ADVANCES IN THE OBJECTIVE MEASURE OF COMPARISON BETWEEN ANTENNA PATTERN FUNCTIONS

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

BAE SYSTEMS[1] UK, Queen Mary, University of London [2] UK

ABSTRACT

The utility of a variety of objective, quantitative and robust methods of assessing similarities between antenna measurement data have already been highlighted in a number of recent publications (1, 2, 3). These techniques involved the extraction of interval and ordinal features from the data sets that can then be effectively compared to establish their adjacency. However, frequently such is the volume and complexities of the data involved that a single comparison methodology is inadequate to effectively classify all types of data.

Within this paper we intend to compare and contrast several techniques for obtaining a quantitative, holistic, measure of similarity between data sets as well as introducing a new hybrid technique. In addition to more conventional interval techniques, *e.g.* cross-correlation coefficient, two newer more sophisticated techniques will be presented. These are an ordinal and an interval-ordinal technique. In addition to these newer statistical image classification techniques, a novel hybrid categorical ordinal technique is developed that retains the advantages of the interval-ordinal technique but removes the requirement for interpolation and facilitates the comparison of two, or higher, dimensional data sets of differing sizes.

These techniques will be illustrated with reference to a number of data sets that will be examined assessed and classified to obtain measures of adjacency that relate global features of the data sets. This data is derived from the output of partial scan techniques (4) that attempt to reduce truncation errors in planar near field antenna measurements by the construction of bespoke polyhedral sampling surfaces that aim to enclose all the current sources.

1. INTRODUCTION

Frequently one is required to compare two or more supposedly identical data sets, for example the far field three-dimensional radiation pattern of the same antenna measured on two different facilities. The requirement for obtaining an objective, quantitative and robust method of assessing such data is clear. Previously, a novel hybrid technique has been developed that modifies the basic ordinal technique (1) to take account of different regions of interest by re tabulating the data in such a way as to attribute more samples to regions of greatest interest prior to ranking the data. This approach enables the interval aspects of the data sets to be considered whilst minimising the impact of numerical instabilities that are more often associated with purely interval assessment techniques. Unfortunately, this technique relied upon interpolation, either approximate piecewise polynomial or rigorous sampling theorem, to re-tabulate the respective data sets. This could introduce inaccuracy as in general it is always possible to pick a function that can make a mockery of any such scheme. Furthermore, this scheme presented additional complications when extended to two, or higher, dimensions. However, some of these difficulties can be eased if instead the data is categorised before the ordinal coefficient of adjacency is calculated. Although there are a great many ways of categorising a given data set, one of the simplest is to divide the interval data set into a number of bins and to count how many elements fall within each bin. Thus, each data, set will provide a single histogram that can be compared using the ordinal measure of correspondence, described in (5), irrespective of the number of dimensions and the number of elements each data set contains.

The level and size of the categorising bins can be chosen freely which clearly enables a preference to be made as to what is to be emphasised, *i.e.* the bins can be distributed in a non-linear fashion. Thus, more bins can be used for large signals than for small signals or vice versa. In the limit, when the bins are sufficiently small and sufficiently numerous, their will be one bin per distinguishable amplitude. In this limit, the categorisation process will be most sensitive and will closely approximate the interval data set. Care must be taken when defining the categories as it is possible to obtain a histogram with most of the samples only occupying a small proportion of the available bins. However, such clustering can almost always be resolved by redistributing the bins.

This categorising process yields a histogram vector of, say, N elements that contain the number of samples occupying each bin. Repeating this categorisation process for a second data set will yield a similar histogram vector that again contains N elements whereupon the closeness of these vectors can immediately be determined by evaluating the ordinal measure of adjacency, or the interval cross-correlation coefficient. When the adjacency of two histograms is determined with the cross-correlation coefficient (6) hereinafter this shall be termed a categorical-interval comparison. Conversely, when the similarity between the histograms is determined with the ordinal measure of association this shall be termed a categorical-ordinal comparison. If the cross-correlation function is utilised to compare histograms that have been derived from data sets with differing numbers of elements, the histograms must first be correctly normalised to unity. This is not a necessity if the ordinal measure of association is utilised. Importantly, the ordinal comparison process can become computationally expensive if the data sets being compared are excessively large. As the total number of bins contained within the histogram vector can be chosen freely, it is always possible to limit the maximum size of the data sets that have to be assessed by the ordinal comparison.

The principal limitation of any categorisation process is that the sorting of samples into bins fails to retain information pertaining to the location of the sample within the original data set. The impact of this can be illustrated with a simple example. Let I_1 be a set of intensity values. Let I_2 be a set of intensity values such that the order of elements in I_2 is reversed to the order of elements in I_1. This can be expressed mathematically as, $I_1(i = 1,2,3,\cdots,N) = I_2(N, N-1, N-2,\cdots,1)$. Clearly, as the two sets contain the same intensities their respective histograms will be identical. Thus, despite the fact that the respective data sets are in reality very different, the ordinal coefficient of correlation, or for that matter any correlation coefficient, will report a perfect match. However, whilst acknowledging this weakness, the object of this work is to compare radiation patterns that are at least visually similar

2. MEASUREMENT ERROR SIMULATIONS

To illustrate the applicability of this, and other, data assessment techniques to antenna measurements a partial scan technique, which attempts to reduce truncation errors inherent within planar near field antenna measurements, will be simulated. This measurement technique occupies a research area that produces data sets that require detailed analysis to assess its applicability and utility as a measurement process. Moving the antenna under test (AUT) between successive partial scans will necessarily involve the disturbance of the reference path of the RF subsystem and introduce further imperfections in the alignment between partial scans and the antenna. It is the impact of these alignment imperfections that will be assessed. A detailed description of these partial scan techniques can be found in (4, 5) and is not the purpose of this paper.

To illustrate the assessment processes a number of such simulated measurements with in built alignment errors where produced. These simulations were designed to replicate the degree of misalignment between adjacent scans that has been observed in practice. A simple physical optics measurement simulation tool was utilised to produce a series of synthetic measurements that could be used to yield a knowledge of the nature and magnitude of two alignment errors that were thought to be particularly pertinent to the auxiliary rotation partial scan technique. Namely, these were angular and range length, *i.e.* AUT-to-probe distance, errors in the alignment of the partial scans.

In the absence of some overriding definitive standard or infallible model, the only practical methodology for assessing the ability of any test facility to make measurements is by way of repetition of these measurements. This repetition can be accomplished without alteration in the measurement configuration, to simply address repeatability, or with the inclusion of parametric variations to assess sensitivity. As repeatability is inherently a statistical process the validity of any conclusions drawn will greatly depend upon the size of the sample. Thus it is preferable in this case to utilise as large a number of simulations as is practical.

To this end, the assessment of each of these errors entailed the simulation of the ninety-nine tri-scan measurements, *i.e.* two hundred and ninety seven individual partial planes. These measurement sets were transformed to the far field using the existing transformation computer code assuming that the data sets contained *no* imperfections in their alignment. The equivalent multipath level (EMPL) was calculated between the ideal pattern and each of the error simulations. This can be though of as the amplitude necessary to force the different pattern values to be equal. The maximum EMPL, *i.e.* the worst case, value at each angle can be found plotted below in figure 1 together with the ideal cardinal cut. Similarly, figure 2 below contains a plot of the optimum cardinal cut together with the maximum EMPL for case of the range length error simulations.

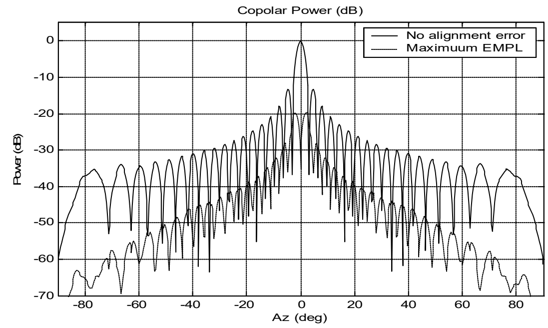

Figure 1 Far field azimuth cut of ideal simulation and maximum EMPL (pointing error).

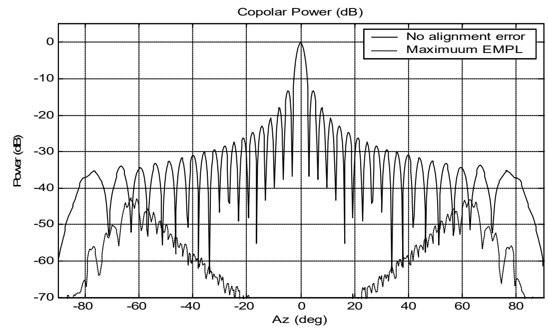

Figure 2 Far field azimuth cut of ideal simulation and maximum EMP (range error) .

Clearly, the natures of the impact of these measurement errors are very different. These plots illustrate that, for the case of the range length error, the greatest errors are observed over a limited, wide out angular range where the field intensities are relatively small. Conversely, the pointing error introduces pattern measurement errors at all angles and at all levels in the far field, *i.e.* even on boresight where the field intensities are greatest. Again, a detailed explanation of these phenomena can be found in (5). Here, it is only the difference in the form of these errors that is of interest.

3. ASSESSMENT OF ERROR SIMULATIONS

The nature of the errors are clearly very different, *i.e.* one error term is independent of angle and present at all signal levels and the other is not, thus, they constitute an ideal test data set for verifying and comparing the utility of various pattern comparison techniques. Unfortunately, although the concept of an EMPL is very useful for highlighting general differences between patterns it fails to deliver a single quantitative metric of similarity that can be used to determine which of these different phenomena is most important. Instead, these error terms have been analysed by computing the following coefficients of adjacency for each of the far field data sets: cross-correlation (6), ordinal correlation (2), hybrid ordinal-interval (1), the novel hybrid categorical-interval, and novel hybrid categorical-ordinal techniques.

A comparison between the ideal "error-free" far field pattern was made with each of the 99 far field angular error simulations using each of the five aforementioned correlation coefficients. This comparison process was then repeated for each of the 99 range-length error simulations. The mean values of these comparisons can be found presented in Table 1 below. Here, all of the correlation coefficients are normalised, thus if k is the correlation coefficient then $-1 \le k \le 1$ where k=–1 represents a perfect negative correlation (*i.e.* opposite signals), k=1 represents a perfect correlation, and k=0 represents no correlation.

Metric	Mean Value	
	Angular Error	Range Length Error
Cross correlation	0.9931	0.9982
Ordinal	0.8132	0.8758
Interval-Ordinal	0.6395	0.8113
Categorical-Interval	0.9991	0.9999
Categorical-Ordinal (a)	0.4792	0.7709
Categorical-Ordinal (b)	0.4275	0.9067

Table 1 Mean Values of Correlation Coefficients

Metric	3 Standard Deviation	
	Angular Error	Range Length Error
Cross correlation	0.0150	0.0039
Ordinal	0.2638	0.1638
Interval-Ordinal		
Categorical-Interval	0.0009	0.0003
Categorical-Ordinal (a)	0.2689	0.1040
Categorical-Ordinal (b)	0.3866	0.1150

Table 2 99% Confidence Intervals for Correlation Coefficients

Here, (a) is consists of 101 bins equally spaced spanning the levels –70 dB to 0 dB, whilst (b) consists of 101 bins equally spaced spanning the levels –30 dB to 0 dB. Table 2 above contains the 99% confidence interval, *i.e.* the 3 standard deviation of the k values about their respective mean values. Importantly, all of the methods employed above for obtaining a quantitative assessment of the adjacency between data sets possess the following desirable features:

- A single coefficient, independent of scaling or shift due to the differences in reference levels,
- Insensitive to the large dynamic range of the data,
- Normalised *i.e.* give correlation value ranging between 1 and –1, and finally,
- Symmetrical or commutative to the operation of correspondence.

Although each of these comparison procedures reveals small but systematic errors introduced into the simulations, the extent with which these variations are reported differ markedly.

4. DISCUSSION OF THE RESULTS

The cross-correlation coefficient is "saturated" by the dynamic range of the data as the coefficient diverged from unity in the third decimal place. Although this coefficient reported that, on average, the range length error is less critical than the angular error, the discrimination observed is small. In general, such purely interval techniques are highly sensitive which in part results from the huge dynamic range inherent in the data whilst more pronounced differences between patterns can yield numerical instabilities within the technique.

The results of the ordinal measure clearly shows that the small but systematic errors introduced into the simulations can be accurately quantified in the calculation of the k value. However, the ordinal process of ranking the data to produce permutations takes no account of either the absolute amplitude or spatial angles at which the data is found. Thus, every region of the pattern is judged to be equally important in the calculation of k. This is clearly illustrated by comparison of the mean average values of k determined from the two different error simulations as their values are very similar, *c.f.* note the larger 99% confidence interval. This is despite the fact that the range length error principally produces differences only in the wide out, small signal, sidelobe region.

The interval ordinal technique aims to address the principal deficiency inherent within the ordinal technique whilst minimising the numerical instabilities that can be associated with purely interval techniques. This technique clearly shows that the angular error is of greater importance as it affects large, as well as small, fields intensities. Thus, this hybrid approach is better able to isolate errors in the data that display amplitude specific traits and thus validates the concept. The extent with which the hybrid interval-ordinal method discriminates between differences in element corresponding to signal magnitudes can be readily varied on a case by case basis to emphasise or de-emphasise the particular feature under investigation. Unfortunately, it could perhaps have been expected that as less importance is being placed upon differences within small signals that the *k* value for the range length error would remain constant, or perhaps even have increased further towards unity. As the converse is observed it could be concluded that the piecewise polynomial interpolation scheme that is utilised within the re-tabulating process is introducing an additional source of error. If this is true, it is clearly intolerable and is the subject of further work.

The categorical-interval/ordinal schemes remove the requirement for pattern re-tabulation by categorising the elements into a vector of predefined bins. Unfortunately, the categorical-interval method that relies upon the conventional cross-correlation coefficient to determine the similarity between the histograms has yielded correlation coefficients that are so close to unity that they are essentially useless. Despite this, the technique still managed to distinguish between the two forms of error source although this discrimination is observed in the fourth decimal place. This is most probably a result of the fact that the histogram contains significantly fewer elements than the data set from which it was abstracted thus the comparison algorithm has fewer elements with which to work.

In contrast, the hybrid categorical-ordinal technique yields results that clearly discriminate differences between small and large signals, does not introduce errors from interpolation, and yields a sensitive correlation coefficient. The effectiveness of the novel categorical-ordinal technique is illustrated by comparing the results labelled (a) and (b) presented within Table 1 above. In each case, the data was divided into an equal number of bins, however, for case (a) the bins were equally distributed throughout a 70 dB dynamic range whilst in (b) the data was similarly distributed over a smaller, 30 dB range. As the range length error is principally concentrated about a relatively small angular range within which the field intensities are relatively small, it would expect that the agreement reported in case (b) would be better than that reported in case (a). Encouragingly, this is indeed the case. Furthermore, as the number of bins devoted to large signals is greater for case (b) than case (a) the agreement obtained for the angular error case has, as expected, worsened. Here, the bins were equally distributed in a logarithmic scale so that the number of empty bins could be minimised. In general, an infinite number of distributions can be chosen. Although this yields great deal of flexibility, such a wide choice could lead to a choice of bins that could obscure other features within the data sets. Finally, the broad characteristics of these comparison techniques can be found summarised in Table 3 below.

5. CONCLUSION

Two principal sources of error within auxiliary rotation partial scan measurement systems have been modelled. The effects of these errors on the far field vector pattern functions have been analysed using five metrics of adjacency that determine repeatability whereupon, their relative merits have been compared and contrasted.

A new hybrid categorical-ordinal comparison technique has been presented that extends to the ordinal technique an ability to discriminate between differences in elements corresponding to signal magnitudes whilst avoiding certain limitations encountered within other hybrid interval-ordinal comparison techniques. This allows more detailed characterisation and classification of specific error sources in the measurements, allowing the interval nature of the data to influence the ordinal permutations that are abstracted from the data.

REFERENCES

1. J. McCormick, S.F. Gregson, C.G. Parini, "The Development of Objective Measures of Comparison Between Antenna Pattern Data Sets", IEE Professional Networks, Technical Seminar on Antenna Measurements and SAR, 28-29 May 2002.
2. S.F. Gregson, J. McCormick, C.G. Parini, "Measuring Wide Angle Antenna Performance using Small Planar Scanners", ICAP 2001.
3. S.F. Gregson, J. McCormick, "Image Classification as Applied to the Holographic Analysis of Mis-Aligned Antennas", ESA ESTEC 1999.
4. S.F. Gregson, C.G. Parini, J. McCormick, "Wide Angle Antenna Pattern Measurements Using A Poly-Planar Near Field Technique", submitted for publication at ICAP 2003.
5. J. McCormick, S.F. Gregson, C.G. Parini, "The Development of Objective Measures of Comparison between Antenna Pattern Data Sets", IEE Professional Networks, Technical Seminar on Antenna Measurements and SAR, 28-29 May 2002.
6. E.C. Ifeachor, B.W. Jervis, "Digital Signal Processing", Addison-Wesley, 1993.

Metric	*Interval*	*Ordinal*	*Single Coefficient*	*Domain*	*Holistic*	*Robust*	*Sensitivity to outlying points*	*Absolute ref*
Cross-correlation	Yes	No	Yes	$-1 \le k \le 1$	Yes	No	Yes	Yes
Ordinal	No	Yes	Yes	$-1 \le k \le 1$	Yes	Very Stable	No	No
Interval-Ordinal	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	Stable	No	Yes
Categorical- Interval	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	No	Yes	Yes
Categorical-Ordinal	Yes	Yes	Yes	$-1 \le k \le 1$	Yes	Very stable	No	Yes

Table 3 Qualitative comparison of various pattern comparison techniques.