

Automatic Piano Reduction from Ensemble Scores Based on Merged-Output Hidden Markov Model

Eita Nakamura

Meiji University

eita.nakamura@gmail.com

Shigeki Sagayama

Meiji University

sagayama@meiji.ac.jp

ABSTRACT

We discuss automated piano reduction from ensemble scores based on stochastic models of piano fingering and reduction process. Music arrangement including piano transcription is an important compositional technique, automation of which creates a challenging research field. As a starting point, we aim at a simple case of piano reduction which is playable and sounds similar to the original ensemble score. It is proposed to formulate the problem as an optimisation of fidelity to the original score under constraints on performance difficulty. First a model of piano fingering is presented to quantify performance difficulty. Next we construct a stochastic model for piano reduction based on the fingering model and probabilities to describe how notes in ensemble scores are likely to be edited, from which a piano reduction algorithm is derived. The models are constructed with merged-output hidden Markov model, which is a recently proposed model suited to describe a musical process involving multiple voice parts. It is confirmed that the constructed algorithm can control the performance difficulty of output reductions taking into account the density of notes and chords, the tempo, and the rhythm of the input ensemble score. The proposed formulation can be applied for more general music arrangement.

1. INTRODUCTION

Piano transcription from ensemble scores in various instrumentations including orchestral, chamber, and choral scores has traditionally been an important compositional technique, which is also widely used in the field of popular music. The purposes include extending the repertoire of piano performance, appreciation and diffusion of original pieces, using piano in place of orchestra accompaniment or other ensemble accompaniment, and exercise for composition and analysis. Automation of piano transcription (and music arrangement in general) should have wide applications and thus become an important musical technology. It is also interesting in the aspect of music research as an empirical test of analyses and models to understand the process of composition, arrange-

ment, and performance. Provided recent developments in information processing and computational power, it is now attractive to study these problems.

There are various possible piano transcriptions even for one ensemble score, reflecting musicality and techniques of the arranger and the intended performance difficulty. Since it is difficult to automate transcriptions with high musicality from the outset, we discuss in the following transcriptions that are as faithful to the original score as possible and playable at the same time. This sort of transcriptions, usually called piano reductions, are often seen in music practice, and their automation is important for direct applications as well as a basis for more advanced techniques of music arrangement.

Automation of piano reduction from ensemble pieces has been discussed in Refs. [1, 2] (see also Ref. [3] for a related work). In these studies, simple constraints such as the maximal number of simultaneous notes and the maximal interval played with one hand are used as conditions for playable reductions, and the necessity to control performance difficulty depending on players' skill is emphasised in Ref. [1]. While these simple constraints are effective for passages with a relatively slow tempo, which these studies take as examples, it is often necessary to change the playability condition according to tempos etc., and more detailed discussions are necessary. In addition, when much editing of notes is necessary, a complex problem of competing rules has to be solved.

To tackle these problems comprehensively, we formulate the problem of piano reduction as an optimisation problem of fidelity to the original ensemble score under constraints on performance difficulty (Sec. 2). Concretely, we first propose a method to quantitatively describe performance difficulty based on a stochastic model of piano fingering [4] (Sec. 3). We then develop a method of piano reduction based on a stochastic model which binds the fingering model and probabilities to describe how notes in ensemble scores are likely to be edited (Sec. 4). A method for guitar transcription based on a similar principle was recently proposed [5], and we extend the model for transcription so that more general transcriptions such as omission of chords in the original scores can be included. The fingering model and the model for piano reduction constructed in this study are based on merged-output hidden Markov model (HMM) [6], which describes a musical process involving multiple voice parts, and it is reviewed in Sec. 3. Finally, results and evaluation of the piano reduction algorithm are given in Sec. 5.

2. FORMULATION OF AUTOMATIC PIANO REDUCTION WITH PROBABILITY

There are various sorts of piano transcriptions from ensemble scores in music practice. Whereas there are transcriptions with high technicality and musicality such as the examples by Liszt [7], there are also transcriptions that are mostly faithful to original scores, which are seen in piano transcriptions of orchestral accompaniments (e.g., compare the overture in Ref. [8] and the corresponding piece in Ref. [7]). The most faithful transcription would be obtained by putting all the notes in an ensemble score into a piano score. However the result is often too difficult to play, and editings such as deleting less important notes are in order. The aim of this study is automatic production of piano transcription that is as faithful to the original score as possible and playable at the same time, which we call plain reduction¹.

In general, a score with less performance difficulty is obtained by deleting notes or changing the registers of notes, etc., but then the fidelity of the score to the original score decreases. Thus the problem of plain reduction can be restated as a problem of retaining as much fidelity to the original score as possible with the performance difficulty kept in a certain range. If we can quantitatively define performance difficulty and fidelity to the original score, we can formulate the problem as optimising the fidelity under constraints on the performance difficulty.

There are several aspects of performance difficulty such as the frequency of performance errors made in sight-reading or after sufficient practices, and the required time to master a piece [9, 10]. Relevant factors include not only the difficulty of making movements to play notes but also the way of representing the score. For simplicity, we mainly consider the difficulty of performance movements, particularly the difficulty of piano fingering, which is presumably the most relevant factor. The treatment of other factors is left for future research. The problem of finding the optimal fingering has been studied using the fingering cost, which is related to the difficulty of fingering [11, 12]. However, the principle to determine the fingering cost has not been established. Alternatively, we pay attention to the naturalness of fingering in a statistical sense, which can be defined from fingering data. A stochastic model of piano fingering can be used to determine the most natural fingering and quantify the difficulty of fingering in terms of statistical naturalness. Details will be explained in Sec. 3.

On the other hand, it is also difficult to set up a general definition of fidelity of a reduced score to the original score. In general, it is related to symbolic elements in the score as well as realised acoustics². For definiteness, we here focus on symbolic elements in scores including pitch and note value and consider their concordance as a factor of the fidelity. As explained in Sec. 4, musicians commonly delete

¹ This is similar to what is conventionally called “piano reduction.” We use the term plain reduction to avoid confusion by different conventions and stress that the idea is not restricted to piano transcription.

² For example, a chord in an orchestral score is often written as a tremolo in its piano transcription, the purpose being to compensate for the decaying sound of piano rather than to have a faithful symbolic representation of score.

notes or transpose them in octaves in piano reductions. Score reduction can be regarded as a process of editing notes with these operations, and a less amount of editing operations generally means more fidelity to the original score. To find an optimal plain reduction, it is necessary to compare multiple possibilities and this can be done with the introduction of an editing cost, which quantitatively describes what sort of editing operations are more allowable. The editing cost can be set up by an arranger, or it can be determined with editing probabilities obtained from some data of reductions. The latter way of using probability is advantageous to connect the editing cost with the performance difficulty, which we consider in this study.

With the quantification of performance difficulty and fidelity to the original score, a method of automatic plain piano reduction can be constructed based the above principle. As explained in detail in Sec. 4, we can build a stochastic model for piano reduction by binding the model of piano fingering and editing probabilities, and a piano reduction algorithm can be derived from the model.

3. PIANO FINGERING MODEL AND PERFORMANCE DIFFICULTY

Recently a quantitative measure of performance difficulty derived from a stochastic model of piano fingering was proposed [4]. For the purpose of using the model and the measure in Sec. 4, we review the main results in this section. For details, we refer the readers to Ref. [4].

3.1 Piano fingering model based on HMM

Let us begin with a fingering model for one hand. A passage is represented as a sequence of pitches $p_{1:N} = (p_n)_{n=1}^N$ (N is the number of notes). A fingering assigned to the score is represented as a sequence of finger numbers $f_{1:N} = (f_n)_{n=1}^N$ ($1 =$ thumb, $2 =$ the index finger, \dots , $5 =$ the little finger), each f_n being assigned to note p_n ³. A stochastic model of piano fingering describes the naturalness of a fingering $f_{1:N}$ to a passage $p_{1:N}$ in terms of probability $P(f_{1:N}|p_{1:N})$.

An explicit model for one hand based on HMM was proposed in Ref. [13]. In the model, we consider two probabilities. One is the probability that a finger would be used after another finger $P(f_n|f_{n-1})$, which we call transition probability. The other is the probability that a pitch would result from succeeding two used fingers $P(p_n|p_{n-1}, f_{n-1}, f_n)$, which we call output probability. With these probabilities, the probability of notes and fingerings is given as

$$P(p_{1:N}, f_{1:N}) = \prod_{n=1}^N P(p_n|p_{n-1}, f_{n-1}, f_n)P(f_n|f_{n-1}), \quad (1)$$

where initial probabilities are written as $P(f_1|f_0) \equiv P(f_1)$ and $P(p_1|p_0, f_0, f_1) \equiv P(p_1|f_1)$. The conditional probability $P(f_{1:N}|p_{1:N})$ is also given by the model accordingly.

³ We here implicitly assume that each note is played with only one finger. An exception is finger substitution, which we do not consider in this study.

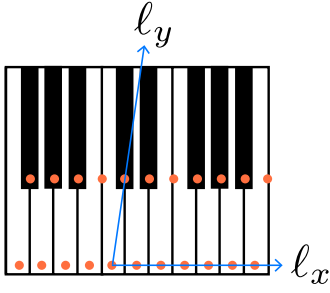


Figure 1. Representation of position on the piano keyboard with a two-dimensional lattice.

Values of the transition and output probabilities can be determined from fingering data. Because there are a great number of parameters of output probabilities and it is difficult to fully determine their values from data of practical size, we put reasonable assumptions and constraints to reduce the number of parameters. First it is assumed that the probability depends on pitches only through their geometrical positions on the keyboard which is represented as a two-dimensional lattice (Fig. 1). We also assume the translational symmetry in the x -direction, the time inversion symmetry, and reflection symmetry between hands. Then the output probability has a form $P(p'|p, f, f') = F(\ell_x(p') - \ell_x(p), \ell_y(p') - \ell_y(p); f, f')$ and satisfies certain conditions, where $(\ell_x(p), \ell_y(p))$ is the coordinate on the keyboard. The above model can be extended to including chords, by converting a polyphonic passage to a monophonic passage by virtually arpeggiating the chords [12, 14]. Here, notes in a chord are ordered from low pitch to high pitch.

3.2 Merged-output HMM and fingering model for both hands

Whereas the left and right hand parts are usually indicated with different staves in a piano score, the separation between hand parts is not given in reducing ensemble scores. We thus need a piano fingering model for both hands.

A hand part (left or right) associated to a note p_n is indicated with an additional variable $\eta_n = L, R$, and a fingering for both hands can be represented as $(\eta_n, f_n)_{n=1}^N$. One could consider an HMM with a latent variable (η_n, f_n) , but such a model cannot effectively describe the fingering process, which has the structure of stronger dependence among notes in each hand and weaker dependence across hands.

Recently a model, called merged-output HMM, is proposed that is suited for describing such a process with multiple voice parts [6]. The basic idea is to construct a model for both hands by starting with two parallel HMMs, called part HMMs, each corresponding to the HMM of each hand, and then merging the outputs of the part HMMs. The state space of the merged-output HMM is given as a triplet $k = (\eta, f_L, f_R)$ of the hand-part indication $\eta = L, R$ and the finger numbers for both hands. With the transition and output probabilities of the part HMMs $a_{ff'}^{L,R} = P_{L,R}(f'|f)$ and $b_{ff'}^{L,R}(\ell) = F_{L,R}(\ell; f, f')$

$(\ell = (\ell_x, \ell_y))$, the transition and output probabilities of the merged-output HMM are given as

$$a_{kk'} = \begin{cases} \alpha_L a_{f_L f'_L}^L \delta_{f_R f'_R}, & \eta' = L; \\ \alpha_R a_{f_R f'_R}^R \delta_{f_L f'_L}, & \eta' = R, \end{cases} \quad (2)$$

$$b_{kk'}(\ell) = \begin{cases} b_{f_L f'_L}^L(\ell), & \eta' = L; \\ b_{f_R f'_R}^R(\ell), & \eta' = R, \end{cases} \quad (3)$$

where δ denotes Kronecker's delta. Here $\alpha_L \sim \alpha_R \sim 1/2$ represents the probability of choosing which of the hands to play the note. Although certain interaction factors between hands can be introduced [6], we confine ourselves to the case of no interactions in this paper for simplicity. By estimating the most probable sequence $\hat{k}_{1:N}$, both the optimal configuration of hands $\hat{\eta}_{1:N}$, which yields hand-part separation, and the optimal fingering $(\hat{\eta}, \hat{f}_{\hat{\eta}})_{1:N}$ are obtained. For details, see Refs. [4, 6].

3.3 Quantitative measure of performance difficulty

A quantitative measure of performance difficulty based on the statistical naturalness of the fingerings can be obtained with the above model. It is given as the time rate of the probabilistic cost:

$$\mathcal{D}(t) = -\ln P(\mathbf{p}(t), \mathbf{f}(t))/\Delta t \quad (4)$$

where $\mathbf{p}(t)$ denotes the sequence of notes in the time range $[t - \Delta t/2, t + \Delta t/2]$, $\mathbf{f}(t)$ is the corresponding fingering, and Δt is a width of the time range to define the time rate, which can take values from a few 10 milli seconds to 10 seconds. It is possible to calculate $\mathcal{D}(t)$ for a score without indicated fingerings by replacing $\mathbf{f}(t)$ with the estimated fingerings $\hat{\mathbf{f}}(t)$ with the fingering model. The difficulty for each hand $\mathcal{D}_{L,R}(t)$ can also be defined similarly. Note that features such as playing speed, pitch entropy, hand displacement rate, hand stretch, fingering complexity, and polyphony rate, which are discussed in previous studies [10], are incorporated, although implicitly, in the measure.

Fig. 2 shows some examples of $\mathcal{D}_{L,R}(t)$ calculated for several piano pieces. Here and in the following, we set Δt to 1 sec. Although it is not easy to evaluate the quantity in a strict way, the results seem reasonable and reflect generic intuition of difficulty. Invention No. 1 by Bach, which can be played by beginners, yields $\mathcal{D}_{L,R}$ that are less than about 10, the example of Beethoven's sonata which requires middle-level technicality has $\mathcal{D}_{L,R}$ around 20 to 30, and Chopin's Fantasia Impromptu which involves fast passages and difficult fingerings has $\mathcal{D}_{L,R}$ up to about 40. It is also worthy of noting that relatively difficult passages such as the fast chromatique passage of the right hand in the introduction of Beethoven's sonata and ornaments in the right hand of the slow part of the Fantasia Impromptu are also captured in terms of \mathcal{D}_R .

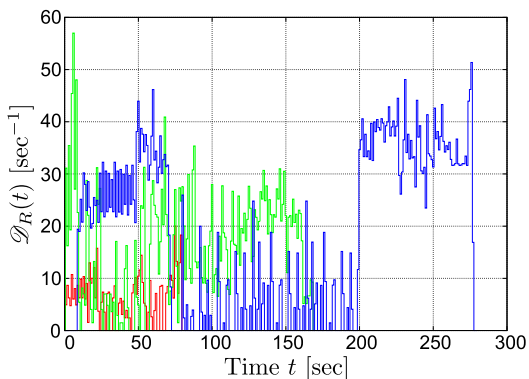
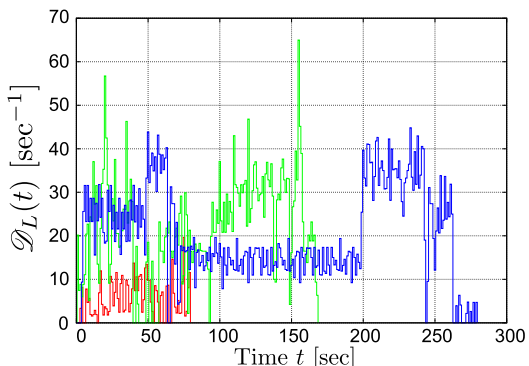

 (a) Difficulty for right hand \mathcal{D}_R

 (b) Difficulty for left hand \mathcal{D}_L

Figure 2. Examples of \mathcal{D}_R and \mathcal{D}_L . The red (resp. green, blue) line is for Bach’s two-voice invention No. 1, (resp. introduction and exposition parts of the first movement of Beethoven’s eighth piano sonata, Chopin’s Fantasic Impromptu).

4. MODEL AND ALGORITHM FOR PLAIN PIANO REDUCTION

4.1 Editing operations and probabilities in piano reduction

As explained in Sec. 2, the fidelity of a reduction to the original score can be defined as an amount of editing operations applied to the notes. The choice of used operations is influenced by music style, structure, melodic and harmonic elements of a piece, and it also reflects the personality of arrangers. Possibility of editing notes is not unique in most situations, and the process of editing a note can be described in terms of probability, which we call editing probability. The fidelity to the original score can then be defined as the accumulation of editing probabilities. Table 1 shows the frequencies of some editing operations of notes in the seventh piece of Tchaikovsky’s Nutcracker Suite for the composer’s piano reduction [15]. Together with the fact that only 1.7% of notes in the reduction have no corresponding notes with the same pitch or notes transposed in octaves in the original orchestral score, we see that deletion of notes and transposition in octaves are the main operations of editing, and there are a few

Table 1. Frequencies of editing operations applied to the notes in Tchaikovsky’s orchestral piece for the composer’s piano reduction.

# Notes	Not edited	Transposed in octave	Deleted and others
3705	47.4%	1.6%	51.0%

newly added notes in the piano reduction. According to this result, we consider deletion of notes and octave transpositions as possible operations in the following.

4.2 Model for plain piano reduction

In the probabilistic formulation, the problem of plain reduction can be defined as the problem of finding the optimal reduction \hat{r} given an ensemble score \mathbf{o} that maximises $P(\mathbf{r}|\mathbf{o})$ among a set R of possible reductions \mathbf{r} ’s. The set R is restricted by the constraints on the performance difficulty. The probability $P(\mathbf{r}|\mathbf{o})$ should reflect the measure of fidelity in terms of the editing probabilities and how \mathbf{r} is a good or natural piano score. This can be clarified by rewriting the probability with the Bayes’ formula as $P(\mathbf{r}|\mathbf{o}) \propto P(\mathbf{o}|\mathbf{r})P(\mathbf{r})$, where we suppressed a factor $P(\mathbf{o})$ which is irrelevant for the maximisation of the probability. If we take \mathbf{r} as the variables (including the outputs) of the fingering model in Sec. 3.2, $P(\mathbf{r})$ can be given with the model. $P(\mathbf{o}|\mathbf{r})$ can be defined with the editing probabilities. The formulation is akin to statistical machine translation (e.g. [16]), and a similar formulation for guitar arrangement is proposed in Ref. [5].

While a chord (or a set of simultaneous notes) is taken as a unit of state transition in Ref. [5], deletions of chords often appear in actual reductions. Such a deletion of a chord can be described by introducing an additional process of selecting whether a note in the ensemble score is played or not in advance to the process of the fingering model. This can be described by extending the variable $\eta = L, R$ to $\xi = NP, L, R$ (NP signifies “not played”), and if $\xi = NP$ a note is not played and both left and right fingers stay at the previous positions. The corresponding note \mathbf{o} in the ensemble score is then supposed to be generated from a uniform distribution $b_{\text{uniform}}(\mathbf{o})$.

To construct a computationally tractable model, we need to simplify $P(\mathbf{o}|\mathbf{r})$. For this, we assume that the editing probability is independent for each note in the ensemble score. Then, if a note \mathbf{o} in the ensemble score is represented as a pitch p in the reduction, the editing probability is given as $P(\mathbf{o}|\mathbf{r})$, which is denoted by $b_p(\mathbf{o})$. $P(\mathbf{o}|\mathbf{r})$ is given as the product of such editing probabilities for all notes.

In summary, the probability $P(\mathbf{o}|\mathbf{r})P(\mathbf{r})$ can be computed with a model with states represented by a set of stochastic variables $\mathbf{r} = (\xi, f_L, p_L, f_R, p_R)$ consisting of the state of the fingering model (f_L, p_L, f_R, p_R) and an additional variable $\xi = NP, L, R$ that determines whether a note in the ensemble score is played or not, and if played, by left or right hand. The probability $P(\mathbf{r})$ and $P(\mathbf{o}|\mathbf{r})$ are given as products of

transition probability $P(r'|r) = a_{rr'}$ and output probability $P(o|r, r') = b_{rr'}(o)$, which are given as

$$a_{rr'} = \begin{cases} \beta_{NP} \delta_{f_L f_L'} \delta_{f_R f_R'} \delta_{p_L p_L'} \delta_{p_R p_R'}, & \xi' = NP; \\ \beta_L a_{f_L f_L'}^L b_{f_L f_L'}^L (\ell(p_L') - \ell(p_L)), & \xi' = L; \\ \beta_R a_{f_R f_R'}^R b_{f_R f_R'}^R (\ell(p_R') - \ell(p_R)), & \xi' = R; \end{cases} \quad (5)$$

$$b_{rr'}(o) = \begin{cases} b_{\text{uniform}}(o), & \xi' = NP; \\ b_{p_L'}(o), & \xi' = L; \\ b_{p_R'}(o), & \xi' = R; \end{cases} \quad (6)$$

Here β_{NP} is the probability that a note in the ensemble score is not played, and $\beta_{L,R}$ is the probability that it is played by left/right hand ($\beta_{NP} + \beta_L + \beta_R = 1$). We see that the model is a merged-output HMM.

Although being implicit in the above equations, the values of the probabilities β_{NP} and $b_p(o)$ can vary according to the corresponding note in the ensemble score to describe tendencies such as a note in a passage in the ensemble score which is difficult to play directly has a high probability to be deleted, and notes in the melodic line and the bass line are less likely to be deleted. In principle, $b_p(o)$ can also describe what kind of editing operation is likely to be applied to a particular note.

4.3 Algorithm for plain piano reduction

Let us explain details of our algorithm for piano reduction based on the above model. We first make a condensed score by collecting all notes in a given ensemble score and putting voice parts with an identical pitch and note value into one voice part. The reduction result is obtained by calculating the most probable state sequence of the model in Sec. 4.2, taking the condense score as the output sequence of the model, which can be done with the Viterbi algorithm.

Control of performance difficulty and options such as retaining particular notes in the reduction can be obtained by tuning the probability values for β_{NP} and $b_p(o)$. We can use the measure presented in Sec. 3.3 for the control of performance difficulty. Since it is not possible to directly control the difficulty of the reduction result, a sub-optimisation method such as iterative optimisation is necessary in general. For the latter options of tuning the parameters, we can apply the methods to extract melody and bass parts given in Refs. [2, 1] and estimate the parameters from score elements, or we can reflect the arranger's preferences by manually tuning the parameters.

5. RESULTS AND EVALUATION

The present algorithm of piano reduction is evaluated with a passage (the last 16 bars) in the seventh movement ‘‘Danse des Mirlitons’’ in Tchaikovsky’s The Nutcracker Suite. Because an iterative optimisation required too much computation time, the difficulty of the reduced result was controlled by setting the probability β_{NP} for each note so that a predetermined value of the difficulty $\mathcal{D}_{L,R}$ is achieved on average. When the difficulty of the condensed score in the neighbourhood of a note is $\mathcal{D}_{L,R}^{\text{condense}}(t)$, the probability is set as

Figure 3. Condensed score of the orchestral piece used for the evaluation (see text).

$\beta_{NP} = \bar{\mathcal{D}}_{L,R} / \mathcal{D}_{L,R}^{\text{condense}}(t)$. Here the time interval to compute $\mathcal{D}(t)$ was taken as $\Delta t = 1$ sec, and the condensed score is preliminary separated into the left and right hand parts (with the fingering model in Sec. 3), and the corresponding difficulty is used. In this study, a melody part (the first flute) and a bass part (contrabass) were determined by extracting the skyline and bassline, and we set $\beta_{NP} = 0.0001$ for notes in these voice parts. The editing probability $b_p(o)$ was set as 0.8 for an unchanged pitch, 0.1 for an octave shift, and 0 otherwise. For the parameter of fingering model, we used the values in Ref. [4].

The reduction results are presented in Figs. 4(a) to 4(c), where cases with $(\bar{\mathcal{D}}_L, \bar{\mathcal{D}}_R) = (5, 10)$, $(20, 30)$, and $(30, 40)$ are shown. In the figures, a chord including an interval larger than or equal to tenth in one hand is indicated with an arpeggio. As a reference, the condensed score (Fig. 3) and a reduction by the composer (Fig. 4(d)) are also shown [15]. Overall, the difficulty of the reductions varies according to the preset values $\bar{\mathcal{D}}_{L,R}$. It is confirmed that the difficulty is also controlled locally: E.g., chords with rapid rhythms have fewer notes in the reductions. There is a tendency that the fidelity to the original score increases as the difficulty increases, but there were cases that some unmusical notes including unresolved discordant notes and ornaments were introduced additionally. The reductions were playable by a good pianist in most parts, but there were also notes that are difficult to per-

Table 2. Evaluation results of the reductions by the present algorithm. See text for the explanation of evaluation measures.

Fig.	$(\mathcal{D}_L, \mathcal{D}_R)$	Difficulty	Fidelity
4(a)	(5, 10)	8	4
4(b)	(20, 30)	8	4
4(c)	(30, 40)	10	8
4(d)	(Composer's reduction)	6	8

form (e.g. the downward-moving chords in the right hand in the 8th bar). There were a few cases of unnatural hand separations (e.g. F#3 in the right hand part in the 14th bar in Fig. 4(a)).

We asked a pianist to evaluate the reductions by the present algorithm in terms of performance difficulty (0: very easy to 10: very difficult) and fidelity to the original score (0: very different to 10: very faithful). Dynamics and articulations were not used for the evaluation. The scores are listed in Table 2, and it was pointed out that there are some notes that are difficult or impossible to perform in all reductions by the algorithm. The cause of the highest difficulty for the result in Fig. 4(a) is the downward-moving chords in the right hand in the 8th bar. Repetitions of chords in 32nd notes were generically judged as unplayable, but it was suggested that deleting one note would make the passage playable in some cases. The evaluation scores for the composer's reduction are also shown in the Table. It has no unplayable notes and is generally musically satisfactory. At the same time, however, there were suggestions for improvement such as adding B4 on the second beat of the fifth bar in the right hand would be more appropriate and the melody on the second beat of the eighth bar would better be in the original register.

Let us discuss the results. The results indicate that automatic piano reduction that can control the performance difficulty and the fidelity to the original score simultaneously is realised by the present algorithm. (Almost) unplayable notes in the reductions can also be recognised with the proposed measure of difficulty. For example, $\mathcal{D}_R(t)$ exceeds 50 around the 32nd notes in the eighth and tenth bars in Fig. 4(b), and it exceeds 80 in the 15th bar in the same reduction. Therefore, better reductions with less unplayed notes are expected to be obtained by iterative optimisation etc. Development of an algorithm and necessary reduction of computational cost are remaining issues. Another important issue is the incorporation of musical constraints such as appropriate voice leading into the model, which is possible in principle but raises a problem of large computational cost.

6. CONCLUSION

We discussed the problem of plain piano reduction from ensemble scores. Based on the proposed formulation of the problem as an optimisation problem of the fidelity to the orig-

inal ensemble score under constraints on performance difficulty, a piano reduction algorithm was developed with a stochastic model. A quantitative measure of performance difficulty was defined in terms of statistical naturalness based on a piano fingering model, and the fidelity to the original score was incorporated into a model for piano reduction with editing probabilities. It is confirmed that the constructed algorithm can control the performance difficulty of the output reductions taking into account the density of notes and chords, tempo, and rhythm, which would be difficulty to achieve with a set of simple constraints as in Refs. [1, 2]. At the same time, the reductions contained some unplayable notes and musically unsatisfactory voice leadings, which necessitates further refinements. We also confirmed that these problems could be improved with refinements on the optimisation algorithm and the model. In order to achieve more human-like reductions, further musical knowledge on harmony and musical forms should be incorporated in the model.

The present formulation of automatic music arrangement can in principle be applied for other forms of arrangement if we replace the piano fingering model with an appropriate performance/score model of the target instrumentation/style, and adapt the editing probabilities for relevant editing operations. It would also be interesting to extend the present method for a system to produce music arrangement with interactions between human and computer. As mentioned in Sec. 5, the reductions made with the algorithm can be much improved with a small amount of editing by human. It would also be effective that a user specifies his/her preference on the relative importance of notes and melodic segments and a particular way of editing notes by tuning the editing probabilities, which would be difficult to achieve in complete automation. Such a human-computer interaction would be an interesting intersection of computer technology and music.

Acknowledgments

This work is supported in part by Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science, No. 26240025 and No. 25880029.

7. REFERENCES

- [1] K. Fujita *et al.*, "A proposal for piano score generation that considers proficiency from multiple part (in Japanese)," *Tech. Rep. Special Interest Group on Music and Computer (SIGMUS)*, **MUS-77**, pp. 47–52, 2008.
- [2] S.-C. Chiu *et al.*, "Automatic system for the arrangement of piano reductions," *Proc. AdMIRe*, pp. 459–464, 2009.
- [3] S. Onuma and M. Hamanaka, "Piano arrangement system based on composers' arrangement processes," *Proc. ICMC*, pp. 191–194, 2010.
- [4] E. Nakamura *et al.*, "Merged-output HMM for piano fingering of both hands," *Proc. ISMIR*, pp. 531–536, 2014.

(a) $(\bar{\mathcal{D}}_L, \bar{\mathcal{D}}_R) = (5, 10)$

(b) $(\bar{\mathcal{D}}_L, \bar{\mathcal{D}}_R) = (20, 30)$

(c) $(\bar{\mathcal{D}}_L, \bar{\mathcal{D}}_R) = (30, 40)$

(d) Composer's reduction [15].

Figure 4. Reductions made by the present algorithm (a)–(c) with composer's reduction (d). The condensed score of the original piece is in Fig. 3.

- [5] G. Hori *et al.*, “Automatic arrangement for guitars using hidden Markov model,” *Proc. SMC*, pp. 450–456, 2012.
 G. Hori *et al.*, “Input-output HMM applied to automatic arrangement for guitars,” *J. Info. Processing Soc. Japan (IPSJ)*, **54(4)**, 2013-04-15, pp. 1–8, 2013.
 G. Hori and S. Sagayama, “HMM-based automatic arrangement for guitars with transposition and its implementation,” *Proc. ICMC|SMC*, pp. 1257–1262, 2014.
- [6] E. Nakamura *et al.*, “Merged-output hidden Markov model and its applications in score following and hand separation of polyphonic keyboard music (in Japanese),” *Tech. Rep. SIGMUS*, **2013-EC-27(15)**, pp. 1–6, 2013.
 E. Nakamura *et al.*, “Merged-output hidden Markov model for score following of MIDI performance with ornaments desynchronized voices, repeats and skips,” *Proc. ICMC|SMC*, pp. 1185–1192, 2014.
- [7] F. Liszt, *Musikalische Werke*, Serie IV, Breitkopf & Härtel, 1922.
- [8] K. Klindworth (piano red.), *R. Wagner: Tannhäuser Vocal Score*, G. Schirmer, 1895.
- [9] C. Palmer, “Music performance,” *Ann. Rev. Psychol.*, **48**, pp. 115–138, 1997.
- [10] S.-C. Chiu and M.-S. Chen, “A study on difficulty level recognition of piano sheet music,” *Proc. AdMIRe*, pp. 17–23, 2012.
 V. Sébastien *et al.*, “Score analyzer: Automatically determining scores difficulty level for instrumental e-learning,” *Proc. ISMIR*, pp. 571–576, 2012.
- [11] H. Sekiguchi and S. Eiho, “Generating and displaying the human piano performance (in Japanese),” *J. IPSJ*, **40(6)**, pp. 2827–2837, 1999.
 M. Hart and E. Tsai, “Finding optimal piano fingerings,” *The UMAP Journal*, **21(1)**, pp. 167–177, 2000.
- [12] A. Al Kasimi *et al.*, “A simple algorithm for automatic generation of polyphonic piano fingerings,” *Proc. ISMIR*, pp. 355–356, 2007.
- [13] Y. Yonebayashi *et al.*, “Automatic decision of piano fingering based on hidden Markov models,” *Proc. IJCAI*, pp. 2915–2921, 2007.
 Y. Yonebayashi *et al.*, “Automatic piano fingering decision based on hidden Markov models with latent variables in consideration of natural hand motions (in Japanese),” *Tech. Rep. SIGMUS*, **81**, pp. 179–184, 2007.
- [14] M. Haruhara *et al.*, “Automatic estimation system for piano fingering applicable to left hand and chords (in Japanese),” *Proc. 72nd National Convention IPSJ*, **2**, pp. 147–148, 2010.
- [15] P. Tchaikovsky, *The nutcracker suite*, P. Jurgenson, 1892.
 P. Tchaikovsky, *The nutcracker suite. Arrangement for solo piano*, P. Jurgenson, 1892.
- [16] P. Brown *et al.*, “The mathematics of statistical machine translation: Parameter estimation,” *Computational Linguistics*, **19(2)**, pp. 263–311, 1993.