**Supporting Information**

Joint sparse representation and denoising method for Raman spectrum

Zheng Fang1,3, Yu Tao2, Wen Wang1,3, Wenxin Zhang1 , Lingfeng Duan4

Ying Liu1,3, Changchun Yan1,3, Lulu Qu5\*, Caiqin Han1,3\*

1 *Jiangsu Key Laboratory of Advanced Laser Materials and Devices, School of Physics and Electronic Engineering, Jiangsu Normal University, Xuzhou, Jiangsu, 221116, China*

2 *School of Physics and Electronic Engineering, Changshu Institute of Technology, Changshu, Jiangsu, 215500, China*

3 *Jiangsu Collaborative Innovation Center of Advanced Laser Technology and Emerging Industry, Jiangsu Normal University, Xuzhou, Jiangsu, 221116, China*

4 *Qianjiang College, Hangzhou Normal University, Hangzhou, Zhejiang, 310018, China*

5 *School of Chemistry and Materials Science, Jiangsu Normal University, Xuzhou, Jiangsu, 221116, China*

*Corresponding authors: hancq@jsnu.edu.cn; luluqu@jsnu.edu.cn*

1. A MATLAB example is shown below:

**cs\_result.m:**

clc;clear

sum = zeros(1,1536)

K=1;

N1=512;

M1=512;

Phi = randn(M1,N1)/sqrt(N1);

a = Lorentzian();

x = [a(1:512)].'

s=Phi\*x;

m=2\*K;

Psi=dwtmtx(512,'coif2',1)./sqrt(N1);

T=Phi\*Psi';

hat\_y=zeros(1,N1);

Aug\_t1=[];

r\_n=s;

for times=1:1;

for col=1:N1;

product1(col)=abs(T(:,col)'\*r\_n);

end

[val,pos]=max(product1);

Aug\_t1=[Aug\_t1,T(:,pos)];

T(:,pos)=zeros(M1,1);

aug\_y=(Aug\_t1'\*Aug\_t1)^(-1)\*Aug\_t1'\*s;

r\_n=s-Aug\_t1\*aug\_y;

pos\_array(times)=pos;

end

hat\_y(pos\_array)=aug\_y;

hat\_x=abs(real(Psi'\*hat\_y.'));

c = Gaussian()

d = c(1:1536)

x\_r = d.';

A = hat\_x;

N2=512;

M2=512;

Phi=PartFourierMtx(M2,N2)/sqrt(N2);

a = Lorentzian()

x2 = [a(513:1024)].';

s2=Phi\*x2;

m=2\*K;

Psi=dwtmtx(512,'coif2',1)./sqrt(N2);

T2=Phi\*Psi';

hat\_y=zeros(1,N2);

Aug\_t2=[];

r\_n=s2;

for times=1:1;

for col=1:N2;

product2(col)=abs(T2(:,col)'\*r\_n);

end

[val,pos]=max(product2);

Aug\_t2=[Aug\_t2,T2(:,pos)];

T2(:,pos)=zeros(M2,1);

aug\_y=(Aug\_t2'\*Aug\_t2)^(-1)\*Aug\_t2'\*s2;

r\_n=s2-Aug\_t2\*aug\_y;

pos\_array(times)=pos;

end

hat\_y(pos\_array)=aug\_y;

hat\_x2=abs(real(Psi'\*hat\_y.'))

N3=512;

M3=512;

Phi=PartFourierMtx(M3,N3)/sqrt(N3);

a = Lorentzian()

x3 = [a(1025:1536)].';

s3=Phi\*x3;

Psi=dwtmtx(512,'coif2',1)./sqrt(N3);

T3=Phi\*Psi';

hat\_y=zeros(1,N3);

Aug\_t3=[];

r\_n=s3;

for times=1:1;

for col=1:N3;

product3(col)=abs(T3(:,col)'\*r\_n);

end

[val,pos]=max(product3);

Aug\_t3=[Aug\_t3,T3(:,pos)];

T3(:,pos)=zeros(M3,1);

aug\_y=(Aug\_t3'\*Aug\_t3)^(-1)\*Aug\_t3'\*s3;

r\_n=s3-Aug\_t3\*aug\_y;

pos\_array(times)=pos;

end

hat\_y(pos\_array)=aug\_y;

hat\_x3=abs(real(Psi'\*hat\_y.'))

A2=hat\_x3;

figure(1);

B = [hat\_x.',hat\_x2.',hat\_x3.']

hold on;

C = 250:1:1785

plot(C,B)

plot(C,x\_r,'r');

%legend('Recovery','Original')

xlswrite('C:\Users\fangzheng\Desktop\1.xlsx',B)

<https://github.com/Fittzzzzz/Raman-Spectroscopy-Sparse-Representation/blob/master/cs_result.m>

**Lorentzian.m:**

function [ya] = Lorentzian()

ak = 500;

fk = 654;

N = 2249;

f = 250:N;

dk = 10;

y1 = ak \* 1./(1.+((fk-f)./(dk/2)).^2);

ak = 1000;

fk = 1002;

N = 2249

f = 250:N;

dk = 20;

y2 = ak \* 1./(1.+((fk-f)./(dk/2)).^2);

ak = 1500;

fk = 1455;

N = 2249;

f = 250:N;

dk =20;

y3 = ak \* 1./(1.+((fk-f)./(dk/2)).^2);

x = y1+y2+y3;

h = awgn(x,5,10)

y=2500./f+4

b=250:1:2249

ya=h

%plot(ya);

%xlswrite('C:\Users\fangzheng\Desktop\Lorentzian1.xlsx',x.')

end

<https://github.com/Fittzzzzz/Raman-Spectroscopy-Sparse-Representation/blob/master/Lorentzian.m>

**dwtmtx.m:**

function [ ww ] = dwtmtx( N,wtype,wlev )

[h,g]= wfilters(wtype,'d');

L=length(h);

h\_1 = fliplr(h);

g\_1 = fliplr(g);

loop\_max = log2(N);

loop\_min = double(int8(log2(L)))+1;

if wlev>loop\_max-loop\_min+1

fprintf('\nWaring: wlev is too big\n');

fprintf('The biggest wlev is %d\n',loop\_max-loop\_min+1);

wlev = loop\_max-loop\_min+1;

end

ww=1;

for loop = ceil(loop\_max-wlev+1):loop\_max

Nii = 2^loop;

p1\_0 = [h\_1 zeros(1,Nii-L)];

p2\_0 = [g\_1 zeros(1,Nii-L)];

p1 = zeros(Nii/2,Nii);

p2 = zeros(Nii/2,Nii);

for ii=1:Nii/2

p1(ii,:)=circshift(p1\_0',2\*(ii-1)+1-(L-1)+L/2-1)';

p2(ii,:)=circshift(p2\_0',2\*(ii-1)+1-(L-1)+L/2-1)';

end

w1=[p1;p2];

mm=2^loop\_max-length(w1);

w=[w1,zeros(length(w1),mm);zeros(mm,length(w1)),eye(mm,mm)];

ww=ww\*w;

clear p1;clear p2;

end

end

<https://github.com/Fittzzzzz/Raman-Spectroscopy-Sparse-Representation/blob/master/dwtmtx.m>

1. The usage and precautions of the method：
2. The function of the method is to recover Raman spectra from high noise.
3. The input parameter is a block of Raman spectra(the whole Raman spectra are divided into several blocks)
4. One block will recover one band. If block include more than one band, the biggest band will be recovered.
5. If the needed band position is in the input range, the needed band can be recovered. If the needed band position is out of the input range, the band can’t be recovered.
6. The band has just a little out of range, the proposed method will recover the band.
7. The band has just a little in range, the proposed method will lead to a wrong band.
8. The band is approximately half in range and approximately half out of range, the proportion of the band decide the result of the proposed method.