Digital Video Processing (EE392J)
Department of Electrical Engineering
Stanford University

Problem Set No. 4

Issued: February 6, 2002
Due: **Monday**, February 11, 2002
Note: This is a short problem set (due on Monday) because of the midterm scheduled on Wednesday.
Announcement: **Midterm** on Wednesday, February 13, 2002
Reading: Chapter 4, and skim papers handed out in class.

Problem 1. Problem 4.2. If you prefer, you can hand-sketch the magnitude of the frequency response instead of using Matlab.

Problem 2. An important area of video processing today is frame rate conversion. An example of this can be seen in our current television system. Motion pictures are typically shot at 24 frames/sec while our television system is based on the transmission of 60 fields/sec. An "upsampling" process is required to couple the 24 frames/sec film to the 60 fields/sec transmission format. To simplify matters, let us assume that it would be sufficient if the upconverted signal was at 60 frames/sec.

Given the original continuous-time video signal, this could be accomplished by sampling at a rate 2.5 times as high as before. However this is not a practical approach. Instead a technique called 3:2 pulldown is typically used (Figure 1). In 3:2 pulldown the first frame is repeated 3 times, then the second frame is repeated 2 times, the third frame is repeated 3 times, etc. Each odd numbered frame is repeated 3 times and each even numbered frame is repeated twice.

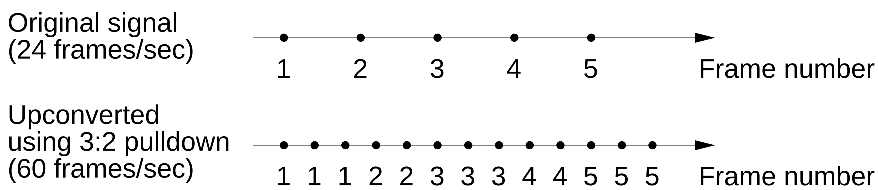

Figure 1: 3:2 pulldown

A. How can we model this 3:2 pulldown process from a DSP point of view. That is, how can we represent this process using upsamplers, downsamplers, and linear filters. Draw an appropriate block diagram, specifying the important elements. Make sure to specify the intermediate lattice, and the equivalent linear filter applied on that lattice.

B. What are the artifacts typically seen in 3:2 pulldown? What produces these artifacts? (Be brief.)

(see back)

Improved frame rate conversion can be achieved by carefully examining the spatio-temporal characteristics of a video signal. For example, if we have a continuous-time video signal $s(x, y, t)$ that is stationary (i.e. no motion) then when we examine its 3-D spatio-temporal Fourier transform $S(F_x, F_y, F_t)$ all of the energy is concentrated in the $S(F_x, F_y, 0)$ plane, i.e. in the plane formed by $F_t = 0$. If the motion is in the form of a constant global shift from one frame to the next (global motion with constant velocity for the continuous video) the plane containing the energy becomes tilted in the 3-D Fourier domain. If the motion is not exactly in the form of a global shift, the energy spreads out around this tilted plane. Note that this can be an appropriate model for local regions of the video signal.

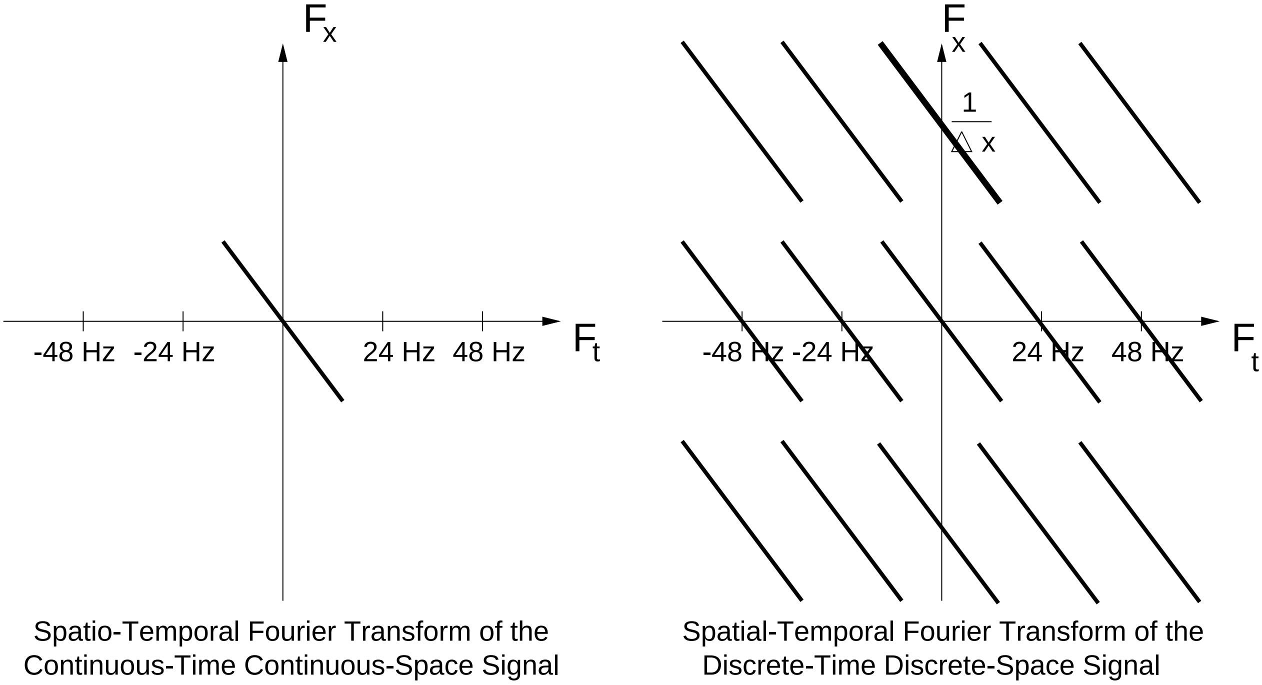

Figure 2: Spatio-temporal Fourier transform of sampled signal.

Assume that there is no motion along the y-direction, this allows us to restrict our analysis to the (F_x, F_t) plane. If we have a motion picture moving along the x-direction with global motion with constant velocity, and sampled at 24 frames/sec, the support of its spatio-temporal Fourier transform is as shown in Figure 2. Note that the replications of the motion picture's baseband appears at multiples of 24 Hz.

C. Can the original continuous video signal be reconstructed by applying a separable continuous filter? That is, by applying a filter with impulse response $h(x, y, t) = h_x(x)h_y(y)h_t(t)$? If yes, explain. If not, why not, and what type of filter may be used? In either case please be explicit describing the filter's passband and stopband, and how this filter support may be achieved.

D. Consider an arbitrary television video signal (assume progressive scanning at 60 Hz), that is subsequently sampled at 24 Hz (frames/sec). Can we prefectly reconstruct the 60 Hz video signal such that it is identical to if we had sampled the video at 60 Hz? If yes, briefly explain. If no, give counterexample.