

# 視覚障害者のための写真撮影・閲覧・共有インタフェース

浅川 智恵子 原田 丞 佐藤 大介 石原 辰也 高木 啓伸\*

**概要.** 写真は視覚的なメディアではあるものの、多くの視覚障害者が記録や晴眼者とのコミュニケーション手段として興味を持っていることが明らかとなっている。そこで本研究では視覚障害者のニーズ調査の結果に基づき、写真と音声メモ・周辺音を結びつけて記録することで、非視覚的な撮影・閲覧・共有を可能にするスマートフォン・アプリケーション“Accessible Photo Album”(APA)を開発した。日常的な状況での評価のため、5人の視覚障害者被験者が実験者とペアで外出して写真を撮影したうえで、別の実験者に写真を説明するという実験を行った。その結果、すべての被験者が音声メモと周辺音を活用して撮影・閲覧・共有を楽しむことができた。

## 1 はじめに

視覚障害者にとって写真は本質的にアクセス困難なメディアである。しかし、ソーシャルネットワークの普及によりコミュニケーション手段としての興味が高まってきた[6]。米国では、視覚障害者を対象とした写真教室もすでに実施されている[14]。視覚障害者が写真を撮影するためには、例えば構図の決定・露出・ブレ等の情報を非視覚的にどう伝えるか、撮影した写真を独力でどのように管理するか等の技術的課題が山積みであり、既に多くの研究が報告されている[6, 11, 1]。しかし、写真を共有するインタフェースの研究は例がなく、我々が実施したアンケート調査では、回答者の半数以上が、写真の共有に問題を感じていた[5]。

このメディアのギャップを埋めるためにはアクセス可能な情報、例えば音声を同時に記録する方法が考えられる。既にいくつかの音声録音可能なカメラやカメラ・アプリケーションが実用化されている[8, 9, 13, 16, 17, 18]が、これらは晴眼者を対象としているためアクセシビリティ上の多くの問題がある。ビデオで代替する方法も考えられるが、撮影時に対象物や人に対して一定時間カメラを向け続けなければならない、瞬間的に撮影可能な写真と比較して困難な操作が要求される。

そこで本研究では音声メモ付写真に着目し、視覚障害者による撮影・閲覧・共有を実現するためのスマートフォン・アプリケーション“Accessible Photo Album”(APA)を開発した。APAはスクリーンリーダを用いて操作することができ、撮影時に周辺音と音声メモを録音する機能を備えている。APAでは閲覧時に音声に加えて日付、時刻、GPS情報も非視覚的に参照可能なため、写真を容易に識別するこ

とが可能であり、さらなる解説を加えての共有が可能である。

本論文では関連研究に続いてAPAのデザインと実装を紹介する。その後、評価実験とその結果について報告したのちに、視覚障害者が独力で情報を発信する手段としての可能性について考察する。

## 2 関連研究

視覚障害者の中で写真撮影に対する興味が高まっている[25, 10, 19, 20]。Jayantらのアンケート調査[6]によると、アンケートに答えた70%のユーザがきれいに写真を撮影することに興味があると答え、ほぼ同数が最近写真を撮影したことがあると答えた。写真の識別には携帯型点字メモ機を併用して写真を撮影した順序に合わせて説明文を記録する、写真のファイル名に意味を持たせる等の方法をとっている。

### 2.1 音声付き写真

FrohlichとTallynが提案した音声付き写真は、写真とそれに付加された録音音声を1つのメディアとしたものである[3, 4]。この音声付き写真の概念は、特に視覚障害者のために考案されたものではないが、視覚障害者が写真を撮影・閲覧・共有する上で非常に有効であると考えられる。

音声付写真を撮影するためのスマートフォン用アプリケーションは、すでにいくつか実用化されているが[8, 9, 13, 16, 17, 18]、アクセシビリティ上の多くの課題がある。例えば、スクリーンリーダ機能を用いると1枚の写真を撮るために、35回のジェスチャ入力が必要となる。また音声付写真を撮影できるカメラも市販されているが、視覚障害者が操作可能な製品は筆者らの調査では見つからなかった。

### 2.2 視覚障害者のための写真撮影支援

iPhoneのVoiceOver[15]やAndroidのTalkBack等、スマートフォンのアクセシビリティ向上に伴い、

Copyright is held by the author(s).

\* Chieko Asakawa, Susumu Harada, Daisuke Sato, Tatsuya Ishihara, Hironobu Takagi, 日本アイ・ビー・エム株式会社 東京基礎研究所

視覚障害者向けのアプリケーションが各種開発されている。写真関連では、画像認識やクラウドソーシングを用いて、視覚障害者が撮影した写真の内容を認識するサービスも利用が広がっている [2, 22, 24]。これらのサービスを利用することにより、ワイン、飲み物の銘柄、葉の種類、郵便物の内容等を遠隔からのサポートにより知ることができる。写真のブレの検出 [7]、視覚情報を音情報に変換する技術 [23]、写真のブレ、暗さ、傾きを音声で伝達する研究 [12] 等も報告されている。

### 3 Accessible Photo Album (APA)

視覚障害者が写真を撮影・閲覧・共有することを目的として APA を開発した。APA は以下の 3 つの方針に基づき設計した。

1. 素早く撮影し、音声メモと周辺音を録音できるように、非視覚的なジェスチャ操作を最適化する
2. 閲覧時に素早く写真を識別可能にするために音声メモの再生を効果的に取り入れる
3. 主体的に晴眼者と写真を共有可能にするために専用のモードを設けてインタフェースを最適化する

#### 3.1 実装

APA は iPhone のアプリケーションとして実装し、読み上げには iOS 付属のスクリーンリーダーである VoiceOver を用いるように設計した。APA はカメラモードとアルバムモードの二つのモードを持つ (図 1)。起動直後にカメラモードに遷移し周辺音の録音が始まり、1 本指で下スワイプすることにより写真が撮影される。1 本指でタップして押し続けること (タップ&ホールド) で音声メモの録音が始まり、指が離れるまでの間記録することができる。録音終了直後に録音した音声メモが自動で再生される。

アルバムモードでは日付別にソートされた写真のリストが表示される。リスト間の移動には 1 本指で左右スワイプ、日付間の移動には 1 本指で上下スワイプを用いる。次の写真へ移動する際に、対応する音声メモが自動的に再生される。再生と同時に写真が撮影された日時と位置情報も音声合成により読み上げられる。位置情報については、写真が保存された時点の GPS 情報を地名に変換して読み上げる。晴眼者と写真を共有するときには、リスト上の写真の位置でダブルタップすることでフルスクリーン表示することができる。フルスクリーンビューでは VoiceOver の標準ジェスチャ機能である 3 本指の左右スワイプを用いて写真間を移動することができる。新しい写真が表示されると、音声メモの再生に続いて、写真の向き (縦方向または横方向) が読み上

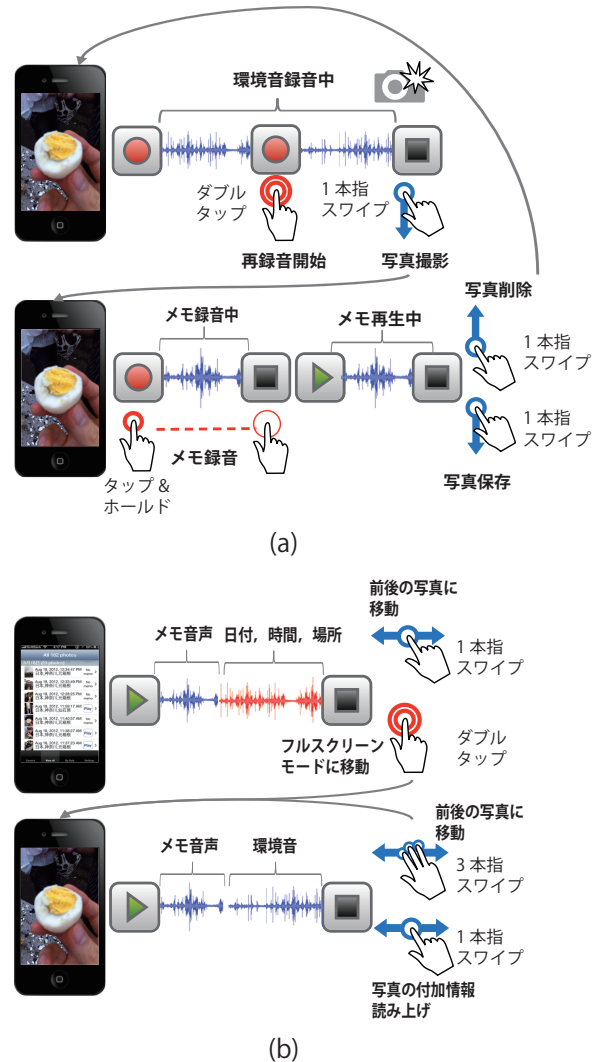


図 1. (a) カメラモードの操作方法, (b) アルバムモードの操作方法

げられるので、デバイスの方向を主体的に変更することが可能である。

### 4 ユーザ実験

5名の視覚障害者被験者の協力により APA の有効性を評価した。実験では以下の 3 つの点に着目した。

- (a) 写真の撮影、音声メモや周辺音の録音を素早く実行できたか?
- (b) それぞれの写真の識別をしながら、素早く閲覧できたか?
- (c) 晴眼者に対して主体的に写真を共有できたか?

#### 4.1 被験者

被験者は、男性 3 名、女性 2 名、年齢は 25 歳から 36 歳までの 5 名で、謝礼を支払った (表 1)。

表 1. 被験者プロフィール.

	年齢	性別	視力	撮影経験	共有経験
P1	25	女性	弱視	有り	有り
P2	34	男性	全盲	無し	有り
P3	35	男性	全盲	無し	無し
P4	34	女性	弱視	有り	有り
P5	36	男性	全盲	有り	有り

#### 4.1.1 写真撮影の経験

P1, P4, P5 は、カメラ付携帯電話で写真を撮った経験があり、画像のファイル名を変更することで写真を管理していた。P1 と P4 は、その方法に比較的満足していたが、P4 は、処理があまりにも単調であることから、日付や時刻で管理するようになり、結果としてあまり写真を撮らなくなった。P2 と P3 は、実験的に写真を撮った経験があるのみである。P3 は、自分のカメラで友人に写真を撮ってもらった経験があった。全ての被験者が独力で良い写真が撮れるかどうか不安を持っていた。撮影経験のある被験者も全員が晴眼者のサポートを得ていた。

#### 4.1.2 写真共有の経験

P1, P4, P5 は携帯電話あるいは PC から e メール添付により他者と写真を共有したことがあった。P4 は写真をブログに投稿した経験があった。P2 は自分で写真を撮った経験はないが、彼が共有したい写真を友人に撮ってもらいその場で SNS に投稿した経験があった。

### 4.2 実験手順

実験は 1 人 1 日 (約 8 時間) で行い、以下の 4 つのセッションで構成される。

#### 1) 導入セッション

被験者は、まず筆者らの研究所にて事前インタビュー・APA の操作方法のトレーニング・実験手順の説明を受けた。この導入セッションは平均 1 時間半であった。

#### 2) 写真撮影セッション

第 1 セッション終了後、被験者は実験者と共に写真を撮るためあらかじめ被験者が行きたいと興味を示した場所に出かけた。そのためそれぞれの被験者は異なる場所で写真を撮影した。これは可能な限り日常に近い状況で実験を行うためである。目的地を決定後、実験者によって道順が確定される。これは単に美しい風景だけを写真に撮るのではなく、被験者が興味を持てるもの、例えば触って楽しめたり直接経験できるものに出会えるよう考慮したためである。実験者は、被験者に同行するだけでなく会話をしながら各地点を回り、周辺の環境の説明を行っ

た。あらかじめ決められた訪問地はあったものの、いつ移動するかについては、被験者に決定させた。時には、元々の道順から大きく外れることになったが、被験者が別の興味を見つけたときにはそれを優先した。また写真を撮るタイミングも彼ら自身が決定できるよう考慮した。被験者は好きな場所で自分自身の写真を取ってもらうよう実験者に依頼できることとした。この場合、実験者が写真を撮り音声メモは被験者が録音した。被験者がカメラの向きを合わせる手助けを必要とした時には、実験者が手助けを行った。カメラの向きを手で修正する前に、まず言葉で向きの修正を手助けした。これは、実際に写真を取っているのは被験者自身であると感じられるように考慮したためである。

#### 3) 写真共有セッション

約 5 時間の外での実験終了後、被験者は研究所に戻り、写真共有セッションを行った。写真の撮影経験に関する簡単な質問の後に、被験者は観察者 (別の実験者) と撮影してきた写真を共有するためのセッションを実施した。被験者は、写真を共有しながらその日経験したことを観察者に共有する。音声メモを他人に聞かれることをためらった被験者がいたため、必要であればイヤフォンの使用を認め、被験者は観察者に聞かせたいと思ったときのみ音声を再生できることとした。観察者は被験者と写真を見ながら、その日の経験について自由に会話するよう指示された。被験者と外出した実験者も、写真共有セッションに同席した。観察者は、被験者が写真についての説明を始めるまでは、コメントや質問を待つよう指示された。これは、被験者が晴眼者の助けなしに、一人で写真の識別できるか確認するためである。観察者は、被験者が写真の識別が困難な時のみ、手助けするよう指示された。写真共有セッションは、平均 30 分前後を要した。

#### 4) 実験後のインタビューセッション

1 日の実験の最後に、その日の経験や APA に関する被験者からのコメントを収集した。

### 4.3 実験結果

それぞれの実験の目的について、ユーザ実験により得られた知見を述べる。

#### (a) 写真の撮影、音声メモや周辺音の録音を素早く実行できたか?

VoiceOver の経験があったのは P3 のみであったが、他の被験者も、導入セッション中に APA を使用するために必要な基本的ジェスチャを習得することができた。写真撮影セッション中に撮影された写真は、13 枚から 20 枚であった (表 2)。

7 段階のリッカート尺度によるアンケート結果を

表 2. 写真の撮影場所と枚数.

	場所	枚数
P1	ショッピングモール/レジヤ施設	17
P2	ランドマークタワー/レジヤ施設	18
P3	ワインショップ/レジヤ施設	20
P4	史跡/公園	13
P5	中華街/ショッピングモール	15

表 3 に示す。「写真の撮影は容易であったか」との問いに対して、2名の被験者が7を、その他は5で回答した。「音声メモの録音は容易であったか」の問いに対しては、2名が7、2名が6、1名が5と回答した。携帯電話で写真撮影の経験があった3名の被験者に対して、以前の方法と、APA とを比較し、音声メモの詳細度、スピード、使いやすさについて評価を依頼した。その結果、3つの項目全てに対して、全員がAPA が優れていると評価した。

(b) それぞれの写真の識別をしながら、素早く閲覧できたか?

全ての被験者が問題無く閲覧操作を行え、撮影した写真を識別することができた。アンケートにおいても「閲覧は簡単だった」、「目的の写真を見つけることができた」、「撮影した状況を思い出せた」という問いに対して、P2を除くすべての被験者が、「全くそう思う」または「そう思う」と返答した(表3)。

P2だけは2枚の写真を混同したため、「目的の写真を見つけることができた」という問いに対して「どちらかというそう思う」と返答している。この2枚は、シーンが同じで1枚は実験者が撮った被験者とランドマークタワーの写真であり、もう一方は被験者が撮ったランドマークタワーだけであったが音声メモがほぼ同じであったため、P2はその区別ができなかった。

(c) 晴眼者に対して主体的に写真を共有できたか?

「写真共有セッションを主体的に行うことができたか」という問いに対して、P2を除くすべての被験者が、「全くそう思う」または「そう思う」と返答した(表3)。P2は、前述の1組の写真の識別ができなかったことから、「どちらともいえない」と答えた。この判断ミスが、彼のAPA全般に関する評価にどの程度影響を及ぼしたかを判断することは難しいが、他のすべての写真については、問題無く詳細を説明できていた。

#### 4.4 観察

##### 周辺音の有効性

P5が中華街の門で撮影した写真を共有していた時、周辺音として水しぶきが再生された。「自分のすぐ隣でだれかが水しぶきをあげていたことに気づ

表 3. アンケート結果

(1:全くそう思わない, 7:全くそう思う)

	P1	P2	P3	P4	P5
写真撮影セッション					
撮影は楽しかった	7	4	6	7	6
撮影は簡単だった	5	7	5	7	5
メモ録音は簡単だった	6	7	6	7	5
独力で写真を撮りたいと思うようになった	3	4	7	7	4
以前の方法との比較					
より簡単だった	7	-	-	6	7
より早くメモを残せる	7	-	-	6	5
より詳細なメモを残せる	6	-	-	7	6
写真共有セッション					
閲覧は簡単だった	7	7	6	7	7
目的の写真を見つけることができた	6	5	7	7	7
撮影した状況を思い出せた	7	6	6	6	7
写真を見せることは楽しかった	7	7	7	6	6
写真共有セッションを主体的にできた	6	4	6	6	7
実験全体について					
写真共有は楽しい	7	6	6	7	6
同様のアプリを使って写真を共有したい	6	7	6	7	3
実験前より写真を撮りたいと思う	6	6	6	7	7
実験前より写真を共有したいと思う	7	7	6	7	4

いていなかった」と、新たな発見をコメントした。P1とP4は、当初観察者に音声メモを共有することをためらっていたが、2枚の写真について周辺音を共有したところ、その価値に気付いたとコメントした。このときの周辺音は、噴水とブラスバンドの音であった。またP4は、周辺音を利用すれば視覚障害者の友人とも写真を通して経験を共有できるとコメントした。

##### 主要なコメント

すべてのコメントは紙面の制限から掲載できないため、以下に主要なものを紹介する。

P1:「音声メモによって、自分で写真の管理ができるようになった。これはすばらしい進歩である。」

P4:「とても楽しかった。うまく写真がとれているかどうか不安だったため、写真共有についての楽

しさの評価は6としたが、それを除けば素晴らしい経験だった。これまでは自分の経験を言葉でしか伝えられなかったが、写真を使うことで視覚的に情報が伝えられることの楽しさを知った。こんな経験は初めてだった。」

P5:「写真が表示されたとき、すぐにそれについての話ができることは素晴らしい。周辺音に録音された周りの声を聞くと、写真を超えた情報を得ることができたと感じられて、とても良い気分である。」

## 5 考察

### 5.1 主体的な写真共有

APAの目的の1つは、視覚障害者が主体的に晴眼者との写真共有を可能にすることであった。その点でP3は最も主体的に、演出も交えて写真共有セッションを楽しんでいた。ある写真を見せる前に、まず観察者に音声メモが聞こえないようイヤフォンを装着し、次に画面が観察者に見えないようiPhoneを裏返した。「多分私がどんなボトルを偶然みつけたか想像できないと思いますよ」と話した後に、画面を元に戻し「コンコルド型のウスキーのボトル」の写真を観察者に見せた。このように手に持つことのできるデバイスだからこそ可能な視覚的演出を、その効果を判断することのできない視覚障害者自身がとったという点で実験者らの予想を超えた行動であった。今後は、このような“演出”も含めた写真共有の表現力向上を目指したインタフェースの工夫や知見の共有が必要となるであろう。

### 5.2 ユーザインタフェースの改善

音声メモや周辺音をどのタイミングで録音するかは、APAのようなアプリケーションを設計する際に重要な要因となる。実験に使用したAPAでは音声メモと周辺音を別々に取り扱った。シャッターが押されるまでの“周辺音”には晴眼者によるカメラの向きの指示や、周囲の状況説明などの音声が含まれてしまったため、実際には撮影後の“音声メモ”のほうが有用であった。

そこで、実験後にAPAのインタフェースを修正し、シャッター前録音を廃止し、シャッター後にユーザ自身が“タップ&ホールド”ジェスチャにより、周辺音と音声メモを自由に録音できるインタフェースに改良した。この場合周辺音と音声メモの記録の違いはなく一つの音声ファイルとして保存される。これによりシャッターを押す、“タップ&ホールド”でメモと周辺音を録音するというシンプルなインタフェースに変更された。

### 5.3 音声メモとテキストメモ

これまで写真にテキストでコメントをつけていた被験者も、音声メモは簡単に、素早く、そして周辺

音を含めた詳細な説明が付与できるとコメントした。確かに音声メモは作成が容易であるという大きな利点があるが、テキストには写真の検索が容易になる、ネットワーク上で共有するときに既存サービスと親和性が高いなどの利点がある。今後、音声認識の利用も含めてテキストを音声メディアに取り組み必要がある。

また、二人の被験者が音声メモを後で修正して詳細度を向上したいとコメントした。P5は、観察者からのコメントを音声メモに追加することで、さらに写真の回想が容易になると述べていた。このように音声メモの編集に対するニーズは高く、検討の必要がある。

## 6 まとめ

本研究では視覚障害者による写真の撮影・閲覧・共有を可能にするスマートフォン・アプリケーション“Accessible Photo Album”(APA)を開発し、評価実験を行った。APAでは写真撮影と同時に音声メモや周辺音を録音し、閲覧・共有時に自動的に再生することで視覚障害者が独力で写真を識別し共有することが可能になる。実験の結果、晴眼者とであっても主体的に写真の共有が可能になるだけでなく、独自の演出表現までも可能になることを観察できた。写真共有は晴眼者だけでなく視覚障害者にとっても価値がある活動になることがこの実験でも裏付けることができた。

今後、より豊かな写真体験を視覚障害者に提供するためにはメディアのギャップを埋めるためにメタデータの拡大が重要な課題の一つとなる。例えばユーザが室内・外どこにいるか、位置情報、コンパス、データベースを組み合わせて写真に写っている建物の詳細な解説などが付与できるであろう。さらにクラウドソーシングのアプローチにより写真の詳細な説明を付加することも可能になってきている。写真は「ソーシャルメディア時代の新たな言語である」とも言われるほど、その重要性が高まっている。視覚障害者もそのような時代に独自の的方法論を蓄積してコミュニティ参加できるようにするための技術開発が今後も必要となるであろう。

## 参考文献

- [1] Baber, C., Cross, J., Khaleel, T., and Beale, R. Location-based photography as sense-making. *Proc. BCS-HCI 2008, British Computer Society (2008)*, 133–140.
- [2] Bigham, J.P., Jayant, C., Ji, H., et al. VizWiz: nearly real-time answers to visual questions. *Proc. UIST 2010, ACM (2010)*, 333–342.
- [3] Frohlich, D. and Tallyn, E. Audiophotography: practice and prospects. *Ext. Abstracts 1999, ACM (1999)*, 296–297.

- [4] Frohlich, D.M. Audiophotography: Bringing photos to life with sounds. Kluwer Academic Pub, 2004.
- [5] Harada, S., Sato, D., Adams, D. W., Kurniawan, S., Takagi, H., and Asakawa, C. Accessible photo album: enhancing the photo sharing experience for people with visual impairment. *Proc. CHI 2013*, ACM (2013), 2127–2136.
- [6] Jayant, C., Ji, H., White, S., and Bigham, J.P. Supporting blind photography. *Proc. ASSETS 2011*, ACM (2011), 203–210.
- [7] Ko, J. and Kim, C. Low cost blur image detection and estimation for mobile devices. *Proc. ICACT 2009*, IEEE Press (2009), 1605–1610.
- [8] Many Ltd. Picle. <http://www.picleapp.com/>.
- [9] Robinson, S. Com-Phone Story Maker. <https://play.google.com/store/apps/details?id=ac.robinson.mediaphone>.
- [10] UCR/California Museum of Photography. Sight Unseen. <http://www.cmp.ucr.edu/exhibitions/sightunseen/>.
- [11] Vazquez, M. and Steinfeld, A. Helping visually impaired users properly aim a camera. *Proc. ASSETS 2012*, ACM (2012), 95–102.
- [12] White, S., Ji, H., and Bigham, J.P. EasySnap: real-time audio feedback for blind photography. *Proc. UIST 2010*, ACM (2010), 409–410.
- [13] Luo Xu. Speature. <http://itunes.apple.com/us/app/speature/id401887412>.
- [14] Blind with Camera School of Photography. <http://www.blindwithcameraschool.org/>.
- [15] Apple - Accessibility - iOS - Voiceover. <http://www.apple.com/accessibility/ios/voiceover/>.
- [16] StoryMark Life. <http://www.storymarklife.com/>.
- [17] Voicepic. <http://voicepic.me/>.
- [18] Storyrobe. <http://storyrobe.com>.
- [19] Photos by Blind Photographers - Photo Essays - TIME. <http://www.time.com/time/photogallery/0,29307,1897093.1883579,00.html>.
- [20] Seeing Beyond Sight: Photographs by Blind Teenagers. Book, video & TED Talk. <http://www.seeingbeyondsight.org/home/>.
- [21] Blind With Camera. <http://www.blindwithcamera.org/>.
- [22] oMoby. <http://www.iqengines.com/omoby/>.
- [23] Seeing with Sound - The vOICE. <http://www.seeingwithsound.com/>.
- [24] TapTapSee. <http://www.taptapseeapp.com/>.
- [25] 伊藤 邦明. 七重, 光をありがとう. 河出書房新社, 2000

## 未来ビジョン

APA を用いて作成された音声情報群を共有することで、他の視覚障害者の経験を楽しむメディアとして利用できる可能性がある。ユーザ自身が音声情報を写真・地理情報とあわせてソーシャルネットワーク上で共有することで、視覚障害者自身の手でソーシャルなオーディオデータベースを作り出すことができる。このアプローチをジオボイスタグと呼ぶこととする。類似したサービスはこれまでも様々な場面で議論されてきた。しかし、テキストタグやフォトタグと比較してあまり注目されていない。理由の一つはジオボイスタグを利用したアプリケーションに限られるためであろう。視覚障害者にとっては状況が異なる。例えば視覚障害者が晴眼者の友人とジオボイスタグを聞きなが

ら歩けば、視覚的には得られない周辺情報を獲得することができ、より対等な立場から会話を楽しむことができる。晴眼者は意図的に周辺情報の説明をする必要がなくなり、視覚障害者はレストランやショップの情報、それらの評判についても主体的に会話をリードできるようになる。視覚障害者が一人で街中を歩く場合のナビゲーションサポートの可能性もある。このシナリオ実現には周辺音を妨げないでボイスタグを提示する必要がある。視覚障害者は通常、車の音をはじめとした周辺音を音声ランドマークとして街中を移動するためである。共有サービス自体のデザインと並行して、使用状況を考慮したインタフェースや、ソーシャルメディア上の情報との融合、高齢者等の他のユーザ層に向けたユースケースなどを検討する必要がある。