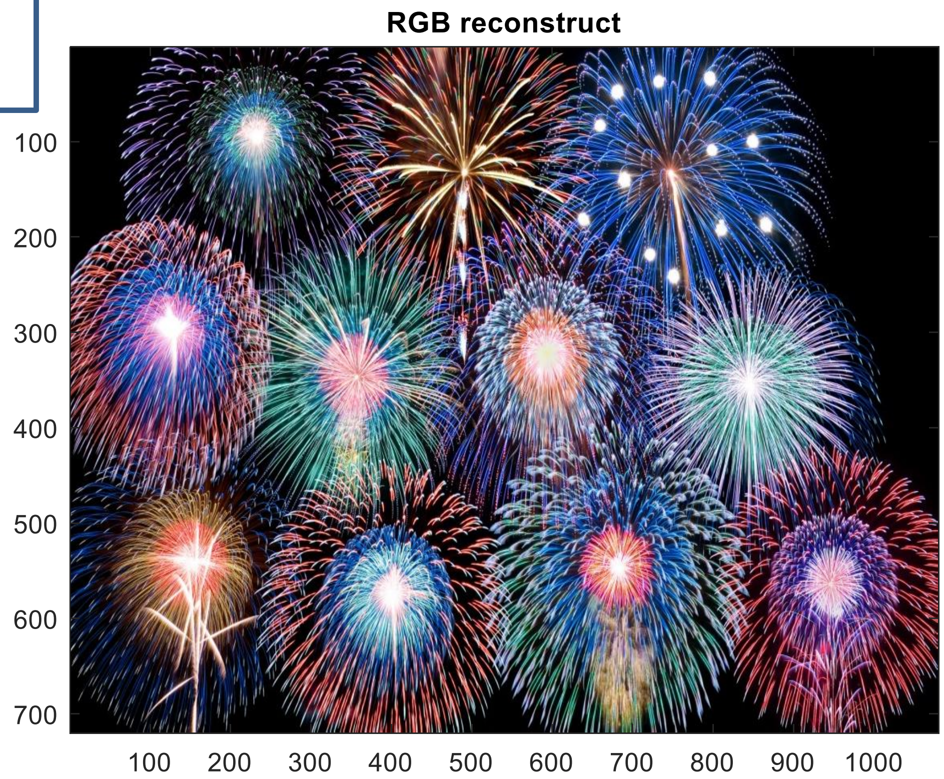
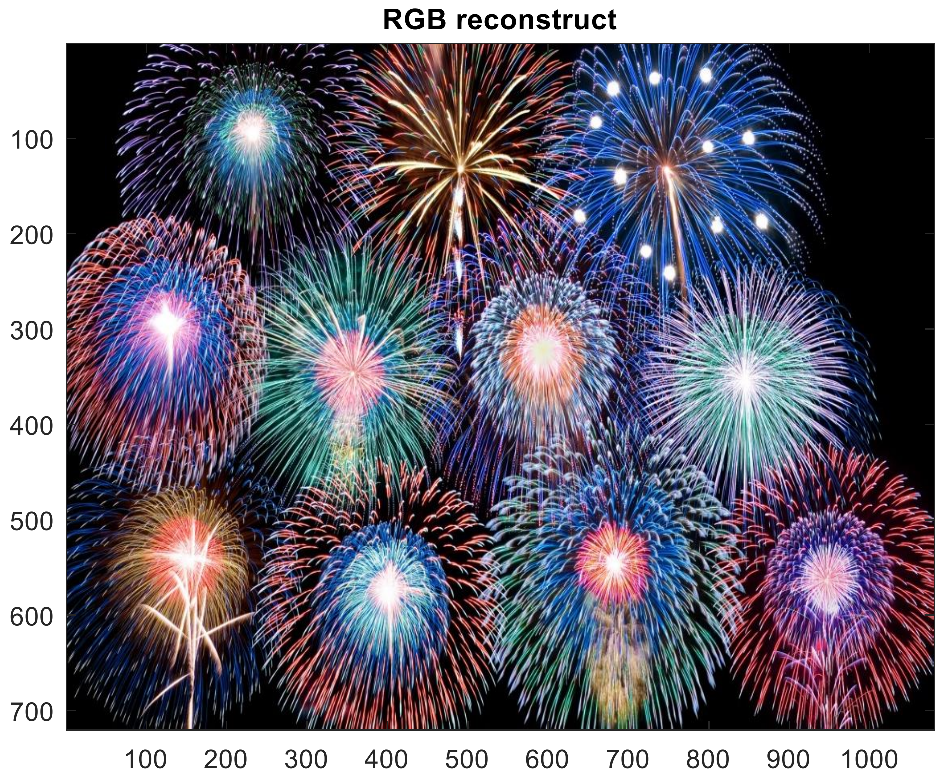
Homework 3 (Due: 5/21st)

(1) Write a Matlab or Python program for the <u>4:2:0 image compression technique</u>.
B = C420(A), A is the <u>input</u> color image and B is the <u>reconstructed image</u>.
Just use the interpolation method for reconstruction. <u>The code should be submitted to</u> [ceiba](). (Note: The command rgb2ycbcr cannot be used.)

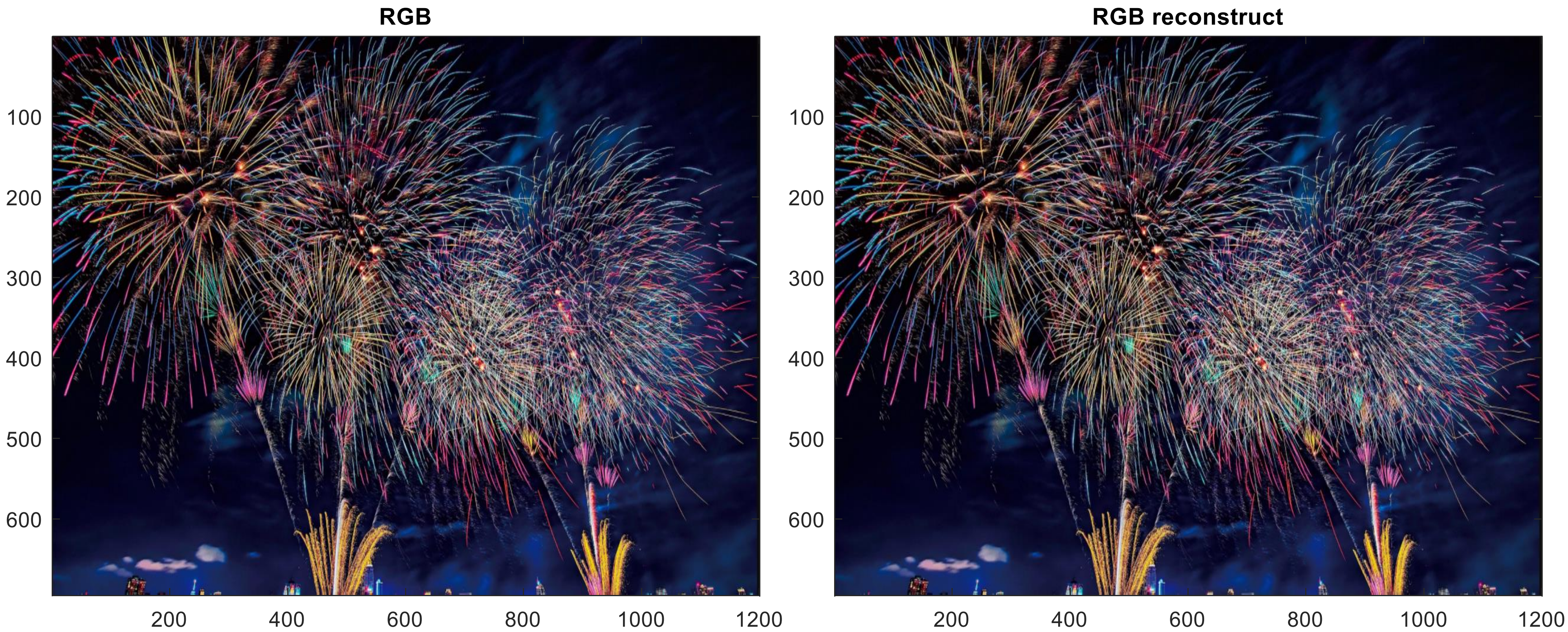
(25 scores)

```
>> m10907314_HW3

PSNR =

   41.5280
```





```
>> m10907314_HW3('fireworks2.png')

PSNR =

   39.7120
```



```
function m10907314_HW3(path_of_image)

    if nargin<1
       RGB = imread('fireworks.jpg');
    else
       RGB = imread(path_of_image);
    end
    RGB_to_YCbCr = [0.299 0.587 0.114;-0.169 -0.331 0.500;0.500 -0.419 -0.081];
    YCbCr_to_RGB = inv(RGB_to_YCbCr);
    image(RGB)
    title('RGB')

    [m,n,d] = size(RGB);
    RGB_transpose = double(permute(RGB, [3 2 1]));
    RGB_transform = reshape(RGB_transpose, 3, []);
    YCbCr_transform = RGB_to_YCbCr*RGB_transform;
    YCbCr = reshape(YCbCr_transform, d, n, m);

    Y = YCbCr(1,:,:);
    Cb = YCbCr(2,:,:);
    Cr = YCbCr(3,:,:);

    Cb_reduce = Cb(:,1:2:n,1:2:m);
    Cr_reduce = Cr(:,1:2:n,1:2:m);

    Cb_reconstruct = zeros(1,n,m);
    Cr_reconstruct = zeros(1,n,m);

    n_index = 1:1:n;
    n_index_resize = 1:2:n;
    m_index = 1:1:m;
    m_index_resize = 1:2:m;
```

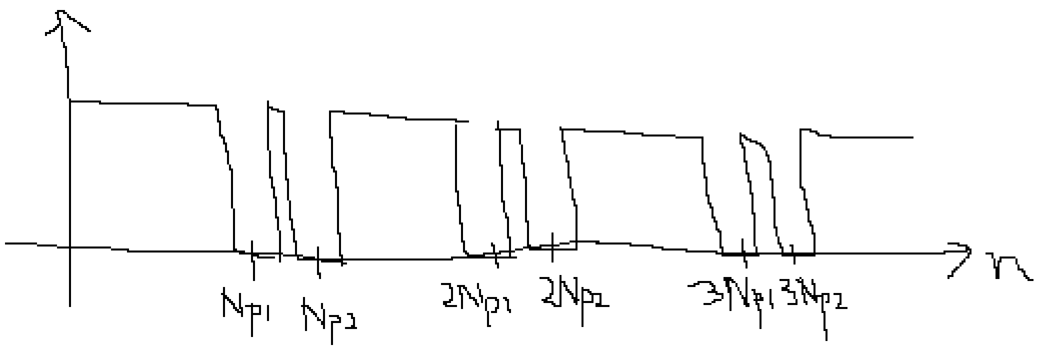
```
    Cb_reconstruct(1,n_index_resize,m_index_resize) = Cb_reduce;
    Cr_reconstruct(1,n_index_resize,m_index_resize) = Cr_reduce;
    for i=1:m/2
        Cb_reconstruct(1,:,m_index_resize(i)) =
            interp1(n_index_resize,Cb_reconstruct(1,n_index_resize,m_index_resize(i)),n_index);
        Cr_reconstruct(1,:,m_index_resize(i)) =
            interp1(n_index_resize,Cr_reconstruct(1,n_index_resize,m_index_resize(i)),n_index);
    end
    for i=1:n
        Cb_reconstruct(1,i,:) =  interp1(m_index_resize,squeeze(Cb_reconstruct(1,i,m_index_resize)),m_index);
        Cr_reconstruct(1,i,:) =  interp1(m_index_resize,squeeze(Cr_reconstruct(1,i,m_index_resize)),m_index);
    end
    Cb_reconstruct(1,n,:) = Cb_reconstruct(1,n-1,:);
    Cb_reconstruct(1,:,m) = Cb_reconstruct(1,:,m-1);
    Cr_reconstruct(1,n,:) = Cr_reconstruct(1,n-1,:);
    Cr_reconstruct(1,:,m) = Cr_reconstruct(1,:,m-1);

    YCbCr_reconstruct = [Y;Cb_reconstruct;Cr_reconstruct];
    YCbCr_reconstruct_transform = reshape(YCbCr_reconstruct, 3, []);
    RGB_reconstruct_transform = YCbCr_to_RGB*YCbCr_reconstruct_transform;
    RGB_reconstruct_transpose = reshape(RGB_reconstruct_transform, 3,  n, m);
    RGB_reconstruct = uint8(permute(RGB_reconstruct_transpose, [3 2 1]));
    figure;
    image(RGB_reconstruct)
    title('RGB reconstruct')

    MAX = 255;
    MES = sum(sum(sum((RGB_reconstruct-RGB).^2)))/(m*n*3);
    PSNR = 10*log10(MAX*MAX / MES)

end
```

(2) Write <u>two concepts</u> you learned from the oral presentation on 4/30. (10 scores)

1.影像重建的模型中第一層產生2倍解析照片 的參數透過share parameter的方式傳遞到下一層產生4倍解析度，就能夠減少模型參數量。
2.Time and frequency masking 兩種方式比較有效，因為能夠讓模型逼迫從一個不完整的Spectrogram學到正確的語音辨識的方法。

(3) Suppose that $y[n] = x[n] + a_1x[n\text{-}15] + a_2x[n\text{-}20]$. How do we use the lifter to remove the multipath problem and reconstruct $x[n]$ from $y[n]$? (10 scores)

將y[n]做倒頻譜
因為在倒頻譜時multipath的訊號就只有在n等於N_{p1}的倍數以及N_{p2}的倍數才會有值，N_{p1}=15、N_{p2}=20，於是將lifter設計成濾除N_{p1}的倍數以及N_{p2}的倍數的樣子如下圖



Convolution lifter之後再轉換回來就能重建回 y[n] = x[n]

(4) Suppose that one of the string of the guitar has the length of 65cm. (a) What are the frequencies of the vocal signal that the string will generate at 15°C? (b) Suppose that the vocal signal generated by the string is Do. What are the string lengths to generate Mi and So? (10 scores)

(a)

$$\lambda = 2d = 2 * 65 = 130cm$$

$$f = \frac{v}{\lambda} = \frac{340}{1.3} = 261.5385\ Hz$$

$$ans:\ 261.5385k\ Hz, k = 1,2,3, \ldots \ldots$$

	Do	升Do	Re	升Re	Mi	Fa	升Fa	So	升So	La	升La	Si
Hz	262	277	294	311	330	349	370	392	415	440	466	494

(b)

$$f_{Do} = 261.5385$$

$$f_{Mi} = 2^{\frac{4}{12}} f_{Do}$$

$$f_{So} = 2^{\frac{7}{12}} f_{Do}$$

$$d = \frac{v}{2f}$$

$$d \propto \frac{1}{f}$$

$\frac{d_{Do}}{2^{\frac{4}{12}}}$ =51.5905 cm

$$the\ string\ lengths\ to\ generate\ Mi = \frac{51.5905}{k}\ cm, k = 1,2,3, \ldots \ldots$$

$\frac{d_{Do}}{2^{\frac{7}{12}}}$ =43.3823 cm

$$the\ string\ lengths\ to\ generate\ So = \frac{43.3823}{k}\ cm, k = 1,2,3, \ldots \ldots .$$

(5) (a) Discuss three possible ways to compress a music signal more effectively.
(b) Discuss two possible ways to compress a cartoon image more effectively. (10 scores)

(a) 1. repeated melody
2. frequency is stable within a beat
3. interval of a note is a multiple of b/2 (b: beat length)

(b) 1. edge information + color in each region
2. edge can be encoded using few parameters

(6) Why it is better to apply the DCT instead of the DFT and the KLT for image compression? (10 scores)

1. DCT比DFT更能把能量集中在0附近
2. DFT要多紀錄虛數不利於壓縮，DCT只需記實數
3. KLT雖然是最佳算法但KLT依賴於輸入不利於壓縮

(7)What are the advantages to apply the techniques of (a) differential coding for DC terms and (b) zigzag for AC terms in the JPEG process? (10 scores)

(a)因為鄰近的區域DC值會差不多所以會將數值都集中分布在0附近

(b)Zigzag會將低頻成分排前面高頻成分排後面，通常數值都分布在低頻，數值都集中在前半部的話，如果之後數值都為0那就可以告訴電腦說不必執行。

(8) Suppose that $x[k]$ is a series of natural number and the length of $x[k]$ is 50000. Also suppose that $P(x[k]=n)=\left(1-e^{-\lambda}\right)e^{-\lambda n}$ for $n = 0, 1, 2$……….. where $\lambda = 0.005$. Estimate the range of the total coding lengths in the binary system when using (i) the Huffman code and (ii) the arithmetic code to encode $x[k]$. (15 scores)

$$entropy = \sum_{j=1}^{J} P(S_j)\log\frac{1}{P(S_j)} = 6.2983$$

Matlab code

```
lamda = 0.005;
n = 0:1:49999
P = (1-exp(-lamda))*exp(-lamda*n);
plot(P)

Entropy = sum(P.*log(1./P));
```

(i)Huffman code

$$ceil\left(N\frac{entropy}{\log k}\right) \le b \le floor\left(N\frac{entropy}{\log k} + N\right)$$

$$ceil\left(50000\frac{6.2983}{\log 2}\right) \le b \le floor(50000\frac{6.2983}{\log 2} + 50000)$$

$$454328 \le b \le 504327$$

(ii)Arithmetic code

$$ceil\left(N\frac{entropy}{\log k}\right) \le b \le floor\left(N\frac{entropy}{\log k} + \log_k 2 + 1\right)$$

$$ceil\left(50000\frac{6.2983}{\log 2}\right) \le b \le floor(50000\frac{6.2983}{\log 2} + \log_2 2 + 1)$$

$$454328 \le b \le 454329$$