ON-LINE TRACKING FOR BLIND SOURCE SEPARATION USING ZERO-POINT PROBABILITY

G. J. Scruby and S. J. Flockton

Abstract

We have developed a novel on-line method for separating instantaneous, linear mixtures of super-Gaussian sources. The method uses a simple, constantly updating estimate of the central part of the probability distributions of the candidate mixed signals which can then be used to update the unmixing coefficients. The method is simple to implement and its concentration on the central part of the probability distribution makes it insensitive to outliers in the data. This is in contrast both with methods involving explicit estimates of higher order statistics, which are very sensitive to outliers, and implicit methods that raise signals to a high order power as part of their estimation process. In this paper we outline the details of the "zero-point probability" as a contrast for source separation and compare its resistance to outliers with standard fourth-order contrasts.

Introduction

This paper deals with the separation of independent signals that have been linearly, instantaneously mixed. This means no temporal information is available and higher-order statistics must be employed to separate the signals. This mixing model can be represented as follows:

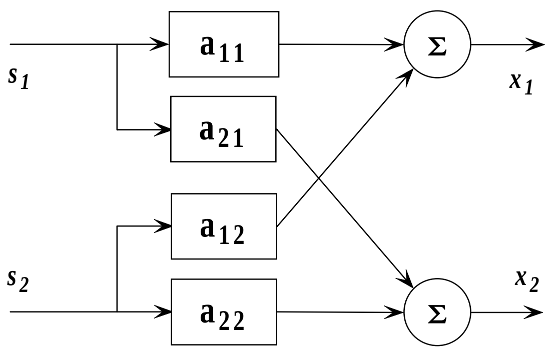

Figure 1: Linear, instantaneous mixture model

where s_1 and s_2 are the original independent signals. a_{11}, a_{21}, a_{12}, a_{22} form the mixing matrix **A** and are the linear coefficients with which the signals are scaled and mixed. x_1 and x_2 are the only signals available to the observer: unknown linear mixtures of the sources. In this paper, as in much of the literature, the observed signals in question are considered to be zero mean, of unit standard deviation and uncorrelated. Decorrelation is usually carried out in tandem with removal of the mean and normalisation of the standard deviation by means of pre-processing. This technique, known as whitening or sphering [3], removes the effects of the first- and second-order cross-moments. Separation is achieved by adapting the coefficients of a separating matrix such that scaled, permuted versions of the original signals are obtained (the inverse of **A**, but possibly scaled and permuted).

G. J. Scruby and S. J. Flockton are both with Royal Holloway College, University of London

A common separation criterion is to use a contrast [1] that is maximised when the signals fulfil some measure of independence. Statistics of fourth order are the lowest that can be usefully employed to measure independence. Second-order statistics merely show correlation and third-order statistics only differentiate between skewed signals. Moments of higher order than fourth are both very difficult to estimate with any reasonable accuracy and have been found empirically to have a relatively small effect on the total accuracy of the result.

Cumulants as separation contrasts

One popular contrast function for signal separation is the cumulant squared sum. In order for separation to be achieved for super-Gaussian signals the following function must be maximised [1].

$$J(y) = \sum_{i=1}^{M} cum^2\left(y_i{}^4\right) \qquad \text{eqn 1}$$

where y_i is the i'th signal, M is the total number of signals and $cum(x^b)$ is the bth-order cumulant of sample x. The fourth-order cumulant used here is defined as [4]:

$$cum\left(y^4\right) = E\left(y^4\right) - 3E^2\left(y^2\right) \qquad \text{eqn 2}$$

which reduces to the following when the signal has been whitened:

$$cum\left(y^4\right) = E\left(y^4\right) - 3 \qquad \text{eqn 3}$$

Consequently, a simplified [3] contrast can be derived with which to separate decorrelated, super-Gaussian signals.

$$J(y) = \sum_{i=1}^{M} E\left(y_i{}^4\right) \qquad \text{eqn 4}$$

When the above contrast has been maximised then separation has been achieved. Since the signals are already white the separating matrix reduces to an orthogonal transformation of the data [6]. This means that gradient ascent of the contrast with respect to a simple rotation of the data can be used as a separation method.

Problems with cumulants

A fourth-order solution to the separation problem has been found empirically to produce good results when operating on a batch of data [1]. This can be understood intuitively by considering that all information up to fourth order has been used optimally and that information above fourth order is both difficult to estimate accurately and has little effect on the solution. A problem arises, however, if an on-line separation system is required that can track changes in the mixing coefficients. Separation then becomes two distinct but equally important problems: the continuous estimation of the contrast function and the continuous update of the rotation. On-line estimation of the contrast causes difficulty due to the fourth-order nature of the function. Large data values have a disproportionate effect upon the contrast estimation because their fourth powers are used. This means that for signals with high kurtosis (very "peaky" signals) the contrast estimate will be very unstable and consequently so will the rotation estimate and thus the solution.

This problem can be shown diagramatically by considering an on-line contrast estimator (the equation above) of the form

$$J_{t+1} = (1-\lambda)J_t + \lambda J_{current\ \ estimate} \qquad \text{eqn 5}$$

where t is discrete time and λ is a learning coefficient. Results using the above algorithm are shown in Figure 2, using a signal drawn from a Gaussian cubed distribution and the contrast function shown in equation 4 (with λ =0.001).

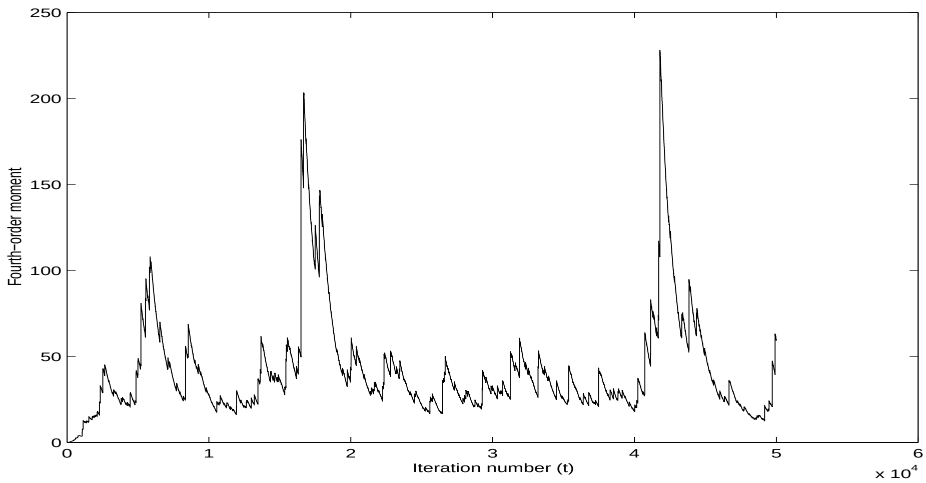

Figure 2: Iterative estimation of fourth-order moment

In Figure 2 the range of the contrast is ~500% of its batch calculated value. Uncertainties of this magnitude can entirely prevent estimation of the orthogonal matrix required to separate the data. For example, in the case where two Gaussian cubed signals are mixed the fourth-order moment contrast range is only ~40, which would be swamped by the uncertainty in the contrast estimation.

Zero-point probability

In an attempt to ameliorate the problem of variability in the contrast estimate we introduce the idea of "zero-point probability" as a contrast function. This contrast is formed by calculating the integral of the probability density function over a very small distance centred on zero for each channel and then forming the sum of these values. When it is maximised, separation has been achieved. In order for this to be a sufficient condition for separation the signals must be zero-mean and white, just as for many other contrasts (including the fourth-order contrast already shown). The principle on which this contrast depends can be explained using the central limit theorem. This states that the limiting distribution of a standardised sum of independent, identically distributed summands as the number of terms increases, is a normal distribution [5]. Therefore, if two signals that are "peakier" than a Gaussian signal are added together, the resultant signal will be more Gaussian than either of the two original signals. Thus by taking the zero-point probability as a contrast and attempting to maximise it, the least mixed version of the signals will be found.

Practical considerations

In practice, we are using a finite sample so exact continuous integration is impossible; an approximation must be made. We choose a certain "window" about zero (of width γ) and use the proportion of data points that fall within this area as an approximation to the zero-point probability (Figure 3).

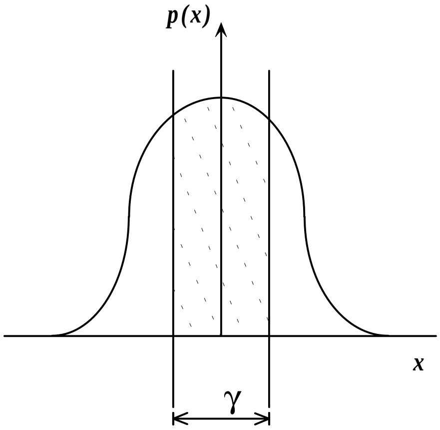

Figure 3: Window about zero for zero-point approximation

The width of γ will, of course, affect the performance of the separation algorithm. This can be seen below in a plot that demonstrates the change in contrast with differing rotations and γ of a two channel mixture. The mixture consists of two 50000 point Gaussian cubed signals. The signals have been whitened and maximally mixed (effectively rotated by 45°). It can be seen that the contrast is greatest at 45° rotation. Thus a 45° rotation applied to the mixed signals would separate them again (as the surface is periodic with a period of 90° [2]).

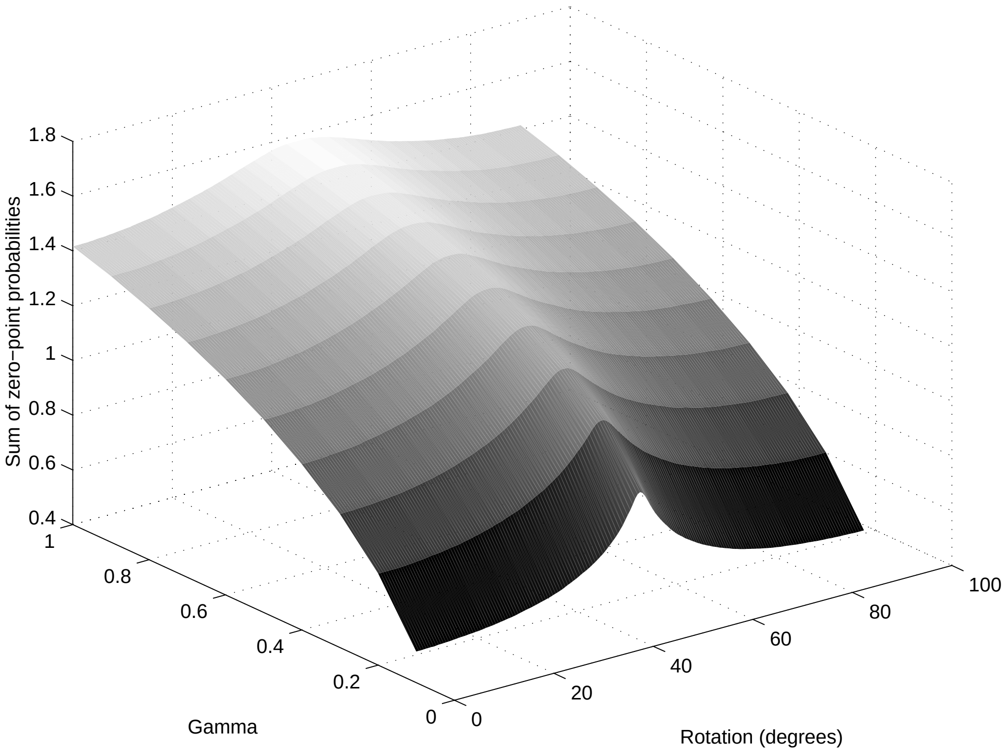

Figure 4: Performance surface of zero-point probability

It is clear that if γ is large a very smooth performance surface will be obtained but with low resolution between minimum and maximum value whereas with small γ good definition will be obtained between minimum and maximum values at the cost of a more statistically unreliable performance surface. Figure 4 also shows that the contrast function is smoothly varying and uni-modal. The main reason for introducing the zero-point probability as a source separation contrast function, however, is its performance in on-line systems when adaptively estimating the contrast. Since no fourth-order statistic is used there are no adverse effects from outliers in the data. Indeed, in this method outliers are ignored altogether whether they are part of the underlying signal structure or simply spurious, as only the probability very near to the zero-point is considered. An example of this method's resistance to outliers can be seen in Figure 5, which shows the running estimate of the zero-point probability for the same data used to create Figure 2.

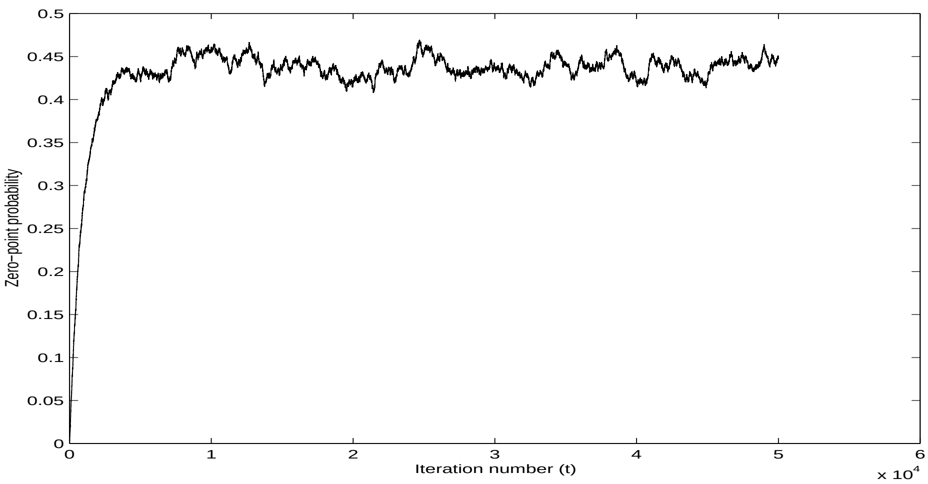

Figure 5: Iterative estimation of zero-point probability, using a value of 0.05 for γ

The contrast range in this example is ~14% of its batch calculated value, a large improvement on the variability compared with the fourth-order estimation methods.

Conclusion

We have introduced a new method for on-line source separation using a numerical contrast function, called the "zero-point probability" which has been derived from the central limit theorem. We have shown that during on-line estimation the zero-point probability is far more tolerant to outliers in the data than higher-order statistical methods.

References

1. Comon, P: "Independent component analysis, a new concept?", Signal Processing, Vol 36, pp 287-315, July 1994
2. Flockton, S, Yang, D and Scruby G: "Performance surfaces of blind source separation algorithms", Proceedings of the International Conference on Neural Information Processing '96, ICONIP '96, Hong Kong, pp 1229-1234
3. Karhunen, J, Wang, L and Vigario, R: "Nonlinear PCA type approaches for source separation and independent component analysis", Proceedings of the '95 IEEE International Conference on Neural Networks, ICNN '95, Perth, Australia , pp 995-1000, 27th November 1st - December 1995
4. Kendall and Stuart: "The Advanced Theory of Statistics", Volume 1, 4th Edition, Charles Griffin and Company Ltd, 1977, ISBN 0 85264 242 3
5. Lindgren, B W: "Statistical Theory", 3rd Edition, MacMillan Publishing Co Inc, 1976, ISBN 0 02 3708301
6. Oja, E and Karhunen, J: "Signal separation by nonlinear Hebbian learning", Computational Intelligence: A Dynamic System Perspective, pp 83-97, IEEE Press 1995, ISBN 0-7803-1182-5