Frequency Characteristics of Moving Images

Brian Guenter
GVU Center
College of Computing
Georgia Institute of Technology
internet: brian.guenter@cc.gatech.edu
phone: (404) 853-9378

1. The Fourier Transform under Affine Transformations

The class of multidimensional signals we will be dealing with is three dimensional: x, y, and time. We want to compute the Fourier transform of a moving image and the easiest way to approach the problem is to begin by examining the behavior of affine transformations under the Fourier transform because image translation can be compactly described by an affine transformation. Given

$$F(g(\mathbf{x})) = \int_D g(\mathbf{x}) \exp(-j2\pi\omega^T \mathbf{x})\, d\mathbf{x} \qquad \text{(Eq. 1)}$$

find F{g(**A**x + **z**)} where **A** is a linear transformation applied to the input signal **x** and **z** is an offset term.

$$F(g(\mathbf{Ax} + \mathbf{z})) = \int_{D^*} g(\mathbf{Ax} + \mathbf{z}) \exp(-j2\pi\omega^T \mathbf{x})\, d\mathbf{x} \qquad \text{(Eq. 2)}$$

Making the substitution $\mathbf{x}' = \mathbf{Ax} + \mathbf{z}$ and using multivariable change of variables [MARSDEN 1976] gives

$$\int_{D^*} |\det \mathbf{J}|\, g(\mathbf{Ax} + \mathbf{z}) \exp(-j2\pi\omega^T \mathbf{x})\, d\mathbf{x} = \int_D g(\mathbf{x}') \exp(-j2\pi\omega^T \mathbf{A}^{-1}(\mathbf{x}' - \mathbf{z}))\, d\mathbf{x} \qquad \text{(Eq. 3)}$$

where **J**, the Jacobian, is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} & \frac{\partial x'}{\partial z} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} & \frac{\partial y'}{\partial z} \\ \frac{\partial z'}{\partial x} & \frac{\partial z'}{\partial y} & \frac{\partial z'}{\partial z} \end{bmatrix} \qquad \text{(Eq. 4)}$$

Writing out the matrix equation $\mathbf{x}' = \mathbf{Ax} + \mathbf{z}$

$$\begin{bmatrix} x' \\ y' \\ t' \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + a_{13}t + z_1 \\ a_{21}x + a_{22}y + a_{23}t + z_2 \\ a_{31}x + a_{32}y + a_{33}t + z_3 \end{bmatrix} \qquad \text{(Eq. 5)}$$

and computing partials in the equation for x' gives

$$\frac{\partial x'}{\partial x} = a_{11} \ , \frac{\partial x'}{\partial y} = a_{12} \ , \frac{\partial x'}{\partial t} = a_{13}$$ (Eq. 6)

The results for the y', and z' partials follow trivially. Now it is clear that **J** is simply **A** and consequently |det **J**| = |det **A**| . This is a constant so it may be factored out from the integral giving

$$F(g(\mathbf{Ax} + \mathbf{z})) = \int_{D^*} g(\mathbf{Ax} + \mathbf{z})\exp(-j2\pi\omega^T\mathbf{x})d\mathbf{x}$$

$$= \frac{1}{|\det \mathbf{A}|}\int_D g(\mathbf{x}')\exp(-j2\pi\omega^T\mathbf{A}^{-1}(\mathbf{x}' - \mathbf{z}))d\mathbf{x}$$ (Eq. 7)

which reduces to

$$F(g(\mathbf{Ax} + \mathbf{z})) = \frac{1}{|\det \mathbf{A}|} \exp(+j2\pi\omega^T\mathbf{A}^{-1}\mathbf{z})\, \mathbf{G}\left({\mathbf{A}^{-1}}^T \omega\right)$$ Eq. 8)

This shows that the Fourier transform of a signal being operated on by an affine transform is equal to the original Fourier transform operated on by the matrix $(\mathbf{A}^{-1})^T$ and multiplied by a scale term and a phase shift term.

1.1. Spectrum of a Moving Image

Now that we have the mathematical description of affine transformations under the Fourier transform we can apply this directly to the problem of determining the Fourier transform of a moving image [WATSON 1986]. Let the static image be **C**(**x**) and assume that the offset term **z** in (Eq. 8) is 0. Then uniform translation of the image with velocity r_x along the x axis and r_y along the y axis can be written as **C**(**Ax**) where **A** is the matrix

$$\begin{bmatrix} x' \\ y' \\ t' \end{bmatrix} = \begin{bmatrix} 1 & 0 & -r_x \\ 0 & 1 & -r_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ t \end{bmatrix}$$ (Eq. 9)

Multiplying **Ax** out and writing it term by term gives

$$\mathbf{C}(\mathbf{Ax}) = \mathbf{C}(x - r_x t, y - r_y t, t)$$ (Eq. 10)

Now applying what we already know about the Fourier transform from (Eq. 8) it is clear that the Fourier transform of a uniformly translating image is

$$\frac{1}{|\det \mathbf{A}|}\mathbf{G}\left(\left(\mathbf{A}^{-1}\right)^{T}\omega\right) \qquad \text{(Eq. 11)}$$

where $\mathbf{G}(\omega)$ is the Fourier transform of $\mathbf{C}(\mathbf{x})$. Writing the Fourier transform out component by component gives $\mathbf{G}(\omega_x, \omega_y, \omega_x r_x + \omega_y r_y + \omega_t)$. Motion in the image plane skews the ω_x, ω_y frequency plane along the ω_t axis. This is easier to illustrate in 2 dimensions than in 3. Figure 1 shows uniform translation with a velocity of r in the x direction.

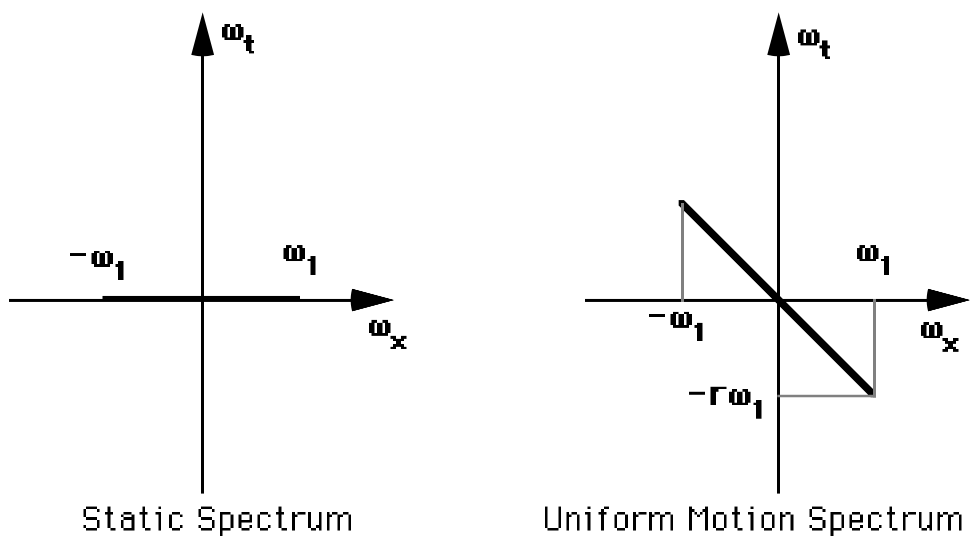

Figure 1

When this signal is sampled in the time dimension with sampling frequency ω_{st} then the spectrum is replicated along the ω_t axis. Only two replicas are shown but there are infinite number in both positive and negative time frequency.

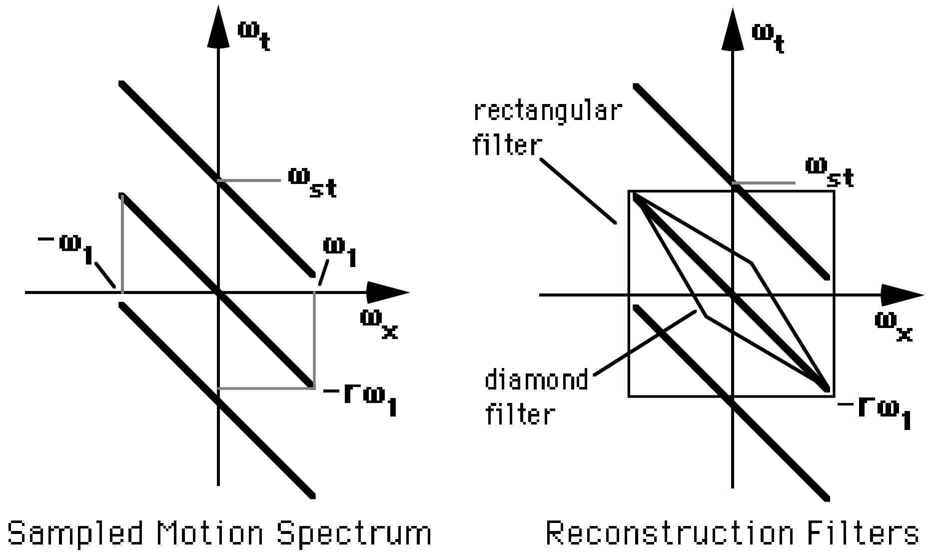

Figure 2

If the original signal is spatially band limited then it can always be reconstructed no matter how low the time sampling rate since the spectral replicas along the time dimension do not overlap. The best reconstruction filter is not necessarily rectangular. Notice that the rectangular filter passes all of the baseband signal but also passes part of the temporal spectral replicas. The diamond shaped filter passes only the baseband signal. The shape and orientation of the low pass reconstruction filter are a function of image velocity. In general different parts of the image will be moving at different velocities so the ideal temporal antialiasing filter will be vary with x, y, and t.

1.2. Spatial and Temporal Filter Equivalence

What is the effect of a given temporal frequency response? Let us start with a moving two dimensional impulse and observe the effect of temporal filtering on it's spectrum.

$$\delta\left((1\ 0\ 0)(x\ y\ t)^T\right) \cdot \delta\left((0\ 1\ 0)(x\ y\ t)^T\right) \qquad \text{(Eq. 12)}$$

where the multiplication of the two delta functions is interpreted in the generalized function sense as the limit of the product of a sequence of functions. This is a line impulse extending infinitely along the time axis. The Fourier transform is

$$\iint_{-\infty}^{\infty} \delta(x)\delta(y)\exp\left(-j2\pi(x\omega_x + y\omega_y + t\omega_t\right)dxdy \qquad \text{(Eq. 13)}$$

which is

$$\delta\left([0\ 0\ 1]\,[\omega_x\ \omega_y\ \omega_t]^T\right) = \delta(\omega_t) \qquad \text{(Eq. 14)}$$

Applying the linear transformation **A** to (Eq. 12) gives

$$\delta\left([1\ 0\ 0]\mathbf{A}[x\ y\ t]^T\right) \cdot \delta\left([0\ 1\ 0]\mathbf{A}[x\ y\ t]^T\right) \quad \text{(Eq. 15)}$$

which reduces to

$$\delta(x - tr_x)\delta(y - r_y t) \quad \text{(Eq. 16)}$$

This is line impulse lying in the x,t plane with equation $x = tr_x$. The Fourier transform of this line impulse from (Eq. 8) is

$$\delta\left([0\ 0\ 1]\left[\left(\mathbf{A}^{-1}\right)^T[\omega_x\ \omega_y\ \omega_t]\right]^T\right) \quad \text{(Eq. 17)}$$

or

$$\delta(\omega_x r_x + \omega_y r_y + \omega_t) \quad \text{(Eq. 18)}$$

which can also be written

$$\delta\left([r_x\ r_y\ 1][\omega_x\ \omega_y\ \omega_t]^T\right) \quad \text{(Eq. 19)}$$

Eq. 19 describes a plane impulse with normal $[r_x\ r_y\ 1]$. If this plane impulse is filtered by a temporal filter with impulse response $f(t)$ and Fourier transform $F(\omega_t)$ then the inverse transform of the filtered plane impulse is, assuming that $r_y = 0$,

$$f(x, y, t) = \iiint_{-\infty}^{\infty} \delta(\omega_x r_x + \omega_t)\, F(\omega_t)\, \exp\left(j2\pi\left(x\omega_x + y\omega_y + t\omega_t\right)\right) d\omega_x\, d\omega_y\, d\omega_t \quad \text{(Eq. 20)}$$

Integrating first with respect to ω_t

$$\int_{-\infty}^{\infty} \delta(\omega_x r_x + \omega_t)\, F(\omega_t)\, \exp(j2\pi t\omega_t)\, d\omega_t = F(-\omega_x r_x)\exp(-j2\pi\omega_x r_x t) \quad \text{(Eq. 21)}$$

Then integrating with respect to ω_x and ω_y

$$\int_{-\infty}^{\infty} F(-\omega_x r_x)\, \exp\big(j2\pi\omega_x(x - r_x t)\big)\, d\omega_x \int_{-\infty}^{\infty} \exp\big(j2\pi y\omega_y\big)\, d\omega_y \qquad \text{(Eq. 21)}$$

which gives

$$\frac{1}{|r_x|} f\left(t - \frac{x}{r_x}\right)\delta(y) \qquad \text{(Eq. 22)}$$

What does this equation mean? Motion effectively projects the temporal impulse response onto the image plane yielding an equivalent spatial impulse response. The temporal impulse response is scaled in duration by r_x and in amplitude by $1/r_x$. As r_x increases the spatial impulse response increases in length and vice versa. Assume that the low pass temporally filtered signal is sampled at a time t_0. Then the equivalent spatial impulse response at time t_0 is

$$\frac{1}{|r_x|} f\left(t_0 - \frac{x}{r_x}\right) \qquad \text{(Eq. 23)}$$

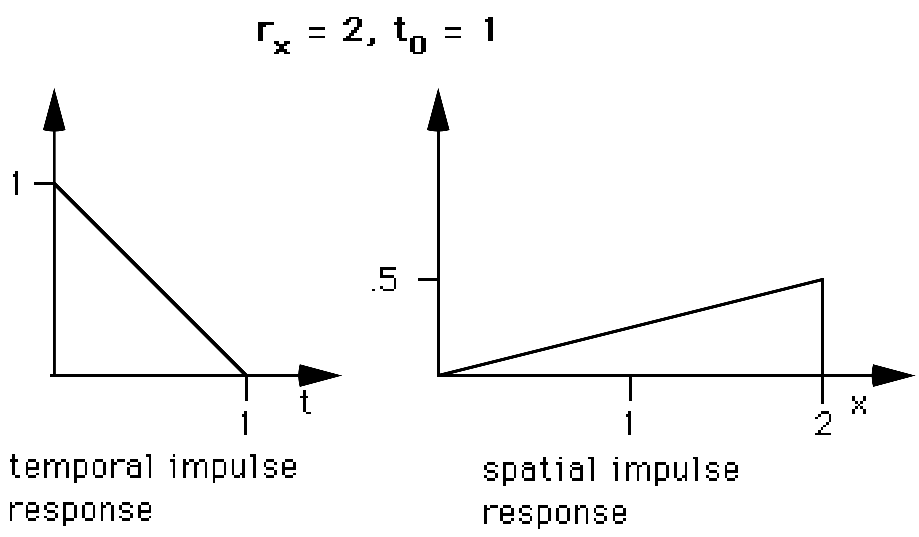

Figure 3

So to remove temporal aliasing artifacts from a moving image at time t_0 the moving image can either be temporally filtered or a static image displaced by $r_x t_0$ can be spatially filtered with the projected filter function. This can easily be generalized to motion in an arbitrary direction. Rewrite the initial spatio-temporal motion function as

$$\delta\left([1\ 0\ \text{-}r_x][x\ y\ t]^T\right)\delta\left([0\ 1\ 0][x\ y\ t]^T\right) \qquad \text{(Eq. 25)}$$

The vectors $[1\ 0\ \text{-}r_x]$ and [0 1 0] are the normals to the plane impulses $\delta(x\text{-}r_x t)$ and $\delta(y)$. To rotate the plane impulses we choose a transformation **B** such that the normal is rotated counterclockwise about the t axis by an angle θ

$$\mathbf{B} = \begin{bmatrix} \cos\theta & \sin\theta & 1 \\ -\sin\theta & \cos\theta & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(Eq. 26)}$$

which gives the rotated impulse planes

$$\delta\left([1\ 0\ \text{-}r_x]\mathbf{B}[x\ y\ t]^T\right)\delta\left([0\ 1\ 0]\mathbf{B}[x\ y\ t]^T\right) \qquad \text{(Eq. 27)}$$

From (Eq. 8) F(g(**Ax**)) has Fourier transform $\mathbf{G}\left((\mathbf{A}^{-1})^T\omega\right)$. Rotation matrices are orthonormal so $\mathbf{G}\left((\mathbf{A}^{-1})^T\omega\right)$ = G(**A**ω). If the direction of motion is rotated by the matrix **A** then the transform is also rotated by **A**. Since the rotation of the impulse planes is about the t axis the rotation in the frequency domain will be about the ω_t axis. Because of this the operations of temporal filtering and rotation of the spectrum are commutative. In other words the unrotated spectrum could be temporally filtered and then rotated or could be rotated and then temporally filtered and both would yield the same result.

If different parts of the image are moving at different velocities then the projected spatial filter will vary as a function of x and y and t. Occlusion of one object by another makes the operation nonlinear and the correct temporal filter can only be approximated by a simple projected temporal filter.

1.3. Relationship Between Spatial and Temporal Bandwidth

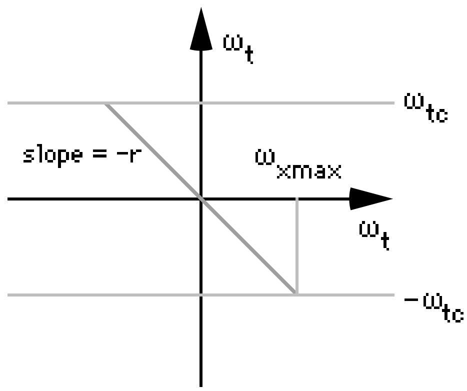

Figure 4

The relationship between temporal bandwidth, image velocity and spatial bandwidth is now easy to calculate. Assume that the signal is band limited in the x dimension to ω_{xmax} and in the time dimension to ω_{tc}. If the image is moving with velocity r then until $r\omega_{xmax} = \omega_{tc}$ there will be no reduction in the spatial resolution of the image. When $|r\omega_{xmax}| > \omega_{tc}$ or equivalently when $|r| > \omega_{tc} / \omega_{xmax}$, then the image will be spatially band limited in the x dimension with bandwidth $\omega_{xc} = \omega_{tc}/| r |$. Let r_c, the critical speed, be defined as $r_c = \omega_{tc} / \omega_{xmax}$. Spatial resolution is thus a hyperbolic function of image velocity once the critical speed has been passed and is independent of velocity below the critical speed. As an example assume the signal is to be sampled at the normal television rate of 30 frames per second and that the spatial bandwidth in x and y is 256 cycles per picture height and width. If the image is to be temporally band limited to 15 cycles per second - the maximum possible without introducing aliasing - then the critical velocity is 1 pixel per frame or 30 pixels per second. Table 1 shows the maximum possible spatial resolution in the direction of motion for various velocities.

image velocity		maximum spatial resolution
pixels/frame	pixels/sec.	cycles/picture height
2	60	128
4	120	64
16	480	8

Table 1

Spatial resolution falls rapidly as image velocity increases and is quite low even at low speeds. If the signal is not temporally low pass filtered before sampling and if w_{tc} is half the sampling frequency then r_c is the maximum image speed which can be reconstructed without aliasing if a rectangular low pass reconstruction filter is used. This implies that high temporal sampling rates are necessary if no temporal prefiltering is performed.

1.4. Temporal Filter Design for Antialiasing

Since image motion causes the temporal impulse response to be projected onto the image plane the temporal filter must meet many of the same subjective criteria required of spatial filters. An ideal low pass spatial filter produces visually unacceptable ringing artifacts at high contrast boundaries in the image. When objects are moving slowly enough to be tracked by the eye the image will be static on the retina and the projection of the temporal impulse response will be perceived solely as a spatial variation in brightness. This means that an ideal low pass temporal filter will not be perceptually acceptable for image motion slow enough to be tracked.

Generally the wider the transition band of a low pass filter the less ringing is present in the impulse response. A wide transition band causes either the loss of high frequencies in the pass band or imperfect rejection of spectral replicas or a combination of the two. The low pass filter required for temporal antialiasing will therefore be a tradeoff between eliminating spectral replicas, maximizing bandwidth, and minimizing ringing. The greater the temporal bandwidth the sharper moving objects will appear. The sharper the cutoff the more ringing will be projected onto the image plane. If elimination of spectral replicas must be accomplished and a wide transition band is used to minimize ringing then the bandwidth will necessarily be reduced. A compromise solution might center the transition band at half the sampling frequency so that maximum bandwidth is achieved and spectral replicas are still substantially reduced.

---------- References ----------

MARSDEN 1976 Marsden, Jerrold E., and Anthony Tromba. Vector Calculus. W.H. Freeman and Company 1976

WATSON 1985 Watson, Andrew B. and Albert J Ahumada Jr. Model of Human Visual Motion Sensing. Journal of the Optical Society of America A, vol. 2, no. 2, February 1985, pp. 322-342.