

Movies of data-lag histograms with application to deconvolution research

Rick Ottolini and Fabio Rocca

Data-lag histogram movies

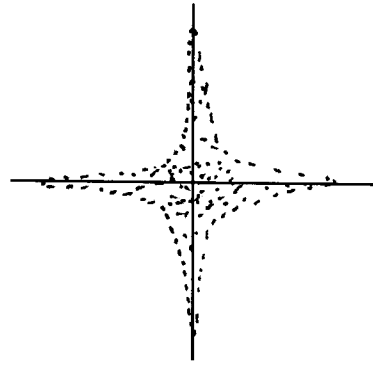
A data-lag histogram graphs points at the coordinates defined by the value of two data samples separated by a given lag. The graph may be contoured or colored by the density of points per unit area to make the display easier to interpret.

A histogram image allows one to directly visualize certain statistical aspects of a dataset such as correlation, phase, and distribution function. Data samples at large lags are independent and can be modeled by the generalized Gaussian

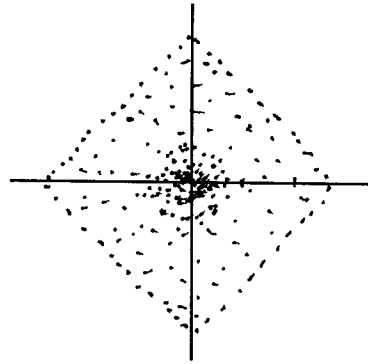
$$\frac{\alpha}{2\beta\Gamma(\alpha^{-1})} \exp\left(-\frac{|x|}{\beta}\right)^\alpha.$$

As the shape parameter α deviates greater than or less the normal distribution, $\alpha=2$, the histogram becomes more boxy or pointy (figure 1a-1d). Correlation appears as an elongation along the diagonal (figure 2). The phase is a rotation about an axis (figure 1f). Sometimes phase is difficult to distinguish from the shape parameter.

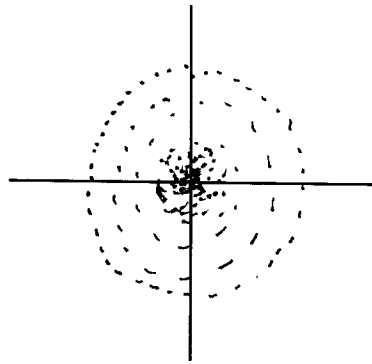
Animating data-lag histograms makes it easier to visualize these and other patterns. Such movies are being used to estimate statistical parameters of a dataset and study the performance of data processing algorithms. Gray (1979) and Godfrey (1979) used multi-dimensional histograms for seismic data. Walden & Hosken (1985) looked at data-lag histograms of reflectivity series.



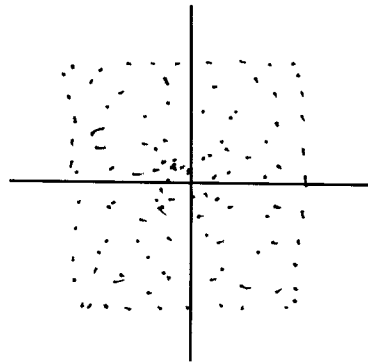
(1a) $\alpha < 1$



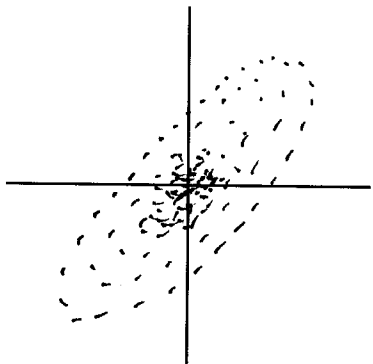
(1b) $\alpha = 1$



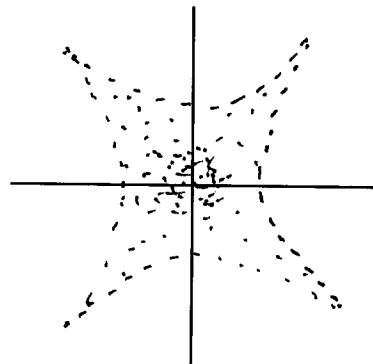
(1c) $\alpha = 2$



(1d) $\alpha > 2$



(1e) correlated



(1f) phase shift

FIG. 1. Schematic joint histograms illustrating the effect of the shape parameter, correlation, and phase.

Characteristics of field data movies

We constructed histograms of the forty Western Geophysical field profiles at the first fifty lags. We applied a time power gain correction and combined all the traces in each profile. We also looked at a sonic log and synthetic distributions.

Figure 2 illustrates a typical result. Each field data histogram has a high density inner core representing small amplitudes correlated with small amplitudes and a low density outer halo representing the strong amplitude data. A histogram movie through increasing lag appears as wobbling ellipses gradually contracting into steady circles. The wobbling ellipses means the data is highly correlated at small lags.

A transverse slice through the ellipses (figure 3) reveals a sinusoidal 'wavelet'. This is probably the autocorrelation of the source wavelet. The autocorrelation is the integral of the histogram image with each point weighted by the product of its coordinates. The dense core is a partial integration or autocorrelation. The process of watching a movie is yet another partial integration. The small data values in center of the histogram are more likely to be the whitest part of the reflectivities or noise. Then the convolutional model says that the autocorrelation is the wavelet. The wavelet character corresponds with what would be expected from the given source and geology of the forty Western profiles.

The low density outer halo has interesting features of its own. This is where the strong reflectors, or 'signal', may be. Often we see a diffuse halo 'wavelet' closely mimicking the core wavelet. We also see sinusoidal strings, one sample wide. These may be the source wavelet convolved with a strong reflector.

It is difficult to observe the diagnostic Gaussian shape of the histogram when there are wobbling ellipses. A first order 'deconvolution' is to decorrelate the data. We fit the slope of the principle axis and subtract it. This slope is the autocorrelation value at that lag*. *Walden & Hosken, 1985 show the elliptical eccentricity is the autocorrelation. It is easy to should that the linear regressor slope is proportional to eccentricity.)

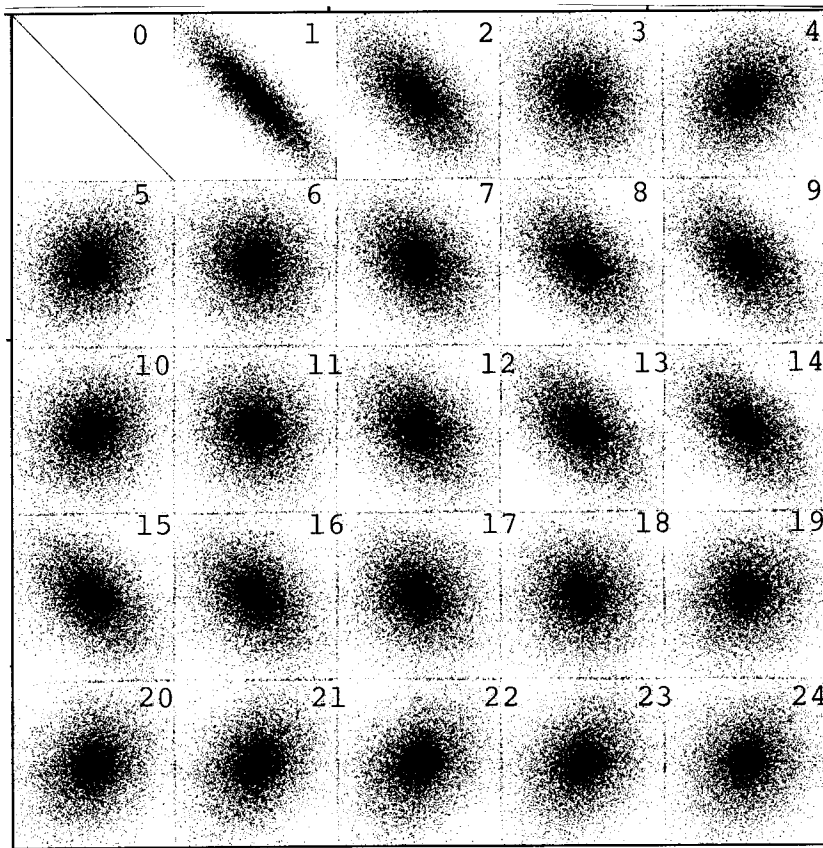


FIG. 2. Lag images of a Vibroseis dataset. The correlation ellipse fluctuates along the positive and negative diagonals. The effect is much clearer animated upon a color graphics terminal.

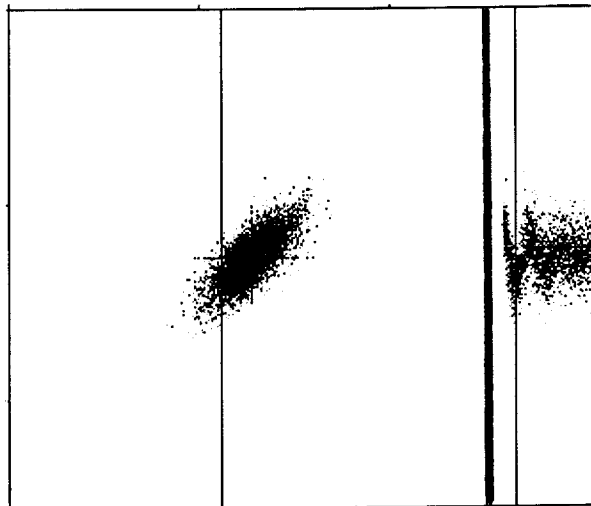


FIG. 3. A small-lag histogram image (left) and transverse image (right) from the same Vibroseis profile. The vertical lines show where the images tie. The elongation in the histogram is a sign of high correlation. The transverse image exhibits a typical 'wavelet'.

After viewing forty field data histogram movies I (Rick) am pessimistic about finding non-Gaussian shapes in reflection seismic data. We did observe a near-unity shape parameter in a well-log sonic (as did Walden & Hosken). Sometimes the out halos (strong reflection events) had a different shape parameter than that of the inner core (weak noise data). Some datasets had ambiguous indications of phase rotations.

In general, histogram images are symmetrical for positive and negative lags. Rocca suspects that non-Gaussian components may appear as asymmetries, particularly in the outer halo. We are developing techniques to better observe such asymmetries.

Deconvolution aid

Various deconvolution methods or decon parameter settings leave 'residual wavelets' in the lag movie. This suggests a technique to measure deconvolution performance. We have not yet systematically compared deconvolution algorithms. The general result of predictive deconvolution is that it doesn't estimate small lags of the wavelet well.

We were able to observe the effects of an artificial phase shift applied to a dataset (figure 4). This encourages us to try extracting wavelet phase from the data (Kostov & Rocca, 1986). Visually, a phase shift closely resembles a change in the shape parameter toward pure Gaussian ($\alpha=2$). An empirical phase estimator would be to apply a series of phase shifts to the dataset and select the one which gives the least Gaussian shape parameter.

One interesting observation is that field data is more correlated than the synthetics we were able to model. Sources of correlation we have not yet modeled include interbed multiples and the blocky nature of geological impedances (O'Doherty & Anstey, 1971).

Discussion

The strength and weakness of the histogram movie is that it permits the eye rather than some mathematical function estimate statistical parameters of the data. Many estimators are integrals which may miss some local structure in the data.

Histogram movies are useful tool for observing the statistical character of field data. We need to perform a more systematic study evaluating the effects of data acquisition and geology.

We attempted to look at the triple joint-density histogram, but have got no additional useful information from the image. It would help if we could display the entire tri-histogram in some transparent fashion and display several of these simultaneously.

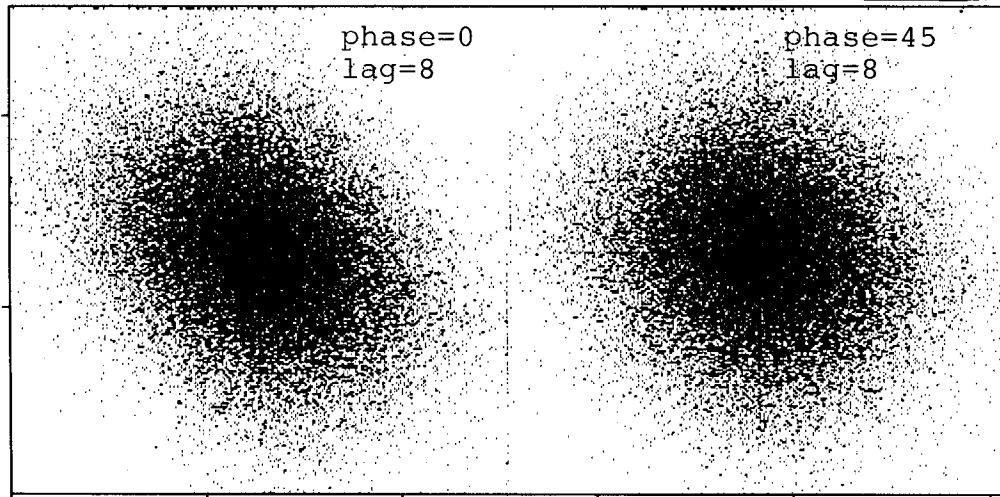


FIG. 4. The same synthetic dataset ($\alpha=1.5$) with the applied phase shift shown. The difference in shape appears more clearly as a phase rotation when a succession of histogram images are viewed.

One could also consider spatial lags in migration focusing, stacking power, and optimization statics (Van Trier, 1986). The latter cases have more complicated histograms to construct, display, and analyze.

Appendix: Histogram construction notes

The proper histogram display reveals the maximum number of features in both the dense core and sparse halo. Too dense or too sparse histogram images can greatly reduce the contrast in some part of the image. The contrast is sensitive to number of bins and density contouring method.

We initially chose a 256^2 histogram array because it gave optimal movie speed and resolution. Fortunately, this also gave high contrast images. An empirical result is that the input data samples should be within a factor of three of the cells in the histogram. This means that most of the cells contain between zero and ten samples. This is for seismic data that has been time gained but not contrast enhanced by agc or a gamma power. We set clip the histogram to the 99th data percentile so the outliers don't compress the histogram.

Color is a quick and adequate method of density contouring. This works because the density, i.e. cell count, gradually increases toward the origin. We have been using four or five hues to span the density range. Too few hues decrease contrast. Too many hues exposes local density fluctuations and confuses the image.

Our histogram displays are constrained by the current movie software. The coloring schemes are biased toward contrasting bipolar, high spatial frequency seismic images. Suggested improvements include:

- (1) Dynamically recontouring the densities by sliding hue boundaries up and down the scale.
- (2) Real time rebinning of the histograms.
- (3) For multi-dimensional lags (i.e. migration focusing applications), display an array of movie cubes rather than the current single movie cube.
- (4) Retain the input coordinates of each histogram contribution so that an interesting histogram region can be identified on the input data.

ACKNOWLEDGMENTS

Clement's comments helped clear up some mathematical ambiguities after Fabio had returned to Italy.

REFERENCES

- Godfrey, R.J., 1979, A stochastic model for seismogram analysis: PhD thesis, Stanford University
- Gray, W. C., 1979, Variable norm deconvolution: PhD thesis, Stanford University
- Kostov, C., and Rocca, F., 1986, Estimation of residual wavelets: this SEP volume
- O'Doherty, R.F., and Anstey, N.A., 1971, Reflections and amplitudes: Geophysical Prospecting, v.19, p.430
- Van Trier, J., 1986, Autofocusing of migrated data: this SEP volume
- Walden, A.T., 1985, Non-Gaussian reflectivity, entropy, and deconvolution: Geophysics, v.50, p.2862-2888
- Walden, A.T., and Hosken, J.W.J., 1985, Nature of the non-Gaussianity of primary reflection coefficients and its significance in deconvolution: 1985 Meeting of the EAEG in Budapest (preprint)
- Walden, A.T., and Hosken, J.W.J., 1985, A investigation of the spectral properties of primary reflection coefficients: Geophysical Prospecting, v.33, p.400-435

Doctorate now takes a decade, study says

WASHINGTON (AP) — U.S. universities awarded doctoral degrees in 1984 to 31,253 people who took an average of 10 years to earn their Ph.D., the National Research Council said.

New doctors of education took the longest time, 14.6 years, while physical scientists completed their studies in the least amount of time, 7.2 years.

In the fields of health sciences, psychology, languages and literature, and education, women Ph.D.s outnumbered the men.

In 1960, it took the average new Ph.D. 8.8 years to win the degree. That fell to a low of 7.9 years in 1970. The 10-year average in 1984 is the longest since the National Research Council began tracking new Ph.D.s in 1958.

The council is an arm of the National Academy of Sciences.
