情報処理学会研究報告
IPSJ SIG Technical Report

Vol.2018-MUS-119 No.56
Vol.2018-SLP-122 No.56
2018/6/17

物を叩く・擦る際の音響信号を活用した実世界向けユーザインタフェースの研究

平井 重行†

概要：音響信号処理を用いた HCI 研究として，人が物を叩いたり，擦ったりした際に鳴る音に着目して，物に取り付けたコンタクトマイクを通した音響信号からイベント検出を行い，様々な用途に活用できるユーザインタフェースの研究を行ってきた．叩く音については，叩く位置や叩き方（音色），叩くパターンを区別でき，様々な用途に利用できるシステムとして構築してある．また，擦り音については水場で活用できるユーザインタフェースとして，擦る指の本数や擦る時間長，往復した擦り動作などの検出ができるシステムとして構築してある．本報告においては，それぞれのイベント検出手法と性能，スマートハウスにおける生活支援や音楽アプリケーションなどの応用例について紹介する．

キーワード：ユーザインタフェース，スマート環境，叩打音，擦り音，信号処理, NMF

1. はじめに

Internet of Things (IoT) の機能を持つ家庭用スマートデバイスが多数登場し，住宅内の情報化が多種多様に進展しつつある．一方で，家具や調度品，住宅設備にコンピュータを組み込んでインタラクティブな住環境とする研究が以前から様々に行われている[1-5]．そこで利用されるユーザインタフェース（以降，UI と表記）は，カメラ画像入力を用いるものや，赤外線センサ，超音波センサ，マイクを利用するものなど，様々である．ただし，浴室やその前室となる洗面脱衣室，トイレなどでは，衣服を脱いだ状態でいることがあるため，プライバシーの観点からカメラ画像を元にしたジェスチャ入力の UI 技術などは適用し辛い事情がある．一方で，近頃では AI スピーカーを用いた音声入力インタフェースが普及しつつあり，カメラ画像を利用せずとも機器やサービスの操作などが可能になってきている．しかし，音声入力インタフェースも，調光照明で明るさを微調整する場合などで，操作により時間がかかることもあり，万能とは言い難い．このように，実世界 UI は，スマートフォンや PC のような端末の閉じた環境で特定の入力手段によって操作が完了するものとは事情が違い，UI は場所や場面，状況，目的によって適した手法・手段があると言える．

筆者らは，このような実世界向け UI に対し，非音声を対象に音響センシングを用いる幾つかの手法について研究している．それらは，浴槽や洗面台，キッチンテーブルなどの住宅設備の内側や表面裏にピエゾセンサ（コンタクトマイク）を設置し，それらの表面を手指で叩いたり，擦ったりした際に発する音を機器やサービスなどの操作手段として利用するものである．これらの手法では，ピエゾセンサを設置した物体の固体振動を計測するため，その周囲の空気中で鳴っている音はほとんど入力されず，SNR が比較的良く音響入力ができる利点がある．筆者らは，このピエゾセンサを設置した場所において，その表面を手指で叩いた際の「叩打音」や，湯水のある場所でこすった際になる「キュッ」という「擦り音」を，検出・識別などすることで，様々な UI を実現している．叩き音については，特に浴槽の縁を操作場所とする RapTapBath [15]として，また擦り音については，キッチンや洗面台，浴室などの湯水で濡れている滑らかな面で利用する Rubbinput [18]として研究している．

本報告においては，これら RapTapBath，および Rubbinput について，そのインタラクション手法と技術概要，システム構成，スマートハウスにおける生活支援応用例について紹介する．

2. RapTapBath：叩打音利用インタフェース

2.1 RapTapBath のインタラクション手法概要

RapTapBath は，浴槽縁の表面を，手指で叩くと鳴る叩打音から，その「叩打位置」と，叩き方によって鳴る「叩打音色」，連続的に叩いた場合の打数やリズムを含めた「叩打パターン」の 3 つの情報を抽出して，様々な操作手段を実現するシステムである（図 1 参照）．

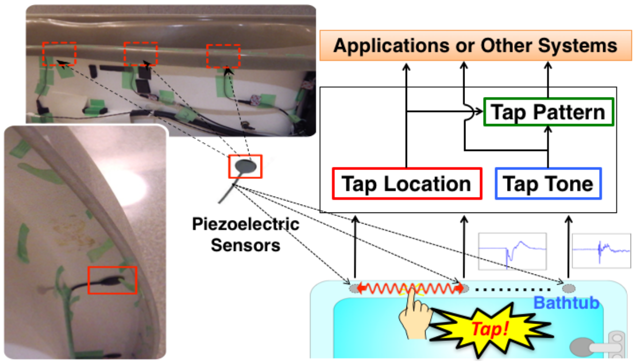

図 1 RapTapBath のシステム概要

† 京都産業大学 情報理工学部
Faculty of Information Science and Engineering, Kyoto Sangyo University

2.2 RapTapBath システム概要

日本国内における一般的なシステムバスの多くは浴槽の横壁（エプロン壁）が開く構造となっており，オプション機器が後付け可能である．図１左側の写真は，エプロン壁を外してピエゾセンサを設置した様子である．本研究では，これを利用してピエゾセンサを浴槽内部へ後付けすることを前提としている．これにより，浴槽表面を叩いた音響を浴槽の固体振動として計測し，その音響信号を元に，各種信号処理などを施して，叩打位置検出，叩打音色識別，叩打パターン識別をリアルタイムで行う．これらの処理結果は個別にもしくは組み合わせたイベントとして，機器操作やアプリケーションに利用する．なお，浴槽縁上の表示は，天井裏に設置したプロジェクタから投影する．天井に浴室乾燥機が設置されている浴室では，天井裏に空きスペースが十分あり，プロジェクタ設置が容易である．

2.3 RapTapBath での叩打音処理手法

2.3.1 叩打位置検出

位置検出は，浴槽裏側に複数のピエゾセンサを一定間隔で設置して，叩打音による振動到達時間差から算出する．信号到達時間差で位置検出を行うためには，到達時間を正確に検出する必要がある．ここでは，Josephら[7]と同様に移動実効値法（Running Root-Mean-Square Method）を用いて，波形そのものではなく振動エネルギー（二乗振幅の一定時間窓での積分値）が設定した閾値を超えた時刻を信号の初動時刻として検出する．現時点では，30cm間隔で設置したピエゾセンサによって，10cm間隔での位置検出が行えている [13].

2.3.2 叩打音識別

図2に示す5種類（Fist, Knuckle, Pad, Tip, Fingers）の叩き方について，パワースペクトルを元に識別する．

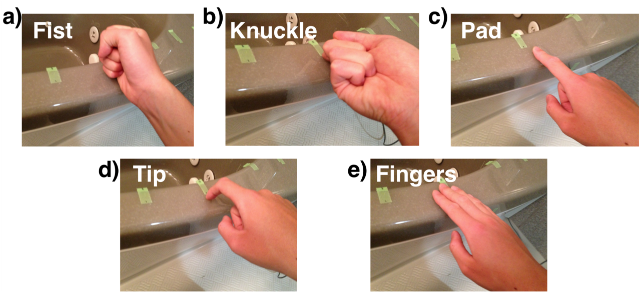

図2　浴槽縁上での叩き方

a) 拳（Fist），b) 指関節（Knuckle），c) 指先の腹（Pad），d) 爪（Tip），e) 複数本指の腹（Fingers）

ここでは，非負値行列因子分解（以下NMFと表記）を用いる手法と，Newral Network（以下NNと表記）を用いる手法の2つの手法を実装している．浴槽縁の上での叩打音は，高周波成分が比較的少なく，低周波成分が大きい傾向がある．しかし，いくつかの叩き方（PadやTip，Fingers）の信号には低周波成分と比較すると弱いが，識別処理において特徴量になりうる高周波成分が含まれている．このことから，NMFを用いる手法では，前処理としてカットオフ周波数を5kHz に設定したハイパスフィルタ処理を施し，低周波成分を削りそれぞれの叩打音色の特徴的な高周波成分を強調することにより，識別精度の向上を行っている．

一方で，NNについては，全結合型で中間層10層のNNを用いており，入力データにはハイパスフィルタ処理を行わず，生波形から得られる振幅スペクトルとそのパワー成分，RMS値，零交差点の特徴を採用している．

識別精度の評価については，5人の被験者によって5種類の叩き方それぞれ100回ずつ浴槽縁を叩打してもらい，そのデータを用いた．個人毎のデータを分けて評価したものを図3に，5人のデータをまとめて評価したものを図4に示す．ここでは，5種類の叩き方に対し，2種類だけ利用する組み合わせ（10通り），3種類だけ利用する組み合わせ（10通り），4種類だけ利用する組み合わせ（5通り），5種類すべてを対象とした場合について，識別する種類の数毎に5分割交差検証の結果を求め，グラフに表している．なお，先行研究として示してあるのは，参考文献[14]に載せたものである．これらから， いずれのデータ区別，叩打音組合せでも，上記の構成のNNで処理することで十分な精度が得られることが確認できた．

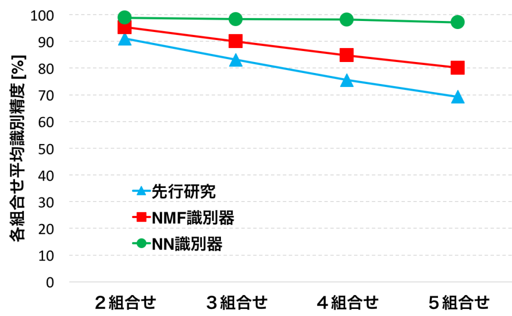

図3　個人毎のデータでの叩打音色識別精度

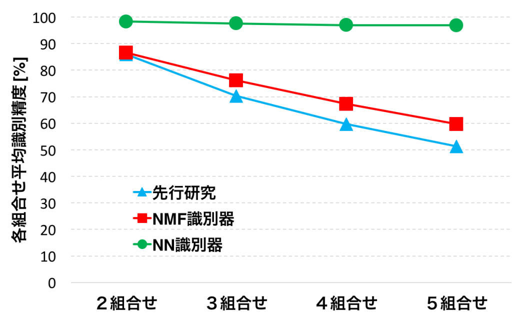

図4　複数人のデータでの叩打音色識別精度

2.3.3 叩打パターン認識

叩打パターンは，一定時間内の叩打音の連打数を元に検出する．ここでは,「タタン」（2連打），「タタタン」（3連打）のような単純な連打だけでなく，「タンタタ」（以

下●・●●と表記）や「タタンタ」（●●・●）のような特定リズムパターンも入力イベントとして認識し，利用できるようにしている．ここでは，叩打音色識別処理から出力される叩打音色IDを用いる．パターンの受付時間，最大叩打回数を設定しておく．そして，叩打音識別処理から送られる叩打イベント（ID）と次の叩打イベントとの間の時間間隔を記録する．最初の叩打から受付時間を超えた場合，もしくは最大叩打回数を超えた場合に，それまで受け付けた叩打を1つのパターンとしてみなしてパターン認識処理を行う（図５参照）．そして，記録した叩打間の時間差の合算値で各時間差を正規化し，その正規化時間差を用いて確率密度関数によってパターン推定率を求める. 3叩打の場合の処理概念図を図6に示す. 4叩打以上のパターンについては，楽器演奏経験がない人の多くが叩き分けられないことが予備実験から判明したため，現状のシステムの実装では3叩打までの処理に留めている．しかし，この手法は4叩打以上のパターンでも適用できる. 3叩打のパターンに対する識別の被験者実験では100%の精度が出ている．

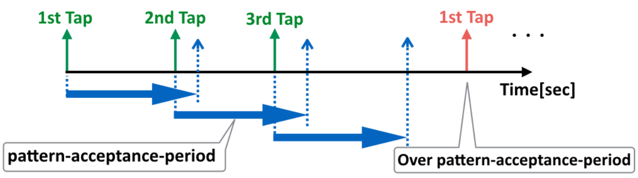

図5　パターン検出時の時間処理

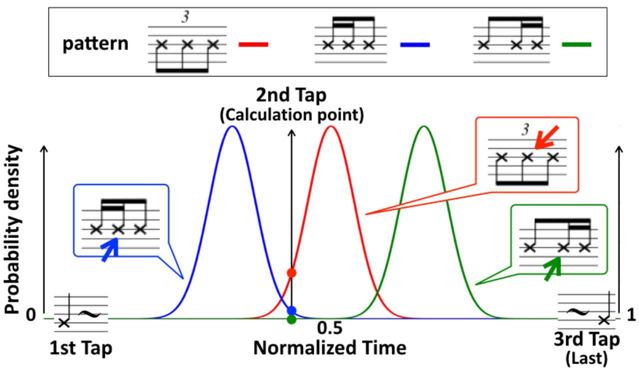

図6　3叩打時のパターン識別処理概念図

2.4 RapTapBath のアプリケーション

前節で述べた，個別の操作イベントを単体もしくは組合せて様々な操作を実現することができる. RapTapBath システムとしては，浴槽縁上でアプリケーションの切り替えや選択ができる主メニュー操作インタフェースを構築してある．そのメニューから様々なアプリケーションを選択・操作する．具体的なアプリケーションとしては，給湯機の給湯温度設定，照明操作，音楽プレイヤーなどの機器操作を叩打音でも行えるような浴室機器の制御を行うものがまず挙げられる．

また，日本の文化では，浴室はリラックスする場，また親子のコミュニケーションの場, 半身浴で読書する場など，様々な側面を持った空間と言えることから，個人もしくは複数人で楽しめるものも含め，図７に示すような様々なエンタテインメントアプリケーションも作成した．

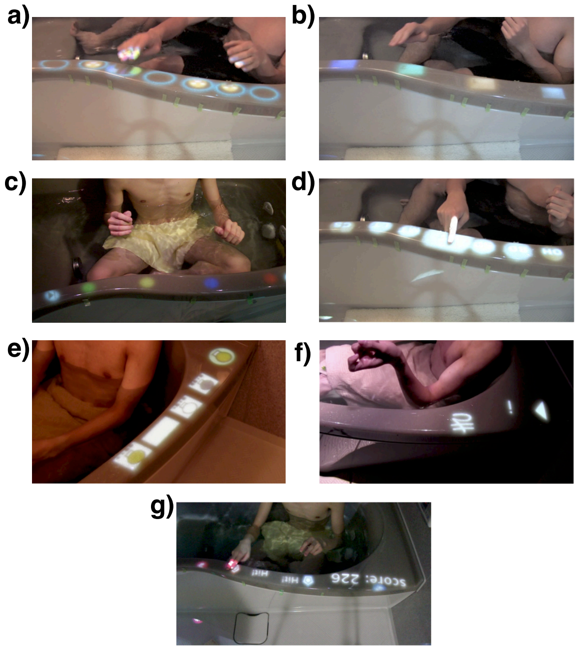

図７　RapTapBath のアプリケーション例
a)　Whac-A-Mole，b)　クイズ早押しボタン，
c)記憶ゲーム Simon，d)　打楽器 BathDrum2，
e)　照明制御，f)　テキスト入力，g)　音楽ゲーム

3. Rubbinput：擦り音利用インタフェース

3.1 Rubbinput でのインタラクション手法概要

Rubbinput は，水がある滑らかな面を指で擦ると鳴る「キュッ」という音を検出し，その鳴らし方に関する情報を抽出して，様々な操作手段を提供する UI である．擦り音を用いる UI には Bathcratch[13]があるが，そこでは擦り音検出の厳密な信号解析はせず，基本周波数の検出とその時間連続性により，擦り音か否かの判定をしていた．Rubbinput では，擦り音としての検出手法から見直しており，かつ一つの擦り音から「時間長」「指本数」「往復動作」の 3 つの情報を抽出し，操作イベントとして利用する（図 8 参照）．

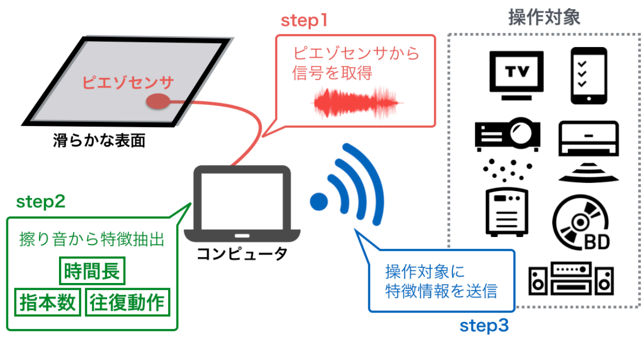

図 8　Rubbinput システム概要

なお，指本数識別については，現状では「一本指」か「複数本指」かの識別に留めている．その理由は，擦る指本数が増えるにつれ，全ての指で擦り音を正確に鳴らすことが困難であることが予備実験により確認できている．また複数本指を擦る際に力を入れやすい（擦りやすい）本数は指の長さや形状に起因して人によって異なることも理由である．図 9 に 1 本指と複数本指の場合のそれぞれのイメージ図とスペクトログラムを示す．

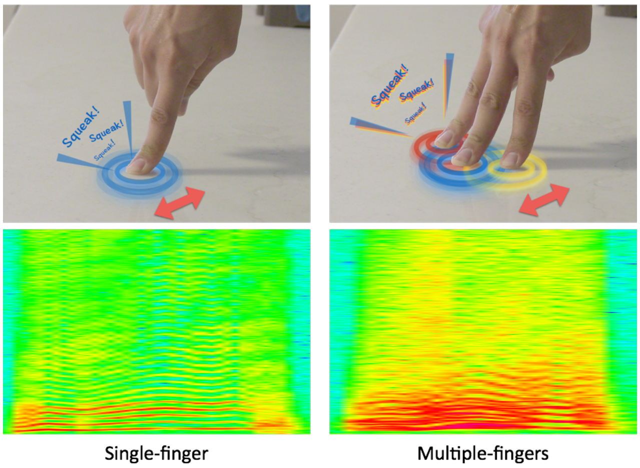

図 9　擦り音のスペクトログラム;
左) 1 本指，右)複数本指.

3.2 Rubbinput システム構成

浴槽の縁の裏や，キッチンの天板の裏などにピエゾセンサを設置し，そこで得られる音響信号を 16bit，44.1kHz サンプリングのデータとしてコンピュータへ入力する．その信号処理については，現時点ではフレーム幅 1024[sample]（約 23.2[msec]），フレーム移動幅 256[sample]（約 5.8[msec]）の移動窓で解析して操作イベントに変換する．窓関数には Hamming 窓を採用している．

3.3 擦り音の検出と各種情報抽出の概要

Rubbinputでの擦り音検出は，RapTapBathと同様に，RMS値を用いて図10のような3つの検出処理を行い，2段階（擦り音の開始と終了）のタイミング検知によって構成される．擦り音開始検知時においては，3つの検出結果を通知することもあり，ここまでに3つの検出処理を終える必要がある．処理開始のタイミングはRMSの変化値に着目し，急激な変化が得られた瞬間としている．これは様々な要因，例えば同じ場所にセンサを設置しても擦る場所（距離）で音圧レベルは大きく変化する．得られる擦り音の入力音圧レベルは常に一定とは言えないため，変化量を採用している．

ここでは3つの検出処理を行い，2段階（擦り音の処理開始から擦り音開始通知までの間に「擦り音検知処理」および「指本数識別処理」を完了し，「往復処理識別処理」は以前の擦り音から処理開始までに完了する．また傾向としてこの開始通知までの時間を長くとることで各検出処理の精度が高くなるが，反面レイテンシが長くなるため操作時の違和感が大きくなる．つまり精度とレイテンシはトレードオフの関係にある．そのため開始通知を適切な時間を設定することが重要となっている．

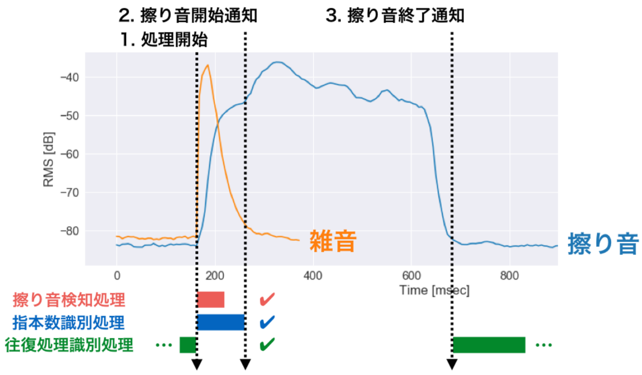

図10　擦り音のRMS遷移での各処理と結果出力のタイミング

3.4 擦り音からの各情報の抽出

3.4.1 時間長

擦り音の時間長は，RMS 値の励起タイミング（図 10 の 2）と，立ち下がりタイミング（図 10 の 3）の時間差を擦り音の時間長として扱う．

3.4.2 指本数

指本数識別は，擦った各指から個別の基本周波数 FO とその調波構造による音が発生することを前提にしている．複数本指の場合は，それぞれの指が FO と調波構造を構成するため，図 9 の右側に示すような複数の調波構造がまばらに重なった複雑なスペクトル成分が観測される．Rubbinput では，RMS の値から擦り音と検出された区間の特定のタイミングからパワースペクトルの有意なピークを検出する．そして，その最低周波数のピークを基準に各ピークを正規化した周波数で換算して，その分布を計算するスペクトルの複雑さを表現する調波構造評価関数を定義した．この関数の出力値を元に，閾値処理によって一本指か複数本指かを判断する．定義した関数の処理フローを図 11 に示す．

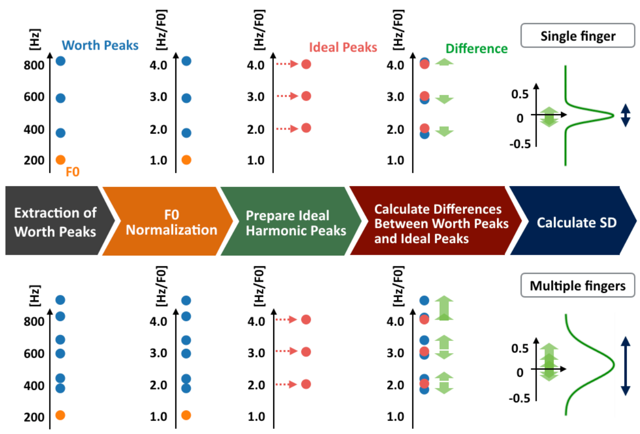

図 11　調波構造評価関数の処理フロー
上）一本指の場合の例，下）複数本指の場合の例

3.4.3 往復動作識別

往復動作の識別には以前の擦り音が鳴り終わってから再び擦り音の励起を検知するまでの無音時間長によって判断する．図12に，5人の被験者による，同一方向に連続的に擦った場合と往復動作で擦った場合の50試行の無音時間長の分布を図示する．これから，すべての被験者において，往復動作と同一方向動作とで有意な差が見られることが分かる．この無音時間長を各自適切な閾値を設けることで識別が可能である．この結果を元に閾値を180[msec]と設定して，それ以下の無音区間であれば2つの擦り音は往復動作であると判断して操作イベントとして用いる．

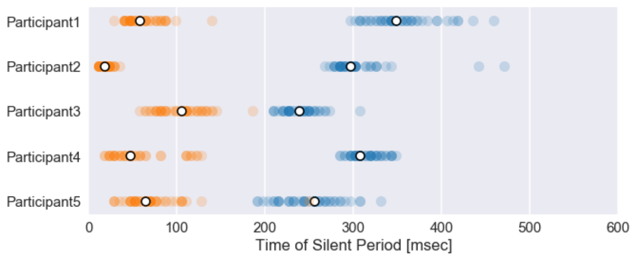

図 12　往復動作（オレンジ）と同一方向（青）の無音時間長分布

3.5 Rubbinput の性能評価

3.5.1 擦り音検知と雑音耐性の評価

擦り音は，水分と滑らかな面が必要だが，面の材質や表面の滑らかさ，また人の指先の皮膚の様子や力加減によっても擦り音の鳴り方が変わる．ここでは，擦り音の開始を検知してから12フレーム（約87.07 [msec]）とし，擦り音検出の閾値を0.3と設定した場合の結果を表1に示す．

表 1　場所や材質の違いによる擦り音検知性能

Environments	Material	Number of Data	Count of detected	Accuracy
Bath	Artificial marble	300	298	99.33%
Kitchen	Artificial marble	300	299	99.67%
Miror	Glass	300	300	100.00%
Small-env	Glass	354	354	100.00%
Window-glass (Long)	Glass	300	300	100.00%
Window-glass (Middle)	Glass	313	308	98.40%
Window-glass (Short)	Glass	314	314	100.00%
All	-	2181	2173	99.63%

また，上記と同じ設定で，擦り音以外の音（雑音）を鳴らした場合の検出についても確認を行った．ここでは，手のひらで叩いた音，握りこぶしで叩いた音，硬いもので叩いた時の音，調理中の様子を1時間半収録した音を雑音として用いた．

表 2　雑音識別性能

Noise type	Number of detected	False detection count	Accuracy
Cooking	857	21	97.55%
Hard Object	32	0	100.00%
Body	32	0	100.00%
Hand	30	0	100.00%
All	951	21	97.79%

3.5.2 指本数識別の性能評価

3.4.2節で説明した調波構造評価関数の値を元にした指本数識別の性能評価を行った結果を表3に示す．ここではRMS励起が安定してから14フレーム（約98.68[msec]）から最後の6 フレーム（約52.24[msec]）を検証に利用した．

表 1　指本数識別性能

Environments	Material	1-finger	2-fingers	3-fingers
			Multiple-fingers	
Bath	Artificial marble	96%	92%	89%
Kitchen	Artificial marble	95%	95%	99%
Miror	Glass	84%	81%	100%
Small-env	Glass	97%	100%	100%
Window-glass (Long)	Glass	99%	100%	99%
Window-glass (Middle)	Glass	87%	94%	100%
Window-glass (Short)	Glass	93%	90%	86%

3.6 Rubbinput アプリケーション

ここではRubbinputのUIを活用したアプリケーションとして，次の三つを示す．

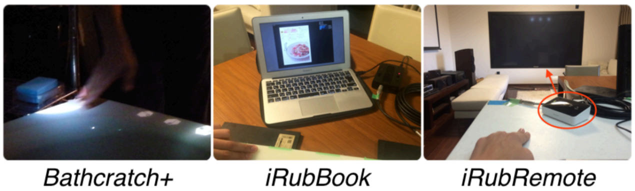

図 13　Rubbinput アプリケーション例.

3.6.1 Bathcratch+

Bathcratch+は，擦り音を利用する先行研究 Bathcratch に対し，指本数識別の機能を組み込んだ DJ スクラッチング演奏を楽しむエンターテイメントシステムである．擦る指の本数で，出力するスクラッチサウンドが切り替わることで新たな演奏の楽しみ方が可能となっている．

3.6.2 iRubBook

キッチンなどに導入する例として作成した擦り音を用いたブックリーダーアプリである．調理中などで手が水に濡れた状態や汚れている状態でも擦る指本数によってページ送り戻しやスクロールが可能である．また，マウスのダブルクリックのような連続操作にも対応しており，連続で擦った回数と指本数によって読む本の切り替えが行える．

3.6.3 iRubRemote

家電のスマート化やスマートデバイスの登場など，様々な機器がネットワークで接続され，リモートコントロールできる状況にある．ここでは高機能学習型の赤外線リモコン機器と擦り音識別を連携させ，様々な機器を操作するシステムを構築した．例えば，キッチンで調理中にキッチンカウンターを擦り，離れた位置にあるテレビのチャンネルを切り替えたり，BGMの音量を変更したりすることができる．

4. おわりに

本稿では，スマートハウス環境などで利用可能で，人が周囲の環境を用いて任意に鳴らすことができる音響イベントを機器操作等に利用できるユーザインタフェース 2 種類について述べた．叩打音を用いる RapTapBath については，音響信号処理に加えて NMF や NN などの機械学習手法を適用したものとなっている．浴槽に適用することを前提としてシステム化が図られているものの，操作イベントとして利用する内容は，浴槽以外の場所でも利用できるものとなっている．一方で擦り音を用いる Rubbinput については，機械学習は用いずに，擦り音の特徴や人の動作の特徴を踏まえた内容を音響信号処理技術のみで実現している．さらに，Rubbinput は Raspberry Pi 3 でシステム構成できており，小型の箱としてキッチンや洗面台など様々な場所へ取り付けて利用可能なである．

これらのシステムは，周囲の環境に直接触れて人が意図的に何かの音を鳴らすことで機能するインタラクション手法である．音声認識技術で言語によって機器操作するものと比べて，値を微調整することや，繰り返し操作したい単純な物事に対してはより短時間で実行・操作が可能なものになると考えている．今後はそれら音声インタフェースとの比較を行う予定である．また，人が任意に行える別の行為による音響イベントを，機器操作に活用する手法を提案・実装し，非言語音を活用するユーザインタフェースについて推進していきたい．

参考文献

[1] 椎尾一郎, Jim Rowan, Elizabeth Mynatt: Digital Decor : 日用品コンピューティング, ヒューマンインタフェース学会論文誌 Vol.5 no.3 p.323-330, 2003.
[2] 上田博唯, 山崎達也: ユビキタスホーム: 日常生活支援のための住環境知能化の試み, ロボット学会論文誌, Vol.25, pp.10-16, 2007.
[3] 平井重行, 上田博唯: 実験住宅Ξ Home(くすぃーほーむ)でのユーザエクスペリエンス研究へ向けて, 計測自動制御学会 SI2011 講演論文集, 2011.
[4] 榊原吉伸, 林宏憲, 平井重行: TubTouch: 湯水の影響や自由形状への適用を考慮した浴槽タッチ UI 環境, 情報処理学会論文集 Vol.54, No.4, pp.1538-1550, 2013.
[5] Shigeyuki Hirai, Yoshinobu Sakakibara and Hironori Hayashi. 2013. Enabling Interactive Bathroom Entertainment Using Embedded Touch Sensors in the Bathtub, In Proc. of ACE2013, 544-547, 2013
[6] Chris Harrison, Scott E. Hudson. Scratch input: creating large, inexpensive, unpowered and mobile finger input surface, Proc. of UIST'08, pp205-208, 2008.
[7] Joseph A. Paradiso, Che King Leo, Nisha Checka, Kaijen Hsiao. Passive acoustic sensing for tracking knocks atop large interactive display, Proc. of the 2002 IEEE International Conference on Sensors, pp.512-527, 2002.
[8] Hiroshi Ishii, Craig Wisneski, Julian Orbanes, Ben Chun and Joe Paradiso. PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play, Proc. of CHI'99, pp.394-401, 1999.
[9] Chris Harrison, Desney Tan, Dan Morris. Skinput: appropriating the body as an input computing systems, Proc. of CHI2010, pp.453-462, 2010.
[10] Chris Harrison, Julia Schwarz and Scott E. Hudson. TapSense: Enhancing Finger Interaction on Touch Surfaces. Proc. of UIST'11, pp.627-636, 2011.
[11] Pedro Lopes, Ricardo Jota, Joaquim A. Jorge. Augmenting touch interaction through acoustic sensing, Proc. of ITS2011, pp.53-56, 2011.
[12] Kazuhiko Yamamoto. Possessing Drums: An Interface of Musical Instruments that Assigns Arbitrary Timbres to Personal Belongings, Journal of Information Processing, Vol.21, No.2, pp.274-282, 2013.
[13] Shigeyuki Hirai and Daiki Ito, Entertainment Applications for Tapping on a Bathtub Edge Using Embedded Acoustic Sensors, Proc. of ACE2015, 2015.
[14] Tomoyuki Sumida and Shigeyuki Hirai, BathDrum2: Percussion Instruments on a Bathtub Edge with Low Latency Tap Tone Identification, Proc. of ACE2016, 2016.
[15] Tomoyuki Sumida, Shigeyuki Hirai, Daiki Ito and Ryosuke Kawakatsu, RapTapBath: User Interface System by Tapping on a Bathtub Edge Utilizing Embedded Acoustic Sensors, Proc of ISS2017, pp.181-190, 2017.
[16] Shigeyuki Hirai, Yoshinobu Sakakibara and Seiho Hayakawa. 2012. Bathcratch: Touch and Sound-Based DJ Controller Implemented on a Bathtub, In Proc. of ACE2012, 44-56.
[17] Ryosuke Kawakatsu and Shigeyuki Hirai, Interaction Technique Using Acoustic Sensing for Different Squeak Sounds Caused by Number of Rubbing Fingers. Proc. of UIST'16 , 2016.
[18] Ryosuke Kawakatsu and Shigeyuki Hirai, Rubbinput: An Interaction Technique for Wet Environments Utilizing Squeak Sounds Caused by Finger-Rubbing, Proc. of PerCom 2018, pp.643-648, 2018.