HISTOGRAM EQUALISATION AS A METHOD FOR MAKING AN OBJECTIVE COMPARISON BETWEEN ANTENNA PATTERNS FUNCTIONS

S.F. Gregson [(1)], J. McCormick [(2)], C.G. Parini [(3)]

[(1)] Nearfield Systems, Inc. 19730 Magellan Drive, Torrance, CA 90502-1104 (US), sgregson@nearfield.com
[(2)] SELEX SAS, Crewe Road North, Crewe Toll, Edinburgh, EH5 2XS, (UK), john.mccormick2@selex-sas.com
[(3)] Queen Mary University of London, Mile End Road, London E1 4NS, (UK), c.g.parini@elec.qmul.ac.uk

ABSTRACT

Attempts to produce robust, objective, quantitative measures of comparison between data sets using statistical methods have been widely reported in the literature [1, 2]. Recently [3], techniques have been developed that require the antenna pattern functions to be converted into histograms before the comparison, *i.e.* the measure of adjacency, is made.

The success of such a tactic can be crucially dependent upon the choice of categorising "bins". As the number, level, and size of these bins can be chosen both a priori and freely, it is possible that the resulting histograms will be sparsely populated with the majority of the samples falling within only a few of the categories. This difficulty can be avoided if the bins are defined in such a way that roughly equal numbers of samples fall within each of the categorising bin.

This paper describes an efficient method for "equalising" any histogram and illustrates the effectiveness of this strategy with example, measured data.

1. INTRODUCTION

As is the case for other measures, the measure of correspondence is the true value of a quantity that characterises the similarities, or differences, between respective data sets. Without the ability to produce such metrics of similarity any assessment as to the integrity of a given data set is necessarily reduced to subjective value judgements. Furthermore, the utility of such comparisons or measures, of adjacency between data sets, lies not only in their ability to determine the degree of similarity between various data sets, but also in their ability to categorise the *way* in which these sets differ.

Attempts have been made to produce objective quantitative measures of correspondence between data sets (for example, on antenna data), that can be used to assess the accuracy, sensitivity and repeatability associated with the production of that data have been reported [1, 2, 3]. Previously, a variety of statistical methods have proved successful in the robust assessment of similarity, or adjacency, between respective antenna pattern functions where the comparison is complicated by three principally factors:

1) The large amounts of interferometric, *i.e.* complex data which is used for convenience to represent amplitude and phase,
2) The large dynamic range of that data, typically greater than 70 dB, and the unavailability of a convenient absolute reference,
3) The antireductionist, *i.e.* integral, relationship that exists between the near and far fields with each point in one domain affecting every point in the other.

Statistical techniques have been found to be ideally suited to the treatment of variables that possess such inconvenient features. Ordinal and hybrid interval ordinal methodologies both overcome many of the disadvantages displayed in traditional, non-statistical, assessment strategies but they do place constraints on the types of data sets that can be compared. The comparison of arrays of data that can be treated as permutations requires the two data sets to be one of the following: 1) identical in terms of sampling interval, extent and number of data points, or 2) for it to be possible to interpolate the sets to arrive at a situation where these conditions hold. For complex, multidimensional data sets containing many different angles and frequencies this is often impossible. One technique that was found to be particularly effective in avoiding these difficulties involved comparing the histograms obtained from two data sets [3].

Although a great many ways of categorising a given data set exist, one of the simplest is to divide the interval data set into a number of amplitude bins and to count how many elements fall within each bin. This is called categorical interval methodology. It is categorical, because the data is split up into a finite number of distinct bins, or categories, and it is interval because the categorisation of the data is based on the interval nature of the data [3]. Each data set that is processed in this way will provide a single histogram that can be normalised before subsequently being processed to provide the measure of correspondence. Here, normalisation would usually be accomplished by ensuring the total summation of the frequencies of the

Proc. 'EuCAP 2006', Nice, France
6–10 November 2006 (ESA SP-626, October 2006)

two sets to be compared was equal while the relative frequencies for the bins in each data set remained constant.

The technique of histogram equalisation is often utilised in the field of image processing to accentuate the useful information in a given image or to improve the visual presentation of an image; *e.g.* to correct pictures that are too dark, too light or that have insufficient contrast. This rather sophisticated technique is therefore being employed as an image enhancement method that is modifying the dynamic range so as to "stretch" the contrast levels within an image so that saturation highlights, *i.e.* regions where information is effectively lost, are minimised or removed altogether. This is accomplished by merely altering that image such that its intensity histogram assumes a desired shape.

For exactly the same reasons, such histogram equalisation techniques are also useful in the preparation of the histograms that are employed within the categorical-ordinal/interval measure of adjacency that have been recently developed. However, in this case, the objective is not to "correct" an image, but merely to prepare a histogram that is essentially continuous with no single bin being more frequently occupied than any other over as large a set of categorising bins as is practicable. This is accomplished not by altering the image itself, *i.e.* the antenna pattern, but rather by adjusting the "width" of the non-overlapping categorising bins. In other words, the intervals within which any data point is categorised to belong are adjusted in order that the same number of data points occupies each bin.

2. METHOD

Fig. 1 below illustrates cardinal cuts through two data sets and the calculated equivalent multipath level, (EMPL) between them. This can be thought of as the amplitude necessary to force the two different pattern values to be equal. If no account is to be taken of the phase of the patterns, as is often the case when assessing far field data, then the EMPL [1] can be expressed in terms of the amplitude of the samples as,

$$\mathrm{EMPL}(\theta,\phi)\big|_{dB} = 20\log_{10}\left(\frac{\big\||E_1(\theta,\phi)| - |E_2(\theta,\phi)|\big\|}{2}\right) \quad (1)$$

Here, E_1 and E_2 are used to denote the field patterns being compared and the factor of one half has been included as it is assumed that the "correct" value lies between the two measured samples. The measurements results Antenna 1 (Ant 1) and Antenna 2 (Ant 2) are of the same antenna with the measurement set-up altered so that the noise level in Antenna 2 is greater at angles off boresight than for Antenna 1.

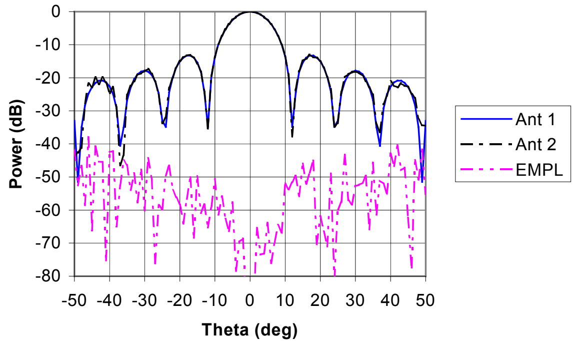

Figure 1. Cuts through Antenna patterns and EMPL.

Clearly, the EMPL in the figure reflects the presence of noise in the measurements. This noise source is not related to the signal level and in fact, is smaller in regions away from boresight where the recorded signal level is greatest, which is often the case for measurements taken in the presence of multipath. Fig. 2 below illustrates two simple histograms constructed from the data sets illustrated in Fig. 1. Here Antenna 1 is Ant 1, and Antenna 2 is Ant 2.

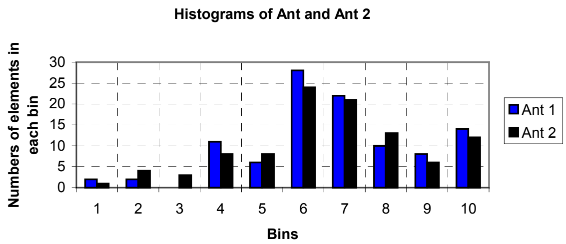

Figure 2. Histograms of data sets for two similar antennas, antennas 1 & 2.

Here, ten bins placed linearly between –50 dB and 0 dB were used to construct the histograms. As one would expect from inspection of Fig. 1, the two histograms are similar. Any of the available statistical image classification techniques could be deployed in order to quantitatively assess the correspondence between these respective data set histograms. However, if the comparison of different data sets can be reduced to a comparison of their amplitude histograms, then a range of highly developed techniques (primarily developed in the areas of image processing) can be deployed to improve the quality of the results attained. Techniques based on the concept of histogram equalisation, are particularly accurate and effective [4].

As outlined above, histogram equalisation is concerned with producing a histogram where there are equal numbers of entries in each bin. Usually this is accomplished by varying the levels and sizes of the bins until an equal number of points are to be found in each.

For example, if the data for antenna one is sorted in terms of the number of data points at given levels then Fig. 3, shown below, will be produced where there are required to be 10 points in each bin. From this figure the levels that will be required to equalise a histogram of the data can be calculated and are shown in the figure below. The levels are calculated by first sorting the samples into a vector of values in ascending order. This vector of values, l_n, can then be divided into n equal parts where n is the number of categorising bins required an l is the value of the element in the n^{th} bin. This value represents the level of the edges of the categorising bin which should contain a list of monotonically non-decreasing values. Essentially then, we are simply counting the number of elements within the data set that fall between the elements in l, *i.e.* for the case of the n^{th} bin, those elements that have a value that is greater than or equal to l_n and less than l_{n+1}.

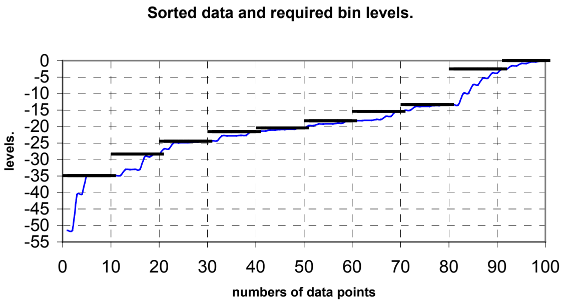

Figure 3. Sorted amplitudes and required categorising bin levels for antenna 1.

If a histogram of antenna 2 were constructed using the *same* bins then it is unlikely that the resulting histogram would be exactly equalised (unless the two patterns were identical). Fig. 4 below shows the equalised histogram for antenna 1 (Ant 1) using the values shown on Fig. 3 to define the bins and the resulting, more even, but unequal histogram of antenna 2 (Ant 2).

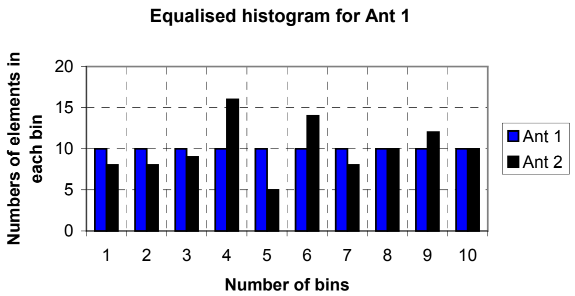

Figure 4. Antenna 1 histogram equalised using the levels shown in figure 3.

From the Fig. 4 above it is clear that the bin levels calculated to equalise antenna 1 have not equalised antenna 2, and thus again there are clear quantifiable differences identified between the two patterns. Since we know the values of the categorising bins, it is also quire clear that the equalised histograms are more evenly equalised for larger signal levels (corresponding to bin numbers 7 through 10 inclusive) than for lower levels (corresponding to bin numbers 1 through 6 inclusive). This is as expected because from inspection of the EMPL plot in Fig. 1 above, it clearly can be seen that the patterns diverge in regions of lower field intensity. Importantly, if the above data is instead plotted as a *cumulative* frequency histogram as per Fig. 5 below, it is apparent that many standard regressive techniques can be deployed on the data to calculate the measure of correspondence.

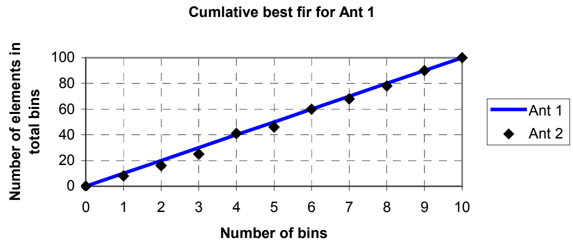

Figure 5. Best-fit line for Ant 1 cumulative data and Ant 2 scatter around it.

The expression of the difference between two antenna patterns as the regression of a linear function allows for a great deal of freedom as to which data assessment methodologies should be used. For example, Fig. 6 below illustrates two very different theoretically possible patterns, Set 1 and Set 2, that will have exactly the same histograms so clearly a great deal of care must be exercised at an early stage in the choice of assessment methodology.

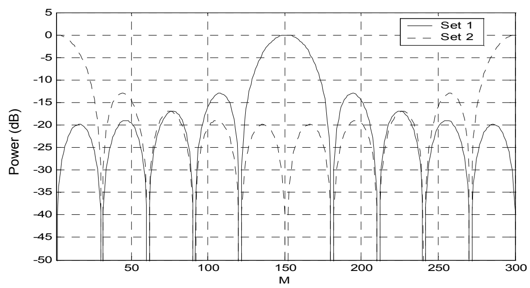

Figure 6. Different patterns with identical histograms

3. EXPERIMENTAL RESULTS – MARS

Reflections in anechoic chambers can limit the performance and often can dominate over other error sources. The recently developed technique, named Mathematical Absorber Reflection Suppression (MARS), is a sophisticated proprietary post processing technique that essentially reduces the detrimental influence of scattering on far-field pattern results. That is, *it* improves the quiet zone performance of a given facility, by utilising a novel filtering process to suppress certain undesirable scattered field components. The results of the MARS processing, which is applied within the conventional spherical near field to far field transformation, are being used herein merely to illustrate the effectiveness of the histogram equalisation technique. A full and detailed treatment of the MARS processing technique is beyond the scope of these discussions and instead, the reader is referred to the literature [5, 6] for further information.

Fig. 7 below contains a comparison of far field cardinal cuts taken from a low gain, linearly polarised open ended waveguide field probe that was acquired in an environment with a large amount of reflections. Here, the blue trace represents the far field pattern without MARS processing and the red trace represents the same pattern after the MARS processing has been applied. Clearly, the red trace is very much smoother than the blue as it is free from the high angular frequency ripples that are so characteristic of range multipath effects. This follows from recognising that a low gain instrument such as this is not sufficiently electrically large to excite the high order modes required by such rapid fluctuations.

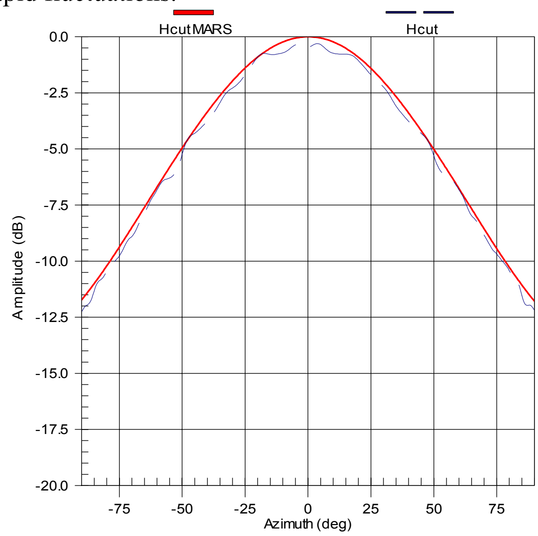

Figure 7. Comparison of azimuth cardinal cuts with and without MARS processing applied

To illustrate the difficulties encountered when calculating histograms from real world data such as this, Fig. 8 contains a comparison of the histograms obtained from these data sets when the levels are chosen so that they span the range of values linearly, whilst Fig. 9 contains the corresponding histograms when the categorising bins are distributed evenly logarithmically.

Here, it is quite apparent from inspection of Fig. 8 and Fig. 9 that when the histogram vectors are sufficiently long for subsequent comparison, the histograms are skewed and very sparsely distributed with many bins containing no observations and a few containing many thousands of observations thus complicating any further analysis. Essentially with these plots it is difficult to see the statistical differences (or relevancies) between the respective histograms. Unfortunately, this is often the case when the categorising bins are chosen in this way and is a direct consequence of the large dynamic range inherent within the data sets. Although by simply choosing a smaller number of categorising bins the number of unoccupied bins can be reduced, this results in histogram vectors containing fewer elements which can at worst compromise, and at best desensitise, any subsequent comparison assessment.

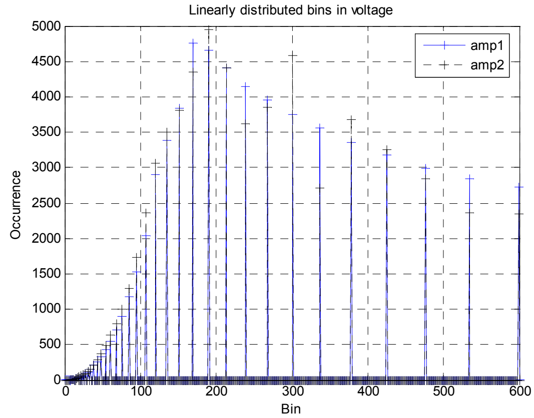

Figure 8. Linearly distributed categorising bins

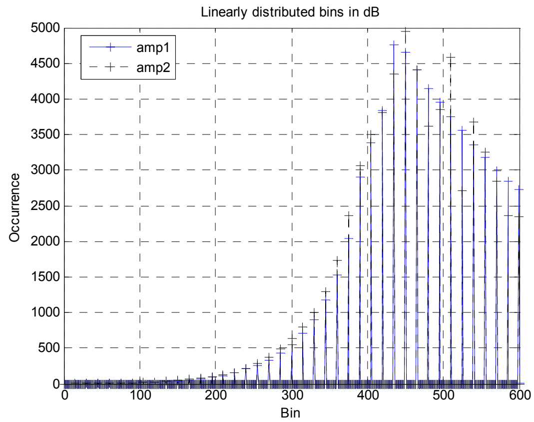

Figure 9. Logarithmically distributed categorising bins

Fig. 10 below contains a plot that shows a comparison of the histograms obtained when the categorising intervals have been chosen to equalise the resulting histograms. Here, the histogram of antenna 1 has been equalised and then the same categorising bins have been used to obtain the histogram for antenna 2. From inspection of Fig. 10 below it is apparent that the categorising bins are no longer sparsely filled, that each one contain roughly 100 samples, and that the highly skewed form observed within the previous histograms have been corrected making the act of comparison between the respective histograms very much easier than was previously the case. The process of histogram equalisation essentially results in the selection of levels for which the number of samples will be welldistributed. Thus, this process is useful in selecting the levels to use in the preparation of two-dimensional iso-level (*i.e.* contour) plots.

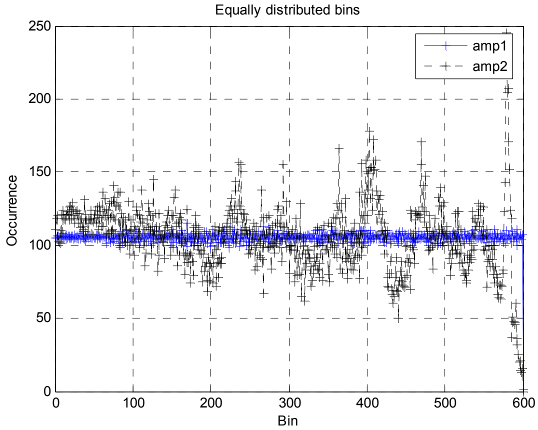

Figure 10. Different patterns with identical histograms

Fig. 10 above illustrates the effect that the MARS processing have at all field intensity levels within the far field antenna pattern function. Clearly, the multipath ripple can be seen to be present at all signal levels (and corresponding to the entire far field pattern). This comparison process both enables the measurement engineer to determine the degree of agreement between respective data sets, but furthermore and rather usefully, it also enables them to determine where the aforementioned differences reside. In this example, counting how many samples had an amplitude value that fell within some chosen interval created the histograms. However, it is possible to form a histogram by categorising some other parameter within the pattern, for example by concentrating on the cross-polar discrimination, axial ratio, or even the tilt angle.

4. DISCUSSION

It is quite clear that the novel antenna measurement techniques being pioneered at present offer an assessment challenge if the large volumes of data these techniques generate are to be quantitatively, effectively and concisely analysed and summarised. Almost all data assessment techniques at route depend on reducing the dimensionally of the data sets to make them more easily accessible. Antenna patterns acquired in test ranges may contain tens or hundreds of thousands of individual data points and the quantitative assessment of such large data sets is close to impossible without distilling the data down to manageable levels. However it should be remembered that all data reduction techniques would involve the loss of some information from the data sets.

Thus, use of a combination of the techniques is recommended so that reliable conclusions are drawn. *For example,* Fig. 6 above shows two plots that would yield identical histograms. If an alternative interval assessment methodology were also to be used - *e.g.* a moments or cross-correlation based assessment - then the exact nature of their differences could be extracted, [3].

All of the above arguments confirm the applicability of such assessment methodologies to the measurement process. These arguments also show how methodologies can be used to enhance the understanding and interpretation of measurement data. .The use of such objective quantitative methods of assessment additionally allows the interpretation of the results by a whole new variety of users which could include intelligent systems, as the decisions may need to be made in the presence of levels of uncertainty, *i.e.* in situations where there is room for doubt.

Within this paper, a technique for the efficient equalisation of antenna histograms has been presented and illustrated with example, measured data. This equalisation process significantly eases the problem of comparing the resulting histograms and is particularly well-suited to the task of obtaining a non-local (anti-reductionalist) measure of adjacency between data sets that comprise large amounts of data that encompassing very large dynamic ranges.

REFERENCES

1. S.F. Gregson, "Probe-corrected Poly-planar Near-Field Antenna Measurements", PhD thesis, University of London, 2003
2. S.F. Gregson, J. McCormick, C.G. Parini, "Advances in the Objective Measure of Comparison Between Antenna Pattern Functions", International Conference on Antennas and Propagation, University of Exeter, 2003
3. J. McCormick, S.F. Gregson, C.G. Parini, "Quantitative Measure of Comparison between Antenna Pattern Data Sets", IEE Proc.-Microw. Antennas Propag, Vol. 152, No 6, December 2005
4. Russ J. C., " The Image Processing Handbook", 3rd Edit., IEEE press, 1999, pp. 233-239.
5. G. Hindman, A.C. Newell, "Reflection Suppression To Improve Anechoic Chamber Performance", AMTA Europe, Munich, 1-4 May 2006.
6. G. Hindman, A.C. Newell, "Reflection Suppression in Large Spherical Near-field Range", AMTA, Newport, 2005.

ACKNOWLEDGEMENTS

The authors extend there gratitude to Mr Donald Gray for his valuable conversations and very detailed comments.