

卒業論文 2016 年度（平成 28 年度）

Gyaon

音声活用の幅を広げる録音/共有システム

慶應義塾大学 環境情報学部

佐竹 紘明

増井俊之研究会

2017 年 1 月

卒業論文 2016 年度（平成 28 年度）

Gyaon

音声活用の幅を広げる録音/共有システム

論文要旨

最後に書く

キーワード

音声, 録音, インターフェース

慶應義塾大学 環境情報学部

佐竹 紘明

目次

第1章 序論	1
1.1 背景	2
1.1.1 音声の有用性	2
1.1.2 音声を活用されない理由	2
1.2 本研究の目的	3
1.3 本論文の構成	3
第2章 システムの提案	4
2.1 設計指針	5
2.2 利用例	5
2.2.1 PC版	5
2.2.2 Android版	9
第3章 実装	10
3.1 システム構成	11
3.2 サーバ	11
3.3 PC版	11
3.4 Android版	11
3.5 その他	11
第4章 先行事例	12
4.1 雰囲気の詳細	13
4.2 長時間音声データの管理	13
4.2.1 ブラウジング	13
4.2.2 タグ付け	13
4.3 音声ログの検索手法	13
4.4 写真との組み合わせ	14
第5章 考察	15
5.1 自分で使ってみた	16
5.2 研究室で使ってみた	16
第6章 結論	17
6.1 本研究の成果	18

6.2 今後の展望	18
謝辞	19
参考文献	20

図 目 次

2.1	録音ボタン	5
2.2	音声リスト	6
2.3	Scrapbox のオーディオ記法	6
2.4	地図機能	7
2.5	著者の PC では、command キー + fn キーを録音専用キーとして設定している。 .	8
2.6	ペダルの例	9

第1章 序論

本章では，本研究の背景と目的，および本論文の構成について述べる．

1.1 背景

録音可能な各種の小型レコーダ装置は古くから販売されており、またスマートフォンでは様々な録音アプリケーションを利用できる。これらは音楽に携わる人やインタビュー/会議を記録するといった用途に重宝されるが、日常的にメモを記録するなどして活用している人は少ない。音声といえば、最近では Siri や Amazon echo といった音声入力システムが普及している。これらは高い認識精度と強力なアシスタント機能を持ち、連携するアプリケーションやハードウェアも多く、利便性は高い。しかし、音声認識によってテキストを生成したり操作したいシステムに対して「電気をつける」「音楽を流す」などのコマンドを発行するような使い方に留まっている。

1.1.1 音声の有用性

音声は優れた情報記録 / 情報交換の手段となりうる可能性を秘めている。ネット上での情報記録 / 情報交換はテキストベースのものがほとんどであるが、録音なら手書きメモや PC / スマートフォンへのキー入力を行う必要がなく、記録したいことを喋るだけで良い。日本の若者を中心に人気のコミュニケーションサービス LINE¹ ではボイスメッセージ機能が利用でき、簡単さとテキストにはない表現力の豊かさから、一部のユーザーは積極的に活用している。長時間の録音でも、始めてしまえば記録に集中する必要がないので、同時に別のことを行える。これはテキストや写真、動画による記録とは異なる点である。また、写真や文章といった他のメディアとの組み合わせによって、互いの内容理解を助ける相互補完性の存在が知られている [6]。MP3 などの音声圧縮技術が発達していることから、データ量を気にせずこういった活用ができるはずである。すでに様々な活用が知られている音声だが、なぜ日常的に利用されていないのか。

1.1.2 音声を活用されない理由

以上のような音声の有用性が知られているにも関わらず多くの人に活用されないのは、便利に使える録音システムが存在しないからである。既存の録音システムは録音の手順が煩雑である。一般的に録音/停止/再生それぞれでボタン操作が必要であり、音声記録の単純さをインターフェースが阻害している。また、撮った音声を管理するのも簡単でない。録音時刻 / 再生時間といった基本的なメタデータや自ら設定したタグ等をもとに音声を検索できるが、音声が増えていくほど、内容理解を助ける仕組みとして十分とはいえない。それらの音声は、スマホアプリでさえもその録音システムの中でデータを管理しないとけないことが多く、外部アプリケーション / サービスでの音声活用の可能性を狭めている。音声を最大限活用するためには、以上の音声記録の不便を解消する新しい録音システムが必要である。

¹<https://line.me/ja/>

1.2 本研究の目的

既存の録音システムに見られる不便を解消し，音声を有効活用できるようなシステムを開発することが本研究の目的である．

1.3 本論文の構成

最後に書く

第2章 システムの提案

本章では、音声利用についての背景を踏まえ、関連研究を概観し、新しい録音システム「Gyaon」を提案する。

2.1 設計指針

「Gyaon」では

- 単純な録音操作
- 音声管理の簡単さ
- 他システムからの音声の利用のしやすさ

を優先した設計を行うことで録音の不便を解消し、音声の有効活用を促すシステムを目指している。

2.2 利用例

様々な利用シーンに対応するため、PC用 Web アプリケーション¹および Android スマートフォン用アプリケーションにて Gyaon システムを構築している。それぞれのプラットフォームにおける利用例を述べる。

2.2.1 PC 版

(1) 基本操作

図 2.1 の赤いボタンが録音ボタンであり、長押しすることで録音を開始する。ボタンを離すと録音が停止し、直ちに音声データがサーバへアップロードされる。アップロードが完了したものは音声リスト (図 2.2) に表示され、マウスカーソルを重ねるだけで再生できる。

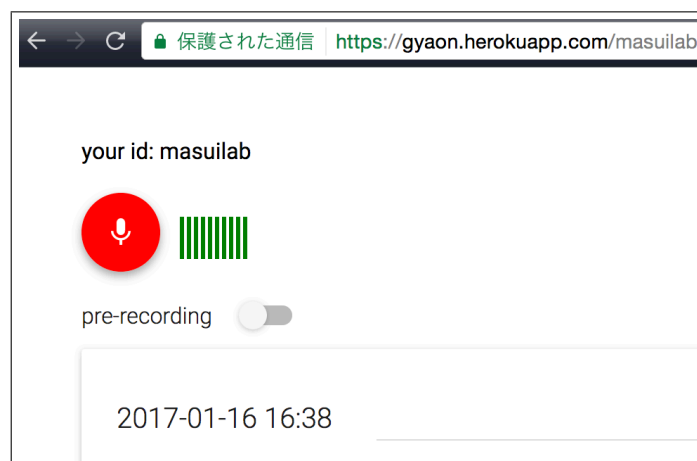


図 2.1: 録音ボタン

¹<https://gyaon.herokuapp.com/>にて試験運用中

2016-10-25 14:42	Hikaruのピアノで録音 at MacMini	00:01	×	🗑
2016-10-25 13:52	うさぎ算(hagino)	01:53	×	🗑
2016-10-25 11:52	ツインピークスも見よう	00:02	×	🗑
2016-10-24 20:37	シリコンバレーを見よ Making the world a b	00:02	×	🗑
2016-10-24 19:43	スクランブル交差点	00:15	×	🗑
2016-10-24 19:41	研究会が盛り上がってます	00:05	×	🗑
2016-10-24 19:04	Gyaonで授業しちゃう (masui)	00:02	×	🗑
2016-10-24 19:04	しゃべらずに生活する (masui)	00:05	×	🗑
2016-10-24 18:55	ペダルはいいぞ (satake)	00:01	×	🗑

図 2.2: 音声リスト

(2) コメントの追加

音声リスト内にコメント入力欄があり、クリックすると入力モードに移行する。音声の内容などを自由に記録することができる。

(3) 音声 URL の取得

音声リストの右側のコピーボタンをクリックすることで、PC のクリップボードに音声データの URL がコピーされる。そのままアクセスすれば音声をダウンロードでき、ローカル環境のアプリケーションなどで利用できるほか、URL を経由して他人に音声を渡すことも可能である。また Notainc²の Wiki システム Scrapbox³ではオーディオ記法を利用でき、決められた書式で音声の URL を貼ると、Gyaon と同様にマウスカーソルによる音声の再生が可能となる。



図 2.3: Scrapbox のオーディオ記法

²<http://www.notainc.com/ja/>

³<https://scrapbox.io/>

(4) プリレコーディング機能

図 2.1 の pre-recording スイッチを ON にするとアプリ起動中は常時録音し、録音ボタンを押す 10 秒前から音声を記録できる。もし突然撮りたい音が聞こえてきた場合でも逃さず録音することが可能である。

(5) 地図機能

音声データに紐付けられた位置情報をもとに、音声を地図上にマッピングして一覧できる⁴。音声の録音場所が分かることから、その場所の雰囲気や、景色がいい/この店は美味しいといったその場所ならではの有用な情報を記録/共有できる。

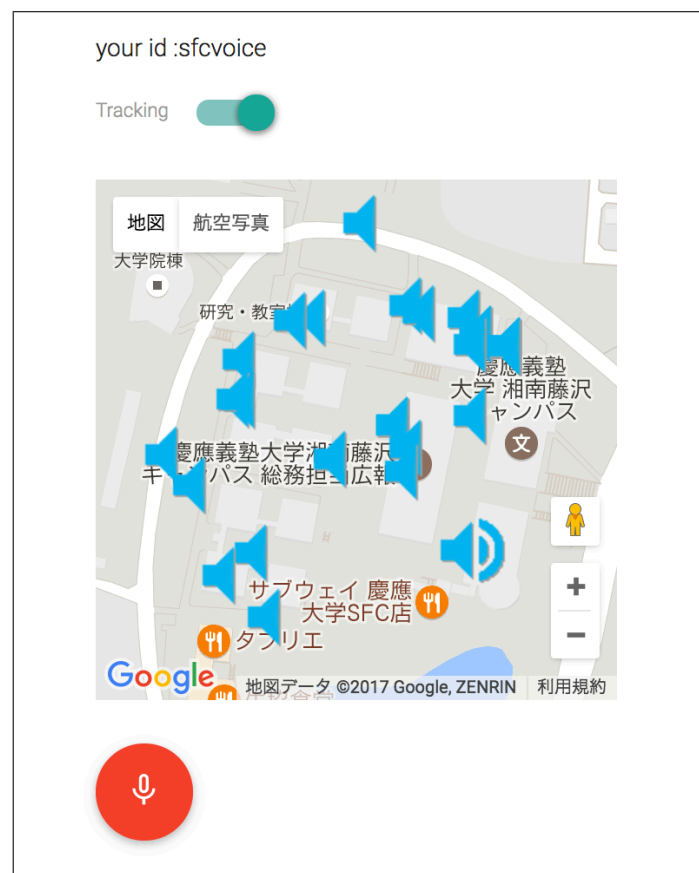


図 2.4: 地図機能

(6) ID の共有

Gyaon では ID ごとに音声を管理しており、URL 末尾でユーザ ID を指定できる。この ID を複数人で共有することで Gyaon を共有タスクリストのように利用したり、ボイスメッセンジャーのような使い方が可能である。

⁴地図機能は <https://gyaon.herokuapp.com/map/>にて試験運用中

ユーザ ID の指定

`https://gyaon.herokuapp.com/{ ユーザ ID }`

(7) Gyaon キー

Gyaon キーアプリケーション⁵を利用すれば、PC の任意のキーを録音ボタンに設定できる。これにより、アプリケーションを起動することなく録音可能である。Web アプリケーションと同様に長押し録音、自動アップロードが行われる。



図 2.5: 著者の PC では、command キー + fn キーを録音専用キーとして設定している。

(8) Gyaon ペダル

USB 接続の汎用なペダル (図 2.6) を接続すれば、楽器演奏/運転中/料理中といった手を使えないシーンでもハンズフリー録音が可能となる。また、ペダルとマイクを接続した Raspberry Pi 等を複数用意し、録音専用コンピュータとしてセットアップすることで、ユビキタスな録音環境を自宅や研究室等に構築可能である。

⁵Mac 用アプリケーションを <https://github.com/stkay/GyaonWithKarabiner> にて配布中



図 2.6: ペダルの例

2.2.2 Android 版

第3章 実装

本章では, 第2章で述べたシステムの設計を受け, 「Gyaon」の実装について述べる.

3.1 システム構成

3.2 サーバ

3.3 PC 版

3.4 Android 版

3.5 その他

第4章 先行事例

本章では、音声利用に関する先行事例を概観する。

4.1 雰囲気の詳細

日常的な体験や雰囲気を記録するための録音デバイスについて研究が行われている [2]. 録音は何かを記録, 説明したり, 思い出すのに重要な手段だが, 日常的に行っている人は少ないと指摘されている. より気軽に録音できるよう日用品にセンサーを埋め込み, 状況の変化に応じて録音/再生するデバイスが提案されている. 日用品に溶け込み, 単純明快な録音/再生インターフェースが実現されているが, 本を閉じる/開くといった二値的な状況変化でしか操作できないため, デバイス単体で複数の音声を管理することが困難になっている.

4.2 長時間音声データの管理

長時間にわたる音声データでは, 聞き返したい箇所を効率良く見つけることが困難となる. この課題に対して, 音声データのブラウジングやタグ付けの手法について研究が行われている.

4.2.1 ブラウジング

人間生活に関わる情報を長期間にわたり記録するライフログ関連研究では, そのログの一部として音声データがよく利用されている [1]. 音声データの特徴量を計算し, 閾値処理するインターフェースを提供することで再生区間を決定する手法 [9] や, 講義ノートといった手書きメモに対して音声データを関連付ける手法 [3] などが提案されている.

4.2.2 タグ付け

研究ノートの補助的な記録手段として音声データを利用するシステムが提案されている [5]. 事前に用意された研究ノートから形態素解析などを利用してキーワードを抽出し, 音声データに付与する手法がとられている.

また, 音声データの特定のタイミングにタグを付与できる録音システムが開発されている [8]. 学生がノートテイキングの補助として使うことを想定しており, 重要な説明等を逃さないよう素早くタグ付けが行えるインターフェースが実装されている.

4.3 音声ログの検索手法

音声認識技術を活用したテキストによる音声検索手法が提案されている [4]. 試作されたアプリケーションでは音声データに含まれる単語からキーワード検索が可能となっているほか, 音声認識の信頼度を文字色に反映させたり, ストップワードを見えにくくするなどしてユーザが録音の要点を思い出しやすくなるよう配慮されている.

4.4 写真との組み合わせ

写真と音声を組み合わせによる体験記録手法が提案されている [6]. 写真と手書きだけでは表現できない雰囲気を音声に記録し, 写真によって音声データに一覧性を持たせる仕組みである. 写真が音声の内容把握を助け, 音声データの価値を高めることも確認されている.

また, 同様の手法によって撮影された写真を, 音声とともに閲覧できる Web サイトが公開されている [7].

第5章 考察

本章では、「Gyaon」についての考察を行う。

5.1 自分で使ってみた

5.2 研究室で使ってみた

第6章 結論

本章では、本研究で得られた成果と今後の展望について述べる。

6.1 本研究の成果

6.2 今後の展望

謝辞

本論文の執筆において、担当教員である増井俊之教授に多大なるご指導と貢献をしていただきました。また、本システムの実装において、研究会 OB の橋本 翔氏、桜井雄介氏に多くの貢献と助言を頂きました。環境情報学部 of 早川 匠氏には有用な音声素材を惜しげもなく提供していただきました。この場を借りて感謝の意を表します。

参考文献

- [1] J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong. MyLifeBits: Fulfilling the Memex Vision. In *Proceedings of the Tenth ACM International Conference on Multimedia*, MULTIMEDIA '02, pp. 235–238, New York, NY, USA, 2002. ACM.
- [2] I. Poupyrev, H. Oba, T. Ikeda, and E. Iwabuchi. Designing Embodied Interfaces for Casual Sound Recording Devices. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, pp. 2129–2134, New York, NY, USA, 2008. ACM.
- [3] L. Stifelman, B. Arons, and C. Schmandt. The Audio Notebook: Paper and Pen Interaction with Structured Speech. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pp. 182–189, New York, NY, USA, 2001. ACM.
- [4] S. Vemuri, C. Schmandt, W. Bender, S. Tellex, and B. Lassey. An Audio-Based Personal Memory Aid. In *UbiComp 2004: Ubiquitous Computing*, Vol. 3205 of *Lecture Notes in Computer Science*, pp. 400–417. Springer, 2004.
- [5] 川西 康介, 大平 茂輝, 松原 茂樹, 長尾 確. 研究活動における音声ログの整理と利用. 第 74 回全国大会講演論文集, 2012(1):723–724, mar 2012.
- [6] 中蔵 聡哉, 角 康之, 西田 豊明. 体験記録システムにおける写真撮影と音声録音の相互補完性. 人工知能学会全国大会論文集, 8:358–358, 2008.
- [7] 増井 俊之. 写録(しゃーろっく): 写真撮影と同時に録音した音を楽しむ. <http://www.pitecan.com/Sharoku/>.
- [8] 藤坂 祐史. Recture - 復習しやすい授業記録アプリ. <http://recture.com/>.
- [9] 河村 竜幸, 中西 英之, 石黒 浩. 連続音声録音を用いた会話体験の探索. 人工知能学会全国大会論文集, 21:1–4, 2007.