琉球大学学術リポジトリ

工工四の楽譜認識に関する研究

メタデータ	言語: Japanese 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): 3-dimensional table method, Character recognition, Hough transform, Musical score 作成者: 祖慶, 実利, 山城, 毅, 伊波, 善清, 渡久地, 實, Sokei, Sanetoshi, Yamashiro, Tsuyoshi, Iha, Zensei, Toguchi, Minoru メールアドレス: 所属:
URL	http://hdl.handle.net/20.500.12000/1466

琉球大学工学部紀要　第53号, 1997年

工工四の楽譜認識に関する研究

祖慶 実利*　山城 毅**
伊波 善清**　渡久地 實**

Study of recognition for okinawa syamisen score KUNKUNSI

Sanetoshi SOKEI*　Tsuyoshi YAMASHIRO**
Zensei IHA**　Minoru TOGUCHI**

Abstract

Handwritten character recognition has been increasing for its importance and has been expanding its application areas. In this paper, we propose the recognition system of musical score, and as one of the applications we tried to input the data to the okinawa syamisen automatic performance robot. The recognition system is used a simple, fast feature extraction method, and an efficient way of combination of fast recognition method(3-dimensional table method) and exact recognition method (multiple simularity method). We demonstarated the performance as the completed system by experiments using real data.

Key Words: 3-dimensional table method, Character recognition, Hough transform, Musical score

1. はじめに

近年, 手書き文字認識の重要性は増加し, その応用範囲も拡張して来ている. 本研究では, その応用としてアミューズメントロボットとして現在研究が進められているサンシン (沖縄三味線) 自動演奏ロボット「あがさー」[1], [2], [3] へのデータ入力用としての手書き文字認識システムを提案する.

サンシン (沖縄三味線) の楽譜は, 工工四と呼ばれ, 音符は 10 数種の漢字で毛筆によって記されている (Fig. 1 参照). 現在,「あがさー」への曲データは使用者が楽譜を読み取り, コンピュータ入力用のデータに変換して入力している. この楽譜の読み取り, ロボット用のデータへの変換という一連の動作を自動化することは, 曲データを逐一入力するという手間を省くだけではなく, 楽譜の知識を持ち合わせていない人にでも気軽に使用できるという利点がある. また, 楽譜さえあれば, さまざまな曲の演奏を楽しむことができる, という手軽さは, アミューズメントロボットとして必要不可欠な要素である.

受理: 1996 年 11 月 29 日
平成八年度 電気関係学会九州支部連合大会にて発表.
* 大学院工学研究科 電気情報工学専攻
(Graduate Student,Electrical and Information Eng.)
** 電気電子工学科
(Dept.of Electrical and Electronic Engineering,Fac. of Eng.)

そこで, 本研究では, 手軽にかつ高速に認識を行なうために, 細線化などの前処理を必要としない, 単純かつ高速に抽出できる特徴量を使用した. また, 認識については, 本研究で提案する高速な認識を行うことのできる 3 次元テーブル法と認識精度の良い複合類似度法のお互いの長所をうまく組み合わせて使用する [4] ことで, 処理速度, 認識率の点から満足の行くシステムを構築した. このシステムは, Fig. 2に示すように前処理部, 特徴抽出部, 認識部, 出力部の 4 つの part から構成されている. 第 2 節で, 前処理部, 第 3 節で, 特徴抽出法部, 第 4 節で認識部について述べ, 第 5 節では, 実際のデータに対する認識実験結果を示す.

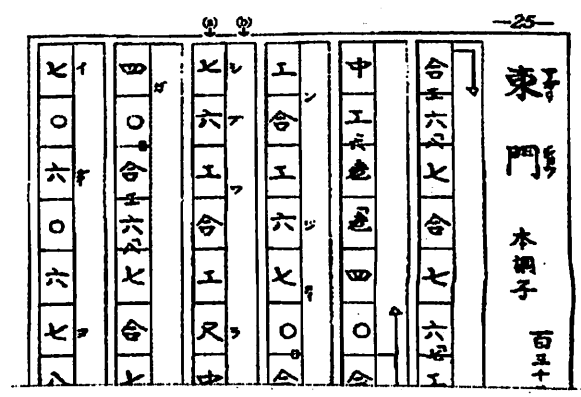

Fig. 1. サンシンの楽譜:工工四

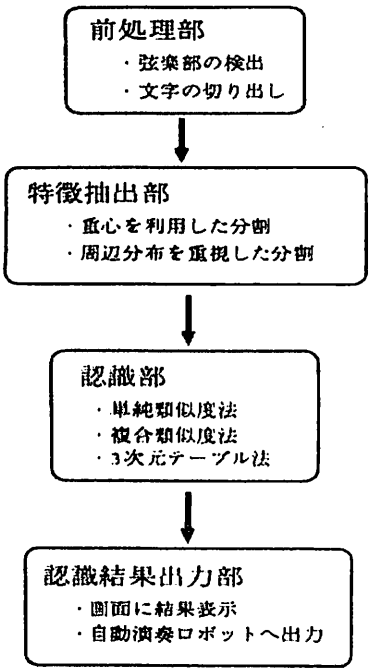

Fig. 2. 認識システムについて

2. 前処理部

認識に使用する楽譜 (B5 版) を，スキャナによって 65536 色で，コンピュータに取り込む．このとき，各文字を大きく取り込むために Fig. 3 に示すように，上下 2 分割して取り込む．取り込んだ各画像は，512x512 ドットの画像として，保存される．また，楽譜画像は，しきい値決定式 [5] によって，自動的に 2 値化を行なっている．

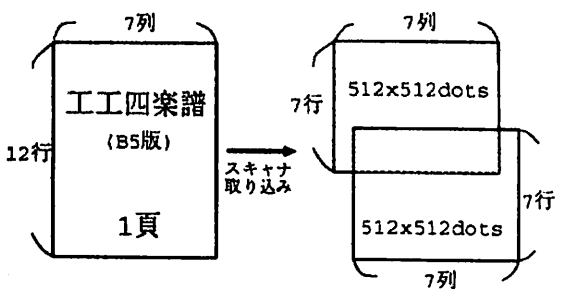

Fig. 3. 楽譜画像の取り込み

次に，文字の切り出しについて述べる．

2.1 弦楽部の検出と文字の切り出し

工工四（クンクンシー）の楽譜は，縦線によって仕切られた弦楽部 (a) と声楽部 (b) が対になって存在し，定まった書式で書かれている (Fig. 1 参照)．

認識に使用するのは，弦楽部に記述される文字 (音符) であり，弦楽部を構成する縦線の対を式 (1) に示すハフ変換 [6] により検出する．

$$\rho_0 = x_i \cos\theta_0 + y_i \sin\theta_0 \qquad (1)$$

ハフ変換には，途切れた直線でも一本の直線として検出できるという利点がある．しかし，使用する楽譜には，直線を構成している画素だけでなく，文字を構成している画素も多数存在している．このような場合に，文字画素をハフ変換をしてしまうと，あたかもそこに直線があるように検出してしまう．よって，これを避けるために走査線上の一本あたりの平均の黒画素数を算出し，この平均値をしきい値として，平均値以上の値をもつなら文字上を走査している可能性が高いと判断して，ハフ変換を行なわない．平均以下の走査線についてのみハフ変換を行なうことで，直線の誤検出を防ぐことができる．尚，数ドットおきに走査をすることにより検出時間も大幅に短縮できる．

また，ハフ変換は，精度を上げるために，膨大な処理時間と記憶領域が必要となるが，検出対象となる直線の傾きがほぼ垂直であるため，ハフ変換の探索範囲を $-2.0 \leq \theta \leq 2.0$，ステップを 0.1 度刻みで行なうことで，計算時間，記憶領域の簡略化を行なっている．

ここで検出された直線の間隔の大小により弦楽部を決定する．決定された弦楽部において，水平方向の黒画素分布を調べることで，文字の高さを決定することができる．ここで切り出された文字の高さにおいて，垂直方向の黒画素分布を調べることで，文字の幅を決定する．

以上の一連の処理によって，文字を切り出すことができる．Fig. 4 に切り出し後の楽譜を示す．

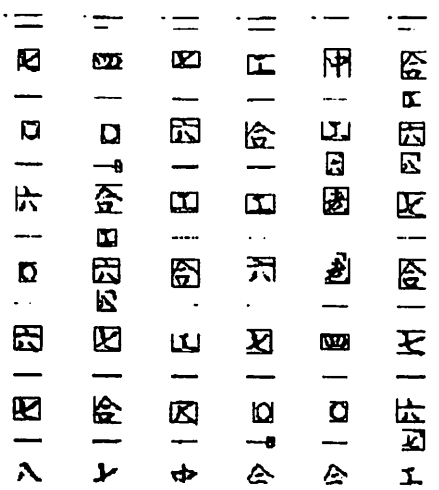

Fig. 4. 認識対象文字の切り出し

3. 特徴抽出法

認識対象とする文字は，毛筆によって記述されているため，各文字を構成している線分は，ある程度の幅を持っている．従来用いられているいくつかの手法によって，文字認識のための特徴抽出をするためには，計算量の多い，細線化を行なわなければならない．

本研究では, この細線化の処理を行なわず, 単純に, しかも高速に抽出できる黒画素の分布を特徴 [7], [8] として用いることにした. この時文字領域を文字の特徴が出るように 2 種類の方法で分割し各小領域について黒画素割合を求める. この各小領域には, Fig. 5(a) に示すように番号をつける. このように各小領域毎に黒画素割合を求めることで, 位置的な情報と構造的な情報を併せ持った特徴量を抽出することができる.

以下にその抽出手順を述べる.

3.2 重心を利用した分割法

文字の重心の位置は, 文字特徴と何らかの関係があると考えられる. よって, 一つ目の分割法では, まず, 切り出した文字の重心が領域の中心にくるように領域を拡張する. そして, 拡張した領域を 3x3 の小領域に分割しその中心の領域の黒画素割合を算出する (Fig. 5(c) 参照).

3.3 周辺分布を重視した分割法

2 つ目の分割法では, 周辺分布に注目する. 漢字の特徴はその周辺の黒画素の分布に良く表れることが報告されている. 従って, 周辺の領域の面積が大きくなるように, 切りだし枠を各辺の 2:1:2 の比で 9 つの領域に分割し, 中心の領域以外の 8 個の領域について黒画素割合を算出する. Fig. 5.(d) に示す.

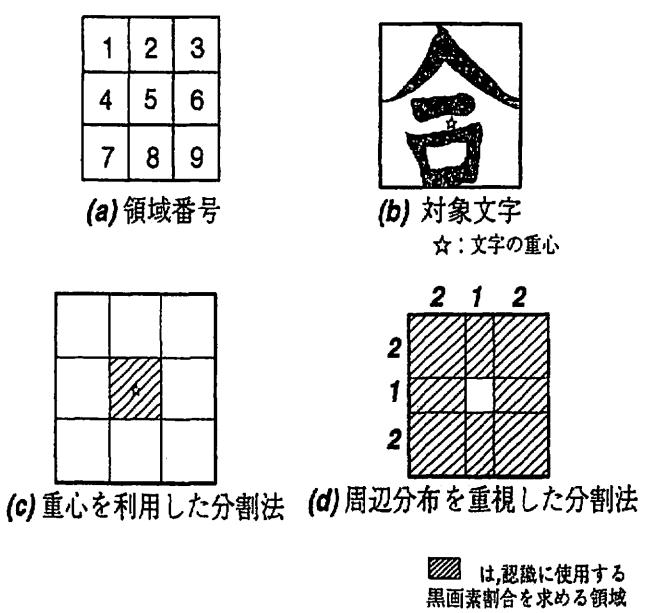

Fig. 5. 特徴量抽出法

4. 認識手法について

第 3 節で抽出した計 9 個の黒画素割合を, 9 次元の特徴ベクトルとして認識に使用する. 本研究では, 複数のアルゴリズムを組み合わせることで認識精度の向上を計っている. まず, 認識に使用するアルゴリズムの候補として, 単純類似度法, 複合類似度法そして, 本研究で提案する 3 次元テーブル法の 3 つのアルゴリズムについて述べる.

4.1 単純類似度法

単純類似度法とは, ある特定のカテゴリに属する標準パターンを一つ選び未知パターンと標準パターンとのパターン空間でのベクトル間の角度を認識の尺度とする方法である. 標準パターンを $\mathbf{X}$, 未知パターンを $\mathbf{Y}$ とすると, パターン空間での単純類似度 $S_S(\mathbf{X},\mathbf{Y})$ は,

$$S_S(\mathbf{X},\mathbf{Y}) = \frac{(\mathbf{X},\mathbf{Y})}{||\mathbf{X}||\cdot||\mathbf{Y}||} \tag{2}$$

で表される. 但し, $(\mathbf{X},\mathbf{Y}) = \mathbf{X}^t\cdot\mathbf{Y}$ とする. 本研究で使用する特徴ベクトルは, 10 次元 (n=10) である.

4.2 複合類似度法

複合類似度法は, 単純類似度法のように一つのカテゴリに対して一つの標準パターンを用いるのではなく, 複数の標準パターンを用いて, パターン空間での平均のベクトルの角度を認識の尺度として用いる方法である. 標準パターンとして m 個のパターンを使用すると, 複合類似度 $S_M(\mathbf{X},\mathbf{M})$ は,

$$S_M(\mathbf{X},\mathbf{M}) = \left\{\frac{1}{m}\sum_{i=1}^{m}\frac{(\mathbf{M}_i,\mathbf{X})^2}{||\mathbf{M}_i||^2\cdot||\mathbf{X}||^2}\right\}^{\frac{1}{2}} \tag{3}$$

となる. 本研究では, 10 個の標準パターンを用いた ($m = 10$).

4.3 3 次元テーブル法

本研究で提案する 3 次元テーブル法による認識について述べる. 3 次元テーブル法とは, Fig. 6に示すように, 横軸に認識対象文字, 縦軸に特徴量を対応させた表を各領域毎に 9 枚用意したものである. (このとき, 特徴量は, 0 から 80[%] までの間を 5[%] 刻みに 0 から 14 まで対応させ, 80[%] 以上は, 14 という 15 段階の値へと変換してある)

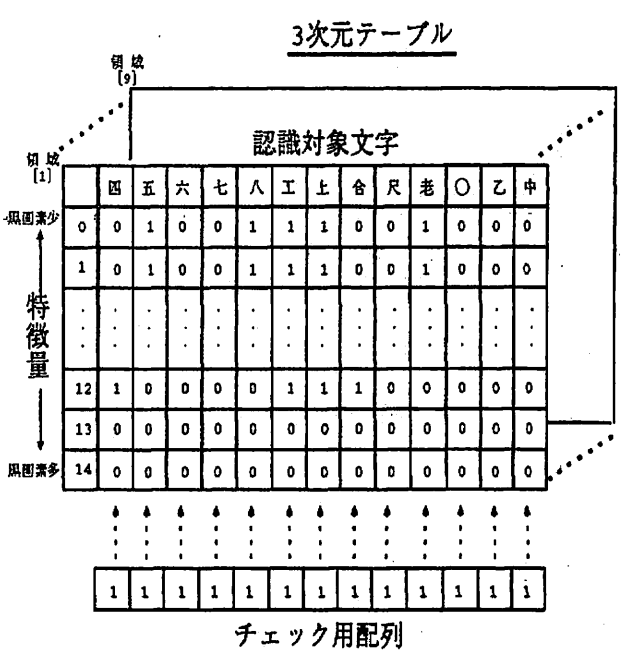

Fig. 6. 3 次元テーブルについて

表の各要素には, ある領域情報について考えた時, 認識対象文字がある特徴量を持つなら 1, 持たないなら 0 の 2

値の値が格納されている．この表は，あらかじめ抽出した特徴データを用いて作成する．

この作成された3次元テーブルを用いて認識を行なう．認識を行なう場合には，まず，認識対象文字に添字が対応している1次元配列を，チェック用配列として用意する．このチェック用配列の各要素には，添字に対応する文字が候補として考えられるなら1，候補ではないなら0が格納される．従って，最初は全ての文字が候補と考えられるので，初期値として1を与える．このチェック用配列と，未知パターンの特徴量に対応する，3次元テーブルの行をそれぞれ論理積を取っていく．この例をFig. 7に示す．

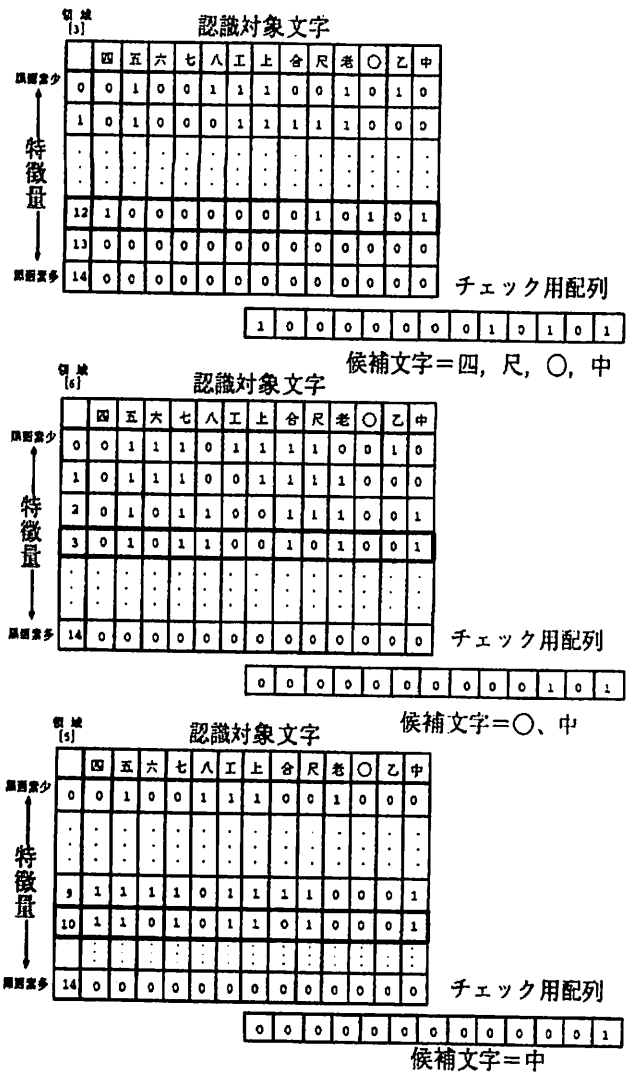

Fig. 7. 3次元テーブルによる認識過程

この例では，未知パターンの特徴量が12, 3, 10の順に出現している．この例より，論理積を取って行く過程で，チェック用配列の候補文字（要素が1に対応）が，(四，尺，○，中）の4文字から（中）の1文字に限定されて行く認識過程が分かる．この最終過程で，チェック用配列の要素が全て0になってしまった場合は，候補文字が消滅したことを示し，リジェクト文字とする．また，全ての領域情報を用いても，候補文字を1文字に限定できない場合にも，リジェクト文字とする．

このように，3次元テーブル法による認識では，単純な演算（論理積）により，認識を行なうことができるため，高速な認識を行なうことができる．また，使用する領域情報の順番を変更することで，計算量が変化する．実験により，どの文字を認識するのにも有効な情報と，ある特定の文字を認識するのに有効な情報に分けられることが分かった．この結果より，使用する領域情報の順番を以下に定めた．

$$\text{領域情報：}\underbrace{3 \rightarrow 6 \rightarrow 9 \rightarrow 5}_{\text{重要な情報}} \rightarrow \underbrace{4 \rightarrow 2 \rightarrow 1 \rightarrow 8 \rightarrow 7}_{\text{特定の文字に必要な情報}}$$

Fig. 8. 領域情報の重要性

このように使用する領域情報に順番をつけることでより高速な認識を行なうことができる．

5. 認識実験

この節では，前節で述べた3つの認識アルゴリズムを単体で用いた時の認識実験と，その実験結果より効果的に2つのアルゴリズムを組み合わせた時の認識実験を示す．認識には，26枚の楽譜画像を用い，学習楽譜としてその中から，6楽譜（529文字），未学習楽譜として20楽譜（1481文字）を使用した．また，今回は認識アルゴリズムの精度を比較することが目的なので，どのアルゴリズムに対しても共通な特徴抽出までの処理時間は，後述する実験結果には含まれていない．

TABLE 1に各認識手法毎の実験結果を示す．

TABLE 1　各認識手法についての実験結果

認識対象楽譜 [文字]	単純類似度法 [%]	複合類似度法 [%]	3次元テーブル法 [%]
学習楽譜 529	71.6	90.5	92.2
未学習楽譜 1481	77.0	90.9	87.8
全楽譜 2010	75.6	90.8	89.0

このTABLEより，認識率の点では，複合類似度法が最も良いことが分かるが，本研究で提案した3次元テーブルによる認識もかなり高い認識率となっている．処理速度を比較すると，全楽譜2010[文字]に対して認識を行なった場合には，単純類似度法で約90 [sec]，複合類似度法で約840 [sec]，3次元テーブル法で約12 [sec]であった．この処理速度の点からは，各種類似度法とは比較にならないぐらい3次元テーブル法が高速であることが分かる．

次に，認識対象文字別の認識結果をTABLE 2に示す．

TABLE 2　各認識手法の認識結果

対象文字 [文字]	各認識手法		
	単純類似度法 [%]	複合類似度法 [%]	3次元テーブル法 [%]
四: 234	94.9	**99.6**	97.9
五: 172	74.4	**90.1**	68.6
六: 73	**89.0**	87.7	75.3
七: 107	98.1	**100.0**	86.9
八: 19	57.9	**78.9**	47.4
工: 283	83.7	96.1	**97.2**
上: 230	29.1	49.6	**90.4**
合: 153	41.8	**93.5**	85.0
尺: 218	50.9	**94.0**	91.3
老: 314	97.8	**99.4**	90.8
○: 55	96.4	**100.0**	76.4
乙: 19	**94.7**	**94.7**	**94.7**
中: 133	98.5	**99.2**	95.5
合計: 2010	75.6	**90.8**	89.0

このTABLE中で, 太文字で示した値は, 各認識手法中最も高い認識率となっている値である. このTABLEより, 複合類似度法は, どの文字についても比較的高い認識率を持っているが, いくつかの苦手な文字に対しては, 3次元テーブル法による認識が良い値となっており, その逆も成り立っている. このことより, この2つのアルゴリズムはお互いの弱点を補うことができるといえる.

よって, 本認識システムの認識部をFig. 9に示すように, first step, second stepの2つのstepに分け, 2種類のアルゴリズムを併用する.

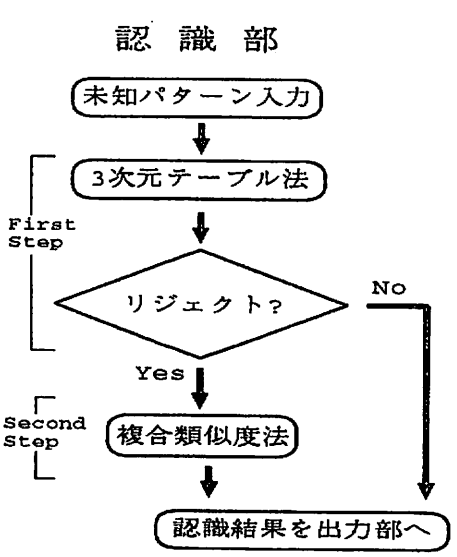

Fig. 9.　認識部について

first stepにおいて, 高速で, 比較的高い認識率をもつ3次元テーブル法によって, 高品質の文字を高速に認識を行ない, second stepでは, first stepにおいてリジェクトされた文字についてのみ認識を行なう. このように認識アルゴリズムを組み合わせることで, 高速な, 認識率の高い認識システムを構築することができる. 3次元テーブル法と複合類似度法を組み合わせた時の認識実験結果をTABLE 3に示す.

TABLE 3　2つの認識手法を組み合わせた時の実験結果

認識対象 楽譜 [文字]	本認識システム [%]	処理時間 [sec]
学習楽譜 529	97.4	約 20
未学習楽譜 1481	93.6	約 52
全楽譜 2010	94.6	約 72

このTABLEより認識アルゴリズムを組み合わせることで, 認識率は4[%] 近く上昇した. また, 処理速度においては, 3次元テーブル法単体で用いる時よりかなり遅くなってはいるが, 複合類似度法単体の場合よりは, まだ10倍以上速い結果となっている.

6. 自動演奏実験

この節では, 自動演奏実験について述べる. システムの構成図をFig. 10に示す. まず, コンピュータとサンシン自動演奏ロボット [1], [2], [3] のコントローラをプリンターインターフェースによって接続する. そして, 予め定められている制御用のパラメータ（演奏タイミング等）を送出した後に, 認識結果を演奏ロボットで使用している形式へと変換し送出する.

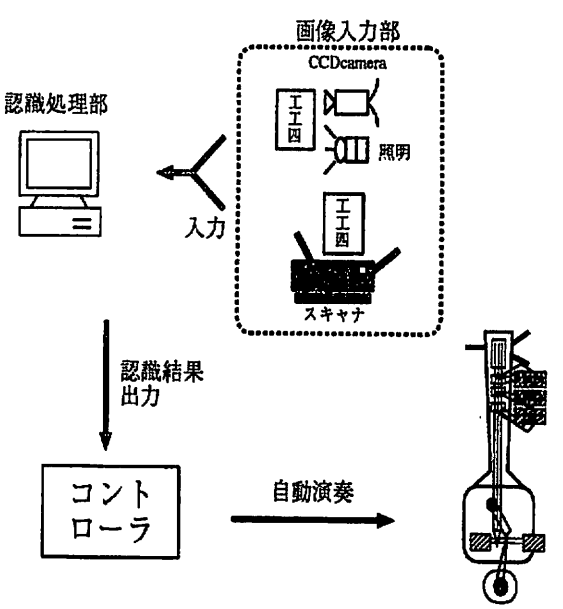

Fig. 10.　自動演奏システム

この自動演奏実験では, 認識率が9割前後の楽譜に対しては, ある程度の演奏ができた.

7. まとめ

本研究では，自動演奏ロボットへのデータ入力に使用する認識システムの構築を試みた．そのため本研究で提案する3次元テーブル法の高速性と精度の良さを実験により証明した．また，更に高精度な認識システムにするため，3次元テーブル法と複合類似度法を効果的に組み合わせることで，平均94.6 [%] の認識率を得ることができ，認識アルゴリズムの併用が，有効であることを示した．自動演奏実験では，認識結果を自動演奏ロボットへプリンターインターフェースを介して，接続することによりある程度の自動演奏を行なうことができた．

今後は，Fig. 10に示すように、楽譜をスキャナではなく，より高速な取り込みができるCCDカメラを利用することを現在検討中である．また、認識部についての改善点として特徴量の追加による，3次元テーブル法の精度の向上を目指す．工工四には，現在認識対象としている文字以外に，弾弦法を規定する記号である，“┐:掻音”,“、:打音”と呼ばれる各種記号が出現する．これらの弾弦法を規定する記号認識の対象とすることが，より人間に近い自動演奏を目指すために必要となるだろう．

参考文献

[1] 伊波善清，山城毅:
“サンシン演奏ロボットに関する研究(1)”，平成6年度日本ロボット学会全国大会学術講演会予稿集,No2(2377)(1994-11).

[2] 伊波善清，山城毅:
“サンシン演奏ロボットに関する研究(2)”，平成7年度日本ロボット学会全国大会学術講演会予稿集,No3(5A2-0-2)(1995-11).

[3] 伊波善清，山城毅:
“サンシン演奏ロボットに関する研究(3)”，平成8年度日本ロボット学会全国大会学術講演会予稿集,No2(2A2-1-4)(1995-11).

[4] H.Yamagata, H.Nishida, T.Suzuki: “A Handwritten Character Recognition System by Efficient Combination of Multiple Classifiers”,IEICE TRANS.INF.&.SYST, Vol.E79-D, No.5,pp.498-503,May,1996.

[5] 長嶺，幸地，山城，安冨祖:
“CCDカメラで取り込んだ文字画像の2値化に関する研究”，信学技報,vol.96,No.116,(1996-6).

[6] 奥水，丹羽:
“一次元配列を用いたHough変換による直線パターン検出法”，信学論(D),J63-D,No.7,(1984-7).

[7] 山城，仲宗根，松堂:
“工工四の自動読み取りに関する研究”，平成5年度電気関係学会九州支部連合会大会論文集.

[8] 祖慶，山城，渡久地，前原:
“工工四の楽譜認識に関する研究”，平成8年度電気関係学会九州支部連合会大会論文集,1685,p889.