Vol.2015-MBL-77 No.6
Vol.2015-ITS-63 No.6
2015/12/2

スマートフォンにおけるリズムパターン認識ライブラリを目指して

平野睦[†1]　伊藤信行[†2]　小林幸彦[†2]　梶克彦[†1]　内藤克浩[†1]　中條直也[†1]　水野忠則[†1]

概要：人がスマートフォンに意思伝達するための新しい手段として利用可能なリズムパターン認識ライブラリを開発している．スマートフォン自体や周辺の机などを指でタップすると，その音響信号からタップされたタイミングが検出される．複数の登録済みリズムパターンと入力パターンとの比較によって，どのリズムパターンと一致するかを判定し，対応した機能を動かすことが可能になる．近年着々と増えているスマート家電に着目し，本ライブラリを用いて家電制御システムを試作した．

キーワード：リズムパターン認識，スマート家電制御

1. はじめに

近年では，スマート家電，カーナビ，ゲーム機器など，様々な機器にインターネット機能を備えている．それに伴い，スマートフォンと連携することが可能な制御アプリが多く発売されている．しかし，それらのアプリでは一つの対応した機器しか制御できず，複数の機器を制御するにはその数と同じアプリが必要になる．また，多くの機能を備えている制御アプリは，使用者によっては余分な機能もあり，電源のオン/オフなどの単純な操作は，手軽にコントロールできることが望ましい．

そこで，本稿では使いたい機器の機能を自ら登録でき，簡単な操作が可能なリズムパターン認識ライブラリを提案する．これを用いることによって，様々な機器の使いたい機能を一つのアプリで制御することが可能になる．

リズムパターン認識とは，簡単に述べると音声認識の言葉がリズムに変わったものである．スマートフォンに内蔵されているマイクを用いて，スマートフォン自体や周辺の机をタップ，または手拍子などの音響信号を検出し，タップされたタイミングからどのようなリズムパターンが叩かれているのかを判別し，対応する機能を動かすことが可能である[1]．このリズムパターン認識制御は，家電制御，鍵の開閉など様々なものに応用できると考えられる．リズムパターン認識のメリットとして，操作が簡単という点が挙げられる．テレビを例に取ると，よく使う機能（電源オン/オフ，選局，音量調整など）の操作のリズムパターンを予め登録しておき，実際に使うシーンではスマートフォン周辺の机をタップするだけでリモコンと同じ操作が可能になる．また，照明のオン/オフをリズムパターン制御で行えば，ドア付近にあるスイッチをわざわざ押しにいかずとも，どこからでも切り替えが可能になる．

2. 関連研究

類似の技術として，音声認識が挙げられる．音声認識とは，人の「声」を「文字」に変換する仕組みのことである．今日の音声認識技術は非常に高く，スマートフォン，家電，カーナビなど様々な機器に機能を搭載している．音声認識のメリットとして，手を使わずに操作することができるという点が挙げられる．両手が塞がっている人，体を動かすことが難しい人など，通常の操作を行うことが困難な状況でも，キーボードの代替手段として簡単に使用することができる[2][3][4]．しかし，使用する環境によっては使いづらい場合もある．例えば，電車の中などの公共の場所では，音声認識での操作に抵抗がある人がほとんどであると思われる．また，周りがうるさい場所での使用も誤動作につながり不向きである．このような周りの環境に左右されてしまうデメリットを，リズムパターン認識では克服できると考えている．

また，現在リズムパターン制御機能を搭載した“Knocki”が発表されている[5]．Knocki とは，米国テキサス州のスタートアップ，Swan Solution が開発したスマートホームデバイスである．壁やドアやテーブルなどに取り付けて，表面をノックするだけで，Wi-Fi に接続した機器のリモコン代わりになり，電灯を点けたりテレビを消したりすることができる．最大 10 パターンのノックを認識する事ができ，専用アプリによってプログラムも可能である．Knocki はノックの感知には加速時計を用いており，叩く衝撃の違いを検出できるアルゴリズムに基づいて，偶然の音や振動による起動を避ける仕組みになっている．

本研究と Knocki の大きな違いは，音響信号の検出方法である．Knocki が加速時計で検出しているのに対して，本研究ではマイクを用いている．検出方法に基づく機能の違いを以下にまとめた（表 1）．

†1 愛知工業大学 情報科学部
†2 三菱電機エンジニアリング（株）

ⓒ2015 Information Processing Society of Japan

Vol.2015-MBL-77 No.6
Vol.2015-ITS-63 No.6
2015/12/2

表 1　Knocki とマイク検知の比較表

	Knocki	本研究(マイク検知)
簡単な操作で機器を動かせる	○	○
場所を選ばない	×	○
叩かれた物の素材の違いを検知できる	×	○
誤作動を防ぐことができる	○	○
デバイス一つで全ての機能に対応できる	△	○

2 つの大きな違いは，場所を選ばないという点と叩かれたものの素材を検知できるかどうかにあると考える．Knocki はノックするための壁や机が必要であるため，それらがある場所でないと使用することができない．したがって，使う場所が家の中などに限定されてしまう．対して，本研究ではマイクを使っているので，壁や机のない，空間での使用も可能である．また，マイク検知では，叩いた物の素材の違いを検出することができる．これにより，同じリズムパターンであっても，素材によって機能を変えることも可能になる．これは Knocki ではできない大きなメリットであるといえる．

3. リズムパターン判別アルゴリズム

本研究のリズムパターン認識ライブラリは，登録リズムパターンの録音，編集機能を備えたものを目指している．また，動かしたい機能とそれに対応するリズムパターンを複数登録することができ，スマートフォン 1 台で様々な機器を操作できるものにしたいと考えている．本稿ではそのアルゴリズムについて説明する．

3.1 リズムパターン判別の概要

実際にリズムパターンを叩いた音響信号と，検出した秒数を重ねあわせると，以下の図のようになる（図 1）．録音データから音圧を調べ，音圧が±10000 を超えている時の秒数と値を出力させ，その出力された秒数に着目し，前後の差が 0.01 秒以内であれば叩かれた回数を 1 回と見なしている．10000 という閾値は，自然な力でリズムを叩いた場合を考えたとき，最適な閾値であると考える．図 1 の場合，0.3 秒，1.1 秒，2.0 秒，2.8 秒の合計 4 回が叩かれていると検出された．

このような音響信号から，予め登録されたリズムパターンと入力されたリズムパターンを比較している．比較する際は，3 段階の処理で同じリズムパターンであるかどうかを判定している（図 2）．

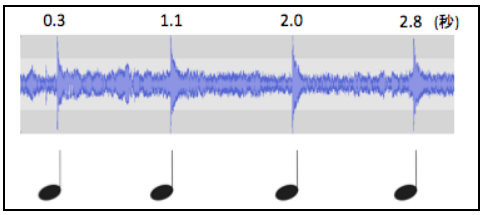


図 1 検出した音響信号とその秒数

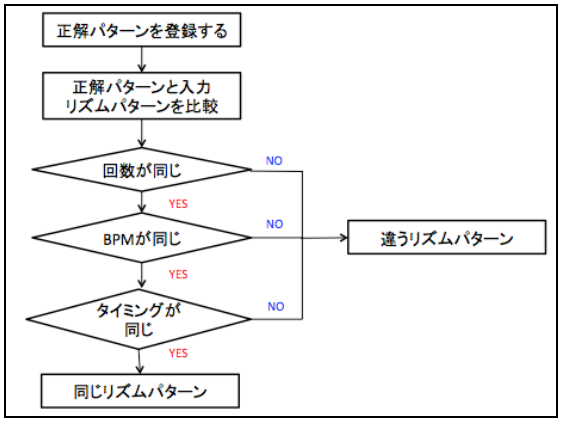


図 2 判別アルゴリズムのフローチャート

以下に 3 段階の判定処理方法の概要を述べる．

3.2 回数判定

この部分では，登録されたリズムパターンと入力されたリズムパターンの叩かれた回数を判定している．2 つのリズムパターンの叩かれた回数が異なる場合，その時点で違うリズムパターンであると判定する．同じであれば，次のステップに進む．

3.3 BPM 判定

この部分では，登録されたリズムパターンと入力されたリズムパターンの BPM にどの程度の差があるかを調べる．BPM 判定を行う前に，登録済みリズムパターンと入力されたリズムパターンの 1 回目に叩かれた秒数を，それぞれ 0 秒に合わせておく．BPM の差を出す式は以下の通りである．

・BPM の差 ＝ endTime1 / endTime2

endTime1 は登録済みリズムパターンの最後に叩かれた秒数，endTime2 は入力リズムパターンの最後に叩かれた秒数である．この計算式によって出された BPM の差が，設定された閾値の範囲内であれば同じ BPM で叩かれていると見なし，次のステップへ進む．

3.4 秒数判定

この部分では，元となる登録されたリズムパターンと，比較する入力されたリズムパターンのタップされたタイミ

ⓒ2015 Information Processing Society of Japan

Vol.2015-MBL-77 No.6
Vol.2015-ITS-63 No.6
2015/12/2

ングを検出し，そのタイミングにどの程度のズレが生じているかを調べる．まず，2 つのリズムパターンの最後に叩かれた秒数を，前途で述べた BPM の差を元にして揃える．秒数を揃えてから，1 回目と 2 回目，2 回目と 3 回目というように差を出していく．これによって出された秒数の差が，設定された閾値の範囲内であれば OK とし，すべて OK であれば同じタイミングで叩かれていると見なした．このように 3 段階の処理を経て，すべてクリアしたものを同じリズムパターンであると判定している．

4. 閾値設定のための実験

予め登録されたリズムパターンと入力するリズムパターンの BPM，叩かれるタイミングを完全に一致させるのは困難である．したがって，判定アルゴリズムの中に，どの程度の誤差まで同じリズムパターンと見なすかという閾値を設定している．BPM 判定及び秒数判定に使用する閾値を設定するための実験を行い，その結果に基づいて閾値の設定を行った．これらの実験の被験者は，実験 1 では 5 人中 2 人，実験 2 では 5 人中 1 人が楽器経験者であり，音楽熟練度が高いと思われるが，それ以外は未経験者であり，音楽熟練度は低いと思われる．

4.1 [実験 1] BPM 判定に用いる閾値の実験

まず，BPM の違いによってズレの大きさが変わるのではないかと考えたため[6]，3 パターンの BPM（遅い：BPM40 普通：BPM70　速い：BPM180）で同じリズムパターンを叩いてもらう．これを正解のリズムパターンとして録音しておく．次に，ベースとなる正解リズムパターンを叩いてから 20 分後，1 時間後，5 時間後，1 日後と時間を置いて，同じ BPM を意識して再度叩いてもらい，BPM にどの程度のズレが生じているかを調べた．この空ける時間は，エビングハウスの忘却曲線を元に設定した．エビングハウスの忘却曲線とは，ドイツの心理学者，ヘルマン・エビングハウスが意味のない 3 つのアルファベットの羅列（例：rit,pek,tas）を，被験者にたくさん覚えさせ，その記憶がどれくらいのスピードで忘れられていくかを実験し調べ，その結果をグラフ化したもののことである．この実験から，20 分後に 42%，1 時間後に 56%，1 日後に 74%，1 週間後に 77%，1 ヶ月後に 79%を忘れてしまうという結果が出た．この結果から，記憶は覚えた直後に半分近く忘れてしまうこと，残った記憶はゆっくり忘れていき，長く保持されるということがわかった[7].

4.2 結果と考察

実験結果はそれぞれ以下の通りである（表 2，3，4，図 3，4，5，6，7，8）．表の数字は全て小数点以下第四位切り上げの数字である．BPM の差を出す式に基づいた計算結果は，1.0 に近い数字ほど，近い BPM で叩かれている．

表 2　BPM40 の結果

BPM40	20 分後	1 時間後	5 時間後	1 日後
A	1.089	1.002	1.071	1.193
B	0.960	0.934	1.010	0.987
C	0.968	0.986	1.210	1.118
D	0.728	0.771	0.849	0.860
E	1.250	1.114	1.283	1.115
平均	0.100	0.961	1.085	1.11
標準偏差	0.171	0.112	0.153	0.118

表 3：BPM70 の結果

BPM70	20 分後	1 時間後	5 時間後	1 日後
A	1.203	1.060	1.246	1.051
B	0.772	0.916	0.858	0.880
C	1.339	1.058	1.475	1.286
D	1.045	1.195	1.360	1.361
E	0.850	0.857	0.916	0.994
平均	1.041	1.017	1.171	1.114
標準偏差	0.212	0.119	0.244	0.181

表 4：BPM180 の結果

BPM180	20 分後	1 時間後	5 時間後	1 日後
A	1.467	1.120	1.200	1.200
B	0.887	0.928	0.922	0.956
C	0.941	0.703	0.922	0.926
D	0.723	0.812	0.766	1.115
E	1.010	0.733	0.619	0.602
平均	1.006	0.859	0.886	0.960
標準偏差	0.249	0.152	0.193	0.205

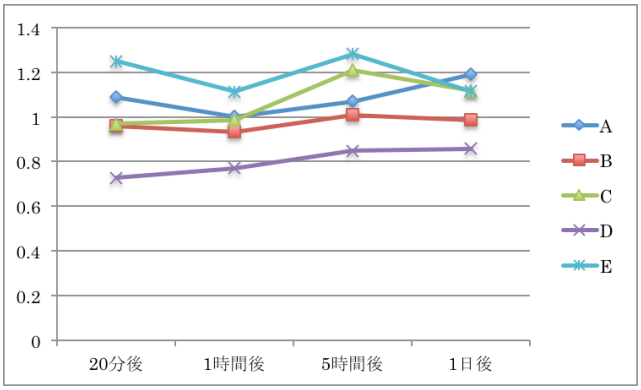


図 3 BPM40 の結果

ⓒ2015 Information Processing Society of Japan

Vol.2015-MBL-77 No.6
Vol.2015-ITS-63 No.6
2015/12/2

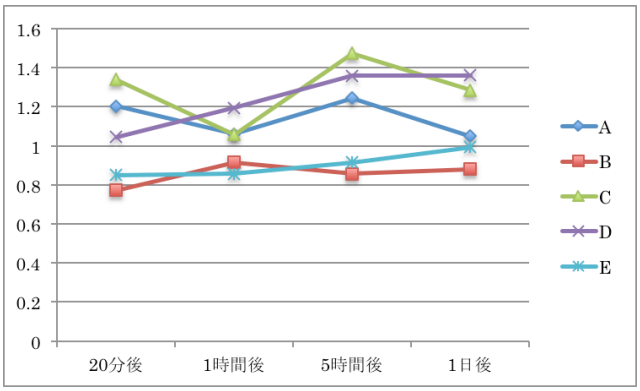


図 4 BPM70 の結果

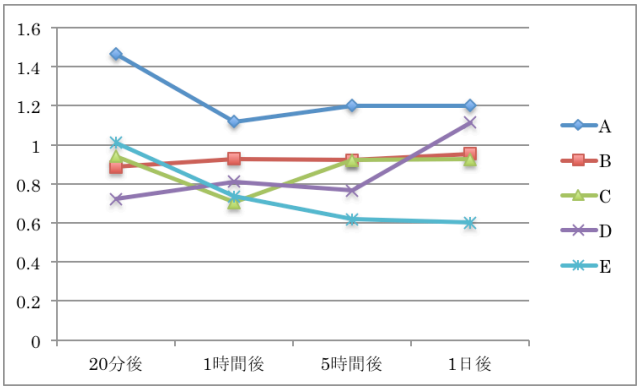


図 5 BPM180 の結果

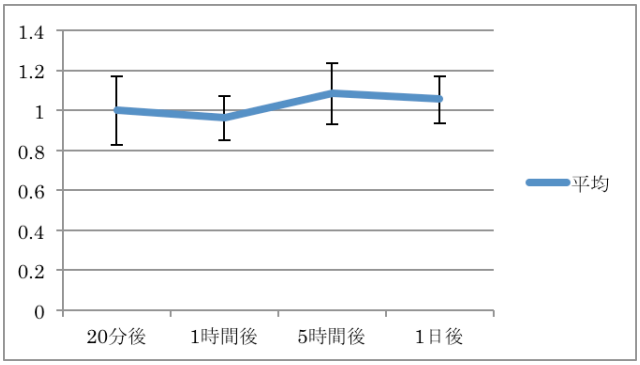


図 6 BPM40 の平均と標準偏差

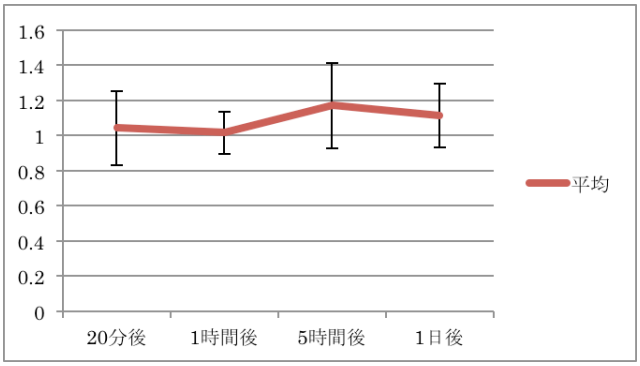


図 7 BPM70 の平均と標準偏差

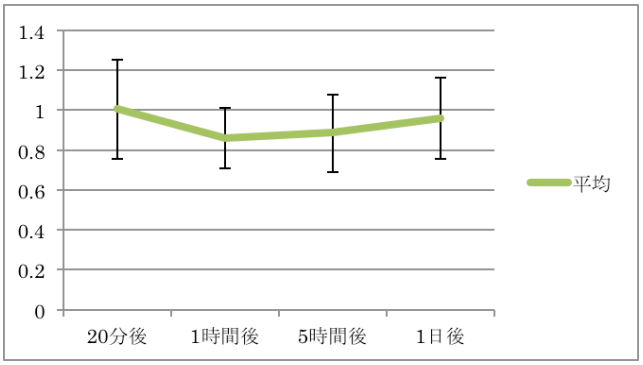


図 8 BPM180 の平均と標準偏差

実験の結果，実験を行った 15 回中 6 回が設定した BPM よりも全て速くなり，5 回が全て遅くなっていたことがわかった．これより，BPM の記憶は，エビングハウスの忘却曲線とはあまり深く関係しておらず，精神テンポによる影響の方が大きいのではないかと考えた．精神テンポとは，歩行や会話時の個人特有の快適なテンポのことである．精神テンポは個人差が大きいが，個人内の変動は少ない[8][9]．よって，自ら登録リズムパターンを録音する本システムでは，2 つのリズムパターンの BPM のズレは小さいと考えられる．したがって，BPM 判定を行う際は 0.8〜1.2 という閾値を設定し，この範囲内であれば同じ BPM で叩かれたと見なすようにした．

4.3 [実験 2] 秒数判定に用いる閾値の実験

予め正解パターンとして 5 つの異なるリズムパターンを録音しておく．5 つのリズムパターンは以下の図の通りである（図 9，10，11，12，13）．これらの 5 つのリズムパターンを設定した理由は，実際にシステムを利用する際には，入力が簡単な短いリズムパターンを使用することが多いことを考え，一小節程度の短いものに設定した．また，個人の音楽熟練度の違いを考慮し，5 段階の難易度を設定した．正解パターンを一度聞いて，それらのリズムパターンと同じように叩いてもらい，叩かれるタイミングのズレがどの程度生じていたのかを調べた．

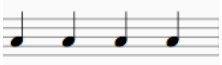


図 9 パターン 1

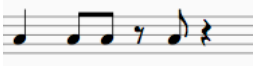


図 10 パターン 2

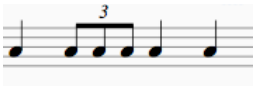


図 11 パターン 3

ⓒ2015 Information Processing Society of Japan

Vol.2015-MBL-77 No.6
Vol.2015-ITS-63 No.6
2015/12/2
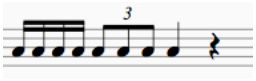


図 12 パターン 4
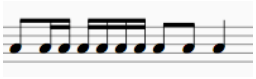


図 13 パターン 5

4.4 結果と考察

秒数判定を行い，すべて OK が出た閾値を調べた．実験結果は以下の表と図の通りである（表 5, 図 14，15）．

表 5 秒数判定ですべて OK が出た閾値

(秒)

	パターン1	パターン2	パターン3	パターン4	パターン5
A	0.4	0.5	0.2	0.4	0.3
B	0.2	0.4	0.1	0.2	0.4
C	0.1	0.4	0.5	0.3	0.4
D	0.2	0.2	0.2	0.1	0.1
E	0.3	0.2	0.3	0.2	0.4
平均	0.24	0.34	0.26	0.24	0.32
標準偏差	0.102	0.120	0.136	0.102	0.117

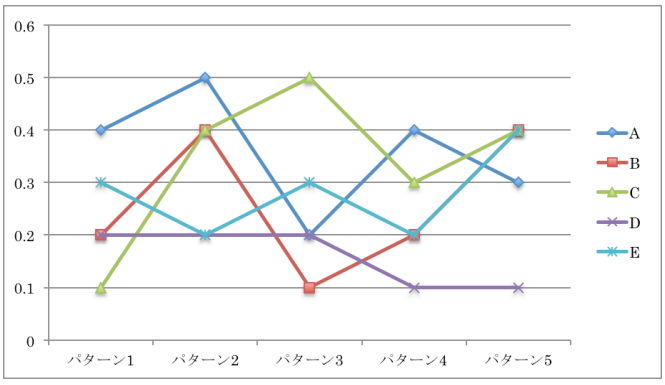


図 14 秒数判定の結果

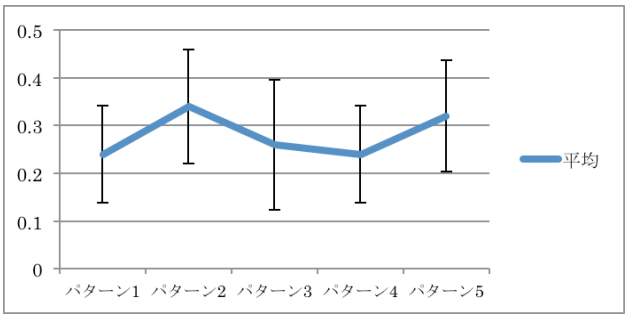


図 15 秒数判定の平均と標準偏差

同時に BPM 判定を行ったが，同じ BPM と見なしたものは，すべて OK が出た閾値が 0.1〜0.3 秒と小さく，違う BPM と見なしたものは，すべて OK が出た閾値が 0.4 秒以上と大きくなった．したがって，BPM 判定をクリアしたものの秒数判定の閾値は小さく設定してよいと考える．よって，閾値を 0.3 秒と設定し，この範囲内であれば同じタイミングで叩かれたと見なすようにした．

この閾値を設定した際，実験データの中で 68%が秒数判定の条件をクリアすることができた．そして BPM 判定と秒数判定を共にクリアしたのは全体の 60%であった．これより，意図して同じリズムパターンを叩いていても，残りの 40%はズレが生じていることがわかった．したがって，リズムパターンの記憶には個人差があり，実際に使用する場面ではうまく認識されない可能性もあると考える．解決策として，登録時に複数回同じリズムパターンを叩き練習させることで，認識率の向上に繋がるのではないかと思われる．また，セキュリティ上問題のない機器を操作するリズムパターンであれば，いつでも登録済みの録音データを聞くことができるような機能を付けることも一つの方法である．

5. 家電制御への応用

リズムパターン認識を家電制御に応用することを想定し，“Philips hue”を用いてリズムパターン認識による制御アプリを試作した．Philips hue とは，1,600 万色を超える色彩表現が可能な LED 電球のことである．iOS，MacOS X，Android 用の SDK が用意されており，スマートフォンアプリから制御することが可能である[10].

このアプリは，Android 端末で利用することができる．録音のためにスタート・ストップボタンを押して，そこで得られた音響信号に対して提案手法の処理でリズムパターンを判定し制御している．録音によって登録されたリズムパターンが入力リズムパターンと一致しているとき，hue の色がランダムで変化する機能を備えている(図 16, 17)．



図 16 アプリ動作前

ⓒ2015 Information Processing Society of Japan

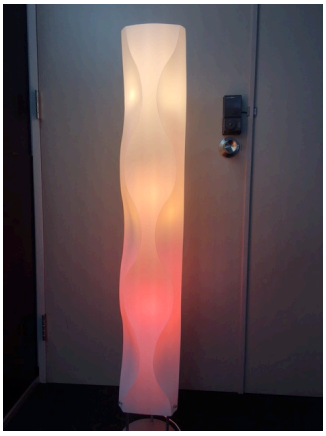


図 17 アプリ動作後

6. おわりに

本稿では，スマートフォンにおけるリズムパターン認識ライブラリに向けての実装と検証を行った．現在試作したアプリは，1 つの機能しか備えていないため，ライブラリというには不十分である．したがって，複数のリズムパターンを登録できるようプログラムを改良し，対応する機能を増やしてより汎用性の高いシステムを作る必要がある．また，最終的にはバックグラウンドでアプリを起動させておき，ボタン操作をすることなくリズムパターン認識のみでの操作を目指している．よって今後の課題として，連続的に録音している音響信号から，リズムの入力開始/終了を自動的に切り出し，リズムパターン判定を行えるようにする必要がある．

参考文献

1) 大槻知史，齋藤直樹，中井満，下平博，嵯峨山茂樹：隠れマルコフモデルによる音楽リズムの認識，情報処理学会論文誌，Vol.43，No.2，pp.245-255，2002.
2) 嵯峨山茂樹，武田晴登，亀岡弘和，西本卓也：音声認識技術を用いた音楽情報処理，日本音響学会誌 Vol.61，No.8，pp.454-460，2005.
3) 古井貞熙，小林哲則，矢頭隆：音声認識実用化技術の展開，電子情報通信学会誌 Vol.93，No.8，pp.725−740，2010.
4) 山出慎吾，馬場朗，芳澤伸一，李晃伸，猿渡洋，鹿野清宏：実環境における頑健な音声認識のための音韻モデルの教師なし話者適応，電子情報通信，学会論文誌，Vol.J87−D−II，No.4，pp.933−941，2004.
5) Make Anything a Remote Control – Knocki
https://knocki.com
6) 荒生弘史：リズム・キープ感と時間知覚，The JpaneseJournal of Psychonomic Science 2010，Vol.28，No.2，pp.249−255，2010.
7) H.Ebbinghaus,，Über das Gedächtnis，1885，On Memory：A Contribution to Experimental Psychology，Translated by Henry A. Ruger & Clara E. Bussenius，1913，originally published by Teachers College，Columbia University.
8) 杉之原正純，相浦義郎，松田俊，平伸二：精神テンポに関する基礎的研究(7)，広島修大論集，第 23 巻 第 2 号 pp.119-134，1982.
9) 堀田晴子，澤村貫太，井上健：被験者の心拍数に応じたテンポによる音楽聴取時の心拍変動について，臨床教育心理学研究 Vol.33，No.1，pp.1-8，2007.
10) フィリップス エレクトロニクス ジャパン「Philips hue」
http://www2.meethue.com/ja-JP

ⓒ2015 Information Processing Society of Japan