TIME-FREQUENCY SIGNAL DECOMPOSITION USING ENERGY MIXTURE MODELS

*Mark Coates**

Dept. of Electrical and Computer Engineering
Rice University,
6100 Main St, TX 77005-1892
mcoates@rice.edu

William Fitzgerald

Dept. of Engineering
Cambridge University
Cambridge, UK, CB2 1PZ
wjf@eng.cam.ac.uk

ABSTRACT

We address the problem of signal decomposition. We specify signal components by the property that their energies are localised and disjoint in the time-frequency plane. Rather than modelling the signal directly, we represent the time-frequency energy of the signal using a finite mixture model. This model is used to develop a partitioning of the time-frequency plane, allowing the application of time-frequency filtering to isolate components. Modelling energy rather than specifying a dictionary of allowable waveforms imposes fewer constraints on what a component may be. We demonstrate how the approach can be applied in the context of vibration analysis, where we wish to isolate the structure of individual bending waves travelling through a beam.

1. INTRODUCTION

Signal decomposition aims to separate individual components comprising an observed signal. There is an underlying assumption that the components can be distinguished in some manner. In performing time-frequency decomposition, we target signals comprised of components whose energy is in some sense localised in the time-frequency (TF) plane. Even with the assumption of clustered energy, we are still not in a position to readily perform decomposition. We need to make the definition of a component more concrete, whilst avoiding imposing unnecessary constraints.

In this paper, we consider the situation of components occupying distinct TF regions. For noiseless signals, signal decomposition is then an exercise in developing a partitioning of the TF plane such that each segment contains a single component. This paper argues that in this scenario modelling the TF energy of a component provides greater flexibility than component modelling in the time-domain.

The majority of methods for performing linear signal decomposition involve over-complete waveform dictionaries. By selecting the (in some sense) optimum set of available waveforms from the dictionary, we can obtain a sparse model of the signal. Such decomposition schemes include matching pursuit [8] and basis pursuit [3], and the chirplet decomposition in [2]. A problem with these decomposition methods is the restricted number of waveforms in the dictionary. While dictionaries containing a wide variety of elements can be employed (at the expense of high computational cost), the representations are not satisfactory unless all signal components are at least reasonably well approximated by dictionary elements. The introduction of parameters in the dictionaries provides greater flexibility, and has been explored in [9] and [5].

The limitations of sparse linear time-frequency models motivate the development of a semi-parametric model of signal time-frequency energy density. We propose a model taking the form of a set of radial basis functions in the time-frequency plane. Although the use of basis functions imposes some limitations on the nature of a component, there is more flexibility than in a linear decomposition approach (unless a vast dictionary is used), because the basis functions are parameterised and the modelling is performed in the time-frequency plane. The purpose of a model of energy density is different from that of a model of a signal. The energy-based time-frequency model identifies how many components there are, describes the nature of their energy structure, and probabilistically associates time-frequency locations with the components. The construction of an energy density model does not, of itself, provide a decomposition of the signal. We develop a partitioning of the TF plane based on the energy model and utilise it to design time-varying filters [7] to decompose the signal.

Section 2 of the paper clarifies the structure of the time-frequency mixture model and discusses two critical issues in its construction: determining the number of components to include and optimising the model parameters. Section 3 details the decomposition approach, and Section 4 discusses an extension to warped time-frequency mixture models. We apply the algorithm to decompose a beam vibration signal in Section 5, and make some concluding remarks in Section 6.

* This work was supported by Texas Instruments.

0-7803-6293-4/00/$10.00 ©2000 IEEE.

2. THE TIME-FREQUENCY MIXTURE MODEL

The time-frequency mixture model consists of a set of basis functions taking the form of two-dimensional Gaussian functions in the TF plane, with full covariance matrices to allow arbitrary orientation and aspect ratio. The model comprised of M-components is

$$F_M(\mathbf{x}|\theta,\lambda) = \sum_{i=1}^{M} \lambda_i f_i(\mathbf{x}|\theta_i) \tag{1}$$

where $\mathbf{x}$ are time-frequency vectors $[t,f]^T$, $\lambda = \{\lambda_1,\ldots,\lambda_M\}$ are mixing weights, constrained to be positive, and $\theta = \{\theta_1,\ldots,\theta_M\}$ is the set of parameter vectors. The parameter vector θ_i of each basis function includes the mean $\mu_i = [t_{\mu_i}, f_{\mu_i}]^T$ and the covariance matrix Σ_i of the normalised Gaussian. Each basis function is considered to correspond to a single signal component, so components are constrained to occupy a Gaussian-shaped region in the TF-plane.

Training this model requires information about the distribution of signal energy in the time-frequency plane and is based on an energetic time-frequency representation (TFR) $P_s(t,f)$ of a signal $s(t)$. The resulting mixture model is dependent on the choice of the TFR; as soon as $P_s(t,f)$ fails to represent some portion of the energetic nature of the signal, it is impossible for it to be recovered. Adaptive TFDs, particularly AOK distributions [6], provide satisfactory distributions for the majority of signals. Since the energetic time-frequency model is only capable of generating a positive value at any time-frequency location, we use only the positive and real part of a time-frequency representation, and denote it $C_s(t,f)$.

2.1. Fixed-order model parameter estimation

Since an energetic TFR is an imperfect and non-pointwise portrayal of signal energy density, we don't view the TFR as a definitive density description. It is useful to consider it as a sample from a true underlying density, a 'data set' based on which density estimation can be performed. The parameter estimation exercise is then equivalent to probability density estimation using a finite mixture model. Maximising the likelihood of the calculated TFR given the parameters of the model is a sensible optimisation technique.

The likelihood of the TFR C_s given the M-component model with parameter set θ and weight set λ is

$$p(C_s|\theta) = \prod_{\mathbf{x}} [F_M(\mathbf{x}|\theta,\lambda)]^{C_s(\mathbf{x})} \triangleq \mathcal{L}_{C_s}(\theta) \tag{2}$$

The optimisation problem for fixed model size is then

$$[\theta_{C_s,\mathrm{opt}}, \lambda_{C_s,\mathrm{opt}}] = \arg\max_{\theta,\lambda} \mathcal{L}_{C_s}. \tag{3}$$

The optimisation problem (3) has been investigated at length in the context of probability density estimation. A variant of the expectation-maximisation (EM) algorithm is a practical and powerful method of determining model parameters that correspond to a local maximum of the likelihood $\mathcal{L}_{C_s}$. If this algorithm is initialised carefully using the information provided by the initial TFR, then the local maximum is often a global maximum or a close approximation. The form of this variant and its application to the optimisation problem (3) are discussed in [4].

2.2. Training the model

Training the model involves determining the most suitable number of components. Functional merging [4] is a useful technique for adapting the number of Gaussians in a finite mixture model. It is based on the principle that Gaussians modelling the same mode should be closer together than those modelling separate modes. Critical to its operation is the choice of a distance measure to gauge the distance between two densities. We use the *arc-cosine* [4] distance because it is related to maximum likelihood and has a closed-form expression when used to measure the distance between mixtures of Gaussians.

The functional merging operation consists of the following steps: replace the closest two Gaussians in the model with a single Gaussian; store the distance τ between the original mixture of two Gaussians and the replacement; train the new fixed-order model; and return the distance τ and the new model order M.

Combining functional merging and the fixed-order optimisation strategy leads to the following training algorithm:

1. Propose an over-represented model of N components and optimise the model parameters θ and λ. Choose a maximum distance threshold $\tau_{\max}$ beyond which merging should not occur. Set $\tau_0 = 0$, and $i = 1$.
2. Repeat until $\tau_i > \tau_{\max}$ or $M_i = 1$: Perform the merging operation on the current model to determine τ_i, M_i. If $\tau_i < \tau_{i-1}$, then set $M_{i-1} = M_i$, otherwise increment i.
3. Choose the final model F_{opt} as that which maximises the model probability function:

$$P(F_i) \triangleq \frac{\tau_{i+1} - \tau_i}{\tau_{\max}}, \quad i = 1,\ldots,i_{\max}-1$$

The probability function is based on the principle that the model order should be consistent over the largest threshold range when Gaussians representing the same mode have been merged, and no Gaussians representing separate modes have been merged. After training, it is possible that some basis functions in the selected model do not cover a sufficient TF area to represent true signal components; either there has been insufficient merging (so that a genuine

component's energy is described by a number of Gaussians) or noise terms or cross-components present in the initial TFR are being modelled individually. We test this by checking that all of the basis functions satisfy a bound imposed by the uncertainty principle. If some do not, we replace F_{opt} with the most probable model comprised of fewer components, all of which satisfy the uncertainty constraint.

3. TIME-FREQUENCY DECOMPOSITION

We desire a *segmentation* based on the energetic model: a strict classification of each time-frequency location as a member of the region R_k corresponding to component C_k is appropriate. Based on the mixture model, a probability function for the membership of time-frequency location to a particular component can be constructed

$$\begin{aligned} P(C_k|\mathbf{x}) &= \frac{P(\mathbf{x}|C_k)P(C_k)}{P(\mathbf{x})} \\ &= \frac{\lambda_k f_k(\mathbf{x}|\theta_k)}{\sum_{j=1}^{M} \lambda_j f_j(\mathbf{x}|\theta_j)}. \end{aligned} \tag{4}$$

The classification is then performed using the Bayesian multi-class decision procedure. The decision rule is simply

$$\mathbf{x} = [t,f]^T \in R_k \text{ where } P(C_k|\mathbf{x}) = \max_j P(C_j|\mathbf{x}). \tag{5}$$

The classification results in a partitioning of the TF plane. We use the partitioning to develop perfect-reconstruction filter banks according to the method in [7].

4. WARPED MIXTURE MODELS

The mixture model is restricted to representing a signal's energy structure using Gaussian basis functions, so is suitable for signals displaying approximately linear instantaneous frequency (IF) behaviour. The Wigner distribution and many related TFRs are also poorly suited to non-linear components; time-frequency analysis of such signals is best performed using alternative "warped" representations. In this section, we outline how warped mixture models can be generated using such warped TFRs.

The **U**- and **VU**-Cohen's classes of distributions [1] are suitable for signals whose components all have the same non-linear time-frequency behaviour. The class is constructed by applying a unitary transform **U** to a signal prior to generating a Cohen's class TFR. The resulting representation portrays the distribution of energy in a warped time-frequency plane (the $\widetilde{T}$-$\widetilde{F}$ plane). By constructing a mixture model using this warped energy representation, we can model components that have decidedly non-Gaussian energy shape in the TF plane, but approximately Gaussian shape in the warped plane. The choice of a distribution (or **U** operator) that matches the signal components depends on knowledge of the specific signal structure, but in many cases even a rough approximation can result in a much improved modelling (as evident in Section 5). In order for the method to succeed, all components must display approximately linear behaviour in the warped plane.

One can determine a post-processing operator **V** that performs an inverse warp, mapping the $\widetilde{T}$-$\widetilde{F}$ energy representation, back to the TF plane. This same operator can be applied to the Gaussian mixture model generated in the warped plane, leading to a warped mixture model:

$$\widetilde{F}_M(\mathbf{x}|\theta,\lambda) = \sum_{i=1}^{M} \lambda_i \mathbf{V} f_i(\mathbf{x}|\theta_i)$$

In this model, the basis functions occupy non-Gaussian shapes in the TF plane, but the same approach as before can be used to generate a segmentation and perform decomposition.

5. VIBRATION ANALYSIS

This section applies the mixture model decomposition algorithm to an experimentally-obtained set of data. We analyse the impact response of a beam. A 7.2 m long mild-steel beam was suspended horizontally on light cords, and one end was lightly tapped with a soft-tipped hammer. The impact response was measured using an accelerometer attached to the beam close to the point of impact. Since the group velocity of the bending waves is dependent on frequency, the high-frequency groups travel faster than the low-frequency groups.

Figure 1(e) shows the measured response, and Figure 1(d) shows an adaptive TFR of it. We generate a **VU**-Cohen's class TFR using the transformation:

$$(\mathbf{U}s)(t) = \left|4t^3\right|^{1/2} s(t^4). \tag{6}$$

and the post-processing transformation **V** to warp back to ordinary time t. The resultant representation, displayed in Figure 1(a), clearly indicates that the first three components have been substantially 'linearised'. The warping is matched to some of the components better than others, both because the warping is probably being applied to the wrong variable (time instead of frequency), and the power law is not correct. The second and third components in particular retain a non-linear nature. However, the linearisation is sufficient to make the Gaussian mixture model (Figure 1(b)) an appropriate representation.

The strict classification procedure using the warped mixture model (Figure 1(c)) generates a complete partitioning of the positive half of the time-frequency plane, as shown in Figure 1(d). We use this partitioning to extract the different waves constituting the impact response; Figure 1(f) sketches two of these waveforms.

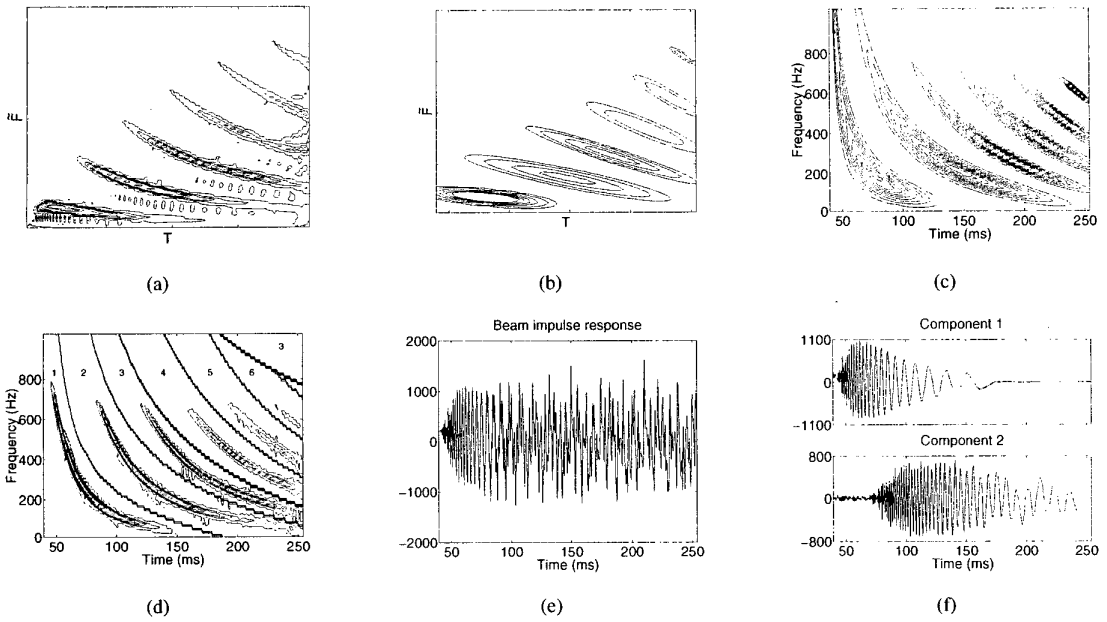

Figure 1: Time-frequency decomposition of the beam impulse response. (a) The **VCU**-smoothed pseudo-Wigner distribution ($T\tilde{F}$-plane). (b) The mixture model ($T\tilde{F}$ -plane). (c) The unwarped mixture model (TF-plane). (d) The segmentation of the plane (regions labelled 1 to 6). The regionally optimised AOK based distribution [4] is superimposed. (e) The beam impulse response. (f) The extracted components corresponding to regions 1 and 2.

6. CONCLUSIONS

We have presented a method for performing signal decomposition of signals consisting of time-frequency disjoint components. The method involves the development of a mixture model to represent a signal's time-frequency energetic structure. The mixture model identifies the number of components in the signal, and specifies the probability that energy observed at a particular time-frequency location arises from a particular component. This approach allows us to develop a decomposition without requiring that a component belong to a particular dictionary of waveforms. When applied to the analysis of the impact response of a beam, the technique provided insight into the characteristics of the individual bending waves travelling through the beam.

7. REFERENCES

[1] R.G. Baraniuk and D.L. Jones. Unitary equivalence: a new twist on signal processing. *IEEE Trans. Signal Processing*, 43(11):2269–2282, Oct. 1995.

[2] R.G. Baraniuk and D.L. Jones. Wigner-based formulation of the chirplet transform. *IEEE Trans. Signal Proc.*, 44(12):3129–3135, Dec. 1996.

[3] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Scientific Comp.*, 20(1):33–61, 1999.

[4] M.J. Coates and W.J. Fitzgerald. Regionally optimised time-frequency distributions using finite mixture models. *Signal Processing*, 77:247–260, 1999.

[5] P.M. Djúric and S.J. Godsill. Parametric modelling and estimation of time-varying spectra. In *Proc. 32nd Asilomar Conf. Sig. Syst., and Comp.*, pages 292–296, 1998.

[6] D.L. Jones and R.G. Baraniuk. An adaptive optimal-kernel time-frequency representation. *IEEE Trans. Signal Processing*, 43(10):2361–2371, Oct. 1995.

[7] W. Kozek and F. Hlawatsch. Time-frequency filter banks with perfect reconstruction. In *Proc. 1991 IEEE Int. Conf. Acoust. Speech. Signal Proc.*, pages 2049–2052, 1991.

[8] S. Mallat and Z. Zhang. Matching pursuit in a time-frequency dictionary. *IEEE Trans. Signal Processing*, 41:3397–3415, 1993.

[9] J.C. O'Neill and P. Flandrin. Chirp hunting. In *Proc. IEEE Symp. Time-Freq. Time-Scale Analysis*, pages 425–428, 1998.