

楽譜エディタの打鍵レベル模型

渡邊哲史 高田正之 小谷善行
東京農工大学

現在、様々な入力方式の楽譜エディタが開発されているが、その評価は、テキスト（文書）エディタのような定量化された方法がない。本研究では、Cardらの打鍵レベル模型をもとに楽譜エディタにおける利用者模型を求めるとともに、楽譜エディタの定量的評価方法を提案した。実験では、自作の楽譜エディタ emucom を使って、4人の利用者が基本的な音符（和音、単旋律）の入力を行い、その打鍵情報を収集した。エディタの利用者は、その音楽的素養によって異なる音楽的な chunk（認識できる単位）を持ち、それが打鍵動作に影響してくることが明らかになった。

A Keystroke-Level Model for Music Score Editor

Satoshi Watanabe Masayuki Takata Yoshiyuki Kotani

Tokyo University of Agriculture and Technology

2-24-16 Nakamachi, Koganeishi, Tokyo, Japan

Recently music score editors with many input system are developed. These aren't evaluated by the quantitative way. We defined a user model of the music score editor by Card's Keystroke-Level Model. And we propose a quantitative evaluation way of the music score editor. We collected four user's keystroke data of the music score editor "emucom" of our own making. Each user inputted some basic notes (chord, melody). We found that a user has a different musical chunk. A gap of chunk is based on a user's knowledge of music. We found that these chunks affect user's keystrokes.

1. 本研究の目的、背景

近年、楽譜を編集対象として処理するさまざまなシステムが開発されている。その入力方法も、計算機のコンソール、マウス、ジョイスティック、楽器などさまざまなものがある。また、編集作業のコマンド体系も数多くあり、利用者側とすれば混乱するばかりである。一方、テキストを対象とするエディタは、使い勝手が当初から研究され、そのコマンド体系は、より使いやすいものに固定しつつある。

テキストエディタにおいては、KL 模型^[1]などの定量化された評価方法が取り入れられ、入力方式、コマンド体系、入力デバイスにいたるまで研究されてきている。そこで本研究では、こういったテキストエディタで使われている定量的な評価方法を、自作した楽譜エディタ emucom (Editor for MUsic COMposition) に取り入れるとともに、その背景にある利用者模型を明らかにした。

2. 会話型システムの評価方法

テキストエディタにおける打鍵レベル模型は、Card^[1]らによって、開発された模型である(以下KL 模型と呼ぶ)。この模型によって、テキストエディタなどの会話型システムの作業時間が打鍵という立場から予測できる。

利用者の計算機を使った入力作業は、大きく次の二つに分けられる。

(1) 作業獲得時間：利用者が作業を与えられてから、その内容を理解するまでの時間

(2) 作業実行時間：利用者が与えられた作業を、実際に手を動かして実行する時間

KL 模型は、(2) だけについて考えたものである。対象システムの熟練者が、誤りを犯すことなく、一気に作業することがこの模型の条件となっている。訂正などの作業も一つの編集動作として模型化しようという考えである。

打鍵レベルのオペレータ(操作名)は図 2.1 のように定義される。

- ・ 運動神経のオペレータ
 - K : Keystroke (キーボード、マウスなど)
 - P : Pointing (マウスなど)
 - H : Homing (デバイス間の手の移動)
 - D : Drawing (マウスで線分を描く)
- ・ システムのレスポンスオペレータ
 - R
- ・ メンタルオペレータ
 - M

図 2.1 KL 模型のオペレータ ([1]から引用)

3. 楽譜入力実験

3.1 楽譜エディタ emucom

実験に用いた emucom は、楽譜データを入力、編集するツールであるとともに、使用中

の打鍵情報（キーの種類、キーが押されるまでの時間）を自動収集する。emucom の環境、構成を表3.1に、楽譜の入力方式を図3.1に、音符類の入力に利用するキーの配置を図3.2に示す。

表3.1 emucom の環境、構成

計算機	パーソナルコンピュータ PC-9800 シリーズ (E以降)
OS	MS-DOS ver. 3.30
使用言語	c 言語、8086 機械語
プログラム規模	105kB (実行形式)

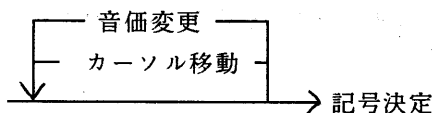


図3.1 emucom の入力方式

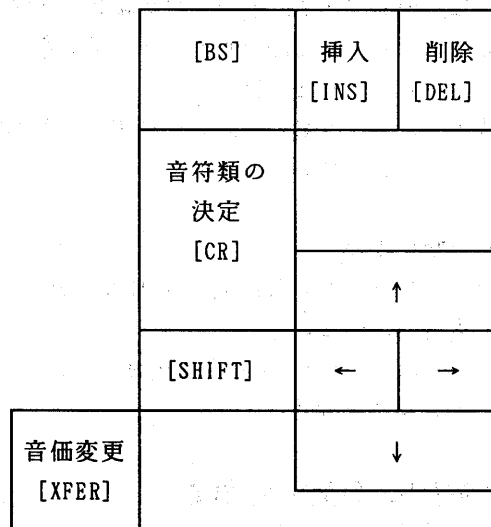


図3.2 基本的なキー配置 (PC-9800シリーズ)

3.2 打鍵情報収集実験

楽譜エディタの利用者の行動を、2章で紹介したKLモデルと同じレベルのモデルで定式化し、楽譜エディタの利用者モデルを明らかにするために次のような打鍵情報収集実験を行った。

3.2.1 実験のための作業内容

作業は、楽譜上の基本的な記号（音符、休符、調号、拍子記号、臨時記号、縦線など）を編集するもののうち、emucom で実現可能なものを5パターン用意した（表3.2参照）。

表3.2 実験に用いる作業内容

番号	実験内容
exp1	総合的な楽譜の入力
exp2	前後の音との音程があまり離れていないフレーズの入力
exp3	その場で覚えて楽譜を見ずに入力
exp4	付点音符と規則性のあるフレーズ（アルペジオ）の入力
exp5	2度音程を含まない和音と、2度音程を含む和音の入力

3.2.2 被験者

被験者は、東京農工大学数情報工学科の教官1名（以下Tとする）、筆者を含む学生3名（以下A、D、Eとする）の合計4名である。このメンバーは音楽情報処理に従事し、emucomを日ごろから使用している。また、いずれもパーソナルコンピュータ上の作業という点では熟練者である。音楽的な知識量や技量は、次の二つのタイプに分けられる。

被験者A、D：一度聞いたメロディを楽器で弾ける。相当数の曲を知っており、自分で作曲を行う。

被験者T、E：音楽には大いに興味があるが、楽器演奏、読譜、作曲などの技術は有していない。

3.2.3 実験の手順

実験の手順は、次の(1)～(3)の要領で行い、15分程度の休憩をはさんで、2度行う。

- (1) 実験の説明を行う。
- (2) emucomの出力による楽譜を与える。
- (3) 編集作業を行う。

4. 実験結果と打鍵レベル模型

4.1 KL模型による予測

まず、それぞれの実行時間をKL模型で予測したところ、実測値の2倍～3倍の過大評価となった。図4.1のような和音の入力を例にとると図4.2のような予測となる。ここで予測における打鍵速度は予備実験による実測値を用い、Mオペレータに要する時間はすべてCard^[1]による1.35秒とした。

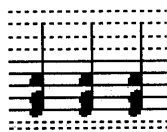
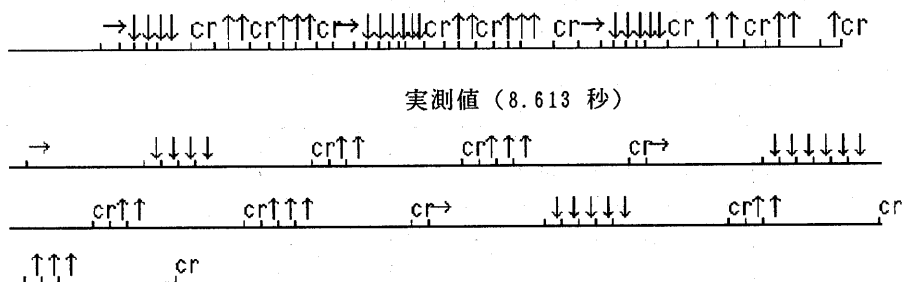


図4.1 exp5 で入力した和音



KL 模型による予測値 ($41K+12M=41*0.200+12*1.35=24.400$ 秒)

図4.2 和音入力の時軸チャート (被験者A、1行10秒)

この模型では、打鍵以外の「ふっ」と考える時間を1つのオペレータで表しているため、実際にはもっと短く済む思考が正しく予測されていない。

4.2 モデル人間プロセッサの導入

Card^[3]らによると、人間の脳の中の働きを次の三つのプロセッサに分け、定義している。

- (1) 知覚プロセッサ t_p
- (2) 認知プロセッサ t_c
- (3) 運動プロセッサ t_m

(1) は目などからの情報を取り込む働き、(2) は作業記憶をもとに、長期記憶を参照、更新する働き、(3) は筋肉を動かす働き、を表す。それぞれに要する時間は次のように定義されている。

- (1) 100 (50~200) ミリ秒
- (2) 70 (25~170) ミリ秒
- (3) 75 (30~100) ミリ秒
- () 内は個人差。

4.1で述べたように、音符入力に関して、KL 模型をそのまま当てはめたのでは、まったく予測は不可能である。そこで、このモデル人間プロセッサの考え方を音符入力の場面に導入してみる。

被験者が日頃楽譜を入力しているときの打鍵情報を観察すると、「ふっ」と考える以外に、現在のカーソル位置を視認する時間が存在することがわかった。これらの時間は、KL 模型におけるメンタルオペレータ (1.35 秒) よりはるかに短い。これらの場面では、利用者は、カーソルの位置を視認していると考え、モデル人間プロセッサを導入すると、

- (1) 現在のカーソルの位置を目に取り込む t_p
- (2) 目で見た位置と頭の中の位置との照合を行う t_c
- (3) (2) にもとづいて、次のキーを選択する t_c
- (4) 指が動く t_m

となる。(1)~(4)の合計時間は、平均的には、

$$W = t_p + 2 t_c + t_m = 100 + 2 \times 70 + 70 = 310 \text{ (ミリ秒)}$$

となる。Wを用いて和音入力の実行時間を予測したものを表4.1に示す。

表4.1 1小節の和音入力

被験者	実測（秒）	予測（秒）	打鍵速度（秒）	予測式*
A	8.613	11.920	0.200	41K+12W
D	7.136	9.393	0.203	31K+10W
T	14.655	10.323	0.233	31K+10W
E	12.477	13.806	0.246	41K+12W

*予測はそれぞれの被験者が実際に用いた入力方法で行った。

表4.1を見ると、被験者A、Dに関してはまだ過大評価になっている。両名は他の2名と比べて音楽的能力（楽器演奏、作曲経験）がかなり高いことがわかっている。この場合、楽譜上の和音を見たとき、それを一つの単位として捉え、入力の際には、一気に入力しているのではないだろうか。つまり、音楽的な面でも「認知のまとまり（chunk）」というものが存在し、この2名がそれをより明確に、より多く持っている、と考えられる。この現象は、通常の楽譜入力の際の打鍵にも現われている。そこで、chunkによるカーソル視認時間の省略を加味した予測を図4.3、表4.2に示す。

表4.2 1小節の和音入力の
chunkを加味した予測

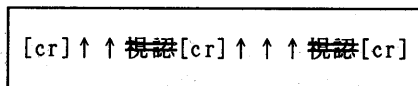


図4.3 和音入力時の視認の省略

被験者	実測値（秒）	予測値（秒）
A	8.613	10.060
D	7.136	7.533

4.3 本模型による予測のまとめ

これまでの考察をもとに、本模型による音符入力の打鍵時間の予測を図4.4～6にまとめた。

1小節の単旋律入力

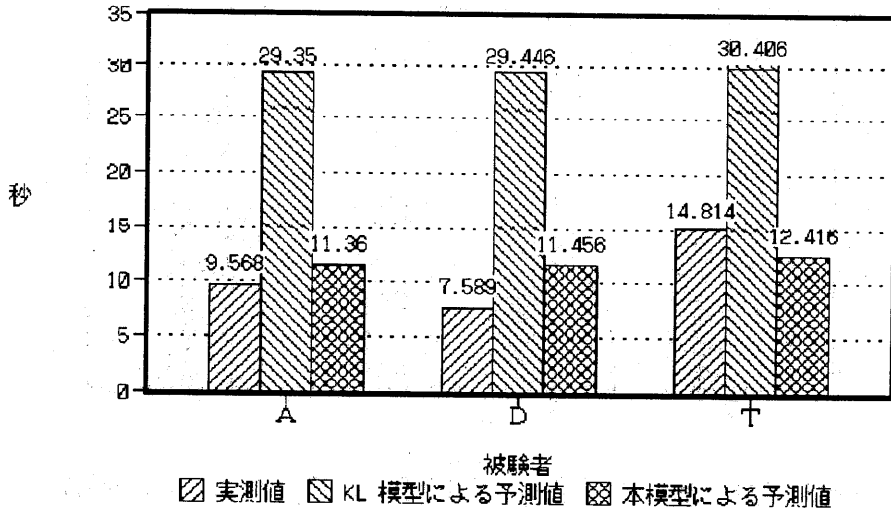


図4.4 1小節の単旋律入力への予測

和音の入力

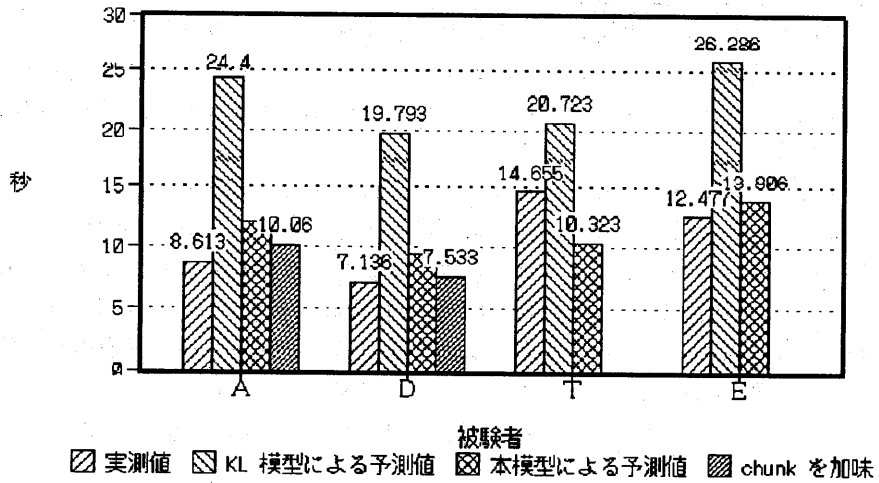


図4.5 和音入力への予測

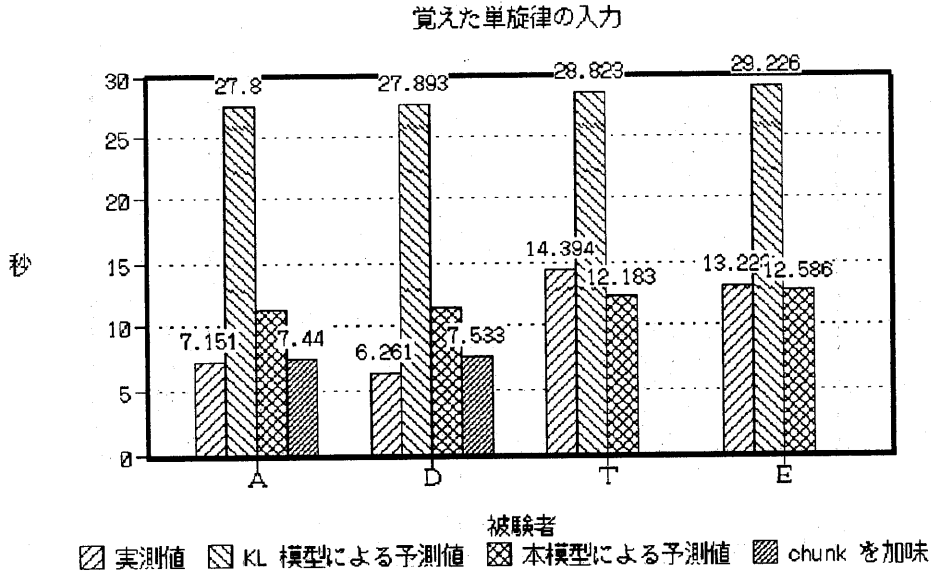


図4.6 覚えた単旋律入力への予測

5. 本打鍵レベル模型の検証、評価

今回の予測に用いた打鍵速度は、単純な打鍵予備実験によって被験者ごとに求めたものを用いた(0.2秒前後)。しかし、通常の楽譜入力の際の、打鍵速度の最頻区間を求めると、0.15~0.2秒と多少速くなっている。これは、[矢印キー]などの連続打鍵が原因していると考えられる。

カーソル視認時間がはっきりと現われるのは、記号を入力する際の決定キー([cr]キー)打鍵の直前である。この値と、今回求めた模型値を比較したものを表5.1に示す。

実際の打鍵には個人差があったり、chunkは、音楽的な素養が大きく関係する。和音というchunkを一つの単位として認識すれば、視認は省略される(これは模型上の考え方であって、実際には視認がかなり速い時間で済むと考える)。打鍵レベル模型の基本的な考え方は、「細かい部分(個人差)はともかく、ある作業(unit task)全体として、どれくらい予測できるか。」であるので、この310msecを一応の目安とする。

表5.1 [cr]打鍵まで時間の最頻区間

被験者	[cr]打鍵までの最頻区間
A	[200-250]msec
D	[200-250]msec
E	[250-300]msec
T	[200-250]msec

視認時間 = 310msec

6. まとめ

本研究により、emucom を中心とした環境（編集、演奏）は、PC-9801 上ではほぼ整備され、他の研究（自動作曲、自動編曲、楽曲分析など）とのインターフェースも実現された。今後は、ワークステーションへの移植を行い、より統合化された環境への整備が望まれる。

楽譜エディタの打鍵情報をもとに、打鍵レベル模型を定義した。その結果、現在位置を視認するオペレータをモデル人間プロセッサにより定義し、和音、単旋律といった音楽的な chunk の考え方を導入した。

楽譜エディタの打鍵レベル模型を定義する段階で、その背景にある、楽譜の取り込み方などの利用者模型を明らかにした。これらの模型は、他の楽譜エディタとの比較評価や、キー配置などの使い勝手の評価に有効であろう。

7. 参考文献

- [1] Stuart K. Card and Thomas P. Moran: The Keystroke-Level Model for User Performance Time with Interactive Systems, Communications of the ACM, Vol.23, No.7, July 1980, pp.396-410.
- [2] Teresa L. Roberts & Thomas P. Moran: The Evaluation of Text Editors-Methodology and Empirical Results, Communications of the ACM, Vol.26, No.4, April 1983, pp.265-283.
- [3] Stuart K. Card and Thomas P. Moran: The Psychology of Human-Computer Interaction, Lawrence Erlbaum Associates, Hillsdale, New Jersey(1983).
- [4] R.B.Allen and M.W.Scerbo: The Details of Command-Language Keystrokes, Acm Trans. O.A., Vol.1, No.2(Apr. 1983), pp.159-178.
- [5] 木村、粕川：変換型日本語ワープロ向き打鍵レベル模型、第27回プログラミングシンポジウム報告集(1986.1)、pp.95-106.
- [6] 木村、粕川：日本語ワープロ向き新打鍵レベル模型の直接的検証、第28回プログラミ

ングシンポジウム報告集(1987.1)、pp.79-86.

[7] 木村、粕川：日本語ワープロ向け新打鍵レベル模型の検証と応用、情報処理学会研究報告(日本語文書処理研究会)JDP-6-4(1986.5).

[8] 渡邊哲史：作曲支援のための楽譜エディタの開発、東京農工大学数理情報工学科1987年度卒業論文