Audio Engineering Society
Convention Paper
Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

Frequency Modulation Tone Matching Using a Fuzzy Clustering Evolution Strategy

Thomas J. Mitchell[1], J. Charles W. Sullivan[1]

[1] *Faculty of Computing Engineering and Mathematical Sciences, University of the West of England, Bristol, UK*

Correspondence should be addressed to Thomas Mitchell (tom.mitchell@uwe.ac.uk)

ABSTRACT
Frequency Modulation parameter estimation has provided a continual challenge to researchers since its first application to audio synthesis over thirty years ago. Recent research has made use of basic evolutionary optimisation algorithms to evolve sounds produced by non-standard Frequency Modulation arrangements. In contrast, this paper utilises recent advances in multi-modal evolutionary optimisation to perform dynamic-sound matching with traditional arrangements. In doing so, a technique is developed that is not synthesiser dependent, and provides the potential for alternative methods of synthesis control.

1. INTRODUCTION

There is now an abundance of (complex) synthesis methodologies, each of which are capable of producing a diverse assortment of timbres. Often the synthesiser interface is a reflection of the underlying synthesis process, and rarely do its parameters relate to sound in human terms. Consequently, there is often a large discrepancy between the dimensions of the synthesiser parameter space and the perceived sound space. When such a discrepancy exists, synthesiser control is often unintuitive and difficult to learn. Inexperienced users/programmers would benefit from a procedure that relates to their mental picture of timbre. For example, it may be desirable to specify a 'brassy' tone, the spectral characteristics of which are well known [1]. A process is therefore required that is able to map known sound qualities onto sound synthesis parameters: a matching technique that can efficiently search a synthesis parameter space for specific spectral requirements.

There have been numerous attempts to transfer synthesiser control into a more intuitive domain [2] [3]. The most promising recent developments utilise the optimisation principles of Evolutionary Computation (EC) for sound navigation and exploration [4] - [10]. When EC is used, assistance is generally provided in one of two forms: interactive evolution, where the user controls the direction of the search as evolution takes place; or sound matching, where the evolutionary search explores the space to find a close match to a given target sound. It is the later that is of interest here.

This work applies an advanced form of Evolution Strategy (ES) to optimise simple dynamic-spectra Frequency Modulation (FM) parameters that synthesise good matches to randomly generated target sounds. It builds upon the previous work, presented by Horner [11] [12], and has wider implications as a platform for a generic synthesiser interface that is not specific to the underlying synthesis type.

Ultimately, this technique will allow users to specify target spectral characteristics, which a synthesiser (of any type and known form) will approximate. At present, that technique is able to match target spectral forms with an FM synthesiser, where it is possible to accurately match the target. That is, when the target spectrum was originally generated with an FM synthesiser.

The next section of this paper will introduce FM audio synthesis, the dynamic-spectra FM models implemented for sound matching, and early attempts at parameter estimation. Section 3 will introduce previous evolutionary sound matching work. Section 4 will provide some background information on EC, specifically for the ES. Section 5 will outline the optimisation engine utilised for this work, and Section 6 will outline FM sound matching procedure, followed by results and plans for future developments.

2. FREQUENCY MODULATION SYNTHESIS

FM audio synthesis, presented originally by Chowning [13], provides a synthesis method by which complex spectra can be created simply and efficiently. In what is termed *simple* FM, the instantaneous frequency of one oscillator is modulated by another, to produce a tone with multiple frequency partials Fig. 1(a). The amplitude function for *simple* FM is given by the formula

$$e = Asin(\omega_c t + Isin\omega_m t). \quad (1)$$

where e is the modulated carrier amplitude, A the peak amplitude of the carrier, ω_c and ω_m the carrier and modulator angular frequency, and I modulation index, given by ratio of the frequency-deviation to the modulating frequency.

Modulation produces side-bands in the frequency domain, with partials deviating from the carrier at integer multiples of the modulating frequency. The amplitudes of frequency partials are determined by the Bessel functions of the first kind and nth order. The bandwidth of the output signal increases as the modulation index is raised, as can be observed in Fig. 2. Notice that as I is raised the amplitude of each partial varies according to a non-linear (Bessel) function. This can make it hard to achieve a target sound when altering parameters by hand. For further reading into the spectral decomposition of FM signals, the reader is referred to [14].

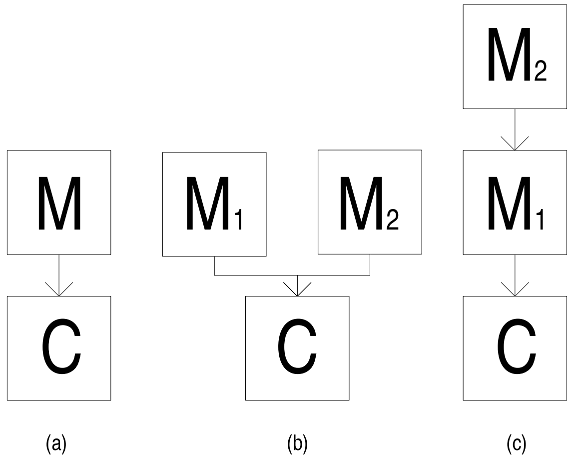

Fig. 1: Simple, double-modulator and nested FM arrangements

Schottstaedt [15] extended Chowning's basic FM arrangement developing models that allowed string instrument tones to be simulated, specifically the piano. In doing so, he introduced two important forms of FM synthesis with complex modulation.

The first of these models, namely *double-modulator* FM (Fig. 1(b)), instantaneously modulates the carrier frequency with the sum of two modulating sinusoids; given by

$$e = Asin(\omega_c t + I_{m_1} sin\omega_{m_1} t + I_{m_2} sin\omega_{m_2} t) \quad (2)$$

with I_{m_1} and I_{m_2} the modulation indices, and ω_{m_1} and ω_{m_2} the two modulator angular frequencies.

With *double-modulator* FM, the spectrum that results is as though the partials produced by the mod-

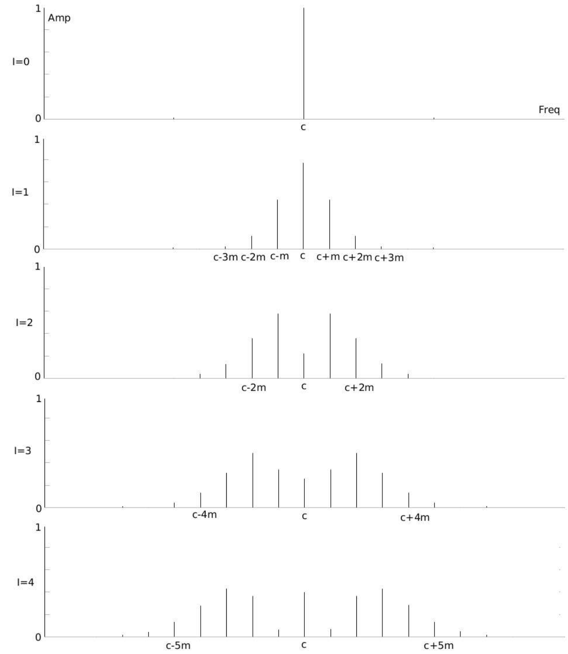

Fig. 2: Synthesised FM spectra with increasing modulation index I

ulation of the carrier by one modulating oscillator, are modulated (as carriers) by the second. For the mathematical derivation of the *double-modulator* FM spectrum, in terms of the Bessel Functions, the reader is referred to [14] and [15].

The second arrangement, *nested-modulator* FM (Fig. 1(c)), modulates the carrier with a modulating sinusoid, which is itself modulated by a third modulating sinusoid; given by the equation

$$e = Asin[\omega_c t + I_1 sin(\omega_{m_1} t + I_2 sin\omega_{m_2} t)] \quad (3)$$

The carrier side-bands are positioned as they would be in *simple* FM, however, the partials each have their own side-bands at frequency intervals of the second modulator frequency. The full mathematical derivation of the nested-modulator spectrum can be found in [12].

The *double-modulator* and *nested-modulator* FM arrangements are of great significance to the work presented in this paper, as they are both implemented within an algorithm that is able to decompose target sounds into FM parameters, using an evolutionary computation process. Details of this matching method will be outlined in Section 6. First we shall look at some of the early FM matching methods that have been documented.

2.1. Frequency Modulation Sound Matching

As can be observed in Fig. 2, there is a complex mapping between FM parameters and the spectral form of the corresponding sound. As such, FM parameter control may appear random and complicated to inexperienced users. The desire to control the sonic diversity that FM provides has motivated a series of studies intent upon providing a systematic means by which FM synthesis can be used to simulate real acoustic instruments. Chowning's original paper initiated interest in this direction, providing example parameters that simulate brass, woodwind and percussive tones.

The earliest attempts at automating sound design with FM were made by Justice [16], creating a phase-analysis procedure, based upon the Hilbert transform, that attempts to decompose signals into *nested-modulator* FM parameters. To verify the success of his algorithm, Justice presented some successful experimentation extracting the parameters of *contrived nested-modulator* target FM signals. However, if the signal does not conform exactly to a product of the nested FM model (eq. (3)), unavoidable error in the approximation is observed. Some theoretical analysis is provided for the application of the model to general signals, but further experimentation is left as future work. Justice suggests that such a system may provide users with a means by which they can jump to approximate regions of the synthesis sound space, leaving finer adjustment to be performed by hand. The process is specific to nested-modulator FM, works only with slow moving oscillator envelope shapes, and does not allow for the FM phenomenon of spectral wrapping[1].

Justice's analytical process of parameter derivation was later extended by Payne [17] to process multiple-carrier nested-modulator FM arrangements. The paper outlines in detail numerous restrictions to

[1] where side-bands synthesised with negative frequency are mapped onto their positive frequency values with inverted phase

which the target signal should conform, but, even when all are met, the process is not always successful. A comparable technique of FM parameter decomposition was proposed by Delprat [18]. FM *law-extraction* interprets formations in the Gabor transform coefficients of the analysed signal to estimate FM parameters. Experimental results are presented that show promising partial interpretation of the coefficients, however the system is not complete and, like many of the procedures outlined in this section, full development is left as future work.

Recent advances, matching target signals with FM, have applied the robust optimisation techniques of evolutionary computation to match time-variant sounds. This is the discipline within which the present work resides.

3. EVOLUTIONARY FREQUENCY MODULATION SOUND MATCHING

Research at the intersection of artificial intelligence and music has provided a collection of studies intent on providing intuitive synthesiser control with evolutionary computation [4] - [10]. The most relevant to the presented work, and indeed the work on which it is based, are the evolutionary FM matching systems of Horner [11] [12].

To facilitate the matching of acoustic instrument tones, Horner's algorithm optimises a set of static basis-spectra generated via FM, which are dynamically recombined to simulate a given harmonic target tone. The synthesis process is therefore very close to that of wavetable synthesis, with FM used only in the production of basis-spectra. The basis-spectra is generated using a simple FM arrangement where the modulator is tied to the fundamental frequency, and the carrier frequency is set to integer multiples thereof, known as *formant* FM. This arrangement is excellent for use in conjunction with wavetable synthesis, as regions of the target spectrum can be reproduced by individual basis-spectra, allowing the optimum spectral envelopes to be established. The restriction of the carrier frequency to an integer multiple of the modulating frequency ensures that all of the basis-spectra are harmonic, and supports the use of a Genetic Algorithm (GA) for optimisation purposes: GAs perform their genetic operations on bit-strings, which make them ideal for integer based combinatorial search domains, such as this.

Whilst Horner's synthesis method provides a means by which static FM spectra can be combined to produce dynamic sounds, dynamic-spectra FM synthesis arrangements have existed for many years. A simple model is provided in Chowning's original paper, where the modulation depth is controlled by a simple envelope to produce dynamic sounds (Fig. 1(a)). The models presented by Schottstaedt [15], also allow such spectral control (Fig. 1(b) and (c)). Combined, Chowning and Schottstaedt's models have formed the basis on which commercial FM has evolved. Consequently, Horner's model cannot be applied directly to explore the sound space of regular FM, as it exploits an alternative synthesis paradigm.

In [11] Horner does experiment with non-static modulation indices interpolating parameters across fixed spectral matches that are made throughout the duration of the target tone. This method proved problematic, as parameter variation from one spectral match to the next is non-continuous, and the perceived tone travels through harsh transitions between the points at which the static matches are made.

The work presented here utilises the basic dynamic-spectra FM configurations outlined above, representing parameters with real valued numbers. When the synthesis variables are not limited to integer numbers, the search is performed across the entire parameter space to locate optimal solutions. This operation is a non-trivial process, as the FM object landscape is extremely complex and multi-modal[2]. Early attempts with simple evolutionary optimisers, like the *simple-GA* and basic ES proved insufficient for the FM matching problem. As such, a specialised optimisation algorithm known as FCES is employed for this work that is designed to operate within such harsh conditions. Before the FCES can be described, the next section will provide a brief overview of evolutionary computation, introducing the basic evolution strategy, the model on which the FCES is built.

4. EVOLUTIONARY COMPUTATION

Evolutionary computation is now a well established research field that has derived inspiration from biological evolution. Three main branches of this discipline, developed independently and contempora-

[2] where the object landscape contains many peaks.

neously, are concerned with function optimisation: evolutionary programming [19], genetic algorithms [20] and evolution strategies [21]. The genetic algorithm has been mentioned already as it was the optimisation engine chosen by Horner in his matching work. For further information on evolutionary programming or genetic algorithms the reader is directed to the associated references; this section will expound only on the basic evolution strategy as this is the general model to which the FCES conforms.

4.1. Evolution Strategies

The Evolution Strategy was originally developed in the 1960's by two students of the Technical University of Berlin [21] [22]. Presented as an automatic engineering design optimiser, shown to outperform traditional gradient oriented techniques, evolution strategies have since undergone numerous modifications and enhancements.

The original evolution strategy used a simple iterative mutation-selection mechanism in which a single 'parent' generates a single offspring, the latter is subject to a mutation operator and the stronger individual forms the parent of the next generation. This formed the precursor of a series of more elaborate strategies which allowed first multiple offspring, then the concept of limited life-span, and finally a population of 'parents' to form the gene pool from which the next generation inherits.

Search points in ESs are usually n-dimensional vectors (object variables) of real valued (in practice fixed-length floating point approximations) parameters. Additionally, each individual normally includes a vector of strategy variables which evolve together with the object variables in a process which has been termed *self-adaptation* by Schwefel [22].

Mutation consists of the addition of a normally distributed (zero mean) random number to each component of the object variable vector, corresponding to a step in the search space. The variance of the step-size distribution is itself subject to mutation as a strategy variable.

Recombination normally takes one of two main classes:

- Intermediate - the genotype/phenotype vector of each offspring is obtained by taking the mean vector of its parents' vectors.
- Discrete - dynamic n-point crossover: each component of the genome of the offspring is produced by choosing either the vector component of the first or the second parent with equal probability.

Selection operators choose the best individuals from the current generation to act as parents for the next, those not 'fit' enough to reproduce, are simply removed from the population through the natural replacement processes of the ES. A range of strategies are possible here, dependent upon the maximum allowable age of the parents. The extremes (and the most usual in practice) are:

- *extinctive* strategies - parents live for a single generation only.
- *preservative* strategies - selection operates on the joined set of parents and offspring, very fit individuals may survive indefinitely.

4.2. MultiModal Evolution Frequency Modulation Object Landscape

The compact FM synthesis model is well suited to evolutionary optimisation, as a broad range of timbres can be accessed via a relatively small number of object parameters. That is not to say that the task of matching tones with FM is trivial, the varying success of the work reviewed in sections 2.1 and 3 informs us that this is not the case. The FM model, in fact, presents an extremely complex real world multi-modal optimisation problem that, until recent developments, would, in all likelihood, be insurmountable by early optimisation engines.

Since the earliest applications of evolutionary computation to numerical optimisation problems, a recurrent theme has been the problem of reliably finding the global optimum in a multi-modal fitness landscape. Selection, acting on a finite population, will tend to cause stable convergence on a single peak; the consequent loss of genetic diversity prevents further exploration except as the result of random mutation. If there are equal peaks, the choice will be random because of the stochastic variations inherent in the genetic operators. Even in the case of unequal peaks, there is no guarantee that the algorithm will always converge on the global optimum.

In order to prevent such premature convergence and stagnation, a number of operators have been designed with the aim of adapting the search strategy to evolve stable separate sub-populations, each of which can converge on a different sub-domain. Examples of such techniques include crowding [23], Restricted Tournament Selection [24], and fitness sharing [25]. The method chosen for this work is an extension to the standard ES known as FCES, developed by Sullivan [26], which uses fuzzy clustering and recombination operators to exploit the partitions of the ES population. The choice has been made as the FM model is both real-valued and multimodal, domains within which FCES is designed to function.

5. FUZZY CLUSTERING EVOLUTION STRATEGY

FCES combines the powerful local search properties of the evolution strategy with the strengths of Fuzzy Clustering, by partitioning the search population into fuzzy sub-populations that locally recombine and progress. With a sufficient number of clusters, and an adequate population size, all of the locally optimal peaks can be identified and thus, a global optimum is consistently found.

Clustering, as a tool for global optimisation [27], was previously utilised to provide multiple start points for a local hill-climber optimisation. FCES follows essentially the same framework but uses a stochastic population-based search (the ES) in place of the local optimisation algorithm and proceeds by alternate application of optimisation and clustering. The aim is to achieve the reliability of clustering methods with the efficient self-adaptive search behaviour of the ES approach.

The basis of the approach is that a clustering algorithm is used to form a partition of the parent population in a regular ES. The algorithm, therefore, is consistent with the standard generational model of an Evolutionary Algorithm with global selection. Subsequent recombination blends genetic material from all parents in proportion to their degree of membership of a particular cluster (fuzzy clustering). This allows clusters, within the population, to form independently at regions of high fitness within the object landscape, preventing premature global convergence at locally optimal peaks.

6. FREQUENCY MODULATION SOUND MATCHING WITH FUZZY CLUSTERING EVOLUTIONARY STRATEGY

This section will provide an overview of the FCES tone matching procedure providing details relevant in its application to dynamic-spectra FM arrangements.

6.1. Target Selection

For the matching procedure to commence, the algorithm requires a target. It is possible to insert any sound into the model at this point; however, for testing purposes, it is useful to follow the methodology presented by Justice [16] and Payne [17]: matching *contrived* target sounds produced by a FM model identical in structure to the matching synthesiser. In such circumstances, a successful match will yield parameters equal to those with which the target tone was produced and, with repeated tests with a variety of targets, demonstrates that any point within the sound space is accessible via the matching process. The alternative, if non-FM target sounds are matched, would be to perform an enumerative search in parallel with with the evolutionary match, to ensure that the 'fittest' evolved solution is indeed globally optimal. Even at a very low search resolution of 0.5 (in the range 0-5), models with 18 dimensions (the maximum that will be optimised at this stage) would require assessment in the order of some 10^{18} potential solutions, an infeasible task. In contrast, where a large number of *contrived* target sounds are consistently and accurately matched, it may be postulated that an optimal match for *any* target sound is feasible. For non-FM target sounds the optimal match may not be an exact simulation, rather, it would be the best match that could be found within the synthesis sound space.

6.2. Synthesis Components

Each dynamic-spectra FM circuit is constructed from two or three sinusoidal oscillators. The parameters of each FM oscillator are provided in Table 1.

The 'Frequency' parameter controls the oscillator frequency expressed as a multiple of the of the synthesiser fundamental[3], 'Attack' is the time taken for the output of the oscillator to reach the 'Level' value from its starting point of zero, and the 'Sustain' is represented as a percentage of the 'Level' parameter.

[3]which is set to 220Hz for the presented experimentation

Parameter	Range
Frequency	0.0 - 5.0
Level	0.0 - 5.0
Attack(time)	0.0 - 0.5s
Decay(time)	0.0 - 0.25s
Sustain(level)	0.0 - 1.0
Release(time)	0.0 - 0.25s

Table 1: Oscillator Component Parameters

In practice the envelope parameters are all scaled onto the range 0-5 so that all parameters operate within the same limits.

The output from each sinusoidal component can be connected to either the audio output, or the input of another oscillator, to modulate its instantaneous frequency. In the latter connection, the 'Level' parameter forms the Modulation Index.

6.3. Fitness Assessment

The FCES requires a means by which good and bad solutions can be differentiated. A metric is required to provide the 'distance' between (the synthesis of) a potential solution and the target sound[4]. The objective function identifies strong offspring, facilitating their selection as parents from which subsequent offspring can be produced.

Within this work, the 'distance' is measured by accumulating the squared error that is measured between the target and candidate spectra at multiple points throughout their duration. This error measure has proved effective in previous studies [8] [9] [11] [12] and offers an excellent balance between detail and execution speed.

The squared error is given by the equation

$$error = \sum_{b=0}^{N_{bin}} (T_b - S_b)^2 \qquad (4)$$

Where T is a vector of the target spectrum amplitude coefficients, S a vector of synthesised candidate spectrum amplitude coefficients and N_{bin} the number of frequency bins produced by spectrum analysis.

For the matching of instrument tones Horner obtained good results by placing ten snap-shots throughout the perceptually critical initial transient and another ten equally spaced throughout the remainder of the sound [12]. In this work, matching *contrived* dynamic tones, twenty spectral snap-shots are taken at fixed intervals throughout the target sound. This may be reduced by placing snap-shots at regions of the target sound where fast changes can be observed, however, further investigations into 'intelligent' snap-shot placement is, at this point, a future plan for this work.

A complete cycle of the objective function is as follows:

1. Insert candidate solution into the FM model,
2. Subject the corresponding synthesised waveform to spectral analysis,
3. Calculate SSE between target and synthesised candidate spectra.

A variant of the squared error has been developed that allows error to be measured across a weighted band, details of this process can be found in [28]. The *windowed* squared error smooths the surface of the fitness landscape. Consequently, the search complexity is reduced, lowering the population size that is required for a successful match. However, this error metric addresses issues specific to the FM problem and has, thus, not been been implemented in the results presented here as it compromises the generic nature of the matching procedure.

6.4. Processing

With the target sound, matching synthesiser and fitness function in place, the matching process of Fig. 3 can begin.

The target tone is first analysed and the extracted spectral snap-shots are passed into the fitness function. The initial population is seeded with randomly generated solutions, which are assessed for similarity with the target tone. The best solutions are then selected to form the parents of the first generation. The parent population is partitioned into clusters, and the membership, to which each solution belongs to each cluster, is calculated. Recombination then blends the genetic information of the parents to create the offspring which, like their random ancestors, are assessed for fitness. Parents are

[4] where a good match is positioned 'close' to its target

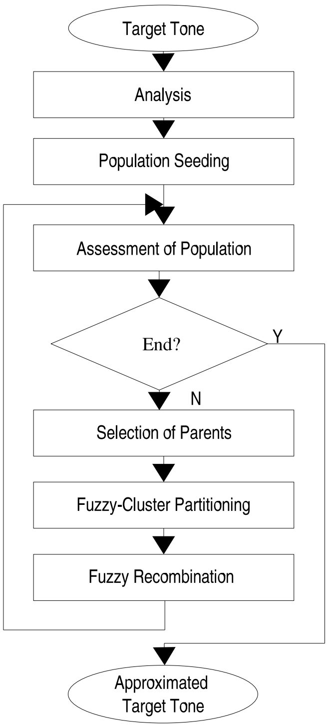

Fig. 3: Matching Procedure

again selected and partitioned, their cluster membership calculated, recombined, and so forth. The loop continues until some end criterion is met: either the generational limit is reached or a sufficiently strong solution located.

7. RESULTS

This section provides the early (promising) results obtained when a FCES algorithm is applied to match *contrived* sounds with dynamic-spectra FM arrangements. One single run of the matching process is as follows:

1. Generate a *contrived* target sound.
2. Extract spectral snap-shots from the target.
3. Match the target tone by searching the relevant FM parameter space for an optimal match.

Contrived target sounds are produced by randomly generating N_{dim} synthesis parameters (where N_{dim} is the number of input parameters to the model - the dimensionality of the problem) in the range of 0-5 (the parameter limits for the current arrangements), and inserting them into the matching model to produce a target tone. For experimental practicality, the target tones are limited to one second in duration.

Parameter	Target	Match
Carrier Parameters		
Frequency	2.290531	2.289480
Level	4.830795	4.816228
Attack(time)	1.045151	1.039380
Decay(time)	4.228095	4.248294
Sustain(level)	2.264006	2.279245
Release(time)	3.952239	3.961980
Modulator Parameters		
Frequency	3.270365	3.268940
Level	4.356868	4.334636
Attack(time)	3.565976	3.510057
Decay(time)	0.772121	0.860160
Sustain(level)	4.254643	4.307577
Release(time)	2.135209	1.499094

Table 2: Target and Matched parameters of a successful run

7.1. Chowning's Dynamic-Spectra models

Chowning's [13]*simple* dynamic-spectra FM model Fig. 1(a), presents a 12 dimensional matching problem. Fig. 4 provides the waveform and spectrogram of a typical *simple* FM target sound, and the corresponding match. The FCES is able to match the target tone to this accuracy in 50 generations, with parent size of 300, offspring size of 1500 and 100 clusters, which currently takes approximately 2 minutes on a 2.4GHz Pentium 4 processor. The two plots display remarkable similarity, and indeed sound identical. To confirm the accuracy of the match, the target tone parameters can be directly compared to the matched parameters. Table 2 provides the carrier and modulator parameters for the waveforms in Fig. 4.

For the carrier oscillator, the parameters are matched to within approximately 0.01 of the variable range. The modulator parameters are not matched to this degree, but are still very good. The 'Release' parameter on the modulator exhibits the most error. The most likely reason for this is that its impact on

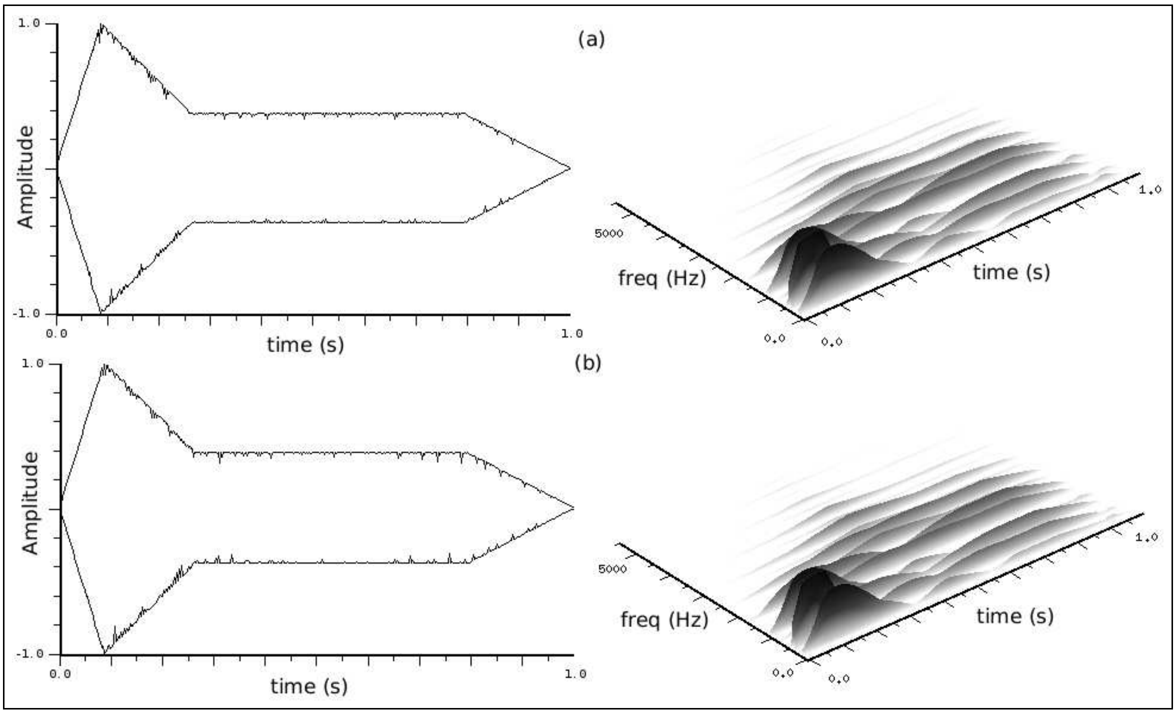

Fig. 4: (a) target sound and (b) matched sound; time and spectrogram plots

the spectral error is minimal.

Not all matches are quite as successful. As the target tone dictates the complexity of the search, in some circumstances the match may be poor. This is because the surface of the object landscape varies significantly from one target to the next. In the example provided in Fig. 4 the population sizes are well suited to the complexity presented by the target tone. With a more complex target, larger populations may be required, and conversely, simple tones may be adequately matched with a smaller population.

Tone matching is usually more difficult when certain parameters lie close to the limits of their range. For example, tones with a very low carrier frequency may have many partials that wrap around from the negative frequency range on to the positive, these interfere with positive frequency partials producing a particularly noisy region of object space around the optimum. Whilst the FCES is equipped to deal with such environments, it may require more clusters to ensure that the global optimum is located. An increase in the number of clusters requires an increase in the number of parents, which, in turn, requires an increase in the number of offspring, which all result in a prolonged search time.

Occasionally, the carrier 'Level' or modulator 'Index' parameters may converge on a non-optimal region of the search space; this may be the consequence of a short 'Attack' time in the target that is not represented well by the uniform placement of the spectral snap-shots. Solutions may escape via mutation, but often the 'Sustain' parameter is able to compensate with an opposing error. If the amplitude 'Level' is low the 'Sustain' level will automatically evolve to an increased value, to make up the difference throughout the sustain period.

In repeated runs the target tones are normally matched, with the offspring sizes that have been given. Occasionally, the target may be too complex and the FCES may converge upon a non-optimal point of the object space. Even the incorrectly matched sound can exhibit good characteristics of the target tone, which may be adequate for some

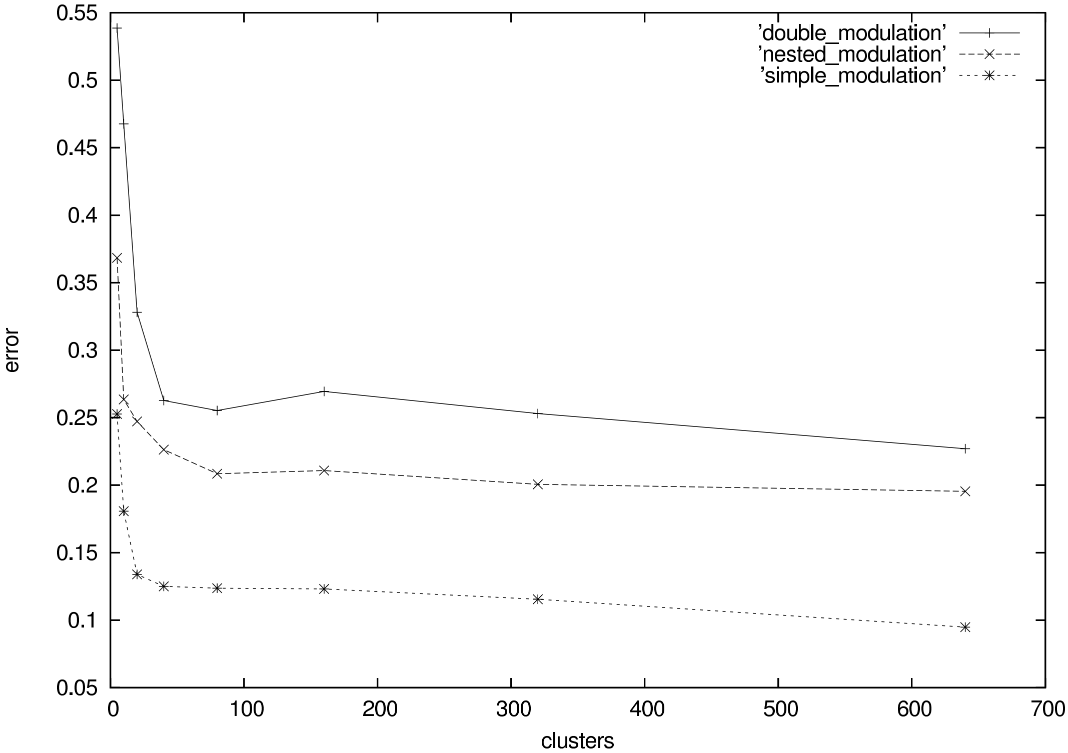

Fig. 5: Normalised error with increasing number of clusters.

applications, but if an exact match is required, the process may be repeated with a larger population size.

7.2. Schottstaedt's Dynamic Spectra models

Schottstaedt's [15] dynamic FM tone arrangements, *double-modulator* and *nested-modulator* (Fig. 1(b) and (c) respectively), each present interesting 18 dimensional optimisation problems. Despite the drastically increased complexity of the search space, early experimentation matching *contrived* FM sounds has produced excellent results; although accurate matching with this model comes at a cost: matches with the current arrangement can take up to 10 minutes (again on a 2.4GHz Pentium 4 processor). To regularly obtain good matches from the model 500 clusters are required with a parent size of 2000, and an offspring size of 10000.

One notable observation in the early matching experiments with *contrived* target tones, is that occasionally an excellent (analytical) match is made, but with vastly different parameters to the target tone. This suggests that as the model becomes more complex, not only is the potential for matching a wider diversity of sounds increased, but there may be many regions of high fitness in the search space. The advantage of using the clustering algorithm is that many of the strong regions are located concurrently, a synthesiser user may benefit from a choice between several sounds that are all similar, in some way, to their target.

Fig. 5 provides the average normalised error that is observed when ten separate randomly generated *contrived* target tones are matched with the FCES using *simple*, *double-modulator* and *nested-modulator* FM synthesis arrangements. Any partials that are not common to both the target and matched tones contribute to the error. An exact match produces an error of zero. The populations are proportioned to allow three parent solutions per cluster, with a offspring population that is five times the size of the parent population. From these early results it appears that the easiest problem to solve accurately is the *simple* FM followed by the *nested-modulator* and *double-modulator* models.

8. CONCLUSION

A new matching method has been presented that applies the FCES optimisation engine to the problem of FM sound matching for dynamic sounds. The early results are promising and, with suitable optimisation, may provide a platform for new, intuitive, synthesiser interfaces. Matching is carried out on three different dynamic-spectra FM configurations: *simple*, *double-modulator* and *nested-modulator*.

9. FUTURE WORK

The work presented here reports upon the initial steps that have been taken to develop a synthesis matching tool. In the future the model will be expanded to operate upon wider parameter ranges and larger, more complex FM arrangements. Work will be done to optimise the process to allow faster convergence that may allow for real-time tone searching. It will also be investigated as to whether the model could be applied to alternative synthesis methodologies. It is planned to arrange perceptual tests with participants, to quantify the accuracy of the algorithm.

ACKNOWLEDGEMENTS

The first author would like to thank the supervisory team, Tony Pipe, David Creasey, Marcus Lynch and Stephen Allan for their constant support, inspiration and help with this work.

10. REFERENCES

[1] D. Morrill, *Trumpet Algorithms for Computer Composition*, Computer Music Journal, 1(1):46-52, 1977.

[2] J. Ethington, D. Punch, *SeaWave: A system for Musical Timbre Description*, Computer Music Journal, vol. 18, no. 1, 1994.

[3] E. R. Miranda, *Striking the right note with ARTIST: an AI- based synthesiser*, M. Chemillier and F. Pachet (editors), *Recherches et applications en informatique musicale*, Paris: Editions Hermes, pp. 227- 239. 1998.

[4] I. Fujinaga, and J. Vantomme. *Genetic algorithms as a method for granular synthesis regulation* Proceedings of the International Computer Music Conference, 1994.

[5] C. G. Johnson, *Exploring the sound-space of synthesis algorithms using interactive genetic algorithms*, Presented at Proceedings of the AISB Workshop on Artificial Intelligence and Musical Creativity, Edinburgh. 1999.

[6] J. P. Mandelis, *Genophone: Evolving Sounds and Integral Performance Parameter Mappings*, Applications of Evolutionary Computing, EvoWorkshops2003, pp538- 549, 2003.

[7] E. R. Miranda, *On the Music of Emergent Behaviour: What can Evolutionary Computation Bring to the Musician?* Leonardo, Vol. 36, No. 1, pp. 55-58. 2003.

[8] K. Wehn, *Using Ideas from Natural Selection to Evolve Synthesized Sounds*, Proc.Digital Audio Effects (DAFX98), Barcelona. 1998.

[9] R. Garcia, *Growing Sound Synthesizers using Evolutionary Methods*, European Conference in Artificial Life ECAL2001. Artificial Life Models for Musical Applications. University of Economics, Prague, Czech Republic. 2001.

[10] J. Manzolli, A. Maia, J. Fornari and F. Damiani, *The evolutionary sound synthesis method*, Proceedings of the ninth ACM international conference on Multimedia, Ottawa, Canada 2001.

[11] A. Horner, J. Beauchamp, and L. Haken, *Machine Tongues XVI: Genetic Algorithms and Their Application to FM*, Matching Synthesis, Computer Music Journal, 17(4), 17-29. 1993.

[12] A. Horner, *Nested Modulator and Feedback FM Matching of Instrument Tones*, IEEE Transactions on Speech and Audio Processing, 6(6), 398-409. 1998.

[13] J. Chowning, *The Synthesis of Complex Audio Spectra by Means of Frequency Modulation*, Journal of the Audio Engineering Society, vol. 21, pp. 526-534. 1973.

[14] LeBrun, Marc. *A Derivation of the Spectrum of FM with a Complex Modulating Wave*, Computer Music Journal, 1(4):51-52, 1977.

[15] B. Schottstaedt, *The Simulation of Natural Instrument Tones Using Frequency Modulation with a Complex*, Modulating Wave. Computer Music Journal, 1(4):46-50. 1977.

[16] J. H. Justice, *Analytic Signal Processing inMusic Computation*, IEEE Transactions on Acoustics, Speech, and Signal Processing,Volume ASSP-27, Number 6, 1979.

[17] R. Payne, *A microcomputer based analysis/resynthesis scheme for processing sampled sounds using FM*, Proc. Int. Computer Music Conf., San Francisco, CA, 1987, pp. 282 289. 1987.

[18] N. Delprat, *Global frequency modulation law extraction from the Gabor transform of a signal: A first study of the interacting components case*, IEEE Trans. Speech Audio Processing, vol. 5, pp. 64 71, Jan. 1997.

[19] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence through Simulated Intelligence*. New York: John Wiley. 1966.

[20] J. H. Holland, *Concerning Efficient Adaptive Systems*, Pages 215-230 of: M. C. Yovits, G. T. Jacobi and G.D. Goldstein (eds), *Self-organizing Systems*. Washington,DC: Spartan Books. 1962

[21] I. Rechenberg,*Cybernetic solution path of an experimental problem*, Tech. rept. RAE Translation 1122. Farnborough, Hants. 1965.

[22] H. P. Schwefel, *Collective phenomena in evolutionary systems*. Pages 1025-1032 of: Preprints of the 31st Annual Meeting of the International Society for General Systems Research, 1987.

[23] K.A. De Jong, *An analysis of the behaviour of a class of genetic adaptive systems* Ph.D. thesis, University of Michigan, 1975.

[24] G. Harik, *Finding multi-modal solutions using Restricted Tournament Selection*. In: Proc. 6th Int. Conf. on Genetic Algorithms, 1995.

[25] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan, 1975.

[26] J. C. W. Sullivan, *An Evolutionary Computing Approach to Motor Learning with an Application to Robot Manipulators*, PhD Thesis, University of the West of England, Bristol, 2001

[27] A. A. Torn, and A. Zilinskas, *Global Optimization*. Lecture Notes in Computer Science, no. 350. Berlin: Springer-Verlag. 1989.

[28] T. J. Mitchell and A. G. Pipe *Convergence Synthesis of Dynamic Frequency Modulation Tones Using an Evolution Strategy*, EvoWorkshops2005: EvoMUSART, 3rd European Workshop on Evolutionary Music and Art, 2005.