CSE 3018 - Content Based Image Retrieval
Lab 3 - CBIR System using Histogram Based Features

Name: Ayush Sharma
Reg. No: 15BCE1335

Question: Implement a CBIR system that uses features derived from Color Histogram such as mean, variance, skewness, kurtosis and entropy. Use histogram with various number of bins such as 8, 16, 32 and 64. Also implement a CBIR system that uses the complete histogram for comparison between images and calculate its precision and recall.

Code for feature extraction CBIR:
For different features and bin size manipulate the variables named 'feature' and 'bin_size'.
Features = mean, var, skewness, kurtosis, entropy
Bin_size = 8, 16, 32, 64

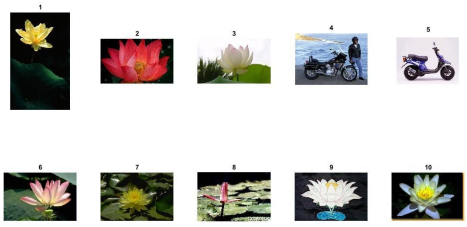
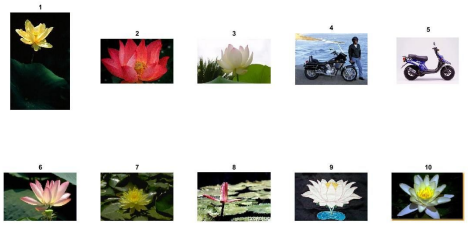
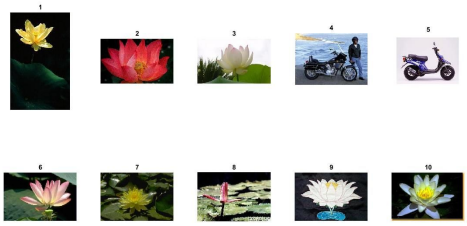
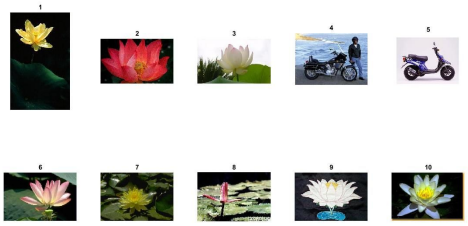
```
queryimg = imread('C:\Users\TEMP\Desktop\query.jpg');
chi_sqr = [];
rhq = imhist(queryimg(:,:,1),bin_size);
ghq = imhist(queryimg(:,:,2),bin_size);
bhq = imhist(queryimg(:,:,3),bin_size);
q= feature([rhq ghq bhq]);
source = dir('C:\Users\TEMP\Desktop\images\*.jpg');

for i=1:99
data = imread(strcat('C:\Users\TEMP\Desktop\images\',source(i).name));
rh = feature(imhist(data(:,:,1),bin_size));
gh = feature(imhist(data(:,:,2),bin_size));
bh = feature(imhist(data(:,:,3),bin_size));
src = [rh gh bh];
m = size(q,1);  n = size(src,1);
mOnes = ones(1,m); D = zeros(m,n);
for j=1:n
  yi = src(j,:);  yiRep = yi( mOnes, : );
  s = yiRep + q;    d = yiRep - q;
  D(:,j) = sum( d.^2 ./ (s+eps), 2 );
end
chi_sqr(i) = D/2;
end
sorted_vector = sort(chi_sqr);
figure (1);
sol = [];
for i=1:10
x = sorted_vector(i);
```

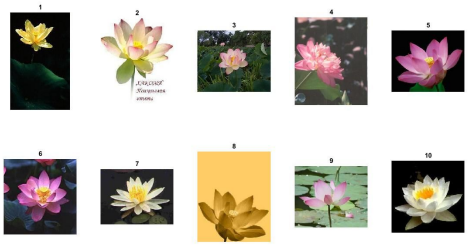
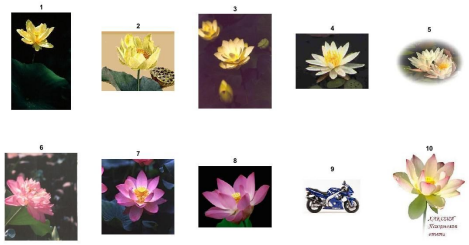
```
for j = 1:99
if(x==chi_sqr(j) && ~ismember(j,sol))
sol(i) = j;
img = imread(strcat('C:\Users\TEMP\Desktop\images\',source(j).name));
subplot(2,5,i); imshow(img), title(i);
break
end
end
end
```

Screenshots:

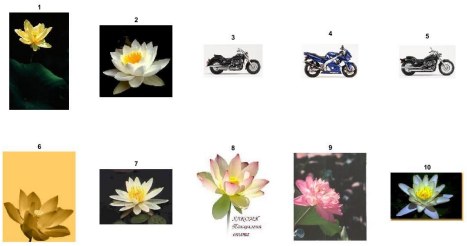
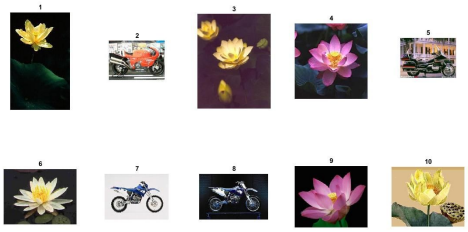
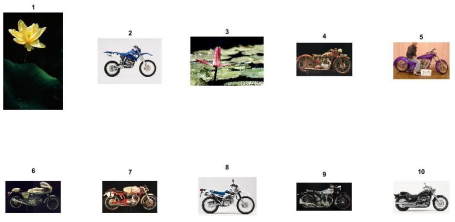
Mean:

8 bins 	16 bins 
32 bins 	64 bins 

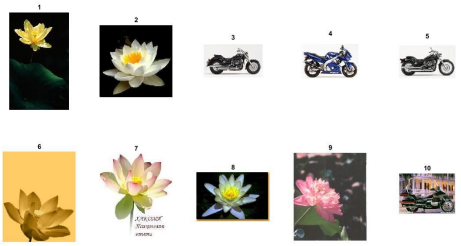
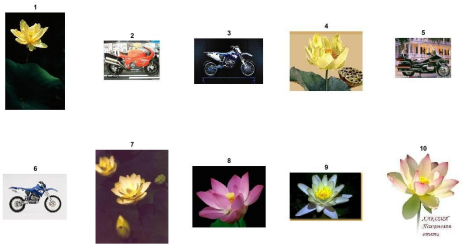
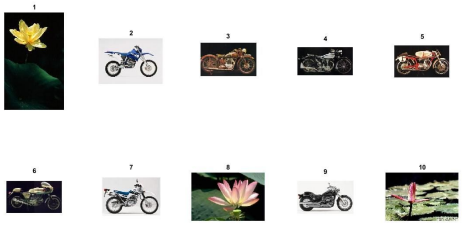
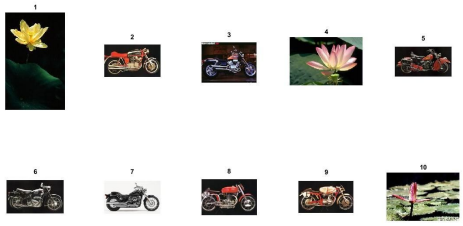
Variance:

8 Bins 	16 Bins 
32 Bins 	64 Bins 

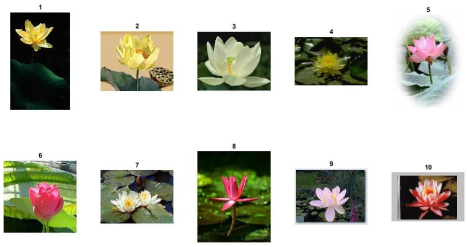
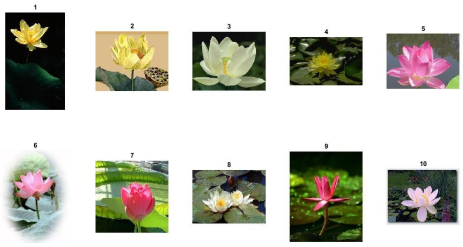
Skewness:

8 Bins 	16 Bins 
32 Bins 	64 Bins 

Kurtosis:

8 Bins 	16 Bins 
32 Bins 	64 Bins 

Entropy:

8 Bins 	16 Bins 
32 Bins 	64 Bins 

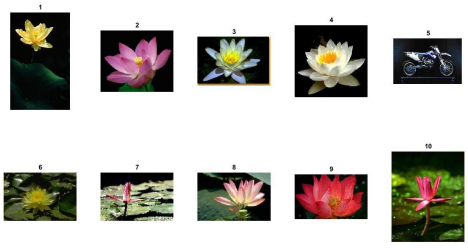
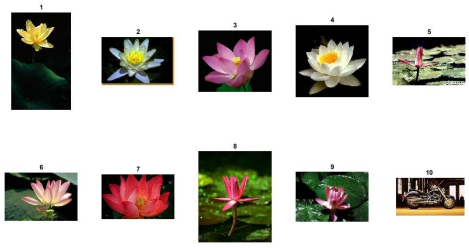
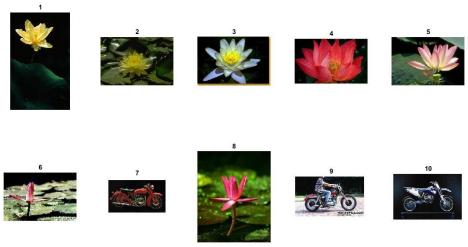
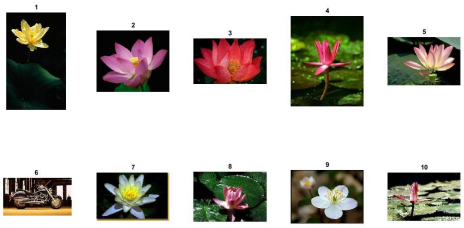
Code for histogram as a whole for comparison:
Change Bin size as 8, 16, 32, 64.

```
queryimg = imread('C:\Users\TEMP\Desktop\query.jpg');
chi_sqr = [];
rhq = imhist(queryimg(:,:,1),bin_size);
ghq = imhist(queryimg(:,:,2),bin_size);
bhq = imhist(queryimg(:,:,3),bin_size);
q= horzcat(rhq',ghq',bhq');
source = dir('C:\Users\TEMP\Desktop\images\*.jpg');

for i=1:85
data = imread(strcat('C:\Users\TEMP\Desktop\images\',source(i).name));
rh = imhist(data(:,:,1),bin_size);
gh = imhist(data(:,:,2),bin_size);
bh = imhist(data(:,:,3),bin_size);
src = horzcat(rh',gh',bh');
sum=0;
for j=1:size(src,2)
x=src(1,j,1)-q(1,j,1);
x=x^2;
sum=sum+x;
end
chi_sqr(i)=sqrt(sum);
end

sorted_vector = sort(chi_sqr);
figure (1);
sol = [];
for i=1:10
x = sorted_vector(i);
for j = 1:85
if(x==chi_sqr(j) && ~ismember(j,sol))
sol(i) = j;
img = imread(strcat('C:\Users\TEMP\Desktop\images\',source(j).name));
subplot(2,5,i); imshow(img), title(i);
break
end
end
end
```

Screenshot:

8 Bins 	16 Bins 
32 Bins 	64 Bins 

Precision and Recall:
Precision is the fraction of retrieved images that are relevant to the query.
Recall is the fraction of the relevant documents that are successfully retrieved.

Considering our dataset to be of two classes: flower and bikes (50 images each).
Taking the top 10 relevant images as the output by the program (let us suppose the rest are considered as images of bike by our system).
Our CBIR system gave us 9 flower images out of 10 for 3 different values of histogram bins (i.e. 8, 16, 64).

This system gave us 10 relevant flower images out of 50 flower images and 9 were actually flowers thus,
Precision = 9/10
Recall = 9/50