CSE 408/598 Multimedia Information Systems

Phase #3

(Due November 29th 2015, midnight)

Description: In this project, you will

- experiment with video data representation,

- feature extraction, and

- experiment with video retrieval.

Tasks:

The input to the project will be a video file.

You will operate in the Y color space.

- **Task I:** Implement a program which divides the frames of the given video into 8-by-8 regions – called *frame blocks*. Then, given a user supplied n, the program performs the following subtasks:

  - **Task I(a):** The program creates a n-bin gray scale histogram, by quantizing the Y color space, for each cellblock of each frame in the given video. The outputs, of the form

    $$\langle frame\_id, block\_coord, gray\_instance\_id, num\_pixels \rangle$$

    are written into a file of the form

    video_filename*hist_n.hst*

  - **Task I(b):** The program applies 2D-DCT on the Y component of each block of each frame in the given video and selects the n most *significant* frequency components. The outputs, of the form

    $$\langle frame\_id, block\_coord, freq\_comp\_id, value \rangle$$

    are written into a file of the form

    video_filename*blockdct_n.bct*

  - **Task I(c):** The program applies 2D-DWT on the Y component of each block of each frame in the given video and selects the n most *significant* wavelet components. The outputs, of the form

    $$\langle frame\_id, block\_coord, wavelet\_comp\_id, value \rangle$$

    are written into a file of the form

    video_filename*blockdwt_n.bwt*

– **Task I(d):** For each frame (except the last) of the video and for each block, the program computes the difference from the same block in the next frame and the creates an n-bin difference histogram, by quantizing the obtained differences. The outputs, of the form

$$\langle frame\_id, block\_coord, diff\_comp\_id, pixelcount \rangle$$

are written into a file of the form

$$\texttt{video\_filename\_}diff\_n.dhc$$

- **Task II:** Implement a program which applies 2D-DWT on the Y component of each image video frame in the given set and selects the m most *significant* wavelet components (m is user supplied). The outputs, of the form

$$\langle frame\_id, wavelet\_comp\_id, value \rangle$$

are written into a file of the form

$$\texttt{video\_filename\_}framedwt\_m.fwt$$

- **Task III** Implement a program which given (a) a query video, (b) a frame id, and (c) n and m values, identifies the best 10 matching frames (except itself) for each of the 5 feature types and displays the original frame as well as the 10 matches (and matching scores) for each feature type.

Deliverables:

- Your code (properly commented) and a README file.

- A report describing your work and the results on a sample image.

Please place your code in a directory titled "Code", the outputs to a directory called "Outputs", and your report in a directory called "Report"; zip or tar all off them together and submit it through the Blackboard.