



# 蓮根：

## めざせ世界一のピアニスト

平賀 瑠美

文教大学

rhiraga@shonan.bunkyo.ac.jp

平田 圭二

NTT コミュニケーション科学基礎研究所

hirata@brl.ntt.co.jp

片寄 晴弘

和歌山大学/科技园さきがけ21

katayose@sys.wakayama-u.ac.jp



ここはある大学の一角。音楽情報処理学の教鞭をとる、るみ子先生の研究室である。この研究室には4月からアフリカからの留学生であるカタヨリが卒論生として配属されることになっている。今日は、期末試験を終えたカタヨリが、配属を前に見学にやって来た。そこに居合わせたのは、現代音楽の作曲家、ケージである。ケージは、るみ子の長年の友人ではあるが、コンピュータを音楽に持ち込むことには懐疑的である。今日もるみ子とケージはホットな議論の最中であった。

### 演奏レンダリング

**ケージ**：音楽をコンピュータ上で表現するなんてできるはずないよ。音楽には、音楽家の天才的なひらめき、芸術的な感性、暗黙の音楽知識などがたくさん詰まっているのだから。

**るみ子**：とても難しいことは確かです。でも今、その難しいことに挑戦しようという試みがホットな話題になっています。それを「蓮根」というんですけど。

**カタヨリ**：れんこん、って、あの穴のあいた？

**るみ子**：うん、といっても、別に蓮根の穴を使って笛のように吹こうという話ではない。「演奏レンダリングコンクール」、略してレンコン。レンダリングという言葉は、画像処理なんかで、きめ細かく本物そっくりに表示することを指して使われることがあるけれど、音楽の場合は、「人間のように情緒ある演奏や表情のついた演奏をすること」と言えるのよ。

**カタヨリ**：ちょっとまだピンと来ないな...

### 素材

**るみ子**：じゃあ、ちょっとこの音楽を

聴いてみて。楽譜がこれ(図-1)。まずは、ブーニンの演奏から(♪)。この演奏を実測してみると、楽譜上では同じ4分音符でも、等間隔、等音量では鳴っていないことがわかります。表情のある演奏というのは、各音の打鍵、離鍵のタイミングや、音量が変化しているのよ。

次に同じ曲を譜面通りに演奏したものです(♪)。機械的に聞こえるでしょう。これは、音価<sup>★1</sup>の等しい音符はすべて同じ長さで演奏されているからです。4分音符の長さはきっちり2分音符の半分だし、演奏のタイミングはメトロノームにぴったり合っています。音の大きさもすべて一律なのよ。

**カタヨリ**：これ、本当にショパンの曲なんですか？全然別の曲に聞こえるのですね。まるでケータイの着メロのようです。

演奏レンダリングシステムは、楽曲情

★1 楽譜上の音の長さ。4分音符、付点8分音符などを指す。



図-1 F. Chopin: Prelude 28-7

報を入力とし、人間が演奏するような情緒や抑揚のついた演奏を出力する。楽曲情報には、音の並びを記述する表面的な楽譜情報から、後述するような深層構造までいくつかのレベルがある。表面的な楽譜情報は標準MIDIファイル(SMF<sup>☆2</sup>)で与えられることが多い。

### 味付け

**カタヨリ**：演奏レンダリングって、簡単にできそうに思えるんですが、こういうメロディだったら、こういうパターンっていう風に決めてしまえばどうですか？

**るみ子**：じゃあ、メロディのパターン認識ってどうするのかな。それにはまずメロディをコンピュータ上で表現しなくてはね。パターン認識風にベクトル化したり、音楽理論に従った木構造のようなもので表現したり。

**ケージ**：そもそも音楽っていうのは、パターン化を乗り越える試みの積み重ねなんだよ。プログラムとしてうまくパターン化、モデル化したつもりでも、すぐに新しいパターンに取って代わられてしまうのさ。創造性

の高い現代音楽もコンピュータが扱えるとは思えないね。

**カタヨリ**：人間の演奏とシステムはどいう関係があるんですか？

**るみ子**：人間の演奏は、分析対象として生成の際の適用事例として利用されます。レンダリングされた演奏も事例として再利用されることがあるのよ。

**カタヨリ**：演奏の雰囲気はどのように扱うんですか？ 僕なんか国でよく「もっとクールに」って先生に言われたもんです。システムにクールな演奏をしてほしい、と思ったたらどうすればいいんだろう。

**るみ子**：「もっとクールに」という演奏表情のルールを作るためには、演奏事例を楽譜と照らし合わせて分析し、打鍵、離鍵のタイミングと音量のパラメタ値の変化として言い直さなくてはならないわね。

**ケージ**：そんなことでレンダリングを正確にモデル化できるわけがないよ。人間なら絶対しないようなヘンテコリンな演奏をするのがおちさ。

**るみ子**：だから結局、音楽を作る現場では、現在でも、人が手作業で一音一音の打鍵、離鍵タイミングと音量を調整する「打ち込み」という作業が主流

なんです。もちろん「打ち込み」のスピードを上げるための便利な入力ソフトは開発されていますけど、基本は人間の勘と経験が頼りなんです。

一音一音についてきめ細かい処理を自動的にやってくれる演奏レンダリングシステムはまだ誰も開発していない。音楽は基本的に階層構造を持っており、低次の階層が楽曲の局所的な構造に、高次の階層がより大局的な構造に対応する。演奏レンダリングでは、この局所的な調節と大局的な調節をうまく協調させなければならぬ。繰り返しが現れる場合には人間は違う演奏を行うということも含め、各音符が持つ演奏のパラメタを、一貫性を保って調節することは難しい。一方、DTM(Desk Top Music)が広まったとはいえ、演奏に満足のいく表情をつけるためには、パラメタ調整のためかなりの汗が必要になる。

### 道具

**るみ子**：今、演奏レンダリングの研究分野では、楽譜の表層の情報だけを頼りにするのは不十分だから、メロディが現れる音楽的な文脈を利用したり、

<sup>☆2</sup> Standard MIDI File. 鍵盤を押さえる、離すなどのコマンドに、時間間隔の情報を加えて演奏情報をファイルとして記録するための形式。

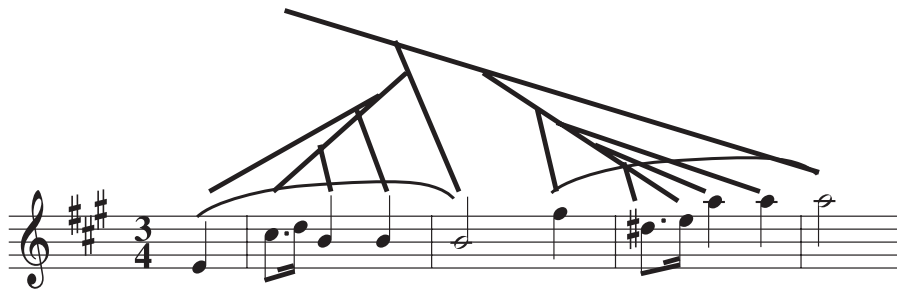


図-2 譜例の解釈とその表現 (GTTMのTime Span分析の結果)

深層構造を分析しようという方向に流れています。

**ケージ**：こういう処理の仕組みは、自然言語処理とよく対比されているようだね。楽器を弾いている人がどんな知識を操っているのか、あるいは、好きな曲を聞いている時、頭の中ではどのようにして曲を認識し理解しているのか、考えたことがあるかい。

**るみ子**：それには、音楽分析の理論が使えろと考えられています。

音楽の深層構造の分析とは、楽譜上では明示的に記述されていない各音の役割を明確にすることであり、その結果を演奏表情付けに利用することができる。図-2は、図-1のメロディをGTTM (Generative Theory of Tonal Music) という音楽理論に従って分析した一例である。図中の木構造は、隣接音どうしを比較して、より重要な音ほど根に近くなるように表されており、Time Span木と呼ばれる。重要な音は、演奏に際し強調されることが多い。解釈の相違で異なる木構造が作られると、それに基づく演奏も違ってくる。GTTM以外にも、先行音による後続音の出現予測やメロディの音楽の継続性、終止感を分析するIR (Implication-Realization) モデルを用いるシステムもある。しかし、いずれの音楽理論も、計算機上への実装、自動化には至っておらず、構造の分析

は手作業で行っているのが現状である。

**ケージ**：どんな音楽理論だってそこで明示的に語られている概念は、音楽家の頭の中にある概念のほんの一部なんだから、それだけで表情の豊かな演奏ができあがるとは思えないね。

**カタヨリ**：でも、演奏表情の手がかりにはなりそうですね。同じ曲でも、いろいろなレンダリングが聞けそうな気がしてきました。どのレンダリングが良くてどれが悪いという決め手はあるのかなあ。

**ケージ**：チェスやサッカーは、まだ勝敗がはっきりしているからいいよ。でも、演奏レンダリングの結果は、主観的にしか評価できないんじゃないの？そもそも他人には分からないことだから主観って言うんでしょ。誰それ風であるとか、明るい、重々しいなどと表される演奏というのも無数にあるんだよ。一体、どうやって正しい演奏、意図した演奏を区別するっていうの！？

**るみ子**：だからコンクールを開きましょって言うっているんですよ。

### 蓮根は泥のなか

**カタヨリ**：それがさっきの蓮根ですね。

**るみ子**：そう。演奏レンダリングの研究は今まで主に日本とヨーロッパで行われてきたのですが、研究発表が行わ

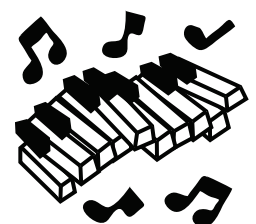
れてもその場にいなければ、後で予稿を読むだけになります。それでは、システムでどのようなレンダリングがなされているか分からないでしょ。

音楽活動が主観的なものであるということは、避けようのない事実なのですが、人間の音楽活動ではそれを承知の上で順序付けがなされています。

**カタヨリ**：演奏レンダリングシステムでも、曲が同じなら、演奏を聞き比べて違いが分かりそうだなあ。

**るみ子**：そうすれば、お互いに刺激になるし、これからの目標も立てやすくなりそうです。専門外の人にも興味を持ってもらえるんじゃないかな。

一般に、演奏レンダリングシステムは、図-3のような構成をしている。ステージ1では、実際の演奏のSMFや楽譜データを与えて分析や学習を行い、レンダリングに必要な情報を事例や規則として取り出す。ステージ2では、レンダリングすべき課題曲の情報を入力し、演奏データを出力する。マニュアルと書いてあるのは、人手によって処



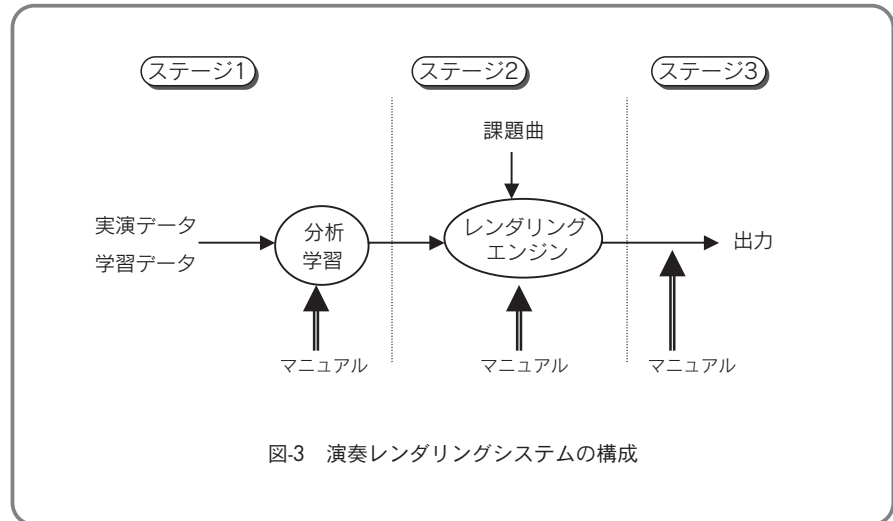
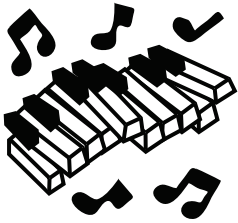


図-3 演奏レンドリングシステムの構成

理に修正や補足を行うことを表す。特に、ステージ1や2の処理を行わず、ステージ3のマニュアルだけで出力を生成する場合を打ち込みと呼ぶ。

**ケージ:**ところで、審査方法はどうか。課題曲はどうするつもりなの。そもそも、システムはまっとうな演奏をレンドリングできるの？

**るみ子:**現状の演奏レンドリングシステムがいきなりチャイコフスキのピアノコンチェルトを弾くことは無理です。システムが自動的にレンドリングした演奏は、人間の子供がピアノを習い始めた時とはまた違う意味で稚拙です。演奏レンドリング研究を育てていく過程で、レベルに合った課題曲や審査方法、審査基準を決めれば良いのではないのでしょうか。

現状では、音楽としてより良い演奏を作るには打ち込みの方が優れているという状況であるから、当分はコンクールといっても、純粋に音楽性を競うものにはならないだろう。研究成果の1つとして、演奏レンドリングのための数学的に美しいモデルが提案されたとして、そのモデルが音

樂的にも美しい演奏を生成してくれるかどうかは分からない。そこで蓮根でも、演技と技術の判断を別々に行うフィギュアスケートや、勝負と技術を別々に扱うロボカップのように、音楽的観点とシステムの観点を独立に評価するところから始める。もちろん、音楽の専門家には当初から審査員として参加してもらおうが、演奏レンドリングが未熟なうちは、蓮根の聴衆による人気投票ということも考えられる。演奏レンドリングの技術レベルに合わせた課題曲を選択しつつ、構造解釈に焦点を当てたり、アウフタクト<sup>☆3</sup>に焦点を当てたりする予定である。

**るみ子:**音楽システムの評価については、かねてから議論されてきました<sup>1)</sup>。蓮根は、新しい評価の方法を具体的に試みる場となります。

### 在来種の蓮根

**カタヨリ:**蓮根は、ピアノ曲に限定されているんですが。僕なんかギターに興味あるんですけど。

**るみ子:**もともとピアノの演奏レンドリングをしているシステムが多いというのも1つの理由です。これは、

MIDI<sup>☆4</sup>によって自動演奏できるアコースティック楽器はピアノが多かったという商業的原因があったかもしれませんが。それに、ピアノ演奏だと、MIDIイベントで打鍵と離鍵に対応するノートオン/ノートオフというイベントを扱うだけで演奏表情を制御できます。ギターの演奏表現をMIDIで制御するためには、MIDI規格に精通する必要がありますし、詳しくなっても、MIDI制御によるチョーキングで思ったとおりの演奏を出すのは機器によって差があり難しいでしょうね。

**ケージ:**ピアノ演奏に限定するにしても、演奏レンドリングが進化するには、ペダルの扱いは避けて通れないよ。当然ハーフペダルも使うんでしょうね。

ピアノの右ペダルを踏み込むと音が残響して聞こえる。このペダルを適切に使うことで美しいレガートをつけることができ、演奏表情が大きく向上する。図-1にもペダルを踏む記号“ $\text{ped}$ ”と離す記号“\*”があるが、譜面通りにペダルを踏んだり離したりしても美しいレガートはつかない。ペダルを最後まで踏み込まずに中間位置で止めて用いるハーフペダルという奏法もあり、こ

<sup>☆3</sup> 図-1の曲のように、小節の開始に先立つ拍で開始する曲を指す。曲の開始だけでなく、第二小節の第三拍の音も次の小節に先立つ音で、アウフタクトと呼ばれる。曲の開始のアウフタクトのある小節は第一小節と数えない。

<sup>☆4</sup> Musical Instrument Digital Interface。電子楽器とコンピュータを接続するための規格。鍵盤を押さえる、離すなどのコマンドが定義されている。



れは連続量の制御を必要とする。ギターのチョーキングのパラメタと同様、イベント数が大量になることや、機器によってどのような効果が出るかが定まっていないという問題がある。

### 新種の蓮根

**ケージ**：それなら、せめて人間の演奏家では実現できないようなコンクールにしてみたいものだね。

**カタヨリ**：課題曲をコンクール当日に公開して、初見演奏で競うってのはどうですか。

**るみ子**：そのためには、課題曲の楽譜を認識する研究、楽譜から適切な構造を導く研究、その構造をどのように利用するかの研究、それらを多面的に組み合わせる研究が必要になりますね。さらに、発音、消音タイミングと音量という信号レベルの処理と、楽曲の解釈という記号レベルの処理を融合する技術が必要となります。これは、ロボットの研究でも取り上げられている情報統合の問題に非常に関連が深いものです。

**カタヨリ**：今まで誰も弾いたことのない曲をいきなり課題曲にするってのはどうでしょう。

**ケージ**：蓮根のために作曲を委託して、当日公開、数分後には世界初演を競い合うということか。それは、人間のコンクールでは無理だな。

### プチ蓮根と本蓮根

**るみ子**：2002年度はATRで行われるICAD (International Conference on Auditory Display) という国際会議で初回蓮根を行う予定です。これはコンクールではなく技術発表とデモンストレーションのつもりで、プチ蓮根と呼んでいます。さらに宣伝普及を兼ねて、

もう1つ大きな会議でもプチ蓮根を開催しようと考えています。まずは、演奏レンダリングの現状を知ること、これまで独立に進歩してきた打ち込みと演奏レンダリングシステムがお互いの結果を利用し合うことの可能性について考えたいと思っています。

**カタヨリ**：打ち込みコンテストの記事、DTM系の雑誌で見ましたよ。結構エントリ数が多いんですね。

**ケージ**：2003年には本当にコンクールをするの？

**るみ子**：人工知能の国際会議では今まで何度か音楽情報処理をテーマとしたワークショップが行われてきました。そこでコンクールを行う予定です。これを本蓮根と呼びます。本蓮根では、プチ蓮根で明らかになった技術的な限界を反映して課題曲を選び、演奏を聞き比べます。課題曲は、1曲全体にはならないかもしれませんが。

プチ蓮根には打ち込み部門と自動生成部門があり、打ち込み部門の課題曲はショパンのピアノ曲から選ぶ予定である。自動生成部門では、図-3のステージ1の処理時間やマニュアル介入には制限を設けず、ステージ2の処理以降をコンクール当日に実施するという規定を設ける予定である。本蓮根ではステージ1にも何らかの規定を課す予定である。

### 蓮根への道

#### 蓮根の花はいつ開く？

**カタヨリ**：演奏レンダリング研究はどのように進んでいくんでしょうか。

**るみ子**：2010年には、3~4分間の小品を演奏できるようになって、打ち込みによる演奏よりも優れた表情が自動的に生成できることを目標としてい

ます。

**カタヨリ**：図-1の曲はレンダリングできるんですね。

**ケージ**：アウフタクト、ペダルも解決済みかい。

**るみ子**：さらに15年かけて2025年には、人間との演奏差がなくなることを目標とします。そうすれば、ショパンコンクールに出場する可能性も出てきます。この年は、ショパンコンクールにシステムがエントリすることに意義があるって感じでしょうか。ロボットの研究もかなり進んでいることでしょうから、ロボットによるレンダリングにも取り組みたいです。ロボットには、楽譜認識をする目のほかにも、滑らかに鍵盤と戯れることのできる指、優雅なペダリングをできる足が必要ですね。それに、本選の課題曲は協奏曲なので、ロボットは指揮者を動画像として認識しなければなりませんし、オーケストラのパートを音源分離できる耳を持ってないといけません。

**ケージ**：筑波の科学万博から約半世紀で、ロボットによる演奏が人間並みになるって★5？

**るみ子**：図-4のように蓮根の目標を立てています。

**カタヨリ**：2050年にショパンコンクール優勝！ ピアノコンクールのグランドスラムってあるんですか？

**るみ子**：チャイコフスキコンクールでも優勝したいわね。

**ケージ**：その頃には、レンダリングシステムが人間に教授を始める？ さらに半世紀でシステムに師事した人間が

★5 1985年の科学万博では、WABOT という名のロボットがオーケストラと競演した。

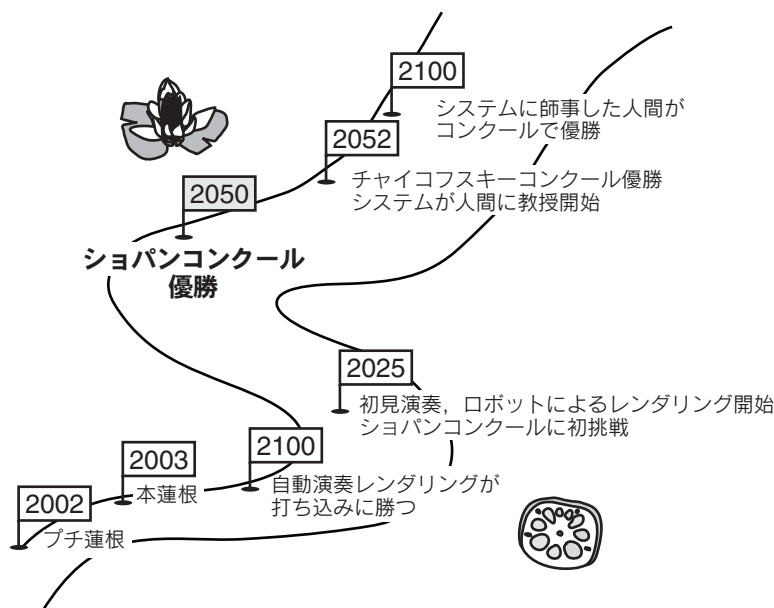


図-4 これが蓮根の進む道

コンクールで優勝する？！

### 蓮根はおいしい

**るみ子**：演奏レンドリングの技術がどんな音楽ビジネスに利用できるかということを少し考えてみましょう。

**カタヨリ**：ケータイの着メロの楽しみ方が増えそうです。カラオケ、駅の発車チャイム、ゲーム…やっぱりエンタテインメント系が多く思い浮かびますね。

**ケージ**：エンタテインメント系はたしかに目立つけど、マーケット自体はまだまだ小さいよ。やはりビジネスの現場で使えないと。

**るみ子**：現時点で、電子的に再生された楽曲のほとんどにはレンドリングが施されていません。ですから、基本的にそれらすべてが演奏レンドリング技術の応用候補です。たとえば、プレゼン資料を自動生成する技術がありますが、プレゼン効果を上げるために、同時にBGMも自動生成して演奏レンドリングしてしまうとか。

それに、これまでは演奏レンドリングのような技術がなかったので、音楽配信はデジタルオーディオデー

タかSMFのコピーをするしかありませんでした。個人の手元で楽曲を加工できるようにになれば、流通する音楽データの質も変わりますし量も増えると期待しています。

いずれにしても、日々の生活の中で音楽を聴いたり作ったりする機会を増やす必要はあるでしょう。

**カタヨリ**：これからの僕らの生活で、音楽に接する時間が減るとはとても思えないなあ。

**るみ子**：P2Pサービスとは相性がいいいと思います。

**ケージ**：それは演奏レンドリングに限定した話というより、音楽情報処理一般の話に近いのでは。

### これからのコンピュータと音楽の世界

**るみ子**：これから音楽の世界では、コンピュータの存在がますます大きくなっていくのではないのでしょうか。

**ケージ**：それでも私はやはり、音楽好きの計算機科学者はプロの音楽家を超えられないと思うね。音楽はゲームとは違うんだから。

**カタヨリ**：卒論テーマは演奏レンドリングで決まりです。「bit別冊コンピュータと音楽の世界」<sup>2)</sup>には、演奏生成システムの記事が載っていました。

**るみ子**：その本には、音楽情報処理に有用な情報や研究も紹介されています。ほかにも音楽情報処理研究に有用な読み物<sup>3)~5)</sup>はありますよ。

蓮根では最近ホームページを立ち上げました<sup>6)</sup>。ブチ蓮根や本蓮根の規定の詳細について随時情報を置いています。本日の話を字面だけで納得するには限界があると思います。ホームページには、実際の演奏レンドリング結果も置かれていますのでぜひ体験してみてください。

#### 参考文献

- 1) 片寄晴弘, 後藤真孝, 堀内靖雄, 松島俊明, 村尾忠廣, 志村 哲, 栗 孝之, 平田圭二: パネル討論会「コンピュータサイエンスとしての音楽情報処理」の報告(第52回全国大会シンポジウム), 96-MUS-15, pp.91-98 (1996).
- 2) 長嶋洋一, 橋本周司, 平賀 謙, 平田圭二編: bit別冊コンピュータと音楽の世界(1998).
- 3) Roads, C. (青柳, 小坂, 平田, 堀内訳・監修): コンピュータ音楽-歴史・テクノロジー・アート, 東京電機大学出版局(2001).
- 4) 平田圭二: 私のブックマーク: 音楽と人工知能, 人工知能学会誌, Vol.17, No.1 (2002).
- 5) 片寄晴弘: 音楽情報処理, 岩波講座 マルチメディア情報学 第4巻「文字と音の情報処理」, pp.163-236 (1999).
- 6) 蓮根HP: <http://shouchan.ei.tuat.ac.jp/~rencon/> (平成13年12月7日受付)