AUTOMATIC EXTRACTION OF TIME-FREQUENCY SKELETONS WITH MINIMAL SPANNING TREES

Olivier Michel, Patrick Flandrin

Ecole Normale Supérieure de Lyon,
Labo. de Physique (UMR 5672 CNRS),
46 allée d'Italie, 69364 Lyon Cedex 07, France
{omichel,flandrin}@ens-lyon.fr

Alfred O. Hero III

The University of Michigan,
Dept. of EECS
Ann Arbor, MI 48109-2122, USA
hero@eecs.umich.edu

ABSTRACT

Theoretical results have recently been established in non parametric entropy estimation, based on asymptotic properties of minimal spanning trees (MST). A new application is proposed for the automatic extraction of time-frequency skeletons in the case of multicomponent chirp-like signals. The proposed method makes use of local maxima of a time-frequency distribution (considered as realizations of a 2D or 3D process), and exploits the efficiency of MST's for density discrimination and clustering.

1. INTRODUCTION

In a recent series of studies [1, 2], we have addressed the problem of estimating the Rényi entropy of a multi-dimensional distribution from a given set of observations. It has been established that *Minimal Spanning Trees* (MST), i.e., acyclic graphs of minimum total length connecting all points of a process sample, allow for a direct estimation of this entropy at a low computational cost. An extension of this result to k-MST's, i.e., subgraphs connecting k points only among all observed realizations, has been shown to permitting a *robust* separation of a statistical mixture. In this paper, we present a new application of those tools to the detection and extraction of structured signal components from a noisy observation. The principle of the approach is to consider local maxima of a time-frequency distribution as realizations of a mixture model ("signal + noise"), onto which a k-MST strategy is applied. In the case of noisy multicomponent chirp-like signals, individual "signal" components can be associated to coherent time-frequency trajectories, as opposed to "noise" contributions whose maxima distribution is incoherent. The rationale of the proposed method is therefore that minimum length trajectories—as identified by k-MST's—are expected to reveal a meaningful signal *skeleton*.

This work was supported in part by CNRS — Programme *Modélisation et Simulation Numérique*, under grant 97N23/0016.

In Sections 2 and 3, we will recall basic definitions and properties of MST's and k-MST's. The question of how to make use of k-MST's in a time-frequency context will be addressed in Section 4, where two different algorithms will be proposed, based either on the 2D (time + frequency) projection of local maxima onto the plane, or on the complete 3D information (time + frequency + energy density).

2. MST ET K-MST

Let $\mathcal{T}_n$ be an acyclic graph (or tree) connecting all realizations $\mathcal{X}_n = \{x_1, x_2, \ldots, x_n\}$ of a point process defined in $\mathbb{R}^d$. Such a graph is indeed a convenient way of coding a set of *vertices* (the points x_i) and *connections* $e_{i,j}$ between them. The total length of the graph being obtained by adding up the lengths of all elementary connections, we will introduce the parameterized quantity :

$$L_{n,\gamma} := \sum_{e_{i,j} \in \mathcal{T}_n} |e_{i,j}|^\gamma, \tag{1}$$

with $\gamma \in]0, d[$.

Given this measure, the *Minimal Spanning Tree* (MST) is, among all possible (acyclic and totally connected) graphs that be constructed, the one with minimum length :

$$\mathcal{T}_n^\star := \arg\min_{\mathcal{T}_n} L_{n,\gamma}. \tag{2}$$

This MST can be exactly computed with algorithms of complexity $\mathcal{O}(n \log n)$.

The above definition (2) can be extended to what is referred to as k-MST's. By definition, a k-MST is a MST connecting k points only among n observed points. Equivalently, a k-MST is the MST associated with a k-points subset $\mathcal{X}_{n,k} \subset \mathcal{X}_n$. In this case, minimization concerns both the identification of the subset $\mathcal{X}_{n,k} := \{x_{i_1}, \ldots, x_{i_n}\}$ and the length of the MST constructed on the points of the subset :

$$\mathcal{X}_{n,k}^\star = \arg\min_{i_1,\ldots,i_k} \arg\min_{\mathcal{T}_n} L_{n,k,\gamma}. \tag{3}$$

In practice, this double minimization is often conducted jointly : this is especially true for the algorithms that we have developed [1, 2]. Of course, the computational cost of k-MST's is increased as compared to simple MST's, and it has been even proved that the problem is NP-complete in $\mathbb{R}^2$ [4]. Ravi *et al.* have proposed an approximate algorithm with polynomial cost in the case of bidimensional distributions. In [2], we have extended this work and proposed an approxmate solution in the d-dimensional case ($d \geq 2$), whose approximation ratio is bounded above by $O(k^{(1-1/d)^2})$. The precise structure of the algorithm, its robustness evaluated by means of influence curves, as well as proof elements of its asymptotic convergence, are detailed in [2] : we will not, here, elaborate further on this very technical part.

3. PROPERTIES

Let $L_{n,\gamma}$ be the quasi-additive euclidean function of order γ defined in (1), and $\mathcal{X}_n$ a set of independent realizations of a stochastic process with Lebesgue density $f(x)$, defined on $\mathbb{R}^d$. Generalizing upon a result by Beardwood, Halton et Hammersley [6], Steele has proved that :

$$\lim_{n\to\infty} \frac{L_{n,\gamma}(\mathcal{X}_n)}{n^{d-\gamma/d}} \overset{\text{a.s.}}{=} \beta(\gamma,d) \int_{\mathbb{R}^d} f(x)^{d-\gamma/d}\, dx. \quad (4)$$

If we now introduce $\nu := 1-\gamma/d$, $\gamma \in]0,d[$ (hence, $\nu \in]0,1[$), and if we define the quantity :

$$\widehat{H_\nu}(\mathcal{X}^*_{n,k}) := \frac{1}{1-\nu} \ln \left(n^{-\nu} L_{n,\gamma}(\mathcal{X}^*_{n,k}) \right) + \beta(\nu,d) \quad (5)$$

as a statistics based on the k-MST length

$$L_{n,\gamma}(\mathcal{X}^*_{n,k}) = \sum_{e_{i,j}\in\mathcal{T}^*_{n,k}} |e_{i,j}|^{(1-\nu)d}, \quad (6)$$

the following central result can be established :

Theorem [2]. *Let $\widehat{L}_{n,\gamma}(\mathcal{X}^*_{n,k})$ be an estimate of the length $L_{n,\gamma}(\mathcal{X}^*_{n,k})$, obtained by the k-MST approximation described in [2], with $k := \alpha n$, $\alpha \in [0,1]$. Plugging this estimate in (5), we end up with a consistent and robust estimate of the Rényi entropy of the density $f(.)$:*

$$\widehat{H_\nu}(\mathcal{X}^*_{n,k}) \overset{\text{a.s.}}{\to} \min_{A:P(A)\geq\alpha} \frac{1}{1-\nu} \ln \int_A f^\nu(x)\, dx, \quad (7)$$

where the minimization is conducted on all Borel subsets A defined on $[0,1]^d$, and whose probability $P(A)$ is such that

$$P(A) = \int_A f(x)\, dx \geq \alpha. \quad (8)$$

It is worth noting that the value β in (5) exactly identifies to the Rényi entropy of a uniform distribution on $[0,1]^d$: it is therefore a function of ν and d only. The parameter k, which controls the size (in terms of connected vertices) of the considered MST, plays a role similar to that of the parameter α in α-truncated mean value estimators : in the presence of outliers, k can be tuned so as to guarantee a form of robustness to the entropy estimator [1, 2]. Finally, one can remark that the proposed method can be extended in a straightforawrd manner to other entropy functionals such as, e.g., the (non-additive) structural entropy of Havrda and Charvàt.

4. MST'S AND TIME-FREQUENCY

In order to apply a MST strategy in a time-frequency context, all local maxima of a given time-frequency distribution $E(t,\nu)$ are first identified. Each of those relative maxima is indeed considered as a realization of a 3D stochastic process, the considered variables being of the type $x = [t,\nu,E(t,\nu)]$, with $t \in T$, $\nu \in F$ and $E(t,\nu) \in \mathbb{R}$. The assumed model is a mixture model "signal + noise", with density

$$f = (1-\varepsilon)g(x|\text{signal}) + \varepsilon g(x|\text{noise}), \quad (9)$$

where $g(x|.)$ is the conditional probability density function of local maxima. The problem of extracting a signal part from the observation reduces therefore to a problem of mixture separation.

A crucial issue consists in defining a relevant norm in the space $T \times F \times \mathbb{R}$. A natural constraint is that such a norm should not depend upon the sampling rate in the time-frequency plane : in other words, the "distance" D_{12} between two energy contributions located at $\{(t_i,\nu_i)\,;\, {\scriptstyle 1} = 1,2\}$ should be independent of the sampling frequency F_e of the time series, as well as of the number N_b of frequency bins. This can be achieved by introducing two normalization constants K and K' (dimensionally homogeneous to time), thus defining :

$$D_{12} = \sqrt{\left(\frac{t_1-t_2}{KF_e}\right)^2 + \left(\frac{K'F_e}{2N_b}(\nu_1-\nu_2)\right)^2} \quad (10)$$

where t_i, ν_i refer to sample indexes. In the following, and for a sake of simplicity, we will take $F_e = 1$ and $K = K' = 1$, i.e., $N_b = N/2$ frequency bins for N time samples. It has however to be remarked that the dynamic range of the third variable $E(t,\nu)$ is totally arbitrary.

A 2D approach. A first possibility is to directly apply two-dimensional techniques which have been previously proposed. The method is based on the construction of a 2D MST in the time-frequency plane (only the locations of the most energetic local maxima are considered), and on its recursive pruning with Banks' algorithm [9]. The set of most energetic local maxima is determined by thresholding, the rejection threshold being fixed by a change point detection criterion applied to the second derivative of the cumulative distribution function of local maxima heights [7]. Alternatively, a detection based on k-MST only is presented, which relies on identifying the most important increase in the entropy as a function of k, see figures 1 and 2.

A 3D approach. A second possibility is to jointly exploit the 3D nature of a time-frequency distribution. In this case, the energy density is normalized so that the dynamic ranges are numerically identical on the three axes. Given $\mathcal{T}_n^\star$, the MST constructed on the total set S of local maxima of the time-frequency distribution, and $\{e_{i,j}\}$ the set of the corresponding segments, the objective is to split the complete MST into two parts : $S = S_1 \cup S_2$, so that S_1 and S_2 are maximally different while being, individually, maximally coherent. In other words, the question is to find a separatrix c on the MST, defined by :

$$c = \arg\min_{e_{i,j}} \max\{H(S_1), H(S_2)\}, \qquad (11)$$

where $H(.)$ is some cost function. If the constraint is to minimize the maximum entropy of the resulting distributions, one can choose for H the Rényi entropy, as estimated by MST's. Such an approach reformulates the problem of detecting time-frequency components as a "clustering" problem on the set of local maxima of a time-frequency distribution. Bidimensional MST's can therefore be applied to each of the resulting subsets. One can remark that, in this case, the usual euclidean norm ($\gamma = 1$) in a space of dimension $d = 3$ leads necessarily to using the Rényi entropy of order $\nu = 2/3$. Determining the best order to use still remains an open problem.

The results obtained by using this procedure are in full concordance with those previously obtained in a 2D context (figure 2), and are therefore not shown again. For evidencing the efficiency of the proposed method, figures 3 and 4 present further results obtained by extending the 3D approach to the example of a two-component signal embedded in noise.

5. CONCLUSION

A novel method has been proposed for the automatic skeletization of spectrograms. The approach, which relies on information-theoretic criteria, presents the advantage of being fully non-parametric and robust. In particular, no a priori knowledge is required concerning statistical properties of the noise distribution in the time-frequency plane.

6. REFERENCES

[1] A.O. Hero, O. Michel, "Robust entropy estimation strategies based on edge weighted random graphs," SPIE'98, San Diego (CA), 1998.

[2] A.O. Hero, O. Michel, "Asymptotic theory of greedy approximations to minimal k-points random graphs," *IEEE Trans. on Info. Theory*, vol. 45, pp. 1921–1938, 1999.

[3] A.O. Hero, O. Michel, "Robust estimation of point process intensity features using k-minimal spanning trees," Proc. of ISIT'97, Ulm (Germany), p. 74, 1997.

[4] R. Ravi, M. Marathe, D. Rosenkrantz, S. Ravi, "Spanning trees short or small," in *Proc. 5th Annual ACM-SIAM Symp. on Discrete Algo.*, pp. 546–555, 1994.

[5] J.M. Steele, "Growth rates of euclidean minimal spanning trees with power weighted edges," *Ann. Prob.*, vol. 16, pp. 1767–1787, 1988.

[6] J. Beardwood, J. H. Halton, J. M. Hammersley, "The shortest path through many points," *Proc. Cambridge Phil. Soc.*, vol. 55, pp. 299–327, 1959.

[7] O. Michel, A.O. Hero, "Pruned MST's for entropy estimation and outlier rejection, " IEEE-IT Workshop on Detection, Classification and Imaging, Santa-Fe (NM), 1999.

[8] R. Hoffman, A. K. Jain, "A test of randomness based on the minimal spanning tree," *Patt. Rec. Lett.*, vol. 1, pp. 175–180, 1983.

[9] D. Banks, M. Lavine, H.J. Newton, "The MST for nonparametric regression and structure discovery," in *Comp. Science and Stat. Proc. of the 24th Symp. on the Interface*, pp. 370-374, 1992.

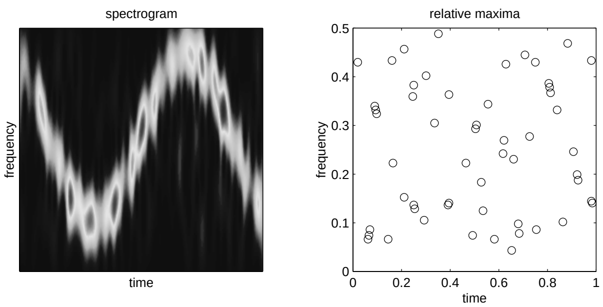

Figure 1: Example of a monocomponent frequency modulated signal, at 5dB SNR. Left : spectrogram; Right : (2D) distribution of the relative maxima

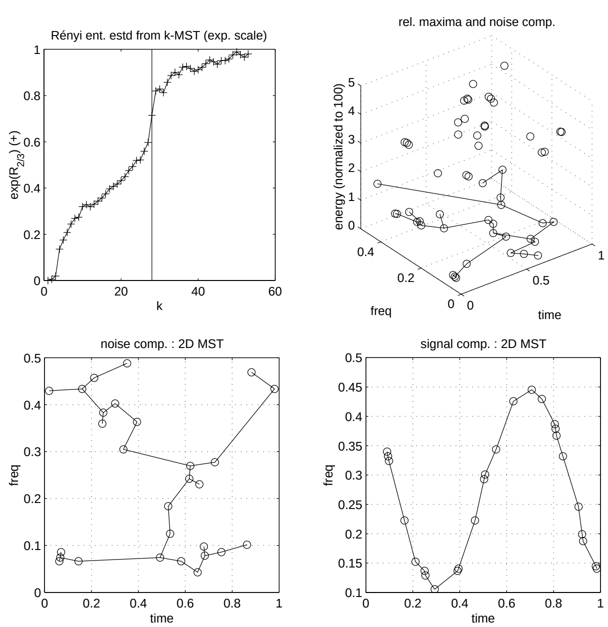

Figure 2: Component separation using Rényi entropy. Top Left : Entropy estimated from k-MST length, and threshold detection (largest entropy increase : $k = 28$). Top right : 3D 28-MST. Bottom : 2D MST's of identified components.

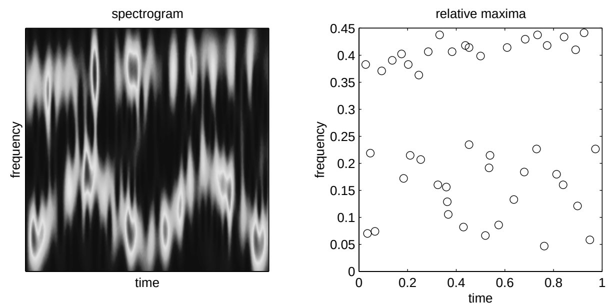

Figure 3: Example of a two-component frequency modulated signal, at 5dB SNR. Left : spectrogram; Right : (2D) distribution of the relative maxima

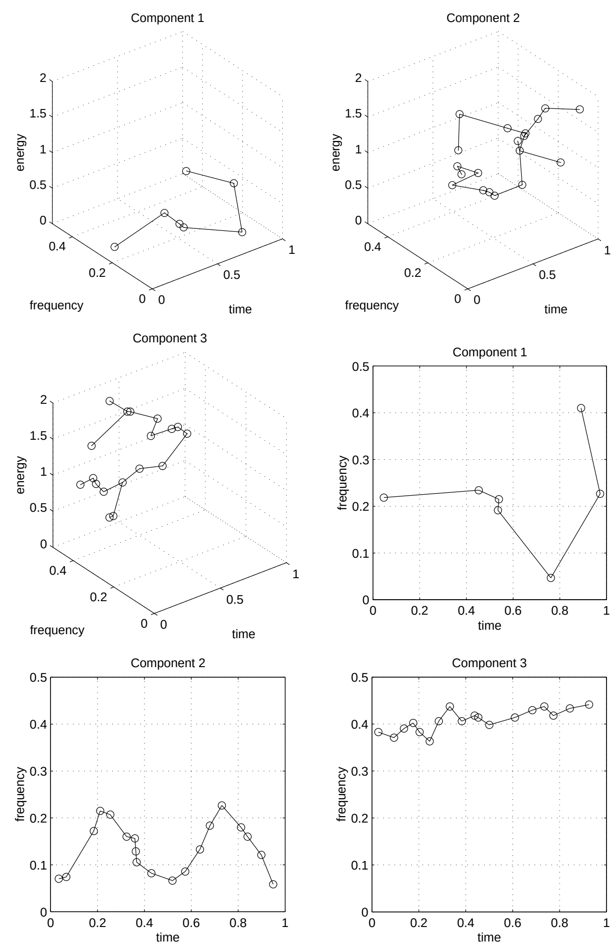

Figure 4: Direct 3D approach in the two-component case : the 3D plots show the 3D MST's of the identified components, whereas the 2D plots show their respective MST's in the time-frequency plane.