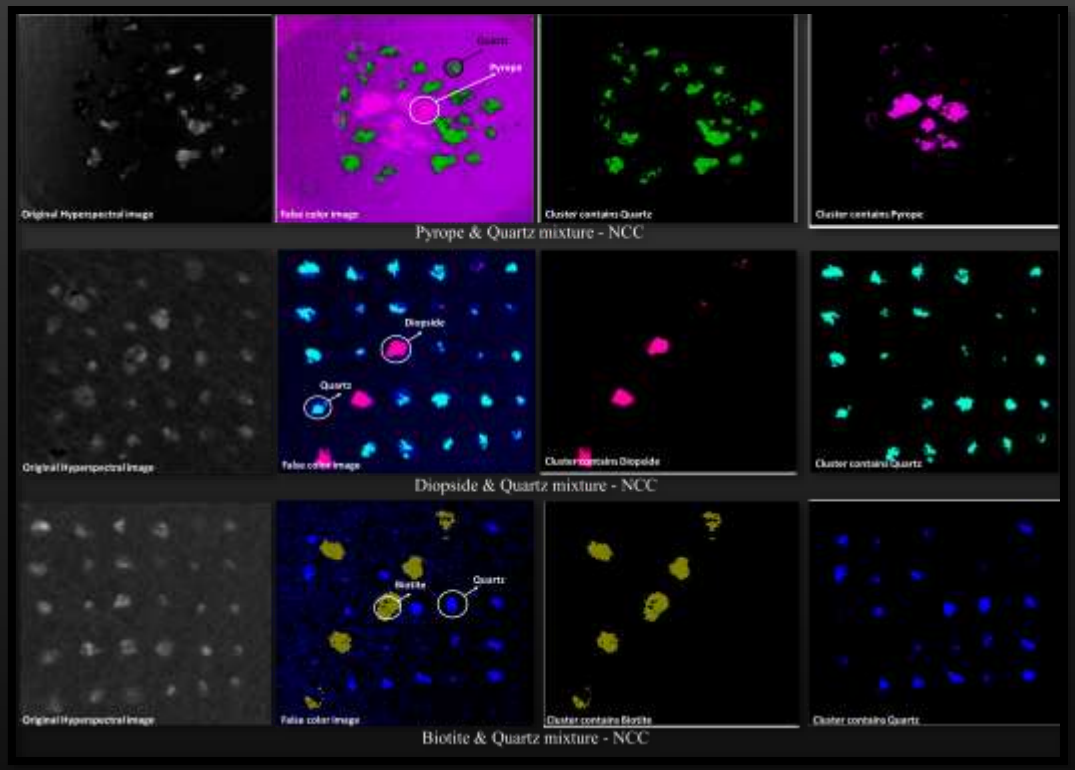
Algorithm I	
Given	Input data I(x,y,z) ∈ $\boldsymbol{R^{N\times M\times Z}}$ which I(x,y) ∈ $\boldsymbol{R^{N\times M}}$ is spatial dimension selected as Region of Interest(RoI) and its unit is pixels, z is spectral resolution and depends on hyperspectral camera and acquisition property.
Step 1	Calculation of reflectance of RoI $\boldsymbol{R(x,y)=\frac{I_{DW_{ON}}(x,y)-I_{DW_{OFF}}(x,y)}{\acute{I}_{IG_{ON}}-\acute{I}_{IG_{OFF}}}}$. $\boldsymbol{\acute{I}_{IG_{ON}},\acute{I}_{IG_{OFF}}}$ are radiance of Infragold while heating source is on and off, respectively (they are constant spectra). $\boldsymbol{I_{DW_{ON}}(x,y),I_{DW_{OFF}}(x,y)}$ are also two down-welling radiation in heating source on and off the in $\boldsymbol{x,y}$ spatial position.
Step 2	Calculation of the spectral comparison techniques: $\boldsymbol{\mathrm{M}_i\,(x,y)=ST_j(R(x,y),\Phi_i)}$ $\boldsymbol{ST_j}$ is represents the spectral techniques and j reveals the number of techniques exploid. $\boldsymbol{\Phi_i\in R^z}$ is shown the reference spectra (i.e. ASTER/JPL) and i is number of spectra.
Step 3	False color generation (Ψ_{RGB}), dependent on i in $\boldsymbol{\Phi_i}$. For every targeted spectrum $\boldsymbol{(R,G,B)}$ amounts would change.
Output	Clustering Ψ_{RGB} to obtain the clusters where reveal the found mineral grains. The output will be $\boldsymbol{C_J}$. C shows the clusters and J representing cluster (J) which related to targeted mineral grains ($\boldsymbol{0\le J\le k}$, k is total number of clusters).

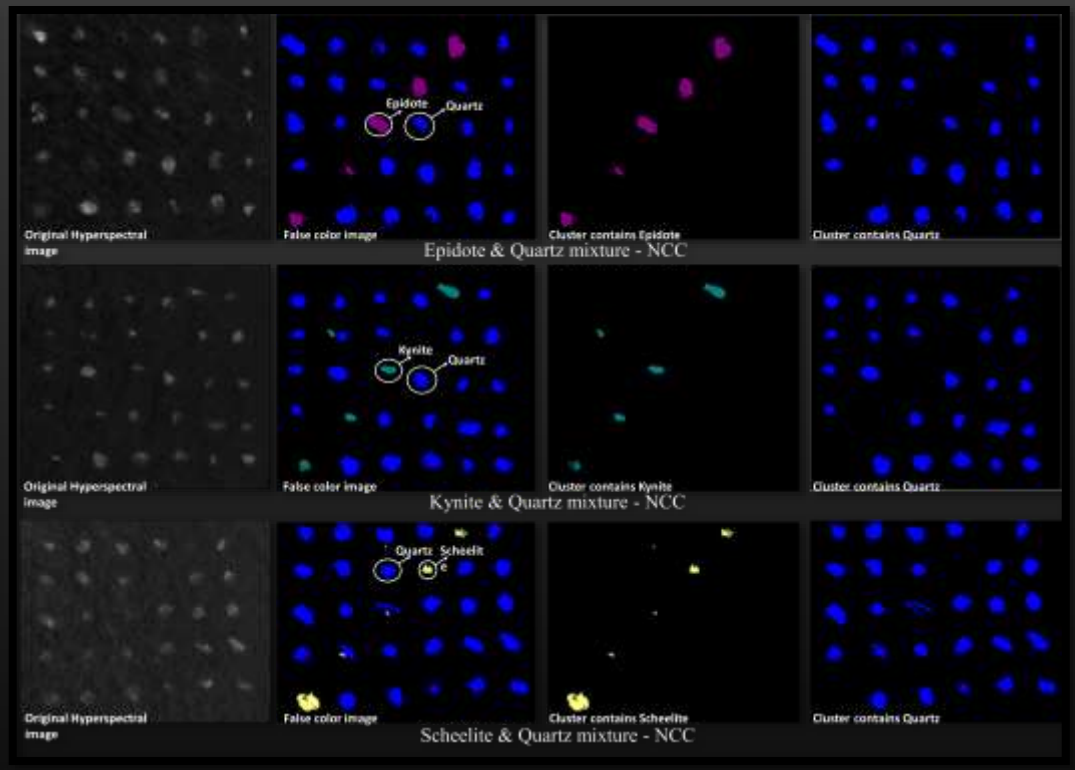
Algorithm 2: **First clustering then spectral comparison.**

Algorithm II	
Given	Input data I(x,y,z) ∈ $\boldsymbol{R^{N\times M\times Z}}$ which I(x,y) ∈ $\boldsymbol{R^{N\times M}}$ is spatial dimension selected as Region of Interest(RoI) and its unit is pixels, z is spectral resolution and depends on hyperspectral camera and acquisition property.
Step 1	Calculation of reflectance of RoI $\boldsymbol{R(x,y)=\frac{I_{DW_{ON}}(x,y)-I_{DW_{OFF}}(x,y)}{\acute{I}_{IG_{ON}}-\acute{I}_{IG_{OFF}}}}$. $\boldsymbol{\acute{I}_{IG_{ON}},\acute{I}_{IG_{OFF}}}$ are radiance of Infragold while heating source is on and off, respectively (they are constant spectra). $\boldsymbol{I_{DW_{ON}}(x,y),I_{DW_{OFF}}(x,y)}$ are also two down-welling radiation in heating source on and off the in $\boldsymbol{x,y}$ spatial position.
Step 2	Clustering $\boldsymbol{I_{(v,z)}}$ into k categories shown by $\boldsymbol{C_k}$, $\boldsymbol{v\in R^V}$ is spatial a resolution in the vector form. Clustering based on spectral different among the clusters ($\boldsymbol{0\le J\le k}$). $\boldsymbol{\dot{C}_k}$ is the best representative of the cluster applying NMF.
Step 3	Calculation of the spectral comparison techniques: $\boldsymbol{\mathrm{M}_i\,(x,y)=ST_j(\dot{C}^s{}_k,\Phi_i)}$ $\boldsymbol{ST_j}$ is represents the spectral techniques and $\boldsymbol{j}$ reveals the number of techniques exploid (e.g. $\boldsymbol{\mathrm{j=1\rightarrow MAM_1=NCC}}$). s represents a sample from the cluster k. $\boldsymbol{\Phi_i\in R^z}$ illustrates the reference spectra (i.e. ASTER/JPL) and *i* is number of spectra.
Output	False color generation (Ψ_{RGB}), dependent on i in$\boldsymbol{\Phi_i}$. For every targeted spectrum, the amount of $\boldsymbol{(R,G,B)}$ would be changed. The output will be, Ψ_{RGB}, an image whose has remarked the materials by different color.

RESULT OF CLUSTERING

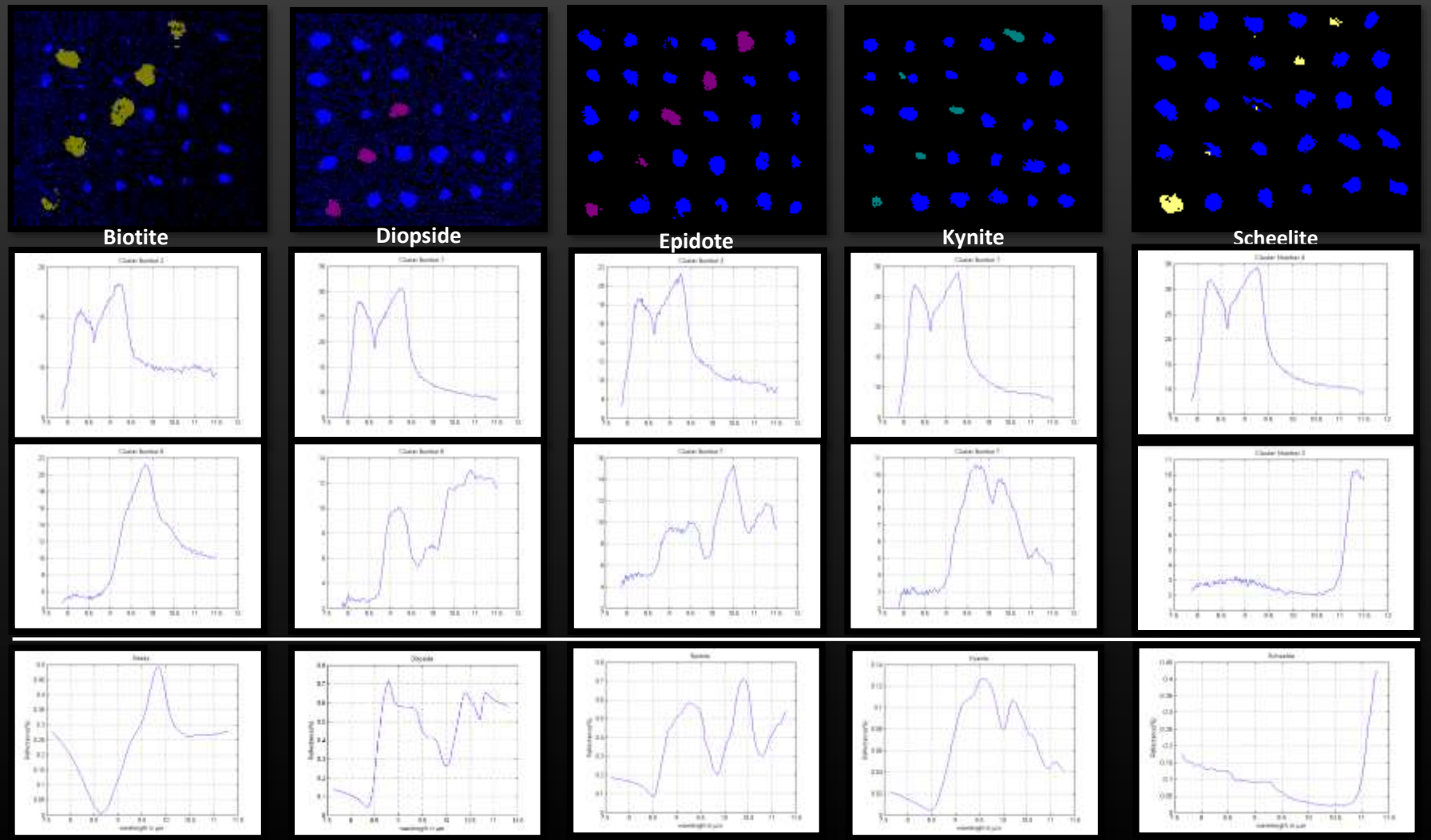
Comparison of two algorithms for clustering: Qualitative comparison **Algorithm 1**





RESULT OF CLUSTERING

Comparison of two algorithms for clustering: Qualitative comparison **Algorithm 2**



RESULT OF CLUSTERING

Comparison of two algorithms for clustering: Quantitative comparison- Accuracy

Accuracy

Minerals	First Algorithm												Second Algorithm											
	Spatial resolution of RoI	NCC						SAM						NCC						SAM					
		Minerals		Number of mineral pixel	Quartz		Number of Quartz pixel	Minerals		Number of mineral pixel	Quartz		Number of Quartz pixel	Minerals		Number of mineral pixel	Quartz		Number of Quartz pixel	Minerals		Number of mineral pixel	Quartz		Number of Quartz pixel
		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$		$FP_{(px)}$	$FN_{(px)}$	
Biotite	123*138	1443	58	495	210	433	638	1353	91	495	526	263	638	388	183	638	229	127	495	229	127	495	392	176	638
Diopside	126*143	64	219	299	339	277	888	1868	93	299	654	411	888	576	177	888	69	142	299	2804	119	299	576	177	888
Epidote	123*148	1281	56	260	158	445	890	234	73	260	1321	434	890	534	116	890	124	69	260	124	69	260	545	116	890
Geothite	118*141	0	235	235	970	373	718	70	210	235	1295	413	718	0	235	235	2961	113	718	0	235	235	561	130	718
Kyanite	123*144	986	21	88	454	159	659	291	30	88	1169	346	659	95	32	88	620	119	659	95	32	88	625	113	659
Scheelite	123*158	1169	93	168	182	498	1006	245	92	168	1231	569	1006	75	102	168	801	106	1006	75	102	168	3862	89	1006
Smithsonite	119*160	1137	353	402	192	692	1117	424	348	402	1087	565	1117	0	402	402	797	219	1117	0	402	402	1066	138	1117
Tourmaline	58*80	130	83	122	28	14	14	132	41	122	292	14	14	153	42	122	0	14	14	153	42	122	0	14	14
Pyrope	159*159	600	69	259	1500	1548	1654	2322	16	259	768	255	1654	764	27	259	544	363	1654	1644	9	259	1031	135	1654
Olivine	172*142	2008	174	433	492	2337	2649	4881	99	433	403	1191	2649	821	105	433	337	1259	2649	1721	96	433	337	1259	2649

RESULT OF CLUSTERING

Comparison of two algorithms for clustering: Quantitative comparison- Computational complexity

Minerals	Computational Cost (time in second)															
		First Algorithm							Spatial resolution of RoI	Second Algorithm						
	Spatial resolution of RoI	NCC	SAM	MF						NCC	SAM	MF				
				MF	PLMF	RMF						MF	PLMF	RMF		
						sum	meanLocal	meanGlobalLocal						sum	meanLocal	meanGlobalLocal
Biotite	131*143	310.3946	273.7491	609.3558	376.2042	376.3622	383.6766	377.8778	123*141	15.2499	15.2310	15.3594	15.2014	15.1994	15.2005	15.6266
Diopside	128*145	288.6295	254.8992	619.6450	421.4566	447.4801	405.8906	380.9209	124*125	14.7789	14.7573	14.8882	14.7392	14.7308	14.7371	15.2249
Epidote	125*157	332.8230	320.9017	608.1568	433.5932	440.4032	459.2130	468.7292	125*157	22.1194	22.1062	22.2272	22.0845	22.0823	22.0851	22.4867
Geothite	124*144	298.0928	261.7532	545.0061	374.3371	374.1298	381.9427	376.2596	120*149	21.7209	21.6984	21.8146	21.6750	21.6753	21.6891	22.0657
Kyanite	129*144	304.6836	264.2952	609.3681	487.1841	664.2808	394.8771	386.9106	126*147	24.3413	24.3288	24.4566	24.3101	24.3105	24.3043	24.7365
Scheelite	136*172	514.3498	462.1700	846.7189	582.2462	621.2716	658.2493	634.7913	125*160	22.9961	22.9642	23.0750	22.9524	22.9500	22.9444	23.3581
Smithsonite	120*163	384.9274	293.9426	641.7419	409.6070	411.5419	417.7018	410.7087	119*160	22.3726	22.3475	22.5123	22.3466	22.3496	22.3423	22.8880
Tourmaline	50*55	211.0122	205.6003	252.7882	213.7044	213.5707	214.1684	213.5159	56*62	7.7948	7.7772	7.8917	7.7520	7.7530	7.7523	8.1223
Pyrope	144*152	362.3834	325.4031	652.9595	349.7858	346.04	356..8165	347.6778	159*170	18.7465	18.7322	18.8346	18.7034	18.7050	18.7047	19.1367
Olivine	157*139	497.7030	331.2886	627.4365	7.8214e+03	369.4791	356.4237	1.2252e+03	159*173	22.1437	22.1061	22.2118	22.0741	22.0687	22.0644	22.5827

C O N C L U S I O N

- The problem of mineral identification has been discussed.
- Hyperspectral infrared in the wavelength (7.7-11.8 µm) provides the signatures for plenty of the minerals. These minerals could be identified in LWIR.
- Some minerals did not have signature in LWIR.
- The possibilities for making a system for automatic identification of the minerals using the hyperspectral infrared imaging have been analyzed.
- Clustering and classification for identification of the mineral has been investigated.

Future Work:

- Further investigate different clustering techniques to improve performance,
- Investigate the application of spectral transformation techniques

This work has also been partially presented at SPIE DCS 2016 in Baltimore:

- Emissivity retrieval from indoor hyperspectral imaging of mineral grains (pp. 98611C-98611C)
- Mineral identification in hyperspectral imaging using Sparse-PCA (pp. 986118-986118)

Bardia Yousefi, Saeed Sojasi, Clemente Ibarra Castanedo, Univ. Laval (Canada); Georges Beaudoin, Univ. Laval (Canada); François Huot, Xavier P. V. Maldague, Univ. Laval (Canada); Martin Chamberland, Telops, Inc. (Canada); Erik Lalonde, Univ. Laval (Canada)

Published in Proceedings Volume 9861: Thermosense: Thermal Infrared Applications XXXVIII, May 2016

THANK YOU

TELOPS NSERC CRSNG MIVIM UNIVERSITÉ LAVAL