ILLINOIS INSTITUTE OF TECHNOLOGY
ECE 565 MULTIDIMENSIONAL DIGITAL SIGNAL PROCESSING (FALL'98)
Project 2: Motion Estimation Using Block Matching

Due Date: December 8, 1998 (Extension will not be granted)

Objectives

The work items of this computer simulation project are:

1. Implement block-matching (BM) motion estimation.
2. Assess the quality of the motion field (MF) for motion compensated prediction.
3. (Bonus) Devise a scheme to adapt the BM block size.

Ground Rules

The ground rules for Project 1 also apply to this project. Your program may call any library functions provided by Matlab and C.

Project Tasks

1. Write a routine to compute, using BM motion estimation, displacement vectors (DVs) for a target video frame. The frame is partitioned into equal-size blocks ("target blocks"), and a DV is computed for each block using BM. The block size is $N \times N$. The DVs have integer resolution, and the DV search range is $\pm M \times \pm M$; in other words, each component of the DV takes a value from $\mathcal{S} = \{0, \pm 1, \pm 2, \ldots, \pm M\}$. The DV value for a target block is found via the following steps:

   (a) Find the DV value $\mathbf{v}_{-0}$ that minimizes the sum of absolute difference (SAD) between the target block and the block in the reference frame whose coordinates are displaced with respect to the target block coordinates:

   $$\mathbf{v}_{-0} = \arg \min_{\mathbf{v} \in \mathcal{S} - \{0\}} \sum_{\mathbf{n} \in \mathcal{B}} |I_t(\mathbf{n}) - I_r(\mathbf{n} - \mathbf{v})|$$

   where $\mathcal{B}$ is the set of coordinates of the pixels in the target block, and $I_t(\mathbf{n})$ and $I_r(\mathbf{n})$ are the pixel intensities in the target and reference frames, respectively. Let the SAD value corresponding to using $\mathbf{v}_{-0}$ be d_{-0}.

   (b) Find the SAD value d_0 corresponding to using the zero DV $\mathbf{0}$ for BM.

   (c) The DV $\mathbf{v}^*$ for the target block is determined using:

   If $d_{-0} < d_0 - c|\mathcal{B}|$ then $\mathbf{v}^* \leftarrow \mathbf{v}_{-0}$, else $\mathbf{v}^* \leftarrow \mathbf{0}$

   where c is a positive thresholding parameter to be determined experimentally, and $|\mathcal{B}|$ is the number of pixels in the target block.

   The total SAD is the sum of the SADs obtained for all the blocks of the target frame, where the SAD for a block is due to using $\mathbf{v}^*$.

2. Given the DV $\mathbf{v}^*$ of a target block, the motion compensated prediction for the target block is given by the block of pixels $I_r(\mathbf{n} - \mathbf{v}^*), \mathbf{n} \in \mathcal{B}$. A motion compensated prediction frame $\tilde{I}(\mathbf{n})$ is formed as a mosaic of such prediction blocks: for every $\mathbf{n} \in \mathcal{B}$, $\tilde{I}(\mathbf{n}) = I_r(\mathbf{n} - \mathbf{v}^*)$. The motion compensated prediction residual or displaced frame difference (DFD) is the frame $e_{mc}(\mathbf{n}) = I_t(\mathbf{n}) - \tilde{I}(\mathbf{n})$, whereas the frame difference is simply $e(\mathbf{n}) = I_t(\mathbf{n}) - I_r(\mathbf{n})$, i.e. obtained without motion compensation. You should be able to plot $I_t(\mathbf{n})$, $I_r(\mathbf{n})$, $\tilde{I}(\mathbf{n})$, $e_{mc}(\mathbf{n})$, $e(\mathbf{n})$, and the MF $\mathcal{V}$. Use the program plotdv.m to plot $\mathcal{V}$, plotting one DV for each block. The origin of the DV should be at the center of the block.

3. Visually inspect the target and reference frames, and the frame difference, to obtain a rough estimate of where motion occurs and in what manner. For $N = 16$ and $M = 8$, find the value c^* for the parameter c that gives the "best" MF. Plot the motion compensated prediction, the DFD, and the MF in order to assess the suitability of a c value. How does choosing a nonzero value for c affect the total SAD? Justify the c^* you chose in terms of your quality criteria.

4. Given c^*, find the MF using $N = 8$ and $M = 8$. Comment on the quality of the MF obtained. Is the c^* value still appropriate now that the block size is halved in each dimension? Are there significant improvements over using $N = 16$? What are they? Carefully justify your observations and conclusions.

Pragmatics

Some of the pragmatics items in Project 1 apply here too. In addition, consider the following:

1. The DFD signal can have negative amplitudes. You will need to find an appropriate method to remap the amplitudes before displaying. Document your remapping scheme in your report.

2. In your BM program, it is necessary to specifically deal with the frame boundaries in the motion search. Certain values of $\mathbf{v}$ would lead to addressing pixels outside of the reference frame in $I_r(\mathbf{n} - \mathbf{v})$. Use a physically meaningful method to deal with this problem. Document this!

3. BM is a very computationally intensive task. Carefully organize your code for the BM computation according to sound programming practice.

Work for Bonus Points

Is there any advantage to varying the BM block size N over the video frame? Suppose you have two sizes to choose from, $N = 8$ and $N = 16$, can you propose a *simple* and yet effective algorithm to choose between them? This means that for each 16×16 block, the algorithm would choose between using one DV for the entire block, or four DVs, one for each of the four 8×8 subblocks of the 16×16 block.

Carefully document the reasoning behind the construction of your algorithm, and any experiments you performed to fine tune its performance. The steps of your algorithm must be clearly defined. Submit all the data and plots that demonstrate the operation and effectiveness of your algorithm. The plotdv.m program can plot the MF obtained from variable block size BM.