

琉球大学学術リポジトリ

画像処理を用いた三線音の自動採譜

| | |
|-------|---|
| メタデータ | 言語: 出版者: 琉球大学工学部 公開日: 2007-08-23 キーワード (Ja): キーワード (En): Sanshin, Pitch detection, Image Processing, Transcription 作成者: 金崎, 賢悟, 山城, 毅, 渡久地, 實 メールアドレス: 所属: |
| URL | http://hdl.handle.net/20.500.12000/1481 |

画像処理を用いた三線音の自動採譜

金崎 賢悟 * 山城 毅 ** 渡久地 實 **

The Automatic Music Transcription Of Sanshin Sound Using Image Processing

Kengo KANASAKI* Tsuyoshi YAMASHIRO** and Minoru TOGUCHI**

Abstract

The Sanshin is three stringed musical instruments in Okinawa. These musical instruments are played with unique score book called "Kunkunsi". In this paper, we proposed how to transcribe the sound of playing Sanshin to the musical note. As preprocessing we attempt pitch detection of Sanshin sound. To detect the pitch we tried to using image processing. The detection with this method was done from an actual consecutive Sanshin sound, and good results were obtained on the single sound.

Key Words: Sanshin, Pitch detection, Image Processing, Transcription

1. まえがき

楽器音や歌声から楽譜に書き出す採譜の作業は音楽的な専門知識や技術, 経験が必要となるが, そのような能力を持った人でも正確な採譜を行うことは困難だとされている. そのため, 採譜をコンピュータ上で自動的に行う自動採譜の研究は音声工学の分野において多くの研究が行われている [1],[2].

沖縄には三線(サンシン)という, 本土の三味線に似ている独特の民族楽器があり, 多くの演奏曲が存在する. これらの曲は工工四(クンクンシー)と呼ばれる三線独自の楽譜を用いて演奏され [3], 現在も新しい楽曲が演奏されている. この工工四を自動採譜により作成できれば, 沖縄民謡などの民族的な楽曲以外にも, 多くのジャンルの楽曲を三線で演奏することが可能となることが考えられる.

本研究では, 三線の実際の演奏音から音高判別を行い, 自動採譜を行うことを目的としている [4],[5]. 自動採譜において必要となるのは, 音高と発音時点である.

音高判別の方法は, 自己相関法, ピークトウピーク法, ゼロ交差検出法などがある [6],[7],[8]. しかし, この方法では多重音などの複数の音が重なった場合には対応ができない. 三線演奏曲は演奏者の唄や他楽器とともに演奏されることがあるため [3], これらの方法が適していないと考えられる.

そこで, 本研究では三線演奏音の周波数スペクトル情報に着目し, 音高を判別する方法を提案した. 周波数スペク

トルを用いて音高を判別する方法は過去にも様々な方法が提案されているが [1],[2],[9], 三線音は基本周波数が小さく, 基本周波数による音高判別は大変難しい. これを解決するため, スペクトル積算値を二次微分処理し, 二次微分処理時の閾値決定に画像処理を用いることにより音高を判別した.

発音時点の測定については, シフトフレームを用いて一定時間ごとに音高判別を行い, 音高が変化した点を発音時点とみなす方法が主流である. しかし, 音の持続時間が短く, 音の立ち上がり鋭い三線音については処理時間のことを考えると非常に無駄が多い.

そこで, 本研究での発音時点測定法には時間波形上における閾値を用いた簡素な方法を用いて, 発音時点測定を行っている.

本稿では, 自動採譜の前処理としての音高判別, 発音時点測定の検討および実際の採譜処理結果について述べる.

2. 三線音の特徴

琉球音楽における調子とは, 三線の絃音進行の運指により分類されるもので, 各絃のチューニング(調絃)により決定される.

三線にはいくつかの調子があり, その中でも基本とされている調子が「本調子」である [3],[10]. 本調子もいくつかの種類があり, 琴と演奏する場合などによっても変化する. 本論文で使用した基本となる本調子は, 第一絃開放を C_2 (ド), 第二絃開放を F_3 (ファ), 第三絃開放を C_3 (ド) に, それぞれ調絃する方法である.

さらに, その他の調子として, 一揚調, 二揚調, 三下調, 一二揚調があげられる. これらの調絃方法は, それぞれの数字にあたる絃を一全音上げる(下げる)ことにより調絃される.

三線で演奏できるもっとも低い音が第一絃の開放音で

受理: 2006年1月6日

平成17年度電子情報通信学会総合大会にて発表.

* 大学院理工学研究科 電気電子工学専攻

(Graduate Student, Electrical and Electronic Eng.)

** 電気電子工学科

(Dept. of Electrical and Electronic Engineering, Fac. of Eng.)

あり、三線独自の楽譜である工工四（クンクンシー）では「合」という符号で表記される。前述の本調子で調絃した場合、 C_2 の音高となり、基本周波数は130.81[Hz]となる。なお、第二絃開放は「四」、第三絃開放は「工」と表記する。

次に、最も高い音についてだが、物理的にはある程度高音領域まで出せるが、一般的な演奏における最高音は第三絃の下部を押して弾く「イ五」という音となり、本調子の場合、 D_4 の音高と対応しており、基本周波数は587.33[Hz]である。また三線演奏の際、工工四に用いられる「工」などの符号は三線の運指（勘所という）に直結しており、どの調でも変化しない音名（C,D,Eなど）ではなく、調性によって変化する階名（ド、レ、ミなど）に近いものである「工」などの音が、どの音高に対応するかは調絃により定まる。なお、下記の表記は、全て本調子の場合であり、特に断らない限りは今後も本調子であるものとする。TABLE 1に本調子の勘所と12平均律での音名、周波数を示す。

TABLE 1 本調子の勘所と音名、周波数

| 勘所 | 音名 | 周波数 [Hz] |
|----|---------|----------|
| 合 | C_2 | 130.81 |
| 乙 | D_2 | 146.83 |
| 老 | E_2 | 164.81 |
| 四 | F_2 | 174.61 |
| 上 | G_2 | 196.00 |
| 中 | A_2 | 220.00 |
| 尺 | $A\#_2$ | 233.08 |
| 工 | C_3 | 261.63 |
| 五 | D_3 | 293.66 |
| 六 | E_3 | 329.63 |
| 七 | F_3 | 349.23 |
| 八 | G_3 | 392.00 |
| 九 | A_3 | 440.00 |
| イ尺 | $A\#_3$ | 466.16 |
| イ工 | C_4 | 523.25 |
| イ乙 | D_4 | 587.33 |

次に、解析対象となる三線音の時間波形の一例を Fig.1 に、また比較を行うために同じ弦楽器であるピアノ音とヴァイオリン音の時間波形をそれぞれ Fig.2, Fig.3 に示す。なお、これら三つの時間波形はすべて周波数 261.6[Hz](C_3 , ド)の波形である。

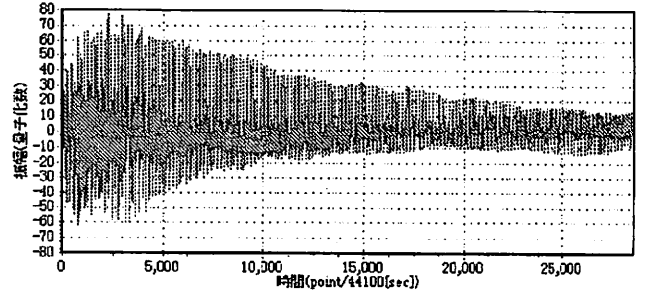


Fig. 2. ピアノ音 (261.6[Hz]) の時間波形

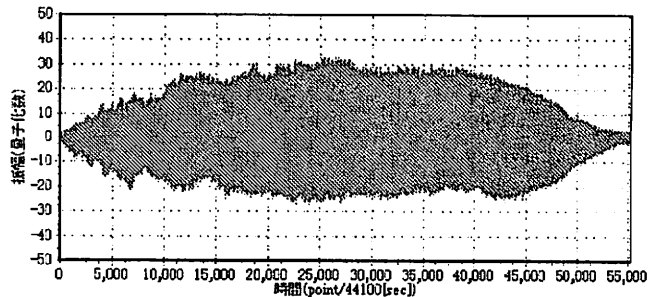


Fig. 3. ヴァイオリン音 (261.6[Hz]) の時間波形

上記の波形について時間比較を行うと、三線音とピアノ音については音圧がピークに達してからの減衰が早く、音の持続時間が短いという特徴がある [11],[12]。三線音はピアノ音よりもかなり早く減衰する特徴を持つ。また、ヴァイオリン音については、時間波形の形態からわかるように持続時間は当然長くなる。また、時間波形から音高に関する情報は得ることができないので、高速フーリエ変換 (FFT) により時間領域のデータを周波数領域に変換する。

変換したスペクトル図, Fig.4~Fig.6 を観ると、基本周波数である 261.6[Hz] の整数倍に周波数スペクトルが分布している。また、三線音は、基本周波数スペクトルが他楽器のスペクトルに比べて大変小さいという珍しい特徴があり、また倍音成分（基本周波数の整数倍スペクトル）が高次まで豊かに出る [4],[5],[9]。

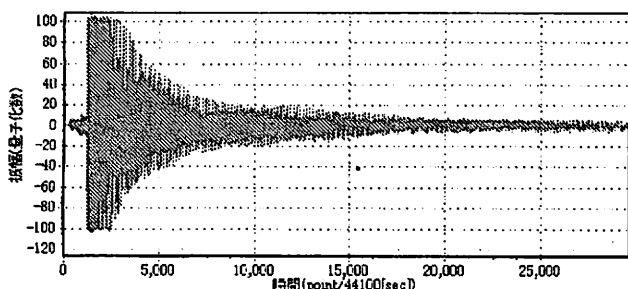


Fig. 1. 三線音 (261.6[Hz]) の時間波形

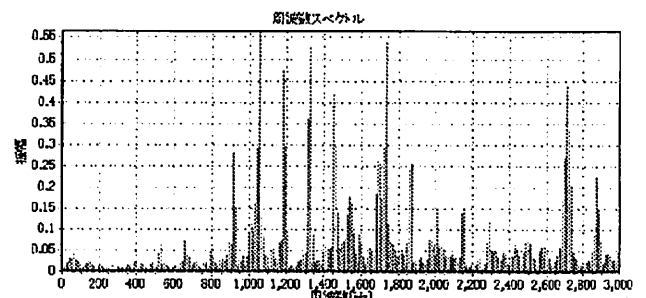


Fig. 4. 三線音 (261.6[Hz]) の周波数スペクトル

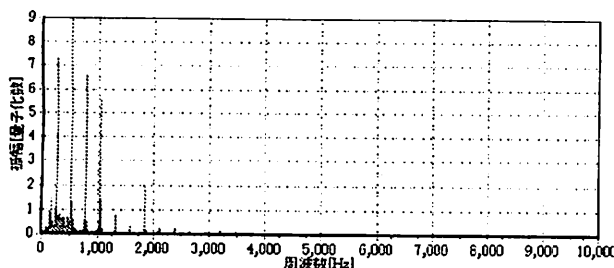


Fig. 5. ピアノ音 (261.6[Hz]) の周波数スペクトル

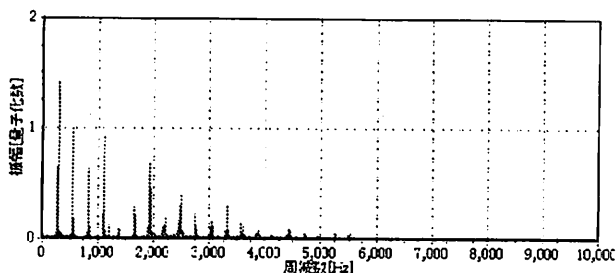


Fig. 6. ヴァイオリン音 (261.6[Hz]) の周波数スペクトル

3. 三線音の音高判別処理

楽器に限らず、全ての音の音高は基本周波数によって決まるため、周波数分布から音高を判別するためには基本周波数スペクトルを調べればよい。しかし、先に述べたように三線音は基本周波数成分がほかのスペクトルに比べて非常に小さいため、基本周波数スペクトルから音高を判別することは大変困難である。したがって、基本周波数成分以外のスペクトル、つまり倍音成分スペクトル分布から音高を判別する必要がある。倍音成分スペクトルは基本周波数の整数倍の分布であるため、その間隔を正しく測定できれば音高が判別できる。そこで、スペクトル積算値の二次微分処理による音高判別方法を提案する。

3.1 スペクトル積算処理

まず、二次微分処理の前処理として周波数スペクトルを積算処理し、さらに三線音の周波数スペクトルの振幅値について100[Hz]~3000[Hz]間で積算処理を行う。例として、Fig.4の周波数スペクトルに対して積算処理を行った結果をFig.7に示す。

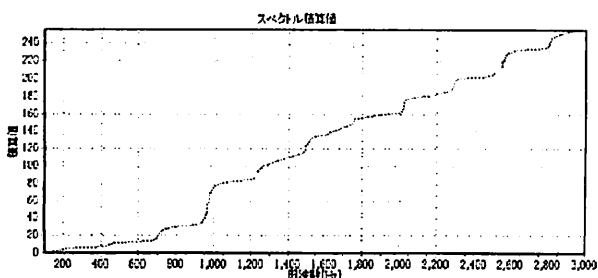


Fig. 7. 三線音のスペクトル積算値

Fig.7を観ると、値の立ち上がり早いところがある。これは、前後間の値の差が大きいほど立ち上がりが早くなるためである。つまり、周波数スペクトルにおいて倍音成分スペクトルの部分で立ち上がっている。そこで、この立ち上がりの間隔を測定することにより音高を判別できる。そこで、このスペクトル積算値を画像処理のエッジ抽出などでも用いられる一次差分フィルタを二回用いることにより二次微分処理を施し、立ち上がり間隔を正しく測定する処理を行う。

3.2 二次微分処理

周波数 f [Hz] の時のスペクトル積算値を $S[f]$ とすると、得られる一次微分値 $d[f]$ は、

$$d[f] = (S[f + \Delta f] - S[f - \Delta f]) \quad (1)$$

(ここで、 Δf は周波数のステップ幅)

となり、この処理を二回行うことにより二次微分値を得ることができる (Fig.8)。得られた二次微分スペクトルを式(2)より閾値処理し、閾値以下のスペクトル振幅を除去した。式(2)において $d2[f]$ は周波数 f [Hz] における二次微分値、 th は二次微分スペクトルの閾値を現す。

$$d2[f] = \begin{cases} d2[f] & (if \ d2[f] > th) \\ 0 & (if \ d2[f] < th) \end{cases} \quad (2)$$

ここで、式(2)における閾値 th を決定するために、スペクトル分布を画像として処理した。これは、音高判別の際に必要なスペクトル本数だけを残すための閾値を容易に決定する。つまり、

1. 得られた二次微分スペクトルの二値画像 (白黒画像) を作成した後、画像を半分のところで走査し、スペクトル本数を数える。この時点で指定した本数 k であれば処理を終了する。
2. 次に指定されたスペクトル本数 k を得るため、上か下に $1/2^n$ ($n = 2, 3, 4, \dots$) だけラスト捜査を行い、指定したスペクトル本数が得られたところで処理を終了し、位置情報 (th) を取得する (Fig.9)。

という処理を行い、これにより閾値が自動的に決定され、高速処理を行うことができる。二次微分スペクトルを式(2)により閾値処理した結果を Fig.10 に示すが、Fig.4と比較するとスペクトル数が減少しているのがわかる。そのため、スペクトルの間隔測定が容易に行え、音高判別が高速で処理できる。

閾値処理で得られた Fig.10 のスペクトルにおいて、スペクトル間隔が三線の最小音高 (C音の130.81[Hz]) 以下の周波数は切捨て、それ以外の最小であるものを音高として判別する。

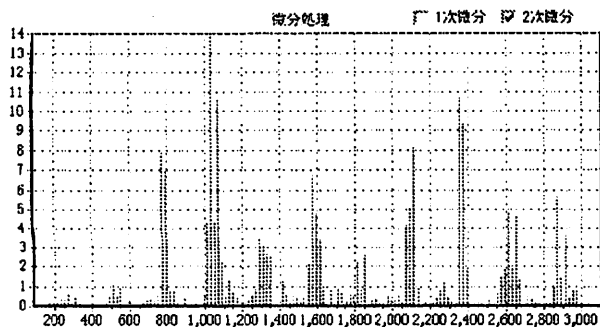


Fig. 8. 二次微分スペクトル

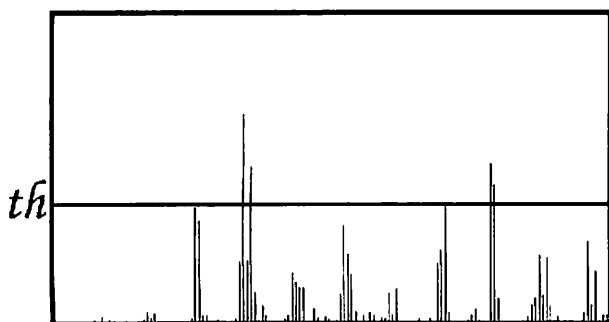


Fig. 9. 二次微分スペクトルの閾値決定処理 (k=5 の場合)

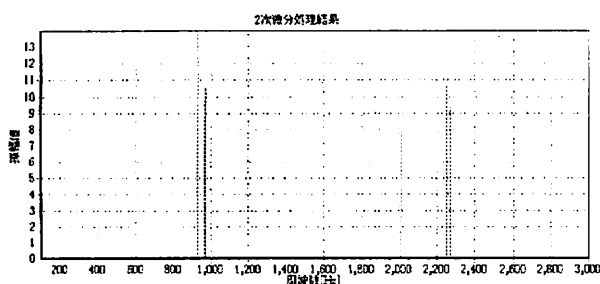


Fig. 10. 二次微分スペクトルの閾値処理結果

3.3 音階の抽出

先ほど述べた二次微分処理によりスペクトル間隔の測定を行うと、実際の音階の周波数とは若干ずれた値が結果として出る場合がある。これは量子化ノイズやデジタル誤差などにより生じるものである。周波数を調べるのみであれば問題はないが、本研究では演奏曲の自動採譜を目的としているため、音階が決定しなければ譜面として成り立たない。そこで、TABLE 1で示した表を元に、測定した周波数に近い音階を抽出する処理を行った。処理の手順を以下に示す。

1. まず、TABLE 1を元に各音高（各周波数）に対応したスペクトル（振幅値一定）の図をあらかじめ用意する。
2. 次に、測定した周波数のスペクトルを前述の図の上に描き、両隣のスペクトルとの距離を調べる（Fig.11）。
3. 距離が最も近いものを音高として抽出する。

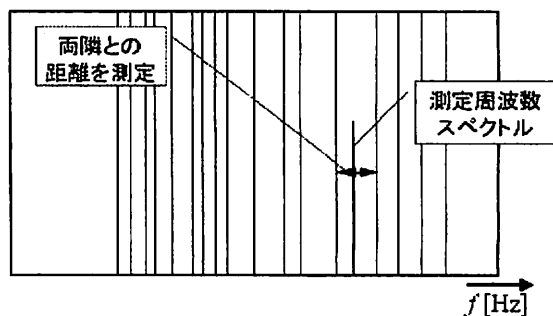


Fig. 11. 音高分布

4. 発音時点測定

自動採譜は、連続音のそれぞれの音について音高を判別するため、各音の発音時点は正しく測定する必要がある。三線音は Fig.1でも示したように、アタック時のピークに達してからの減衰が極めて早い特徴がある。したがって、閾値を音圧ピーク値の80%の値とすると、振幅値は0.1[sec]以内で閾値以下になる。そこで、立ち上がりの時点で閾値を超えた点を発音時点とし、それ以降の発音時点は閾値以下になっている0.1[sec]後に測定するようにした。

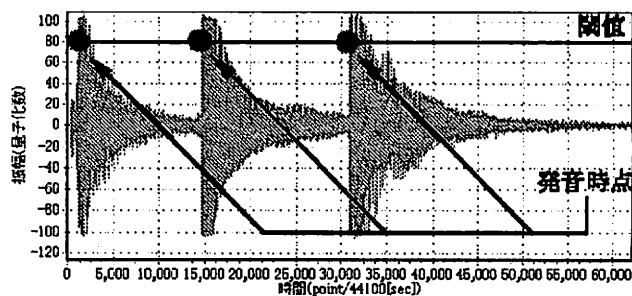


Fig. 12. 発音時点測定

5. 処理結果

処理は、三線で使う領域（ $C_2 \sim D_4$ の26音）について音高判別を行い、実際に三線で演奏された曲について、発音時点測定と音高判別を行った。画像処理による閾値決定時の指定するスペクトル本数kは、実験的に求め、k=5として処理を行った。結果の一例として、Fig.12の連続音に対して、音高判別と発音時点測定を行った結果を示す。

TABLE 2 音高判別及び発音時点測定結果例

| 測定周波数 [Hz] | 勘所 | 音名 | 発音時点 [point] |
|------------|----|-------|--------------|
| 131 | 合 | C_2 | 1737 |
| 260 | 工 | C_3 | 14751 |
| 175 | 四 | F_2 | 30980 |

全体処理時間 281[msec]

TABLE 2の結果を観ると、音高については合 (C_2), 工 (C_3), 四 (F_2) と結果が表示され、全ての音高が正しく判別することが出来た。また、発音時点については、TABLE 2の結果と Fig.12における黒い丸点の位置を比べてもわかるように正しく判別できていることがわかる。また、処理には [msec] オーダーで対処でき、高速処理ができていたことを確認した。

今回、演奏曲についても処理を行ったが、音高、発音時点の両方について良好な結果が得られた。しかし、特殊な奏法を用いた曲や、テンポが速い曲などは音高、発音時点ともかなり精度が落ちていた。音高の違いについては、測定周波数の若干なずれによるマッチングミス、オクターブの違いなどが考えられる。また、発音時点に関しては振幅値の増減が激しい場合に発音時点を取りこぼしてしまうことがある。これらの問題については今後検討していく必要がある。

6. むすび

今回、三線演奏曲の自動採譜の前処理として、画像処理を用いた音高判別法及び発音時点測定法を提案し、良好な処理結果が得られた。しかし、オクターブ違いやテンポの速い曲、特殊な奏法の場合は、判別精度が低下してしまう問題がある。

今後は、上記の問題についていち早く対処し、テンポの速い曲や特殊奏法の場合にも高い判別率を得ることが課題である。また、実際の楽譜作成や三線自動演奏ロボットへの入力データ転送、MIDIを用いた三線シンセサイザなどへの対応が今後の課題として挙げられる。

謝辞

本稿について、有益な御助言と御指導を頂きました琉球大学電気電子工学科の島袋 勝彦助教授に心から感謝いたします。

参考文献

- [1] 尾崎 晃, 斎藤 努, 田所 嘉昭: "マイクロコンピュータを用いた携帯できる音高判別器の開発", 電子情報通信学会総合大会, AS-2-5, (2005)
- [2] 山口 満, 三輪 多恵子, 田所 嘉昭: "並列構成くし形フィルタと特異値分解による多重音の音高推定", 電子情報通信学会論文誌, D-II, pp.1020~1029, (2004,4)
- [3] 山内 秀吉: "琉球音楽の研究-さんしん-", 山内秀吉音楽研究会, (1996)
- [4] 金崎 賢悟, 山城 毅, 渡久地 賢: "スペクトル積算値の二次微分処理による音高判別", 電子情報通信学会九州支部大会学生会講演論文集 (講演奨励賞受賞), D1-23, (2004,9)
- [5] 金崎 賢悟, 山城 毅, 渡久地 賢: "画像処理を用いた三線音の音高判別", 電子情報通信学会総合大会講演論文集, D-14-5, (2005,3)
- [6] 古井 貞照, 酒 善則: "ねっとテクノロジー解体新書⑤ 画像・音声処理技術", 電波新聞社, P19,20,87~89, (2004)
- [7] 金井 浩: "音・振動のスペクトル解析", コロナ社, (1999)
- [8] 中村 尚五: "ヒギナーズデジタルフーリエ変換", 東京電機大学出版局, (1989)
- [9] 和字慶 琢磨, 山城 毅, 渡久地 賢: "倍音スペクトル削除法を用いた三線演奏の自動採譜", 琉球大学工学部紀要第 63 号, P57~62, (2002,12)
- [10] 儀保 榮治郎: "三線のはなし", ひるぎ社, (1999)
- [11] 安藤 由典: "新版 楽器の音響学", 音楽之友社, (1996)
- [12] ジョン・R・ピアーズ 著 村上 陽一郎 訳: "音楽の科学 クラシックからコンピューター音楽まで", 日経サイエンス社, (1989)