AUTOMATIC MUSIC COMPOSITION BASED ON COUNTERPOINT AND IMITATION USING STOCHASTIC MODELS

Tsubasa Tanaka, Takuya Nishimoto, Nobutaka Ono, Shigeki Sagayama
Graduate School of Information Science and Technology, University of Tokyo
{tanaka,nishi,onono,sagayama}@hil.t.u-tokyo.ac.jp

ABSTRACT

In this paper, we propose a computational method of automatic music composition which generates pieces based on counterpoint and imitation. Counterpoint is a compositional technique to make several independent melodies which sound harmonious when they are played simultaneously. Imitation is another compositional technique which repeats a theme in each voice and associate the voices. Our computational method consists of the stochastic model of counterpoint and that of imitation. Both stochastic models are simple Markov models whose unit of state is a beat. We formulate the problem as the problem to find the piece which maximize the product of probabilities that correspond to both stochastic models. Dynamic programming can be used to find the solution because the models are simple Markov models. Experimental results show that our method can generate pieces which satisfy the requirements of counterpoint within two successive beats, and can realize imitations of the theme with flexible transformations.

1. INTRODUCTION

Counterpoint is one of the most basic principles for composition and arrangement. It is a composition technique for tuning voices, which are independent in contour and rhythm. Counterpoint consists of many prohibitions and recommendations such as ones which regulate the progression of a melody or the progression of interval between voices as shown in figure 1. It takes long time to master counterpoint, because finding melodies which satisfy many conditions is difficult. Therefore, realization of automatic contrapuntal composition will be valuable.

Imitation is also a very important technique for contrapuntal music. In a piece based on imitation, a theme melody is exposed at the start of the piece and melodies which imitate the theme repeatedly occur in all voices. As a result, imitation gives sense of unity to the piece. Imitation is an indispensable basis for some musical forms such as canon or fugue. Automatic composition systems based on counterpoint and imitation are useful because whole musical pieces can be obtained by simply providing a short theme and setting the structure of the piece.

Figure 1. Examples of requirements of counterpoint. The left figure shows a requirement for melodies. Conjunct motions (bar 1) should be used frequently. Skips (bar 2) should be used occasionally. The right figure shows a requirement for intervals between voices. Parallel fifths is prohibited [1, 2].

There are some previous studies about automatic counterpoint, such as rule-based approaches [3, 4] and methods based on stochastic models [5, 6]. A method based on hidden Markov models is also proposed in [7]. These studies are dealing with a kind of arrangement which is a composition of counter-melodies played simultaneously with a main melody which is given (cantus firmus).

In this paper, we aim at contrapuntal composition based on imitation of a given theme. We do not aim at generating pieces which are better than human compositions, but generating pieces which are acceptable from the standpoint of counterpoint and imitation.

Cope also proposes an contrapuntal composition system , which is based on the method of re-combination using the fragments of existing music as the components of generated pieces [8]. The limitation of his approach is that results are likely to be similar to particular pieces he uses, and that is his intention. The difference between our purpose and Cope's purpose is on this point. We intend to formulate general model without such limitation of the method of re-combination and generate new pieces.

We focus on two-voice free counterpoint, while there are various types in rhythms and the number of voices in counterpoint. Although the number of voices varies, the most important things are the relations in each pair of voices. Therefore, we focus on two-voice counterpoint, which is the most basic type of counterpoint about the number of voices. Concerning the rhythm, there are strict counterpoint and free counterpoint. The former is developed for pedagogy and has several kinds of fixed rhythms. The latter has variable rhythms and is more practical. We focus on practical composition and therefore deal with free counterpoint. In the following sections, counterpoint means free counterpoint.

2. STOCHASTIC APPROACH TO COUNTERPOINT AND IMITATION

In this section, we discuss a stochastic approach to the process of composition and introduce three assumptions. The formulation for the automatic composition based on counterpoint and imitation is derived from these assumptions.

2.1 Asummption about the Contrapuntal Composition

To generate proper contrapuntal pieces, we should define what is proper and what is not proper for the music we try to generate. For this purpose, it is considered to be effective to adopt a stochastic approach. Although we can use the knowledge of counterpoint represented by explicit rules, all the qualities of music can not be represented only by rules of counterpoint. In actual composition, a substantial part of selecting notes are handled by the composer's intuition. Therefore, it is better to model not only the knowledge of counterpoint but also the tendencies of composers. To model composer's tendencies of composing contrapuntal pieces, we propose a stochastic approach based on the following assumption;

Assumption 1 When an experienced composer write a contrapuntal piece, he or she is subjected to a probability distribution $\mathrm{Pr}(X)$ which is formed by knowledge and training of counterpoint, and realizes composition by a trial from the probability distribution (the variable X corresponds to a piece of music).

In this point of view, existing contrapuntal pieces can be considered as the outputs from the probability distribution $\mathrm{Pr}(X)$ and have high probabilities. In reverse, pieces which have high probabilities are considered to satisfy the requirements of counterpoint and to be acceptable as pieces of music. This probability distribution is considered to be the model of the contrapuntal composer.

2.2 Assumption about Three Steps of Composition based on Imitation

The process of composition based on imitation can be roughly divided into three steps:

1. The first step is to obtain a theme T. It can be composed by composers or taken from existing melodies.

2. The second step is to plan S, the structure of imitations and cadences. The structure of imitations indicates where and from what pitch imitations begin.

3. The third step is to select concrete notes and compose the actual piece X.

In practice, these three steps are not necessarily separated. However, these steps are considered to be carried out step by step from an idealistic viewpoint. S varies with the length and proportion of the piece according to time and circumstances. Therefore, S is better to be planned or selected by the user of the system before the automatic composition is started. For this reason, S should be also given as well as the theme T. These ideas are summed up as the following assumption;

Assumption 2 When a composer makes a piece based on imitation, he or she firstly obtain T, a theme to be imitated. Next, S, the structure of imitations and cadences is decided. Then, composition using the theme T and structure S is started.

If we deal with the style of tonal counterpoint, the structure S should include the plan of code progressions and key modulations. In this paper, however, S do not includes them. To make the problem simple, we deal with modal counterpoint, which is mainly the style of the Renaissance period and do not have code progressions.

2.3 Asummption about the Composition based on Imitation

The process of composition based on imitation of a theme can be viewed from the standpoint of probability in a similar way to assumption 1. Although imitation tends to be similar to the theme, imitation is not necessarily identical to it. For example, pitch shifts on the scale, intention to avoid unnatural harmonies, and tonality often affect imitations and transform them from the original theme. Such flexibility is necessary where strict imitations cause unfavorable results. A probability distribution in the next assumption is useful to realize such flexibility;

Assumption 3 When an experienced composer composes a piece based on imitation of a theme, he or she obtains the theme T and plans the structure S. After that, the composer composes the piece X. At this time, the composer is subjected to a conditional probability $\mathrm{Im}(X|T,S)$, which is formed by experiences of composition.

This probability $\mathrm{Im}(X|T,S)$ is considered as the model of imitation-based composition.

2.4 Formulation Based on the Three Assumptions

From these three assumptions, the problem of the automatic contrapuntal composition based on imitation of a theme is considered as a problem to obtain the piece X which give high probabilities to both the probability of the stochastic model of counterpoint and that of imitation. We show an example of the piece we intend to generate in figure 2. When T and S are given, the best piece $\tilde{X}$ can be formulated as the piece X which maximizes the product of $\mathrm{Pr}(X)$ and $\mathrm{Im}(X|T,S)$:

$$\tilde{X} = \operatorname*{argmax}_{X} \mathrm{Pr}(X)\mathrm{Im}(X|T,S). \quad (1)$$

By this formulation, pieces which realize flexible imitation and satisfy counterpoint are expected to be generated. However, it is difficult to obtain the values of $\mathrm{Pr}(X)$ and $\mathrm{Im}(X|T,S)$ directly from statistics on exsisting pieces of music. The reason is that the number of possible pieces is much larger than that of existing pieces. Therefore, it is necessary to extract the essential information that determine $\mathrm{Pr}(X)$ and $\mathrm{Im}(X|T,S)$, and to approximate these probabilities.

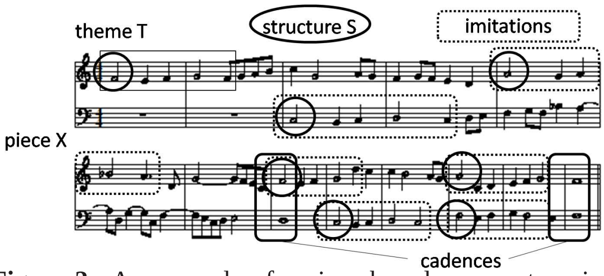

Figure 2. An example of a piece based on counterpoint and imitation [2]. Black circles indicate the starting points and the starting pitches of imitations. Domains bounded by dashed lines indicate imitations.

3. FORMALIZATION OF STOCHASTIC MODELS FOR COUNTERPOINT AND IMITATION

In this section, we propose stochastic models of counterpoint and imitation which approximate equation 1 in the case of two voices.

3.1 Formalization of the Stochastic Model of Counterpoint

To calculate $\mathrm{Pr}(X)$ and find $\tilde{X}$, data sparseness problem and exponential expansion of computational cost must be dealt with. A realistic solution is to approximate the probability by using existing probabilities and transition probabilities of short unit. Meanwhile, in composition and listening music, there are beat transitions behind concrete notes and they are written or listened upon beat transitions. Therefore, it is natural to adopt the length of a beat as the unit length of the states of probabilities. Defining $x_1, x_2, \cdots, x_N$ as information of each beat of the piece X, $\mathrm{Pr}(X)$ is transformed as:

$$\mathrm{Pr}(X) = \prod_{i=2}^{N} \mathrm{Pr}(x_i|x_{i-1}, x_{i-2}, \cdots, x_1)\mathrm{Pr}(x_1). \quad (2)$$

Equation (2) can be approximated using a simple Markov model, which is a favorable model in perspective of the computational cost. The ideas behind the approximation are as follows. Most requirements of the counterpoint are concerned with regulating the transition of pitch and the interval between voices. These requirements tend to be included within two successive beats. Therefore, assuming that long-raged dependences are ignorable, $\mathrm{Pr}(X)$ can be approximated by a simple Markov model whose unit of the state is a beat as:

$$\mathrm{Pr}(X) \simeq \prod_{i=2}^{N} \mathrm{Pr}(x_i|x_{i-1})\mathrm{Pr}(x_1). \quad (3)$$

Dynamic programming, by which the solution of probability maximization can be efficiently searched, can be applied to the simple Markov model. That is why the simple Markov model is advantageous. The details of dynamic programming are explained in section 4.

It is difficult to obtain $\mathrm{Pr}(x_i|x_{i-1})$ and $\mathrm{Pr}(x_1)$ statistically, because x_i contains information of multiple voices and have many possible variations of states, and therefore further approximation is necessary. Referring to the requirements of counterpoint, it is considered that $\mathrm{Pr}(x_i|x_{i-1})$ and $\mathrm{Pr}(x_1)$ are correlated with some elements which consist of them. As such elements, there are;

- Transition of pitch
- Transition of rhythm in each voice
- Transition of interval between voices
- Co-occurrence of rhythms of both voices

It is considered that there is a tendency that the higher probabilities of these elements are, the higher the probability $\mathrm{Pr}(x_i|x_{i-1})$ and $\mathrm{Pr}(x_1)$ are. The transition of pitch in a voice is related to the composition of the melodic progression, the transition of rhythm in a voice is related to the sense of rhythm and the note lengths, and the transition of interval between voices are related to the composition of harmony. Co-occurrence of rhythms of both voices are related to independence of voices and balance between voices.

Therefore, in equation (3), it is appropriate that $\mathrm{Pr}(x_i|x_{i-1})$ and $\mathrm{Pr}(x_1)$ are approximated by products of probabilities related to such elements;

$$\begin{aligned}\mathrm{Pr}(X) \simeq \prod_{i=2}^{N} \{ & \mathrm{Pr}(p_i^1|p_{i-1}^1)\mathrm{Pr}(r_i^1|r_{i-1}^1) \\ & \mathrm{Pr}(p_i^2|p_{i-1}^2)\mathrm{Pr}(r_i^2|r_{i-1}^2) \\ & \mathrm{Pr}(a_{i,i-1})\mathrm{Pr}(r_i^1, r_i^2) \ \} \\ & \mathrm{Pr}(p_1^1)\mathrm{Pr}(p_1^2)\mathrm{Pr}(r_1^1)\mathrm{Pr}(r_1^2)\mathrm{Pr}(a_1)\mathrm{Pr}(r_1^1, r_1^2).\end{aligned} \quad (4)$$

r_i^j and p_i^j are the information of the rhythm pattern and the series of pitch of each voice in the beat i. j means the index of the part. $j = 1$ corresponds to the upper voice and $j = 2$ corresponds to the lower voice. $a_{i,i-1}$ is the series of transition of the interval between the voices from the last interval of x_{i-1} to the last interval of x_i (the information of motions are also included. An oblique motion from major third to perfect fourth is an example of a component of the series). a_1 is the series of transition of interval in x_1.

3.2 Formalization of the Stochastic Model of Imitation

When we try to obtain the value of $\mathrm{Im}(X|T, S)$, data sparseness problem and the problem of computational cost occur as well as when we obtain the value of $\mathrm{P}(X)$. To deal with these problem, it is necessary to approximate $\mathrm{Im}(X|T, S)$. Before formalizing the approximated stochastic model of imitation, we define some variables. Q is defined as the number of imitations. T_n is defined as what T is shifted in pitch to start with the first pitch of the nth imitation. The unit of pitch shift is a semitone. The first pitch of the nth imitation is determined by S. M_n is the part of X which corresponds to the nth imitation. L is the number of beats within T (which is equal to the number of beats within T_n or M_n). t_l^n is information of the lth $(1 \leq l \leq L)$ beat of T_n. m_l^n is information of the lth beat of M_n.

In view of following ideas, we approximate $\mathrm{Im}(X|T, S)$. We can consider that there is a tendency that the more

M_n, which is a part of X, is similar to T_n, the higher the probability $\mathrm{Im}(X|T,S)$ is. Also, it is natural to assume that defferent imitations are independent. On the other hand, it is considered that whatever the notes in the region which is not the part of imitations or cadences are, $\mathrm{Im}(X|T,S)$ do not vary so much and can be regarded as constant. To make the problem simple, we always treat X as a piece which have the cadences determined by S (otherwise, $\mathrm{Im}(X|T,S)$ is regarded as zero). Therefore, $\mathrm{Im}(X|T,S)$ can be approximated as the product of the probabilities of each imitation as;

$$\mathrm{Im}(X|T,S) \simeq \prod_{n=1}^{Q} \mathrm{Im}(M_n|T_n). \quad (5)$$

To maximize both probabilities of the stochastic models at once, it is advantageous if $\mathrm{Im}(M_n|T_n)$, the probability of imitation of each time in equation (5), can be represented by a simple Markov model which have the unit of a beat, similar to the contrapuntal stochastic model. If $\mathrm{Im}(M_n|T_n)$ is represented by a simple Markov model, dynamic programming can be used.

$\mathrm{Im}(M_n|T_n)$ is considered to depend on the similarity of T_n and M_n. This similarity is considered to be determined mainly by four elements discussed later in this subsection. These four elements are reflected in $\mathrm{Im}(m_l^n|m_{l-1}^n, t_l^n, t_{l-1}^n)$, which is a part of $\mathrm{Im}(M_n|T_n)$. $\mathrm{Im}(M_n|T_n)$ can be transformed as equation (6). The relations other than that of the two successive beats in each probability can be ignored as we mention later in this subsection. Therefore equation (7) can be derived from equation (6).

$$\begin{aligned} &\mathrm{Im}(M_n|T_n) \\ &= \prod_{l=2}^{L} \mathrm{Im}(m_l^n|m_{l-1}^n \cdots m_1^n, T_n)\mathrm{Im}(m_1^n|T_n) \quad (6) \\ &\simeq \prod_{l=2}^{L} \mathrm{Im}(m_l^n|m_{l-1}^n, t_l^n, t_{l-1}^n)\mathrm{Im}(m_1^n|t_1^n) \quad (7) \end{aligned}$$

Equation (7) indicates that dynamic programming can be applied, because the stochastic model of imitation is equivalent to the simple Markov model in this equation.

As previously mentioned, the elements which are important in order to measure the similality of T_n and M_n are:

1. similarity of direction of pitch transition (upward and downward skip, upward and downward conjunct motion, and stay in the same pitch)
2. similarity of melodic interval
3. similarity of pitch
4. similarity of rhythm.

Among the four elements, the first and the forth are considered to be most important because the rough character of the melody is determined by these. The second and the third elements are considered to have a secondary role. If these four similarities between T_n and M_n are high enough, the imitation is expected to be successful.

Similarity of direction and melodic interval can be judged by comparison of the transitions from the previous note to the present note. Similarity of pitch can be judged by comparison of only the present note. Similarity of rhythm can be judged by comparison of the onset in each time. Therefore, these four elements are considered to be reflected in $\mathrm{Im}(m_l^n|m_{l-1}^n, t_l^n, t_{l-1}^n)$.

4. DYNAMIC PROGRAMMING

In this section we explain dynamic programming, which is a very efficient algorithm and can be used to find the X that maximizes the probability. Generally speaking, it takes $\mathrm{O}(c^N)$ computing time to search the best answer for series of length N. However, dynamic programming can reduce the computing time to $\mathrm{O}(N)$ by using the locality of a Markov model.

In this paper, the application of dynamic programming is as follows. We define $\mathrm{p}(x_i, x_{i-1})$ as the product of the probabilities across the beat $i-1$ and the beat i relating to both stochastic models of counterpoint and imitation (eqation (4) and (7)). x_i is the information of the beat i in X. $P_{\max}$ is the maximum of the cumulative probability of all $\mathrm{p}(x_i, x_{i-1})$ for each i. $P(x_i)$ is the maximum cumulative probability of $\mathrm{p}(x_i, x_{i-1})$ until x_i appears at the beat i. The Markov property enable us to represent $P(x_i)$ recursively as;

$$P(x_i) = \max_{x_{i-1}} \{p(x_i, x_{i-1})P(x_{i-1})\}. \quad (8)$$

By preserving $P(x_{i-1})$ for every x_{i-1}, we can obtain $P(x_i)$ sequentially. Finally, we can obtain $P_{\max}$ as the maximum of the $P(x_N)$ for every $P(x_N)$. Furthermore, preparing the backward pointer $b(x_i)$ which indicates the optimal path for x_i from beat $i-1$, we can obtain the optimal path which realizes $P_{\max}$. The optimal path can be obtained by going back from the x_N which realize $P_{\max}$ at the final beat to the first beat. $b(x_i)$ is represented as;

$$b(x_i) = \operatorname*{argmax}_{x_{i-1}} \{p(x_i, x_{i-1})P(x_{i-1})\}. \quad (9)$$

5. EXPERIMENT

5.1 Conditions of Experiment

In this section, we report the experiment of generating musical pieces by the proposed method.

Before computing with the algorithm, we manually produced the theme T which has length of 2 bars at the start of the lower voice. We also produced the structure S which start imitation at the 2nd , 8th, and 13th bars with the note B in the upper voice, and the 7th and 12th bars with the note E in the lower voice. Cadences, which are included in S, were also given manually.

In relation to the configuration of the probabilities, further approximations were done as described in the following subsections. As the number of notes in a beat increases, the number of states also increases. When the number of states is large, necessary amount of statistical data is also large. In general, further approximation of equation (4) and (7) which make it easy to set the value statistically is necessary. To determine the values of the probabilities, musical knowledge was also used. Statistical features were taken from "*Invention*" by J.J.Fux, which is shown in the textbook of counterpoint [2] as one of the example of a piece based on counterpoint and imitation.

We adopted eighth note as the minimum unit of time value of a note, and quarter rest as the minimum unit of a rest. Rests were not used except in the parts of exposition of the theme. Dynamic programming was used to find the results.

5.2 Transition Probability of Pitch

Considering that local properties are the most important in counterpoint, it is reasonable to approximate the transition probabilities of pitch $\Pr(p_i^j|p_{i-1}^j)$ by the product of local probabilities. We approximated $\Pr(p_i^j|p_{i-1}^j)$ by the product of unigram probabilities of each transition of pitch from the last pitch of p_{i-1}^j to the last pitch of p_i^j. Unigram probabilities are represented by relative intervals of transition of pitch (such as major third) to deal with data sparseness problem. Correction by geometric mean of unigram probabilities is adopted as equation (10), not to be effected by the number of notes:

$$\Pr(p_i^j|p_{i-1}^j) \simeq \begin{cases} \displaystyle\prod_{k=1}^{N(P_{i,i-1}^j)} \Pr(p_{i,i-1,k}^j)^{\frac{1}{N(P_{i,i-1}^j)}} \\ \quad \text{if } N(p_{i,i-1}^j) \neq 0 \\ P_0 \quad \text{if } N(p_{i,i-1}^j) = 0 \end{cases} \tag{10}$$

$p_{i,i-1}^j$ is the series of transitions of pitch. $p_{i,i-1,k}^j$ is the kth transition of pitch in $p_{i,i-1}^j$. $N(p_{i,i-1}^j)$ is the number of transitions of pitch. Where the number of times of pitch transition is zero (long notes such as half note, ties, etc.), $\Pr(p_i^j|p_{i-1}^j)$ is given by P_0. P_0 is obtained by statistics as the number of times there is no change of pitch within two successive beats, divided by the number of times there are changes of pitch within two successive beats. This division is done to balance lower case with the upper case in equation (10). Normalization is not done, because it does not affect the result. $\Pr(p_1^j)$, which corresponds to the first beat, was approximated within a beat in a similar way.

When the statistics of the unigram probabilities of pitch transition are taken, the probabilities are regarded as the same value whether the transitions are upward or downward and whether the intervals are major or minor. Where the interval is compound interval over augmented eighth, it is regarded as the simple interval (for example, minor tenth is regarded as minor third). If violations of counterpoint (such as succession of skips in the same direction) occur in two successive beats, the corresponding probabilities are lowered in equation (10). The unigram probabilities of pitch transition are shown in table 1.

In the piece from which the statistics are taken, no interval of sixth is included by chance. To correct the value of

Table 1. Unigram probabilities of pitch transition.

interval of pitch transition	$\Pr(p_{i,i-1,k}^j)$
prime	0.000
second	0.824
third	0.095
fourth	0.054
fifth	0.014
sixth	0.014
seventh	0.000
octave	0.000

probability of sixth interval, we regarded the value of probability of sixth interval as the same value as that of fifth interval. It is reasonable that the probabilities of prime, seventh and octave intervals are zero in table 1. That is because prime interval is not favorable and larger skips than seventh interval are prohibited in counterpoint. The transitions of pitch in the regions of imitation are counted only once, because the same patterns which are unique to the theme appear many times.

5.3 Transition Probability of Interval

We approximate transition probabilities of interval in two steps, because there are many combinations in transitions of interval.

In the first step, we approximate $\Pr(a_{i,i-1})$ by the product of the transition probability of interval $\Pr(b_{i,i-1})$ and the probability of motion $\Pr(c_{i,i-1})$ on the assumption that both are independent:

$$\Pr(a_{i,i-1}) \simeq \Pr(b_{i,i-1})\Pr(c_{i,i-1}). \tag{11}$$

$b_{i,i-1}$ is the series of intervals included in $a_{i,i-1}$ and $c_{i,i-1}$ is the series of motions included in $a_{i,i-1}$.

In the second step, $b_{i,i-1}$ and $c_{i,i-1}$ are approximated in a similar way to the case of the transition probability of pitch, and the correction by geometric mean were done as equation (12) and (13):

$$\Pr(b_{i,i-1}) \simeq \begin{cases} \displaystyle\prod_{k=2}^{N(b_{i,i-1})} \Pr(b_{i,i-1,k}|b_{i,i-1,k-1})^{\frac{1}{N(b_{i,i-1})}} \\ \quad \text{if } N(b_{i,i-1}) \neq 0 \\ B_0 \quad \text{if } N(b_{i,i-1}) = 0 \end{cases} \tag{12}$$

$$\Pr(c_{i,i-1}) \simeq \begin{cases} \displaystyle\prod_{k=1}^{N(c_{i,i-1})} \Pr(c_{i,i-1,k})^{\frac{1}{N(c_{i,i-1})}} \\ \quad \text{if } N(c_{i,i-1}) \neq 0 \\ C_0 \quad \text{if } N(c_{i,i-1}) = 0 \end{cases} \tag{13}$$

$b_{i,i-1,k}$ is the kth interval of the series $b_{i,i-1}$. $N(b_{i,i-1})$ is the number of the transitions of interval in the series $b_{i,i-1}$. $c_{i,i-1,k}$ is the kth motion of the series $c_{i,i-1}$. $N(c_{i,i-1})$ is the number of motions in the series $c_{i,i-1}$. Where the number of times of transition of interval is zero, the probability is given by B_0. B_0 is obtained by statistics as the number of times there is no change in interval within two successive beats, divided by the number of times where there are changes of interval within two successive beats. Where the

Table 2. Bigram transition probabilities of interval.

before \ after	perfect	imperfect consonant	dissonant
perfect	0.10	0.40	0.50
imperfect consonant	0.36	0.36	0.27
dissonant	0.09	0.91	0.00

Table 3. Unigram probabilities of motion.

motion	$\Pr(c_{i,i-1,k})$
parallel	0.121
contrary	0.212
oblique	0.667

Table 4. The values of P_0,B_0,C_0.

P_0	0.342
B_0	0.050
C_0	0.050

number of times of motion is zero, the probability is given by C_0. C_0 is obtained by statistics as the number of times there is no motion within two successive beats, divided the number of times there are motions within two successive beats. $\Pr(a_1)$, which corresponds to the first beat, is approximated within a beat in a similar way.

Intervals of the same category (perfect, imperfect consonant, and dissonant) are identified in statistics to reduce the varieties of transition of interval. If violations of counterpoint occur in two successive beats, the corresponding probabilities are lowered in equation (11). The bigram transition probabilities of interval are shown in table 2 and the unigram probabilities of motion are shown in table 3. The values of P_0, B_0, and C_0 are shown in table 3.

5.4 Co-occurrence Probability of Rhythms

Co-occurrence probabilities of rhythms can be determined by inference based on musical knowledge. In counterpoint, there is a tendency that rhythms continue regularly by placing onset in either voice. Where there is a onset in the head of r_i^1 or r_i^2, $\Pr(r_i^1, r_i^2)$ is considered to be high. By such inference, $\Pr(r_i^1, r_i^2)$ is determined as:

$$\Pr(r_i^1, r_i^2) \simeq \begin{cases} 0.9 & \text{if} \quad \text{there is the onset on the head of } r_i^1 \text{ or } r_i^2 \\ 0.1 & \text{if} \quad \text{there is not the onset on the head of } r_i^1 \text{ and } r_i^2 \end{cases} \tag{14}$$

Possible rhythm patterns for a beat within the condition of the experiment are shown in figure 3. The ties in figure 3 are the extension of duration to the next beat or from the previous beat.

5.5 Transition Probabilities of Rhythm

Transition probabilities of rhythm $\Pr(r_i^j | r_{i-1}^j)$ are determined by statistics from the piece mentioned above, because the transition probabilities of rhythm can not be determined by a simple rule.

There are nine possible rhythms in the condition of this experiment as shown in figure 3. The transition probabilities among them are shown in table 5. Each possible $\Pr(r_1^j)$, which corresponds to the first beat, is assigned the same value.

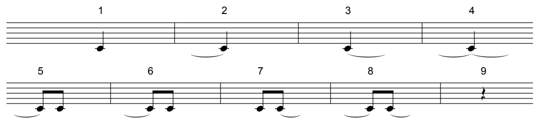

Figure 3. Possible rhythm patterns in the condition of the experiment.

Table 5. Transition probabilities of rhythm.

rhythm pattern befeore \ after	1	2	3	4	5	6	7	8	9
1	0.441	—	0.382	—	0.176	—	—	—	—
2	0.591	—	0.182	—	0.091	—	—	—	0.136
3	—	0.767	—	0.033	—	0.200	—	—	—
4	—	0.500	—	0.500	—	—	—	—	—
5	0.267	—	0.267	—	0.467	—	—	—	—
6	—	—	1.000	—	—	—	—	—	—
7	—	—	—	—	—	—	—	—	—
8	—	—	—	—	—	—	—	—	—
9	0.500	—	0.500	—	—	—	—	—	—

5.6 Probability of Imitation

It is difficult to take statistics of $\mathrm{Im}(m_l^n | \, m_{l-1}^n, t_l^n, t_{l-1}^n)$ in equation (7) because the probability depends on the theme and it has many parameters. On the other hand, $\mathrm{Im}(m_l^n | \, m_{l-1}^n, t_l^n, t_{l-1}^n)$ can be determined by inference based on musical knowledge. Considering that an imitation tend to be similar to the theme, there should be a tendency that where the similarity between (m_l^n, m_{l-1}^n) and (t_l^n, t_{l-1}^n) is high, $\mathrm{Im}(m_l^n | m_{l-1}^n, t_l^n, t_{l-1}^n)$ is also high.

To measure the similarity between the theme and the imitation, the four elements mentioned in section 3 (direction, melodic interval, pitch, and rhythm) can be used. Using the four elements, the similarities $-\Delta_{l,n}^k (k = 1, 2, 3, 4)$ are defined as follows. To compare (m_l^n, m_{l-1}^n) and (t_l^n, t_{l-1}^n), these are sliced into the length of eigths note, which is the minimum unit in this experiment.

We define similarity of direction as $-\Delta_{l,n}^1$. $\Delta_{l,n}^1$ is the sum of the absolute values of the differences of direction between each slice of (m_l^n, m_{l-1}^n) and (t_l^n, t_{l-1}^n). The sum is taken from the last slices of m_{l-1}^n and t_{l-1}^n to the last slices of m_l^n and t_l^n. Directions are represented as follows. the upward and downward skips are represented as ± 3 . Upward and downward conjunct motions are represented as ± 2. Continuation of the same pitch is represented as 0.

Similarity of melodic interval is defined as $-\Delta_{l,n}^2$. $\Delta_{l,n}^2$ is defined as the sum of the absolute values of the differences of melodic intervals (the unit is semitone). The sum is taken from the last slices of m_{l-1}^n and t_{l-1}^n to the last slices of m_l^n and t_l^n.

Similarity of pitch is defined as $-\Delta_{l,n}^3$. $\Delta_{l,n}^3$ is the sum of the absolute values of the differences of pitch (the unit is a semitone). The sum is taken from the first slices to the last slices of m_l^n and t_l^n.

Similarity of rhythm is defined as $-\Delta_{l,n}^4$. $\Delta_{l,n}^4$ is the sum of the absolute values of the differences of onset values. Onset value is defined as 1 for the position where the onset exist, and 0 for the position where the onset do not exist. The sum is taken from the first slices of m_l^n and t_l^n to the last slices of m_l^n and t_l^n.

Concerning $\Delta_{l,n}^1$ and $\Delta_{l,n}^2$, the absolute values of the differences is not included in the sum where the transition of pitch or interval is interrupted by a rest. Concerning $\Delta^3_{l,n}$, the absolute values of the differences is not included in the sum where rests appear.

Using these four similarities, $\mathrm{Im}(m^n_l|m^n_{l-1}, t^n_l, t^n_{l-1})$ are represented as:

$$\mathrm{Im}(m^n_l|m^n_{l-1}, t^n_l, t^n_{l-1}) \simeq \exp(-\prod_{k=1}^{4} \lambda_k \Delta^k_{l,n}). \quad (15)$$

$\lambda_k (k = 1, 2, 3, 4)$ in this equation are the weights of each λ_k, which are tuned manually in preliminary trial runs. Values from λ_1 to λ_4 are set as $12.0, 5.5, 0.5, 40.0$ respectively. $\mathrm{Im}(m^n_1|t^n_1)$, which corresponds to the first bar, is similarly approximated within a beat.

5.7 Result and Discussion of Experiment

From the observation of the results, it is confirmed that the proposed method can generate pieces which largely satisfy the requirements of counterpoint and imitation. As an example, one of the generated pieces is shown in figure 4.

Basic prohibitions of counterpoint such as parallel fifth or hidden fifth, which are included within succsessive two beats are not occurring in the piece shown in Figure 4. In the bar 5, 6, and 15, suspensions, which is important for counterpoint are occurring and dissonances are resolved appropriately. Both melodies are independent in the standpoint of rhythms in both voices at a time and the motions between both voices. However, there are some violations of counterpoint. In the bar 3, octave intervals appear both in the beat 2 and the beat 4. This is a prohibition of counterpoint which is called "successive accented perfect fifths or octaves" [2]. Such violations are sometimes observed. The reason is considered to be that stochastic model of counterpoint is approximated by simple Markov model whose states have the unit of beat and this model can not deal with the requirements of counterpoint which have the length over more than two beats. To cope with this problem, re-evaluation of the N-best solutions might be effective. From the N-best solutions, one which completely satisfy the requirements of counterpoint might be found.

Concerning imitation, adequate treatment is realized by the effect of stochastic model of imitation. In the positions where a imitation is done only in one voice (such as the imitation from the bar 3) the theme is not be transformed. On the other hand, in the positions where imitations in both voices are overlapping (such as imitation from the bar 13.) the theme is slightly transformed. If strict imitation is done there, both imitations will interfere mutually and cause violations of counterpoint. From such observation, we can say that the stochastic model of imitation realizes flexible imitation where transformations are necessary.

5.8 Generation with Given Themes

In the previous subsection, the theme is given by the authors. In this subsection, we give examples of the results in which we use well-known melodies for the theme to demonstrate that the proposing method can generate pieces without user's modification to the theme which may give advantage unfairly to the proposed method.

Figure 4. An example of the result. The lower voice of the first two bars is the theme.

Figure 5. The beginning of the fugue of J.S.Bach (upper) and the result generated with the same theme (lower).

Figure 5 shows the result using the same theme as the fugue of "The Well-Tempered Clavier Vol.1 No.3" by J.S. Bach. In this result, counterpoint is well satisfied and the imitation is also done well. However, in this theme there are skips with seventh interval, which is a prohibition of counterpoint. In the actual pieces like this, violations are sometimes done deliberately. To deal with such cases, we did a smoothing by setting the probability of skip with seventh interval as 0.001, not 0.

Figure 6 shows the result using the melody of "Little Hans" for the theme. In this result, successive accented perfect fifths or octaves are occurring across three successive beats in the bar 9, 17, and 26. In addition, repetitions of a phrase, which are unfavorable, are occurring from the bar 26 to 28. This is also caused by the approximation by a simple Markov model. If we can deal with this weak point by a method such as re-evaluation of the N-best solutions, acceptable results might be generated.

6. CONCLUSION

We proposed stochastic models for an automatic composition based on counterpoint and imitation. The solution is obtained as a musical piece which maximizes the product

Figure 6. A result generated with the theme of "Little Hans".

of probability of the stochastic model of counterpoint and that of imitation. In this formulation, dynamic programming can be used to search the solution. We reported the results of experiment to generate musical pieces by the proposed method. The results showed that proposed method can generate the pieces which satisfy the requirements of counterpoint that are included in two successive beats. The results also showed that flexible imitations can be realized by the effect of the stochastic model of imitation. However, a week point of the proposed method which should be improved was revealed. The week point is that the proposing method can not prevent the occurrences of violations or anfavorable things for counterpoint which can not be included in two successive beats. This week point is considered to be caused by the approximation by a simple Markov model.

There are several future tasks;

A. Dealing with the requirements of counterpoint which can not be included in two successive beats.

B. Extensive statistical learning of the probabilities.

C. Extension of the models for more than three voices.

D. Treating note values which are smaller than eighth note.

E. Introduction of tonal harmony and modulations.

F. Modeling of the characteristics of composers or instruments.

For A, re-evaluation of the N-best solutions might resolve the problem. For B, extensive statistical learning may enable us to obtain the values of probabilities more accurately. Furthermore, it may be possible that we can generate pieces which have diverse tendencies depending on the selection of the pieces for the statistical learning. For C, adding terms which are related to all the voices might realize the models for more than three voices. For D, when we treat smaller note values than eighth note, the problem of the increase of the number of the states or the problem of lack of unity as a piece of music may occur. To avoid the occurrence of these problems, restriction for the rhythm patterns might be effective. The restriction might reduce the number of the states and help to realize the unity as a piece of music. For E, including the plan of code progressions and modulations into the structure S will be necessary to deal with the counterpoint after the Baroque period. For F, introducing the feature quantities of composers or instruments by the musical knowledge or the methods of data mining may realize finer modelings.

acknowledgment: This study was carried out partially by the support of the Ministry of Education, Culture, Sports, Science and Technology Grants-in-Aid for Scientific Research (Fundamental Research A)(Task Number 00303321) and CREST by Japan Science and Technology Agency.

7. REFERENCES

[1] J. J. Fux: *Gradus ad Parnassum*, 1725. Translation by Y. Sakamoto: *Classical Counterpoint*, Ongakunotomo-sha, 1950 (in Japanese).

[2] Y. Hasegawa: *Counterpoint*, Ongakunotomo-sha, 1955 (in Japanese).

[3] William Schottstaedt: "Automatic Counterpoint," in Max V. Mathews and John R. Pierce, editors, *Current Directions in Computer Music Reseatrch*, The MIT Press, 1989.

[4] Hibiki Yoshikawa, Mitsuru Nakai, Hiroshi Shimodaira, Shigeki Sagayama: "Automatic Counterpoint using Dynamic Programming," Proc. of the 2000 Joint Conference of Hokuriku Chapters of Electrical Sosieties, F-82, Sep., 2000 (in Japanese).

[5] Shohei Nakagata: "Automatic Counterpoint based on Dynamic Programming," Master Thesis, The University of Tokyo, 2005 (in Japanese).

[6] M. Farbood and B. Schoner: "Analysis and synthesis of Palestrina-style counterpoint using Markov chains," *Proceedings of the International Computer Music Conference*, pp. 471-474, 2001.

[7] Riyu Tamura, Yasuhiro Tajima, Yoshiyuki Kotani: "Automatic Sub-Melody Generation Using Hidden Markov Models of Pitch and Duration," IPSJ Technical Report, 2007-MUS-69, pp.7-12, 2007 (in Japanese).

[8] David Cope: *Virtual music: computer synthesis of musical style*, The MIT Press. 2000.