

Perspectives actuals en la síntesi digital de sons musicals

Xavier Serra

Quan es genera un so musical electrònicament, és important tenir un bon model amb paràmetres que permetin manipular el so d'una forma intuïtiva. El músic, l'usuari, ha de poder desenvolupar una intuïció musical que li permeti experimentar amb la tècnica de síntesi que utilitzi. En aquest article es presenta una visió general de la síntesi digital, i es concentra en dues línies de recerca que segurament ens permetran trencar les limitacions existents en els mètodes utilitzats fins ara i aconseguir, finalment, fer realitat les promeses que la música amb ordinadors va fer als anys seixanta.

1. Introducció

La generació de sons ha estat el camp musical on la tecnologia digital ha tingut un impacte més gran. Però l'evolució de les tècniques de síntesi digital ha estat més lenta del que en un principi s'esperava, i fins als darrers anys no s'han trobat solucions de futur que poden donar un nou impuls a aquesta evolució.

En un article històric que va aparèixer a la revista *Science* l'any 1963, Max Mathews (Mathews, 1963) mostrava una visió molt optimista de l'ordinador com a instrument musical. L'autor, un pioner de la música amb ordinadors, deia que generar sons a partir de nombres era una forma completament general de sintetitzar sons, ja que l'àmbit de freqüències i dinàmiques de l'oïda estan delimitats i, per tant, qualsevol so que puguem percebre pot ser generat d'aquesta manera. La promesa de la música amb ordinadors era que l'ordinador era capaç de generar qualsevol so que pogués sortir per un altaveu. Uns anys més tard, el mateix Max Mathews, en el seu llibre sobre la tecnologia de la música amb ordinadors (Mathews, 1969), escrivia:

"Els dos problemes fonamentals en la síntesi del so són (1) la gran quantitat de dades necessàries per definir una ona acústica - per tant, la necessitat d'un programa molt ràpid- i (2) la necessitat d'un llenguatge simple i potent amb què descriure una seqüència complexa de sons."

El problema (1) ja ha estat resolt en gran mesura per l'evolució tecnològica, ja que la velocitat dels processadors digitals s'ha anat incrementant de forma exponencial durant els darrers vint anys. El problema (2) continua sense tenir una solució satisfactòria pel fet que és impossible descriure els sons si s'ha de definir cada un dels nombres que representen una ona acústica; s'ha de poder descriure el so a partir de menys nombres o s'ha de partir de sons enregistrats. Afortunadament, la major part d'ones no són musicalment interessants i moltes ones físicament diferents són perceptualment iguals. Per tant, no cal generar totes les ones possibles, i l'objectiu es centra a trobar un grup reduït de tècniques de síntesi i de control que ens permetin explorar tot l'espai tímbric d'interès musical.

L'aproximació tradicional a aquest problema ha estat intentar aconseguir sons combinant elements de síntesi simples. Però en els darrers anys ha quedat clar que és difícil generar d'aquesta manera sons complexos d'interès musical i que per poder competir amb els instruments acústics, en termes de complexitat sonora i control expressiu, es requereix un nou enfocament de la síntesi digital. L'enfocament que aquí proposem parteix de l'estudi de la realitat sonora que ens envolta i busca formes d'estendre aquesta realitat a un món sonor nou, virtual.

2. Tradició de la síntesi digital

Les tècniques de síntesi digital van heretar els coneixements desenvolupats per a la síntesi amb mitjans analògics. El primer sistema de síntesi digital, Music V (Mathews, 1969), desenvolupat per Max Mathews durant els anys seixanta, va introduir el concepte de generador unitari com a versió digital dels mòduls dels sintetitzadors analògics. Un generador unitari accepta entrades numèriques de control i genera un senyal, també numèric, que pot servir com a entrada a un altre generador unitari o ser directament un so. Exemples de generadors unitaris són: oscil·ladors, filtres, sumadors, multiplicadors, generadors d'envolupants i generadors de nombres aleatoris. A partir de la combinació d'aquests elements (figura 1) es poden crear sons sintètics similars als obtinguts amb els mòduls controlats per voltatge dels sintetitzadors analògics, però amb un control molt més precís. La major part de sistemes desenvolupats des de llavors han partit d'aquests conceptes. Els sintetitzadors basats en el protocol MIDI en són un exemple.

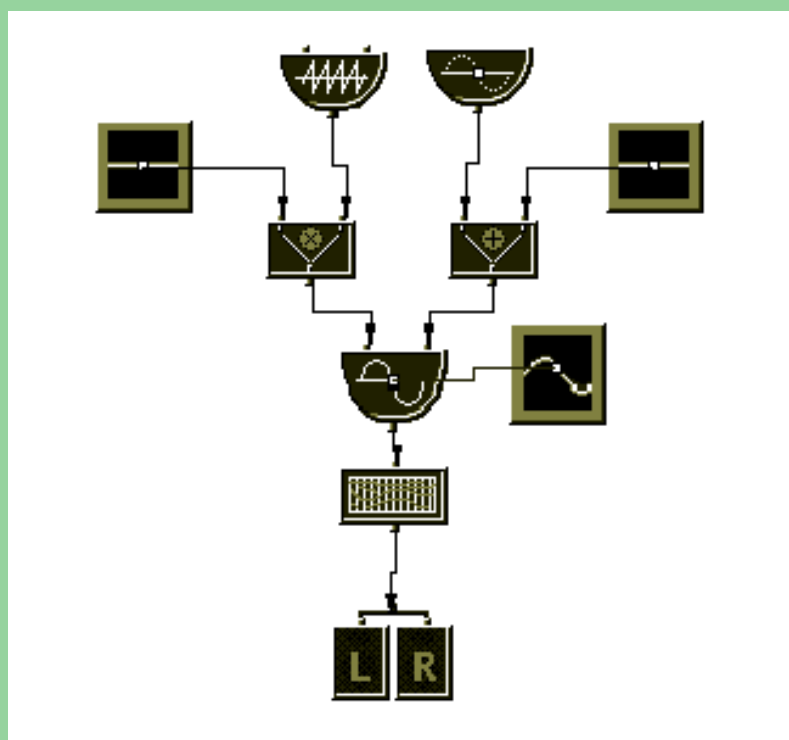


Figura 1: Exemple d'algorisme de síntesi basat en generadors unitaris.

La implementació dels algorismes de síntesi s'ha fet tant en software com en hardware. L'avantatge de les implementacions en software és la seva flexibilitat i la seva complexitat il·limitada, ja que sols estan delimitades pel llenguatge de programació utilitzat, però, per aquesta mateixa raó, difícilment poden funcionar a temps real. D'altra banda, les implementacions en hardware funcionen a temps real, però difícilment poden competir en flexibilitat i en complexitat. Donades les característiques del gran mercat, la major part de sintetitzadors comercials utilitzen implementacions en hardware. De totes maneres, gràcies al constant augment de la velocitat de processament dels ordinadors, la utilització de sistemes basats en software és cada vegada més

eficient, i actualment poden funcionar a temps real un gran nombre dels algorismes més coneguts.

Tradicionalment, les tècniques de síntesi digital s'han classificat en síntesi additiva, síntesi per subtracció i síntesi no-lineal. La síntesi additiva està basada en la suma de sons elementals, cada un d'ells generat per un oscil·lador. La síntesi per subtracció es basa en la idea complementària, és a dir, en el filtratge d'energia d'un so complex. La síntesi no-lineal és un calaix de sastre on s'inclouen un gran nombre de tècniques basades en equacions matemàtiques de comportament no-lineal.

Aquesta classificació està feta amb una visió molt teòrica i no ens porta gaire lluny en el nostre intent de trobar noves solucions d'interès musical. Una taxonomia més útil és la proposada per Julius Smith (Smith, 1991), que classifica les tècniques de síntesi en processament d'enregistraments, models espectrals, models físics i algorismes abstractes.

Una forma d'entendre la classificació de Smith és parlant dels mètodes de síntesi com a tècniques digitals que ens permeten obtenir un continu sonor que va des de la reproducció de sons preexistents (enregistraments) fins a la generació de sons a partir d'una abstracció absoluta (sons imaginats), amb tots els passos intermedis. En aquest context, les tècniques basades en el processament d'enregistraments parteixen de l'extrem dels sons preexistents i intenten crear sons nous, imaginats, retocant directament el so real (aquest és el cas dels aparells anomenats *samplers*, o mostrejadors). En l'altre extrem de la classificació hi ha els algorismes abstractes, que a partir d'equacions matemàtiques generen sons sintètics molt allunyats dels sons "naturals", però manipulant aquestes equacions s'intenta obtenir sons que permetin una comunicació musical determinada (per exemple amb els sintetitzadors basats en la tècnica de modulació de freqüència). Els models espectrals i els físics se situen en la zona intermèdia entre aquests dos extrems, i partint de models, o d'abstraccions, que ens descriuen sons preexistents i objectes generadors de sons, respectivament, permeten explorar gran part de l'espai comprès entre realitats sonores concretes i realitats noves, virtuals. Més endavant aprofundirem en aquests models.

3. Objectius musicals en la síntesi digital

Si preguntem als músics que treballen amb mitjans informàtics què voldrien poder fer en el camp de la generació de sons, molt possiblement, quasi tots estarien d'acord amb les respostes següents: (1) poder crear qualsevol so imaginable i (2) poder manipular qualsevol so preexistent de qualsevol manera que puguem concebre. Aquests objectius són, evidentment, utòpics, no solament per les limitacions tecnològiques, sinó també, i aquest és el fet més important, pels límits pràctics de la nostra imaginació. És difícil poder idear sons sense un referent del món sonor que ens envolta, i en realitat la nostra imaginació sempre parteix d'aquest referent. Això explica el gran interès de les tècniques que tenen la realitat sonora com a punt de partida.

Les limitacions tecnològiques ens obliguen a valorar una sèrie de compromisos que hem de tenir en compte a l'hora de dissenyar o utilitzar una determinada tècnica de síntesi. En concret, voldríem mencionar quatre d'aquests compromisos:

1. Qualitat del so. Per qualitat del so entenem la riquesa interna del so. Un so amb molta qualitat podria ser un so natural, i a l'altre extrem podríem tenir un so simple, sintetitzat electrònicament i sense cap microvariació al llarg de la seva durada.
2. Flexibilitat. Amb aquest terme es descriu la capacitat d'una determinada tècnica de síntesi per modificar el so a partir d'una sèrie de paràmetres de control. Segons aquest criteri, un *sampler*

seria un aparell molt poc flexible, i la síntesi per modulació de freqüència seria molt flexible.

3. Generalitat. Per generalitat s'entén la possibilitat que una mateixa tècnica de síntesi pugui generar un gran nombre de timbres. La síntesi additiva seria una tècnica molt general i l'enregistrament d'un so seria molt específica.

4. Temps de càlcul. El temps de càlcul es refereix al nombre d'instruccions d'ordinador que fan falta per generar cada una de les mostres del so sintetitzat. En aquest sentit, la síntesi per modulació de freqüència és una tècnica molt econòmica i la síntesi additiva requereix molt més temps de càlcul.

Idealment, voldríem maximitzar la qualitat, la flexibilitat i la generalitat, i minimitzar el temps de càlcul. Això no és possible tecnològicament, i en cada cas concret hem de valorar què és el que ens interessa més, i escollir la tècnica d'acord amb aquestes consideracions. Per exemple, si volem la màxima qualitat, segurament haurem de renunciar a alguns aspectes en quant a flexibilitat i generalitat; i si volem una gran flexibilitat, no podrem tenir molta qualitat. Donat que estem mirant més de cara al futur, i és evident que la velocitat dels processadors digitals anirà augmentant, ens permetem el luxe, en la nostra valoració de les tècniques de síntesi, de no preocupar-nos gaire del temps de càlcul. La nostra prioritat serà maximitzar la qualitat, la flexibilitat i la generalitat.

A més de valorar aquests compromisos a l'hora d'escollir un sistema, cal tenir en compte que perquè una tècnica de síntesi sigui útil per a un músic, el seu control ha de ser intuïtiu i, per tant, aquesta ha de partir d'una realitat sonora i musical existent. Hi ha dues realitats sonores que els músics estan acostumats a tractar i que són un bon punt de partida per crear nous sons. Una d'aquestes realitats és la dels objectes físics que generen so mecànicament, per exemple, els instruments tradicionals. Els músics tenen una intuïció clara sobre la relació que hi ha entre els objectes físics i el so que produeixen. Amb aquest coneixement és factible dissenyar nous objectes i, per tant, nous sons. La síntesi amb models físics ens permet partir d'aquesta realitat i crear objectes acústics virtuals que van més enllà de la realitat física que ens envolta.

L'altra realitat sonora d'interès per als músics és la perceptual, és a dir, el so que percep l'oïent. A partir d'aquesta realitat, el músic és capaç de descriure sons i d'imaginar-ne de nous. Amb els models espectrals representem les característiques perceptuals d'un determinat so i alhora podem manipular aquestes característiques per obtenir noves realitats perceptuals.

A continuació presentem amb més detall les tècniques de síntesi digital basades en els models físics i espectrals. Aquestes tècniques donen resposta a molts dels problemes actuals de la síntesi sonora i permeten fer una elecció molt favorable dins dels compromisos abans mencionats, especialment quant a qualitat, flexibilitat i generalitat. Hem de tenir present, però, que pel fet de ser algorismes bastant complexos difícilment s'obtenen temps de càlcul baixos.

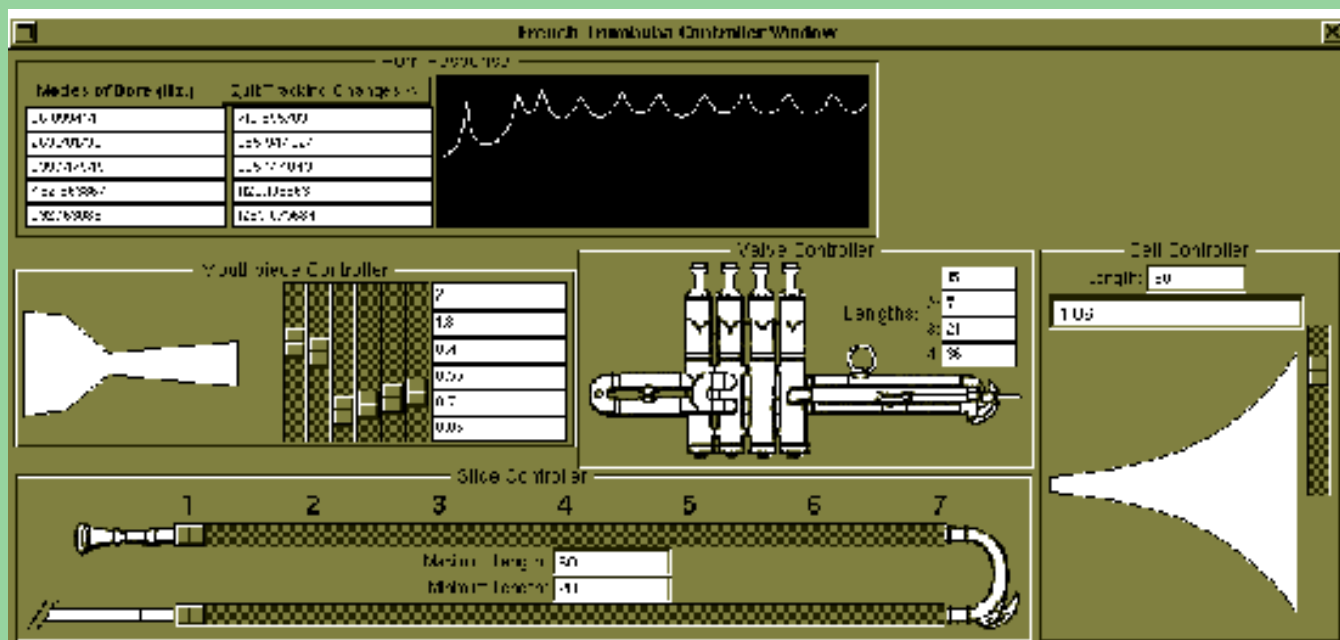


Figura 2: Interfície gràfica per un model físic d'un instrument musical.

4. Síntesi amb models físics

Aquests mètodes de síntesi generen sons descrivint el comportament dels elements que componen un instrument musical, com ara cordes, llengüetes, llavis, tubs, membranes i cavitats resonadores. Tots aquests elements, excitats d'una forma mecànica, vibren i produeixen perturbacions, generalment periòdiques, en l'aire que els envolta. És aquesta perturbació la que ens arriba al nostre sistema auditiu i és percebuda com a so.

L'acústica s'ha preocupat de trobar models, relacions matemàtiques, per descriure aquests sistemes vibratoris. Aquests models s'han utilitzat principalment per entendre fenòmens físics concrets, però des del treball pioner de Hiller i Ruiz (Hiller i Ruiz, 1971) també s'han utilitzat per a la síntesi. El primer pas per implementar aquests models consisteix a definir i mesurar les característiques físiques de l'objecte que es vol reproduir. Per exemple, per generar un so de corda hem de considerar la seva llargada, el seu gruix, la seva densitat, etc. Un cop introduïm aquestes mesures físiques, el model ens permet reproduir numèricament dins de l'ordinador el moviment de la corda i, a la vegada, convertir aquest moviment en so.

Generalment, tots els models instrumentals estan compostos de dos tipus d'elements: excitadors i resonadors (Borin i d'altres., 1992). Els excitadors són els elements que causen, i a vegades mantenen, el fenomen vibratori, i els resonadors són on tenen lloc les vibracions d'interès musical. En el cas d'un violí, l'arc funciona com a excitador i la combinació de la corda amb la cavitat de fusta fa la funció de resonador.

Dintre l'ordinador podem tenir models dels diferents elements, a partir dels quals, combinant-los i detallant les característiques físiques de cada un d'ells, l'usuari crea "instruments musicals". Amb aquests mitjans podem dissenyar i escoltar el so d'un instrument mecànic que sigui físicament impossible de construir. Per exemple, acoblant un model de doble llengüeta de canya per crear una excitació sonora amb un model de membrana de pell de vedell per fer de resonador, generarem so. Dubto que algú pugui construir un instrument com aquest fora de l'ordinador.

El control d'aquests mètodes de síntesi es fa a través dels controls mecànics que tindria un

instrument acústic amb les mateixes característiques. Així, en el cas d'un model d'instrument de vent de metall, hem de donar, a més de característiques de l'excitació produïda pels llavis, les diferents mides dels tubs i de les cavitats que componen l'instrument (figura 2).

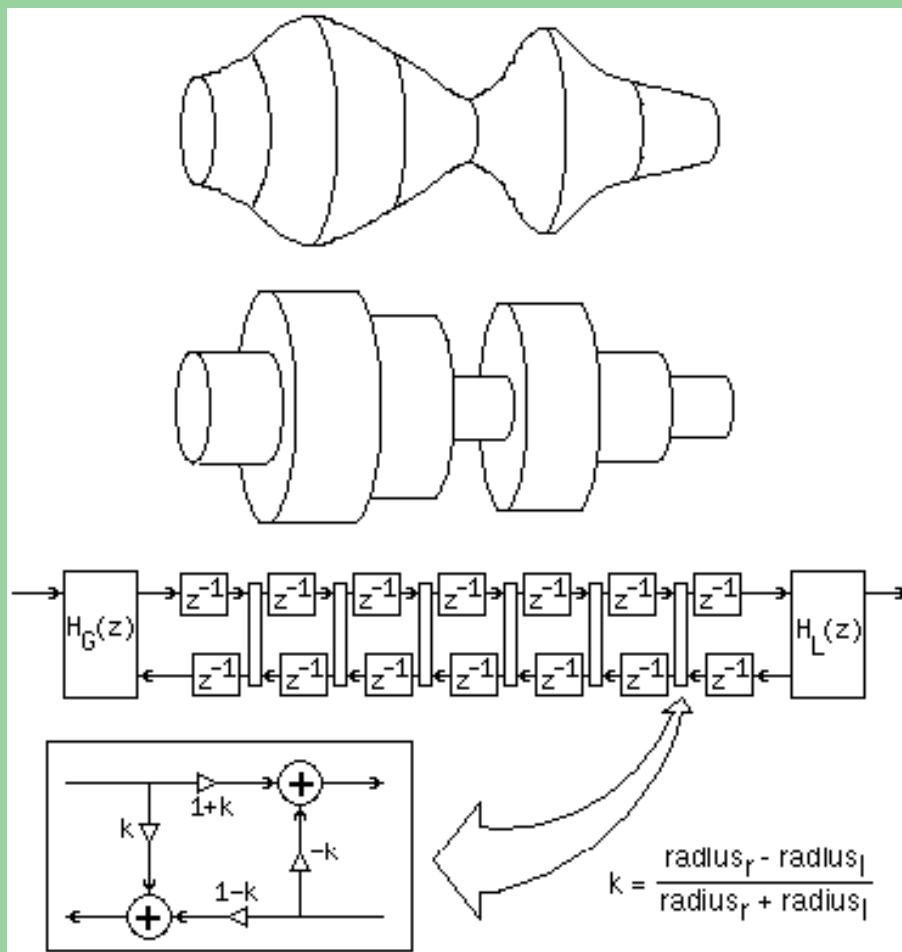


Figura 3: Un tub acústic (primer dibuix) pot ser mostrejat en l'espai al igual que es mostregen els sons (segon dibuix).

El tub mostrejat pot ser directament convertit en un filtre digital (tercer dibuix).

Històricament, els models físics han estat implementats amb algorismes molt complexos que difícilment poden funcionar a temps real amb la tecnologia actual. Aquestes implementacions s'han basat en la integració numèrica de l'equació que descriu la propagació d'ones en un fluid (Fletcher i Rossing, 1991). Recentment, s'han trobat solucions més eficients per a aquest problema (Smith 1992) (figura 3) i han començat a aparèixer sistemes interessants per als músics.

Aquests models físics poden ser considerats com a "generadors de realitat", no solament en el sentit d'imitar els instruments tradicionals, sinó també d'ajudar a conceptualitzar aquesta realitat i de crear estructures que no han de tenir obligatòriament una interpretació física. En aquest cas, la realitat física és utilitzada com a font d'inspiració i no com una referència de qualitat del so produït.

5. Síntesi amb models espectrals

Els models espectrals es basen en la descripció de les característiques sonores que percep l'oient. Per obtenir un so de corda, en comptes d'especificar-ne propietats físiques, descriurem les característiques tímbriques, o espectrals, del so de corda. Després, la generació del so a partir d'aquestes dades perceptuals es fa gràcies a diversos procediments matemàtics desenvolupats en

els darrers anys.

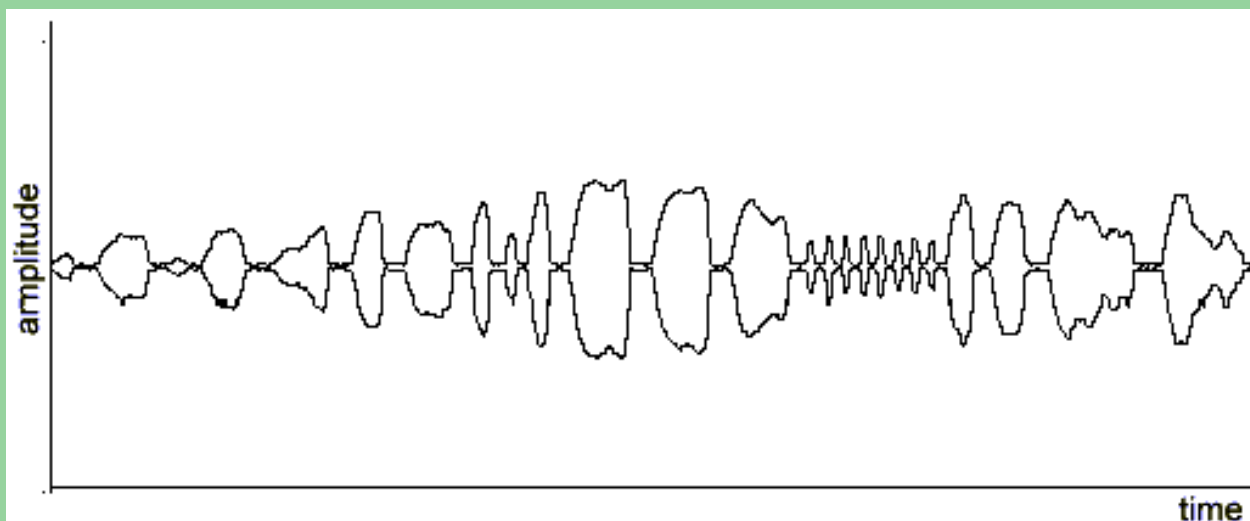


Figura 4: Representació temporal d'un so d'ocell.

Un avantatge d'aquests models és que existeixen tècniques per analitzar sons i obtenir-ne els paràmetres perceptuals corresponents. És a dir, analitzant un so determinat, obtindrem els seus paràmetres de percepció. A partir de l'anàlisi, és possible tornar a sintetitzar el so original, i en el procés, els paràmetres poden ser modificats de tal forma que el so resultant és nou, però conserva aspectes del so analitzat.

Per a un músic, és bastant intuïtiu sintetitzar sons amb aquests tipus de models. Per exemple, podem començar amb un so de clarinet, prèviament analitzat, i anar modificant-lo amb instruccions com ara: fer el so més inharmonic, fer l'atac més percussor, mesclar-hi algunes característiques tímbriques d'un so de veu, fer el timbre més brillant, etc. El so es converteix en un material plàstic que podem anar modelant al nostre gust.

La transformada de Fourier és el primer pas cap a un modelatge perceptual del so. Amb aquesta tècnica, un so (figura 4) és descompost en els seus harmònics, o components freqüencials, dels quals es pot estudiar l'evolució en el temps (figura 5). Un pas més endavant respecte a la transformada de Fourier és el descompondre els sons en sinusoides (parcials) i soroll (component residual), és a dir, analitzar sons amb aquest model i generar-ne de nous a partir de les dades d'anàlisi (Serra, 1994; Serra i Smith, 1990). L'anàlisi detecta els parcials estudiant les característiques espectrals d'un so i els representa amb sinusoides. Aquests parcials es resten del so original i el "residu" restant es representa com a soroll blanc filtrat. La part de síntesi del sistema és una combinació de síntesi additiva per la part sinusoidal i de síntesi per subtracció per la part de soroll. Aquesta estratègia d'anàlisi i síntesi pot ser utilitzada tant per generar sons (síntesi) com per transformar sons preexistents (processament). Per sintetitzar sons, generalment, volem modelar tota una família tímbrica, per exemple, un instrument, i això es pot aconseguir analitzant notes aïllades i transicions entre notes interpretades amb un instrument, i construint una base de dades que caracteritzi tot un instrument o qualsevol família tímbrica, a partir de la qual sintetitzem nous sons. En el cas de l'aplicació del processament sonor, l'objectiu és manipular qualsevol so, és a dir, no estar restringits a notes aïllades ni requerir l'obtenció prèvia d'una base de dades d'anàlisi.

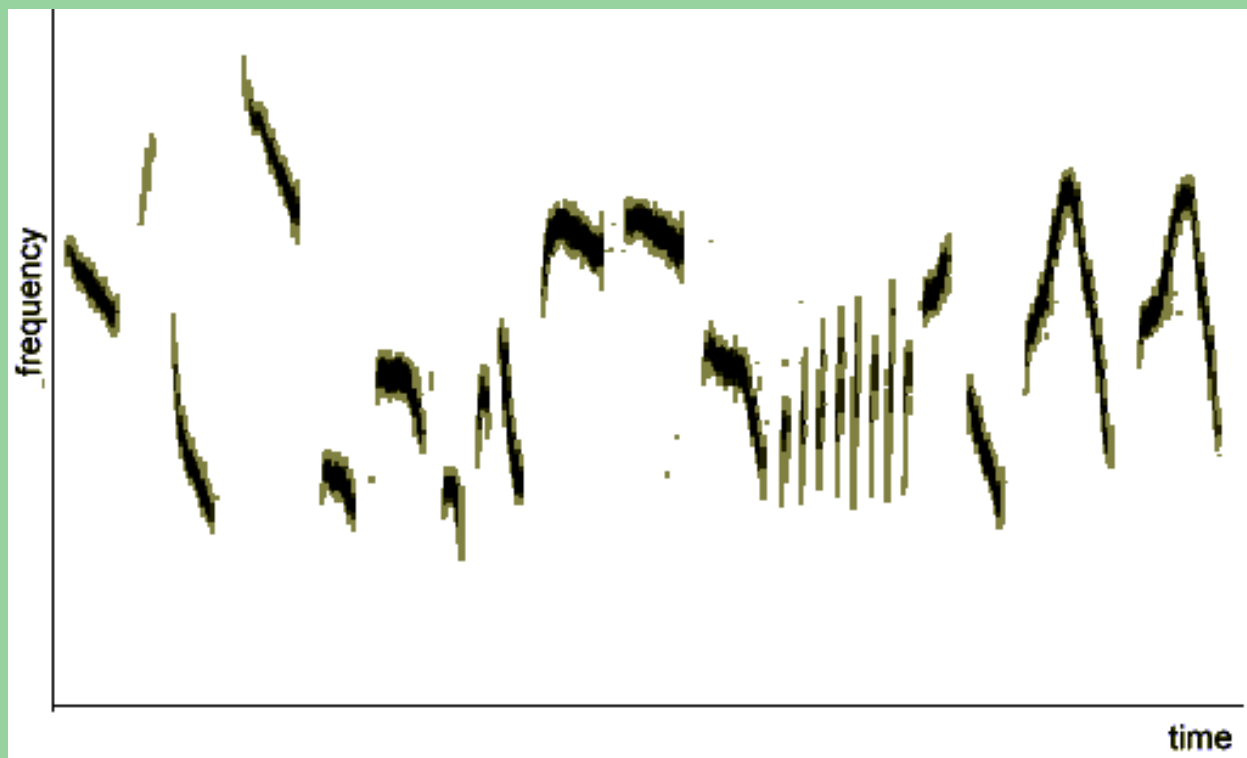


Figura 5: Representació espectral del so d'ocell de la figura 4.

Els models espectrals poden ser considerats com a "transformadors de la realitat". Igual que amb els models físics, podem reproduir una realitat ja existent, en aquest cas perceptual, i modificar-la per obtenir nous sons. La comparació entre aquests dos tipus de models (taula 1) mostra una gran complementarietat entre ells.

<i>models físics</i>	<i>models espectrals</i>
basats en la realitat física	basats en la realitat perceptual
controlats per paràmetres físics	controlats per paràmetress perceptuals
models específics per a cada instrument	models generats per a tots els sons
síntesis independents d'anàlisis	síntesis a partir d'anàlisis
ideal per reproduir instruments tradicionals	menys ideal per reproduir instruments tradicionals
menys ràpid per reproduir sons naturals	ideal per a reproduir sons naturals
temps de càlcul alt	temps de càlcul alt
permet triar entre qualitat i flexibilitat	permet triar entre qualitat i flexibilitat

més a prop de la qualitat que de la flexibilitat	més a prop de la flexibilitat que de la qualitat
--	--

Taula 1: Comparació de les característiques dels models físics amb les dels models espectrals

6. Conclusions

La situació real de la síntesi digital encara no ha arribat a aquests nivells, els sintetitzadors comercials no permeten una aproximació a la generació sonora tal i com s'ha presentat aquí; però ens hi estem acostant. Concretament, no fa gaire que va sortir el primer sintetitzador comercial basat en models físics (VL1 de Yamaha) i hi ha diversos sistemes informàtics, fora dels canals comercials tradicionals, que permeten experimentar tant amb models físics com amb models espectrals, per exemple, el software distribuït per l'IRCAM de París o el software de domini públic distribuït per la Universitat de Stanford. Però fins i tot aquests programes encara no han arribat al nivell que voldríem. Esperem que la investigació en aquest camp continuï i que no triguem gaire a disposar d'aquest tipus d'eines per a la creació musical. Que finalment l'ordinador arribi a la maduresa musical i compleixi les promeses que es van fer als anys seixanta.

7. Bibliografia

BORIN, G.; DE POLI, G.; SARTI, A. (1992) "Algorithms and Structures for Using Physical Models". *Computer Music Journal*, 16(4): 30-42.

MATHEWS, M. V. (1963) "The Digital Computer as a Musical Instrument". *Science*, 142(11): 553-557.

MATHEWS, M. V., i d'altres (1969) *The Technology of Computer Music*. Cambridge (Mass.): MIT. Press.

FLETCHER, N. H.; ROSSING, T. D. (1991) *The Physics of Musical Instruments*. Nova York: Springer-Verlag.

SMITH, J. O. (1991) "Viewpoints on the History of Digital Synthesis". *Proceedings of the 1991 International Computer Music Conference*. San Francisco: International Computer Music Association.

SMITH, J. O. (1992) "Physical Modeling Using Digital Waveguides". *Computer Music Journal* 16(4): 74-87

HILLER, L.; RUIZ, P. (1971) "Synthesizing Musical Sounds by Solving the Wave Equation for Vibrating Objects: Parts I and II". *Journal of the Audio Engineering Society* 19 (6): 462-470 i 19 (7): 542-551.

SERRA, X. (1994) "Musical Sound Modeling with Sinusoids plus Noise". *Models and Representations of Musical Signals*. Cambridge (Mass) MIT Press. [llibre en preparació].

SERRA, X. (1994) "Sound hybridization techniques based on a deterministic plus stochastic

decomposition model". *Proceedings of the International Computer Music Conference*, 1994.

SERRA, X.; Smith, J. (1990) "Spectral Modeling Synthesis: A Sound Analysis/Synthesis Based on a Deterministic plus Stochastic Decomposition". *Computer Music Journal*, 14(4): 12-24.

