Spatiotemporal power spectra of motion parallax: the case of cluttered 3D scenes

Derek Rivait and Michael S. Langer

School of Computer Science, McGill University
3450 University St. Montreal, Canada
Email: {drivai,langer}@cim.mcgill.ca

ABSTRACT

We examine the spatiotemporal power spectra of image sequences that depict dense motion parallax, namely the parallax seen by an observer moving laterally in a cluttered 3D scene. Previous models of the spatiotemporal power have accounted for effects such as a static 1/f spectrum in each image frame, a spreading of power at high spatial frequencies in the direction of motion, and a bias toward either lower or higher image speeds depending on the 3D density of objects the scene. Here we use computer graphics to generate a parameterized set of image sequences and qualitatively verify the main features of these models. The novel contribution is to discuss how failures of 1/f scaling can occur in cluttered scenes. Such failures have been described for the spatial case, but not for the spatiotemporal case. We find that when objects in the cluttered scene are visible over a wide range of depths, and when the image size of objects is smaller than the image width, failures of 1/f scaling tend to occur at certain critical frequencies, defined by a correspondence between object size and object speed.

Keywords: motion parallax, power spectrum, 1/f scaling, natural images

1. INTRODUCTION

This paper addresses image sequences seen by an observer moving laterally relative to a cluttered 3D scene. Examples are an observer walking through the woods and looking off to the side, or a static observer viewing falling snow. Such image sequences are commonly observed in nature, yet are complex since they exhibit many discontinuities. Despite the complexity of such image sequences, however, natural vision systems appear to function well in 3D cluttered scenes. This has been shown formally in psychophysical experiments, for example, observers have little trouble judging direction of heading through a 3D cloud of dots.[1] It is also obvious from observing animals in the wild. Many fast moving animals – in particular, birds – seem to have little trouble navigating through cluttered environments such as the woods.

In this paper, we analyze the spatiotemporal statistics of image sequences of cluttered scenes, in particular, their spatiotemporal power spectrum. In Sec. 2, we review several power spectrum models that have been presented in the literature, and relate these to the specific case in question, namely lateral observer motion in a 3D cluttered scene. In Sec. 3, we present experimental results using computer graphics images.

2. POWER SPECTRUM MODELS OF NATURAL IMAGES

Many studies of statistics of natural images are based either on the image autocorrelation function or its Fourier transform, in particular, the power spectrum. These models capture the fact that nearby pixels tend to have similar intensities, in particular, the autocorrelation function of static natural images is a decreasing function of the distance between pixels.[2,3] Expressing the autocorrelation function by its Fourier transform is convenient for several reasons. It connects the statistics of natural images with linear systems models of image processing[4,5] This allows one to relate the models to psychophysical data such as contrast sensitivity functions and to information theoretic constraints such as efficient coding of the image in the presence of noise.[6–11] In this section, we review several models of natural image power spectra.

2.1. Spatial power spectra

We begin with static images. Let $I(x, y)$ denote the image intensity at pixel (x, y). Let f_x and f_y denote the spatial frequencies in the x and y directions. The Fourier transform of the image is:

$$\hat{I}(f_x, f_y) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) e^{-i\frac{2\pi}{N}(f_x x + f_y y)} \tag{1}$$

and the power spectrum is:

$$P(f_x, f_y) \equiv |\hat{I}(f_x, f_y)|^2. \tag{2}$$

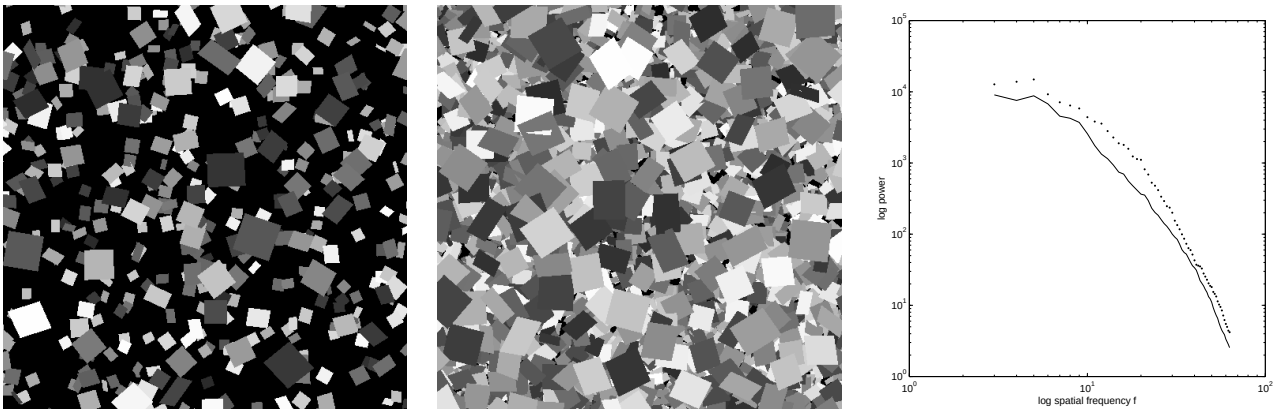
Field[4] and others[12–14] have shown that the power spectra of natural images, when averaged over all orientations* of a given spatial frequency, $f = \sqrt{f_x^2 + f_y^2}$, behaves approximately as a power law:

$$P(f) \sim f^{-m} \tag{3}$$

which is linear with slope $-m$ on a loglog plot. The slope m depends on many factors, such as depth of field.[13, 16, 17]

Although many natural images do have a static spectrum that behaves as f^{-m}, there are exceptions. For example, consider an image of a cluttered scene such as a bush whose surface elements are distributed over a finite view volume. Synthetic examples are shown in Fig. 1, along with a loglog plot of their orientation averaged power spectra. The log-log plots are clearly not linear, as would be predicted from a f^{-m} scaling model. Rather, they become flatter at lower spatial frequencies. The reason for such failures of f^{-m} scaling is that all the objects have a limited width and so the image intensities are statistically independent (and hence uncorrelated) at large image distances.[18, 19] In this paper, we extend this concept to spatio-temporal case.

Before consider the temporal domain, we note that images of Fig. 1 are similar to those of the "dead leaves" model which has been used in recent studies of natural images statistics.[14] One key difference between our model and the dead leaves model is that we are using a perspective projection, whereas the dead leaves model uses orthographic projection. As we will see, perspective projection is crucial for the arguments in this paper, which depend on a relationship between depth, image size, and image speed.



(a) $M = 2000$ (b) $M = 16000$ (c)

Figure 1. (a,b) Two scenes with a different number of squares M, viewed in perspective against a black scene background. (c) A loglog plot of the orientation-averaged power spectrum of several such scenes (dotted line for $M = 2000$, solid line for $M = 16000$. Low spatial frequencies not shown ($f = 0, 1, 2$).

*Static orientation dependencies can occur in some scenes,[13, 15] but we ignore them in this paper.

2.2. Spatiotemporal power spectra

Turning now to image sequences, let $I(x, y, t)$ denote the image intensity at pixel (x, y) and frame t. The 3D Fourier transform is:

$$\hat{I}(f_x, f_y, f_t) = \sum_{x,y=0}^{N-1} \sum_{t=0}^{T-1} I(x, y, t) e^{-i\frac{2\pi}{N}(f_x x + f_y y) - i\frac{2\pi}{T} f_t t} \tag{4}$$

where f_t is the temporal frequency.

Eckert *et al.*[20] modelled the power spectra of individual natural image sequences as products of spatial and temporal frequency. Their model was expressed in the spatial domain by an autocorrelation function,

$$R(r, t) \sim e^{-ar} e^{-bt}$$

where $r = \sqrt{x^2 + y^2}$. This autocorrelation yields a power spectrum which is separable in spatial frequency $f = \sqrt{f_x^2 + f_y^2}$ and temporal frequency f_t,

$$P(f, f_t) = \frac{ab}{(a^2 + f^2)^{\frac{3}{2}}(b^2 + f_t^2)}.$$

This power spectrum model behaves roughly as a power law for large f and large f_t. However, for $f \to 0$ and $f_t \to 0$, the spectra become flat. The constants a, b define spatial and temporal scales above which the spectrum roughly obeys a power law and below which the spectrum becomes flat and behaves more like white noise. A similar flattening of the power spectrum at low spatial frequencies can be seen in Fig. 1.

The Eckert *et al.* model does not explicitly address the motion that occurs in image sequences. As Watson and Ahumada observed, an object translating with velocity (v_x, v_y) in a image produces a spatiotemporal component in the power spectrum with support confined to a *motion plane*,[21]

$$v_x f_x + v_y f_y + f_t = 0.$$

If an image sequence contains many image velocities, then each velocity contributes power to its corresponding motion plane.[†] This is the case we address in this paper.

For the case of an observer moving laterally in a static 3D scene, all these image velocities are in the same direction.[24] For example, suppose the observer is moving horizontally (in the x direction) and the observer's 3D speed in this direction is V. Assuming the projection plane is at depth $z = 1$ unit, the horizontal image speed v_x for a visible object at depth z is

$$v_x = -\frac{V}{z} \tag{5}$$

and the motion plane corresponding to such an object is

$$v_x f_x + f_t = 0. \tag{6}$$

If objects are present at a range of z, then a family of motion planes of various speeds v_x results.[24] All of these motion planes intersect at the f_y axis. Langer and Mann referred to such motion as *optical snow*.

Next, assume the 3D scene is cluttered and scene objects are uniformly distributed over 3D space, with a density parameter α. Since objects are opaque, near objects occlude far objects along any line of sight (or pixel). Van Hateren[7] proposed a model to account for the resulting relationship between depth, occlusions, and visibility as follows. For any pixel, the probability that an object at depth z is visible at that pixel is roughly[‡]

$$p(z) = \alpha e^{-\alpha z}.$$

[†]Strictly speaking, since image regions of a fixed velocity are bounded in space-time, these regions contribute power to a *slab* of motion planes rather than to a single motion plane,[22–24] where the thickness of the slab in inversely related to size of the region.

[‡]For this model to be correct, depth should be measured along a ray rather than strictly in the z direction. To keep the model simple, we ignore this discrepancy. This is not a problem provided the field of view is sufficiently narrow.

With a change of variables, the model relates the probability $p(z)\ dz$ that the visible object lies in the depth interval $[z, z+dz]$ to the probability $p(v_x)\ dv_x$ that a visible object has image speed in $[v_x, v_x+dv_x]$. Since $z = \frac{V}{v_x}$, we have

$$dz = -\frac{V}{v_x^2} dv_x \ .$$

Thus,

$$p(z)dz \sim e^{-(\alpha \frac{V}{v_x})} \frac{V}{v_x^2} dv_x$$

and so

$$p(v_x) = e^{-(\alpha \frac{V}{v_x})} \frac{V}{v_x^2}. \tag{7}$$

This probability density over image speeds v_x is now expressed as a probability of power falling on motion planes of slope $v_x = \frac{V}{z}$. For a fixed f_x,

$$dv_x = \frac{df_t}{f_x}$$

and so, for fixed f_x, f_y, the probability density of a visible object contributing power to frequencies in $[f_t, f_t + dt]$ is obtained by substitution,

$$p(f_t \mid f_x, f_y)\ df_t \sim V e^{-\alpha \frac{V f_x}{f_t}} \cdot \frac{f_x^2}{f_t^2} \frac{1}{f_x}\ df_t \ .$$

This probability is the product of three terms:

- an exponential decay term which depends on the density of objects in the scene. Higher density α leads to more power at high image speeds, since slow distant objects would tend not to be visible.
- an inverse square dependence on image speed f_t/f_x in the direction of motion. Because objects are uniformly distributed in 3D, the probability that a randomly chosen object (possible visible, possibly not) is far away is relatively high. Note that this term works in the opposite direction as the previous term.
- an inverse dependence on spatial frequency f_x, namely a spreading of the power spectrum away from $f_x = 0$. All motion planes intersect at the f_y axis[24] and so power is expected to be concentrated on this axis and to spread away from this axis.[25]

To relate this probability model to the power spectrum, the image intensities must also be considered. In previous work[7,11] which considered orientation averaged spectra, this was achieved by multiplying by a static radially symmetric power spectrum f^{-m}. One could apply this idea to lateral motion only as well. Letting $df_t = 1$, we get a model for the spatiotemporal power at each frequency f_x, f_y, f_t :

$$P(f_x, f_y, f_t) \ = \ f^{-m}\ V\ e^{-\alpha V \frac{f_x}{f_t}} \cdot \frac{f_x^2}{f_t^2}\ \frac{1}{f_x}\ . \tag{8}$$

This model is a separable product of terms that depend on spatial frequency f_x, f_y only, and terms that depend on the speed $v_x = -\frac{f_t}{f_x}$. A similar separability observation was made for orientation-averaged spectrum.[11]

This model is not necessarily appropriate for the cluttered scenes that we are considering, however. As we showed in Figure 1, failures of f^{-m} scaling in the static spectrum can occur in cluttered scenes, when the image width of visible objects is relatively small in comparison to the width of the image. One might be tempted to substitute an alternative static power spectrum which flattens out a low spatial frequencies, for example, a model similar to the spatial frequency component of the model of Eckert *et. al.* As we argue next, however, this does not quite do the trick and a slightly more subtle model is required.

Our argument here is similar to one developed by Langer *et al.*[26] Suppose objects in the scene are all roughly of the same 3D size, and are again seen by a laterally translating observer. Because of linear perspective, objects that are further away will have smaller image width *and* will have lower image speed (see Eq. (5)). Thus, for

any speed v_x, the objects that contribute to the corresponding motion plane will have an image width that is inversely related to their depth. Since the image width of objects defines the spatial frequency f below which failures of f^{-m} scaling occur, *different motion planes should exhibit such failures at different spatial frequencies.*

To be more concrete, suppose[§] that within the motion plane that corresponds to some depth z the power spectrum is flat up to some critical spatial frequency f^* and then falls off as f^{-m} for $f > f^*$. Let f_x^*, f_y^* satisfy $f^* = \sqrt{f_x^{*2} + f_y^{*2}}$. From Eqs. (5) and (6),

$$v_x = \frac{f_t^*}{f_x^*} \sim \frac{1}{z}. \tag{9}$$

How many we interpret the critical temporal frequency f_t^* ? Just as the critical spatial frequency f_x^* is inversely proportional to the image width of objects that contribute to a motion plane, the critical temporal frequency f_t^* should be proportional to the temporal duration over which such objects are visible at a single pixel. Since size is proportional to image speed and $\Delta t = \Delta x / v_x$, it follows that the duration that an object falls on a pixel should be constant[¶]. Thus, we expect that the critical temporal frequency f_t^* to be the same for each motion plane. Another way to make this argument is as follows. Since the critical spatial frequency f^* for a given motion plane is proportional to depth (far objects appear smaller), we have $1/z \sim 1/f^*$. Combining this proportionality relationship with Eq. (9) yields:

$$f_t^* \sim \frac{f_x^*}{\sqrt{f_x^{*2} + f_y^{*2}}}. \tag{10}$$

In the case that $f_y^* = 0$, we have that f_t^* is a (unspecified) constant. This unspecified constant depends on the 3D size of the objects and on the angular width of the image.

Once this constant is specified, Eq. (10) defines a non-planar surface in the 3D frequency domain. For any motion plane, the intersection of this surface with the motion plane partitions the motion plane into two sets, namely those with spatial frequencies below and above f^*. Frequencies well below f^* should have a roughly flat power spectrum and those well above f^* should have a f^{-m} spectrum. In the next section, we present an experiment using computer graphics sequences that examines this model.

3. EXPERIMENTS

3.1. Image sequences

Image sequences were generated using OpenGL. Each sequence consisted of 128 image frames each of size 128×128 pixels. Anti-aliasing was achieved by rendering the sequences at three times this resolution in each XYT dimension, then blurring and subsampling [‖]. Prior to computing the 3D Fourier transform of each sequence, the sequence was multiplied by a Hanning window in each of the x, y, t dimensions.

Each 3D cluttered scene consisted of M squares, randomly distributed in a cube. This cube bounded the union of the view volumes of the laterally moving observer (see below). Each square had random uniform reflectance (between .4 and 1) and was illuminated by constant ambient light. Each square had width 0.05 units and the surface normal was oriented randomly in 3D, subject to the constraint that the maximum slant was 60 degrees with respect to the viewing direction. The camera field of view was 30 degrees wide. The near and far clipping planes of the view volume were located at depths of $z = 1$ units and $z = 4$ respectively. The speed of the observer was chosen so that objects in the view volume had image speeds v_x between 0.5 and 2 pixels per frame. Frames from two sequences with different number of objects M were shown in Fig. 1.

[§] In practice, there is no single critical spatial frequency – see Fig. 1 – and so one would need to define it as a parameter, as in Eckert *et al.*'s model.

[¶] Here we are considering only unoccluded objects.

[‖] We blur with a square box kernel of width 6 pixels, and subsample by a factor of three.

3.2. Results

We study how the power spectrum was distributed both across and within motion planes. The method we use is similar to that of Dong and Atick[11] who studied the distribution of orientation-averaged power as a function of speed $v_x = \omega/f$, and who observed a common scaling behaviors at different speeds. The key difference in our model is that we are considering a more specific scenario, namely 3D cluttered scenes and lateral motion only. This scenario leads to a different analysis and to different results than those found by Dong and Atick.

Like Dong and Atick, we consider temporal frequency and one spatial frequency dimension only. However, unlike Dong and Atick who considered orientation-average spectra, we consider the spatial frequency f_x in the direction of motion and we examine the power in the $f_y = 0$ plane only, that is, on the intersection of each v_x-motion plane with the $f_y = 0$ plane. This intersection is a line in the $f_y = 0$ plane, namely the line $f_t = -v_x f_x$. For any v_x motion plane, we interpolate the power spectrum along this line from frequency samples above and below the line. Let $\hat{P}(f_x \mid v_x)$ denote this interpolated power. We examine how power varies across motion planes by plotting the total** power along the lines $f_t = -v_x f_x$ for various speeds v_x, namely we plot

$$\sum_{f_x=2}^{N/4} \hat{P}(f_x \mid v_x)$$

as a function of v_x. Figure 2 shows this summed power plot for two different object densities, namely $M = 2000, 16000$ as in Figure 1, and averaged over 20 sequences.

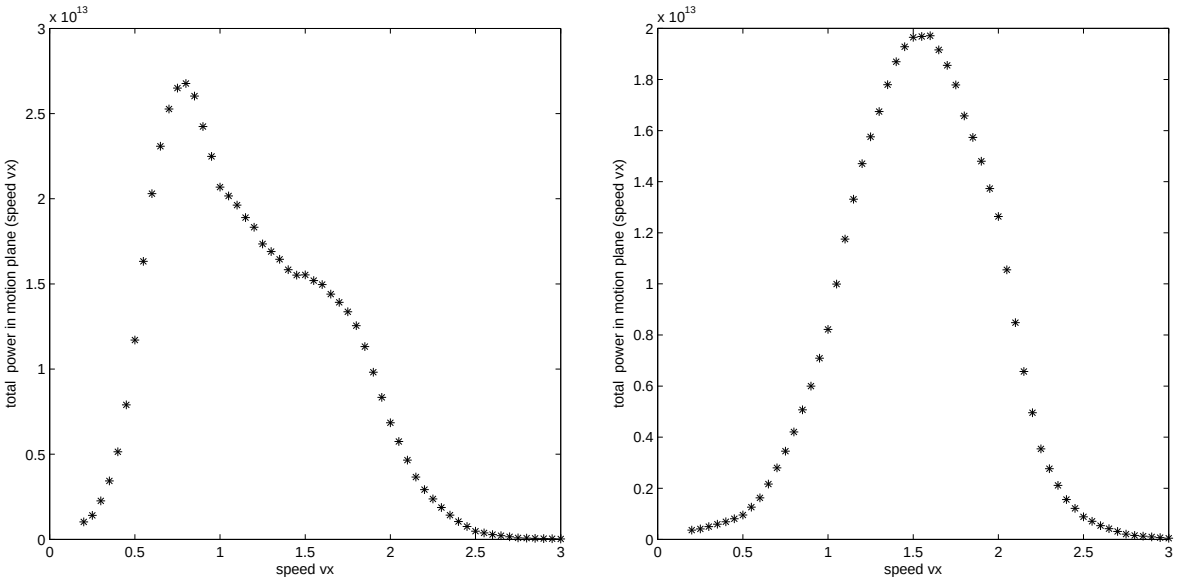


Figure 2. Total power is shown for all frequencies that lie in the intersection of planes $f_t = -v_x f_x$ and $f_y = 0$, for two different numbers M of objects in the scene. For larger M, the power shifts to higher speeds. See text.

The results are qualitatively what we expect. For a smaller number of objects M, there is more power at slower speeds since far objects are more likely to be visible. This is the $(f_x/f_t)^2$ bias of Eq. (8). When the number of objects M is increased to 16000 (twice the density in each xyz dimension), the peak in the curve shifts toward higher image speeds. This shift is entirely due to occlusions. When more objects are present in the scene, each pixel is more likely to see a nearby object. This biases towards faster speeds since nearby objects move

**We do not include $f_x = 0, 1$ in the sum because there is a large dc component in the images which is aliased to $f = 1$ by the Hanning window.

faster in the image (see Eq. 5). Since our arguments about critical frequencies f_x^* and f_t^* ignore occlusions, we would not make as strong a prediction that f_t^* is constant in the more crowded scene condition ($M = 16000$). This is indeed what we will find.

Note that for both values of M, there is some power in motion planes at slopes less than 0.5 and greater than 2.0, even though the speeds of objects is restricted to [0.5,2.0]. This leakage is due to both occlusions[22–24] and to the interpolation function.

We next address the power distribution *within* motion planes. Dong and Atick[11] assumed a f^{-m} model for the static spectrum. If we were to substitute such a model in Eq. (8), we would get

$$\hat{P}(f_x|v_x) \sim 1/f_x^{m+1}$$

since $v_x = -f_t/f_x$ is fixed within each motion plane, and the assumed f^{-m} static power spectrum of Eq. (8) is f_x^{-m} when $f_y = 0$. As we discussed earlier, however, in the case of a cluttered scene a f^{-m} scaling model is not appropriate at low spatial frequencies. What model can we use to better account for the scaling behavior in this case?

According to the arguments we made that led to Eq. (10), we expect the spectrum *within* each motion plane to be flat up to some critical frequency f_x^* (or f_t^*), then to fall off as f_x^{-m} (or f_t^{-m}) for higher spatial frequencies. In particular, recall that we expect the critical temporal frequency f_t^* to be the same for all motion planes. Figure 3(a) shows loglog plots of the product $f_x\ \hat{P}(f_x|v_x)$ for various v_x. (We are multiplying by f_x to cancel the spreading of the power away from the $f_x = 0$ – see bullet 3 above Eq. 8). Figure 3(b) shows the same data, but now it is plotted as a function of f_t, that is, a loglog plot of $f_t/v_x\ \hat{P}(f_t, v_x)$ versus f_t.

In both plots, the curves are flat at low frequencies and then fall off at higher frequencies. However, in Figure 3(a), the falloff occurs at different critical *spatial* frequencies f_x^* for different curves, *i.e.* the curves separate from each other, whereas in Figure 3(b), the critical *temporal* frequencies are roughly the same for all curves. In (b), the curves differ mainly by a vertical shift[††] which is due to different amounts of total power at different speeds v_x (recall Fig. 2).

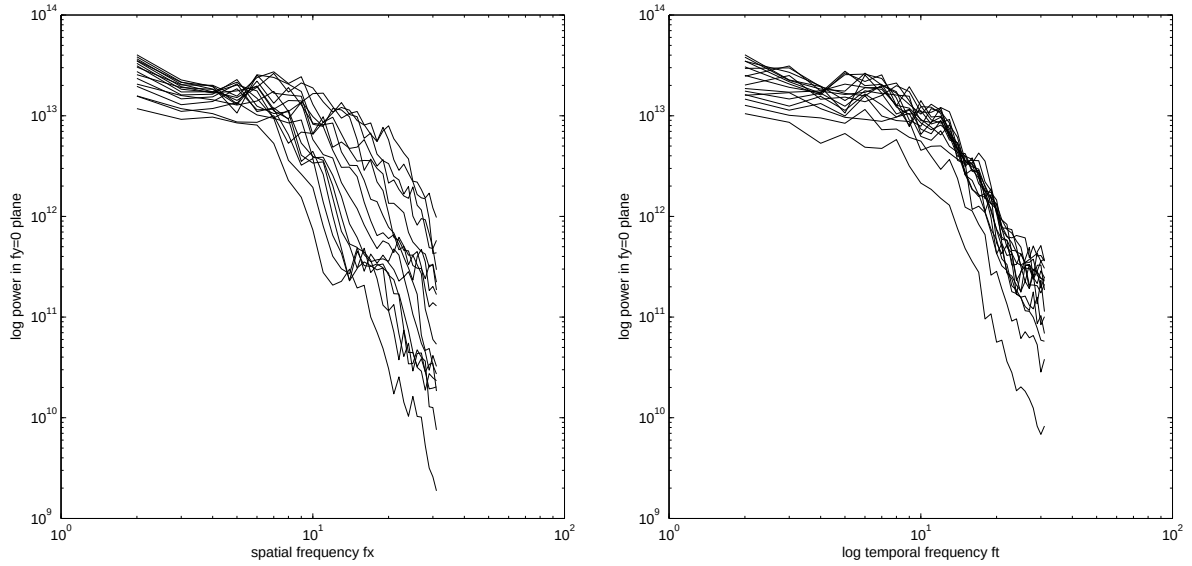
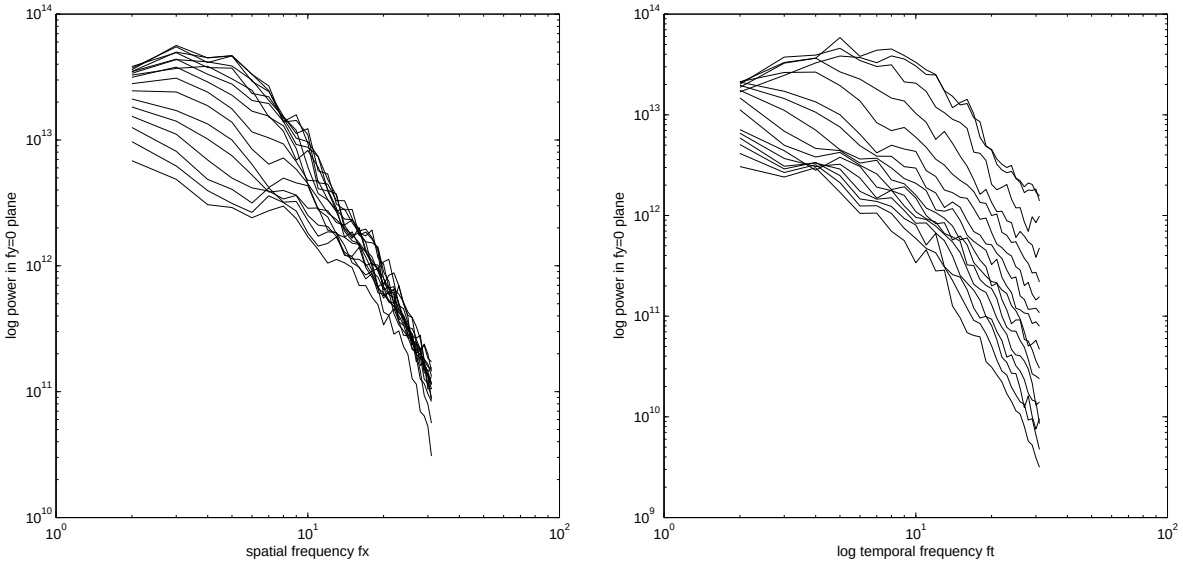


Figure 3. Total power within motion planes is plotted as a function of (a) f_x and (b) f_t for 20 sparse scenes ($M = 2000$). See text for details.

[††]A vertical shift in the loglog domain corresponds to scalar multiplication of $\hat{P}$.

Fig. 4 shows corresponding plots for the dense scene ($M = 16000$). Here the effect is less pronounced, namely the curves in (b) are not merely vertical shifts but rather also seem to have different shapes. We believe the reason for the slightly different effect here is that in the dense scene, occlusions are more severe. Further investigation is required to clarify this case.



(a) (b)

Figure 4. Total power within motion planes is plotted as a function of (a) f_x and (b) f_t for 60 dense scenes ($M = 16000$).

4. CONCLUSION

Power spectrum analysis has a long tradition in the study of the statistics of natural images. In this paper, we have addressed the detailed structure of the spatio-temporal power spectrum of a common class of image sequences, namely those seen by an observer moving laterally in a 3D cluttered scene. We reviewed several models of power spectra, both spatial and spatiotemporal, and applied these to the scenario in question. We showed that these models do not correctly account for failures of f^{-m} scaling in the static spectrum, which can occur in cluttered 3D scenes in which the objects have roughly constant 3D size.

Our main new result is that for cluttered scenes that are relatively sparse – namely sparse enough that far objects are likely to be visible despite occlusions – it is possible to account for the failures of f^{-m} scaling in the spatiotemporal spectrum in a simple way. For any motion plane in the frequency domain, the static power spectrum contributed by that motion plane is flat up to a critical frequency f^* and then behaves as f_x^{-m} beyond that frequency. In particular, the critical temporal frequency f_t^* is roughly the same for all motion planes, whereas the critical spatial frequency f_x^* varies between motion planes. Our arguments were supported with experiments using computer graphics. In future work, we will test the predictions of the model using real image sequences.

Acknowledgements

This research was supported by an NSERC PGS1 grant to Derek Rivait and by an NSERC Discovery Grant to Michael Langer.

REFERENCES

1. L. Li and W. H. Warren, "Perception of heaing during rotation: sufficiency of dense motion parallax and reference objects," *Vision Research* **40**, pp. 3873–3894, 2000.
2. E. R. Kretzmer, "The statistics of television signals," *Bell System Tech. J.* **31**, pp. 751–763, 1952.
3. S. B. Laughlin, "A simple coding procedure enhances a neuron's information capacity," *Z. Naturforsh.* **36c**, pp. 910–912, 1981.
4. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A* **4**, p. 2379, 1987.
5. J. H. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proceedings of the Royal Society of London B* **265**, pp. 359–366, 1998.
6. M. V. Srinivasan, S. B. Laughlin, and A. Dubs, "Predictive coding: A fresh view of inhibition in the retina," *Proceedings of the Royal Society of London* **B 216**, pp. 427–459, 1982.
7. J. H. van Hateren, "Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation," *Journal of Comparative Physiology A* **171**, pp. 157–170, 1992.
8. D. J. Field, "What is the goal of sensory coding?," *Neural Computation* **6**, pp. 559–601, 1994.
9. M. P. Eckert and G. Buchsbaum, "Efficient coding of natural time varying images in the early visual system," *Philosophical Transactions of the Royal Society (London)* **B339**(1290), pp. 385–395, 1993.
10. D. L. Ruderman, "Designing receptive fields for highest fidelity," *Network: Computation in Neural Systems* **5**, pp. 147–155, 1994.
11. D. W. Dong and J. J. Atick, "Statistics of natural time-varying images," *Network: Computation in Neural Systems* **6**(3), pp. 345–358, 1995.
12. D. L. Ruderman and W. Bialek, "Statistics of natural images: scaling in the woods," *Physical Review Letters* **73**, pp. 814–817, 1994.
13. A. van der Schaaf and J. H. van Hateren, "Modelling the power spectra of natural images: Statistics and information," *Vision Research* **36**(17), pp. 2759–2770, 1996.
14. D. L. Ruderman, "Origins of scaling in natural images," *Vision Research* **37**(23), pp. 3385–3398, 1997.
15. E. Switkes, M. J. Mayer, and J. A. Sloan, "Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis," *Vision Research* **18**, pp. 1393–1399, 1978.
16. D. J. Tolhurst, Y. Tadmor, and T. Chao, "Amplitude spectra of natural images," *Ophthal. Physiol. Opt.* **12**, pp. 229–232, April 1992.
17. D. J. Field and N. Brady, "Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes," *Vision Research* **37**(23), pp. 3367–3383, 1997.
18. R. M. Balboa, C. W. Tyler, and N. M. Grzywacz, "Occlusions contribute to scaling in natural images," *Vision Research* **41**(7), pp. 955–964, 2001.
19. M. S. Langer, "Large scale failures of $f^{-\alpha}$ scaling in natural image spectra," *Journal of the Optical Society of America A* **17**(1), pp. 28–33, 2000.
20. G. B. M. P. Eckert and A. B. Watson, "Separability of spatiotemporal spectra of image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(10), pp. 1210–1213, 1992.
21. A. Watson and A. Ahumada, "Model of human visual-motion sensing," *Journal of the Optical Society of America A* **2**(2), pp. 322–342, 1985.
22. D. J. Fleet and K. Langley, "Computational analysis of non-fourier motion," *Vision Research* **34**(22), pp. 3057–3079, 1994.
23. S. Beauchemin and J. Barron, "The frequency structure of 1D occluding image signals," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, pp. 200–206, February 2000.
24. M. Langer and R. Mann, "Optical snow," *International Journal of Computer Vision* **55**(1), pp. 55–71, 2003.
25. V. Chapdelaine-Couture, S. Roy, M. S. Langer, and R. Mann., "Principal components analysis of optical snow," in *British Machine Vision Conference*, pp. 799–808, (London U.K.), 2004.
26. M. Langer, L. Zhang, A. Klein, A. Bhatia, J. Pereira, and D. Rekhi, "A spectral-particle hybrid method for rendering falling snow," in *Rendering Techniques (Eurographics Symposium on Rendering)*, pp. 217–226, (Norrkoping, Sweden), 2004.