視覚情報が音声情報提示の主観的音量に与える影響の調査

渡邉 拓貴[1,2]　松田 裕貴[3,2]

概要：近年の技術発展に伴い，イヤホン型のウェアラブルデバイス（ヒアラブルデバイス）が急速に普及しつつある．ヒアラブルデバイスによって，ユーザはナビゲーションや音声アシスタントからの提案など，システムから常に情報を受け取ることが可能になる．一方で，人間が知覚する音の大きさ（主観的音量）は，ユーザの状況によって異なることが分かっており，システムからの音声情報を常に一定の音量で提示することは適切ではないといえる．そこで本研究では，主観的音量に影響を与え得る要素の中で，ユーザの視覚から得られる情報に着目し，ヒアラブルデバイスでの音声情報の利用を想定した環境で，視覚情報が音声情報提示の主観的音量に与える影響を調査した．視覚刺激として雑踏と公園の 2 種類，聴覚刺激として 3 種類の音量（small，normal，large）の音声アナウンスを用いて主観的音量の変化を調査した．結果として，同じ音圧で音声情報を提示しても，雑踏のような視覚刺激の多い映像よりも，公園のような視覚刺激の少ない映像の方が，被験者は音量を大きく感じる（主観的音量が大きい）傾向が確認できた．また，音量 small の場合には視覚刺激間で有意な差が確認された．

Investigation of the Effect of Visual Information on Subjective-Volume Change of Audio Information Presentation

HIROKI WATANABE[1,2]　YUKI MATSUDA[3,2]

1. はじめに

近年の技術発展に伴い，イヤホン型のウェアラブルデバイス（ヒアラブルデバイス）が急速に普及しつつある．ヒアラブルデバイスによって，ユーザは移動時や運動時など様々な状況下で，ナビゲーションや音声アシスタントからの提案など，システムから常に情報を受け取ることが可能になる．

一方で人間が知覚する主観的な印象は，ユーザの状況によって変化することが先行研究から報告されている．例えば，スマートグラス上に視覚刺激を提示することで主観的な時間の速さが変化するシステムの提案 [1] や，スマートグラス上で再生する動画の再生速度が視聴後の作業速度に与える影響が報告されている [2]．また，新型コロナウイルス感染症といった社会情勢が，ある空間における体感混雑度に影響を与えることが報告されている [3]．音についても，ユーザが感じる音の大きさ（主観的音量）は，ユーザの状況によって異なることが先行研究から報告されている．例えば，ユーザの主観的な音量は，直前までに提示されていた音量によって変化すること [4] や，同音圧の音源でも視覚効果の有無によって主観的な音量や音から受ける印象が変化することが報告されている [5]．一方で，これらの研究では視覚刺激と音とは対応するもの（例えば，車の映像に対して車の音）を用いており，ヒアラブル環境で想定しうる，視覚刺激と音とが対応していない場合（例えば，視覚刺激は公園だが聴覚刺激は音声アナウンス）の影響については，我々の知る限り調査されていない．この影響が明らかにされれば，ヒアラブル環境において従来考慮されていた環境音に応じて情報提示の音量を変化させるだけでなく，視覚刺激にも応じて情報提示の音量を変化させることで適切な音量でのサービス提供が可能になる．また、ユーザとシステムのインタラクション（ユーザによる能動的な音量調整など）を通じて、ある環境がどのような特性を持つ場所であるのかの情報を参加型センシングのアプローチ

1 北海道大学, Hokkaido University
2 科学技術振興機構さきがけ
PRESTO, Japan Science and Technology Agency
3 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

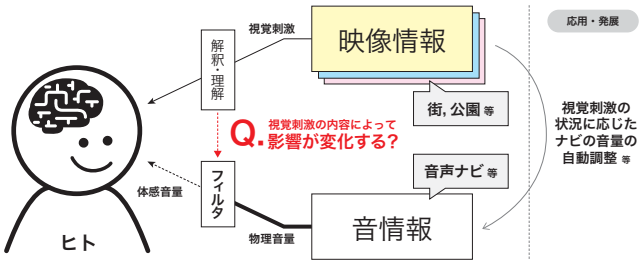

図 1　実験の枠組み

で収集することも可能となると考える。これによって、あらゆる状況（視覚刺激）における体感音量の対応関係を網羅することも可能となる。

従って本研究では，視覚刺激と聴覚刺激との間に関連がない場合に生じる，視覚刺激が主観的音量に与える影響の調査を目的とした評価実験を行った（**図 1**）．ユーザが屋外環境でヒアラブルデバイスから音声情報を得る環境を想定し，ヘッドマウントディスプレイ（HMD: Head Mounted Display）から視覚刺激を，イヤホンから聴覚刺激を与え，ユーザの主観的音量の変化を調査した．視覚刺激として雑踏と公園の 2 種類を，聴覚刺激として 3 種類の音量（small, normal，large）の音声アナウンスを用いた．結果として，雑踏のような刺激の多い映像よりも，公園のような刺激の少ない映像の方が，主観的音量が上がる傾向にあることが確認された．また，音量 small の場合には 2 つの視覚刺激の間に有意差が確認された．

2.　関連研究

2.1　視覚刺激が主観的音量に与える影響

視覚情報と聴覚情報とは相互に影響を及ぼす可能性があることが複数の先行研究から示されている．佐久間らは，視覚情報の変化によって主観的音量が変化するか調査を行った [6]．Web カメラを搭載した HMD を用い，ユーザが指定した箇所に視覚効果を施すことにより，ユーザの主観的音量を制御するシステムを実装した．安倍らは，視覚情報が環境音知覚に及ぼす影響について調査を行った [5]．刺激音を提示する際に，対応する映像も同時に提示して評価実験を行なった結果，音の明るさ因子を除く各因子について，映像の付加による影響が確認されている．Tokunaga らは，自動車走行音を対象として，音のみの場合と，音と映像を与えた場合に，被験者の感じる主観的な音量が変化することを報告している [7]．藤本らも自動車騒音を対象として，現場で実騒音を聴いた場合，実験室内で騒音だけを聴いた場合，実験室内で映像と騒音を試聴した場合において，得られる大きさ，やかましさ，妨害感の評価が異なることを示している [8]．Szychowska らは，聴覚刺激のみ，視覚刺激のみ，聴覚刺激と視覚刺激を与えた場合で，環境騒音から得られる評価が異なることを検証している [9]．Bangjun らは，騒音の音源が視覚的に確認できるかどうかで生じる影響について検討しており，同じ騒音環境であれば，騒音源が視認できる方が不快感が高まることが報告されている [10]．宮川らは日常聞かれる 62 種類の環境音を対象として，音のみを提示した場合と，映像を付加して提示した場合について音の印象を調査した [11]．また，朝倉は都市環境音の主観的印象が視覚情報の有無や種類によって異なることを明らかにしている [12]。結果として、視覚刺激のない場合に比べて滝や飛行機といった視覚刺激がある場合に主観的音量が増大することが示されている。

上記の研究では，視覚情報が聴覚へ与える影響を調査しており，視覚情報と聴覚情報とに関連があることが示されている．しかし，これらの研究では対応のある視覚刺激と聴覚刺激を用いており，本研究で対象とする視覚刺激と聴覚刺激が対応していない状況での影響の調査という点で異なる．

2.2　聴覚刺激が主観的音量に与える影響

聴覚に音刺激が持続的に提示されることで，主観的音量が減少するということは古くから研究されている [13], [14]．矢高らは，騒音等の周囲の状況や会話中などの活動状況の影響に応じてイヤホンから提示する音量を動的に変更するシステムを提案した [15]．同じく矢高らは，主観的音量が直前に聴取した音量の影響を受けることを利用し，システムからの音声情報提示によりユーザの感じる音量が一定とする機構を提案した [4]．

上記の研究では音の提示により主観的音量が変化することに着目しており，本研究で対象とする異なる視覚刺激が音声情報提示の主観的音量に与える影響の調査とは対象が異なっている．

3.　視覚刺激が主観的音量に与える影響調査

3.1　実験方法

3.1.1　実験条件

本研究では，ヒアラブルデバイスの利用時を想定し，屋外環境で音声情報提示を受ける際の主観的音量を調査することを目的とする．実験の枠組みを図 1 に示す．

実験の様子を**図 2** に示す．被験者は，没入型 HMD（Meta Quest 2[*1]）を装着し視覚刺激を，耳にはイヤホン（Apple AirPods Pro[*2]）を装着し，聴覚刺激をそれぞれ受ける．

視覚刺激として，視覚刺激が大きいと想定される雑踏（town 条件）と，視覚刺激が小さいと想定される公園（park 条件）の 2 条件を選択した．それぞれの条件で 6 種類の動画を用意し，視覚刺激とした．利用した視覚刺激の一例を**図 3** に示す．用いた動画は VR180 に対応しており，被験者の頭部方向に応じた視界が得られるため，被験者は没入感を持って視覚刺激を得られる．なお，本研究では視覚刺

*1　https://store.facebook.com/quest/products/quest-2/
*2　https://www.apple.com/jp/airpods-pro/

図 2　実験中の様子

図 3　視覚刺激の一例: (a) 公園，(b) 雑踏

表 1　使用した聴覚刺激

文章	再生時間 [s]
岸田総理は東南アジアとヨーロッパ 6 カ国への歴訪を終え，先ほど政府専用機で羽田空港に到着した．	11
ウクライナ南東部の要衝マリウポリで，ロシア軍は 6 日，アゾフスタル製鉄所に対する攻撃を継続した．	9
半導体不足の影がクレジットカード業界にも忍び寄る．カードの IC チップには半導体が使われる．	10
ウクライナ政府は，南東部マリウポリのアゾフスタル製鉄所から民間人の女性や子供，高齢者全員が避難したと発表した．	11

激による主観的音量の変化の調査を目的とするため，動画の音は無音とした．

聴覚刺激として，3 種類の音量のアナウンス（small, normal，large）を用意した．normal はイヤホンからの出力で約 55dB であり，large は normal の +6dB，small は normal の −6dB である．アナウンスは Text-to-speech で作成されたもので，内容はニュースの見出しとした．時間は 10 秒前後のものを 12 種類（3 音量 ×4 種類）用意した（**表 1**）．

3.1.2　実験手順

実験の流れを**図 4** に示す．実験では，視覚刺激がない（目を閉じた）状態で，まず基準音を提示する．毎回の聴覚刺激の基準とするため，基準音は毎回同じ音量とし，normal と同一のものを用いた．次に，被験者は HMD から視覚刺激を，イヤホンから聴覚刺激をそれぞれ受ける．最後に，被験者は基準音と比較して刺激音がどの程度と感じたか，

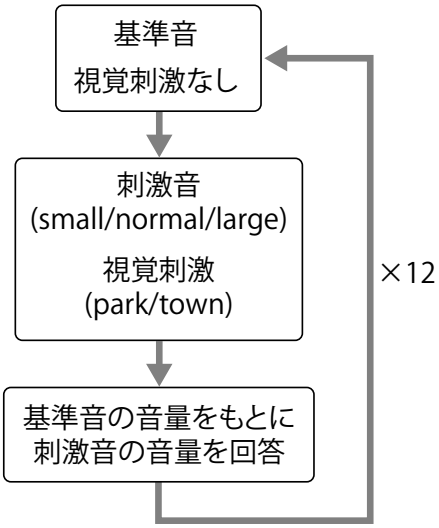

図 4　実験の流れ

数値で解答する．例えば，基準音を 100 とした場合，刺激音が少し大きいと感じたら 120 などと回答するよう指示した．なお，基準音の数値は指示せず，被験者が設定できるようにした．刺激音に対する被験者の回答結果を被験者ごとの基準音の数値で割ることで，各被験者の回答を正規化した．上記の手順を 12 回（視覚刺激：2 条件 ×6 種類，聴覚刺激：3 条件 ×4 種類）行う．視覚刺激と聴覚刺激の組み合わせと提示順は，被験者毎に異なるようにランダムに設定した．

実験では，まず実験環境及び実験手順への慣れを考慮し，テスト動画及びテスト音声を用いて，数回上記の実験手順を繰り返した．その後，本実験として被験者は上記の手順を 12 回繰り返した．被験者には映像，音声ともに条件がどのように異なるかは説明せず，条件名も伝えずに実験を行った．被験者は 20 代から 30 代の男女 10 名である．

3.2　実験結果

各被験者の正規化後の回答結果を**表 2** に，結果を音量ごとにまとめたものを**図 5**，**図 6**，**図 7** に示す．全ての音量の場合で，town 条件よりも park 条件の方が平均値が高い．つまり，park 条件の方が主観的音量を大きく感じる傾向が確認できる．本研究では，物理的な音圧が同じ場合の視覚刺激による主観的な音量の変化に着目している．したがって，音量 small，normal，large それぞれの場合で t 検定を行った．結果として，small 条件の場合，視覚刺激によって主観的音量に有意な差が確認された（$p < .01$）．normal 条件の場合，視覚刺激による有意な差は確認されなかった．large 条件の場合，視覚刺激による有意な差は確認されなかった．

表 2　各被験者ごとの正規化後の回答値

被験者	Small		Normal		Large	
	Park	Town	Park	Town	Park	Town
A	0.70	0.65	1.05	1.00	1.30	1.20
B	0.70	0.65	1.10	0.95	1.35	1.40
C	0.80	0.70	1.00	1.00	1.23	1.20
D	0.88	0.83	1.03	0.94	1.23	1.20
E	0.80	0.85	1.05	0.93	1.18	1.15
F	0.83	0.78	1.05	1.05	1.13	1.13
G	0.95	0.80	0.95	1.10	1.20	1.10
H	0.80	0.70	1.05	1.00	1.25	1.20
I	0.90	0.85	1.10	1.05	1.15	1.10
J	0.90	0.80	1.00	1.05	1.20	1.25
平均	0.83	0.76	1.04	1.01	1.22	1.19

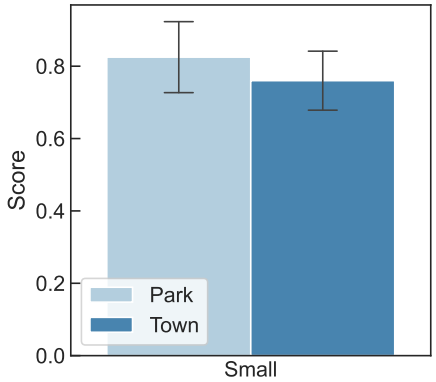

図 5　音量 small の実験結果

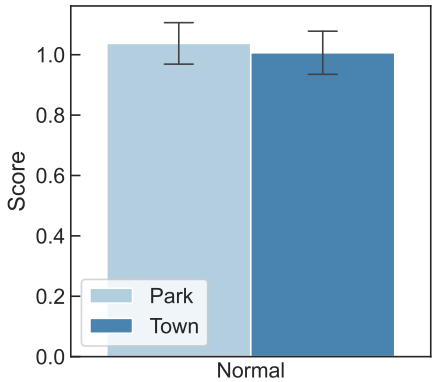

図 6　音量 normal の実験結果

4.　考察

4.1　視覚刺激による主観的音量の変化

実験結果より，音量 small の場合にのみ，視覚刺激によって主観的音量に有意な差が認められ，仮説が一部支持されたといえる．視覚刺激によって主観的音量に変化が生じる原因として，視覚刺激によって期待する音量と，実際に聞こえてくる音量の差異が影響している可能性がある．つまり，雑踏では音量が大きいことを想像するため，聴覚から得られる音量も大きいことを期待しながら音を聞くことになる．一方で公園のような静音な環境では聴覚から得られる音量は小さいことを期待しているため，実際の提示音圧は同一でも雑踏の場合よりも主観的には大きく聞こえたと考えられる．

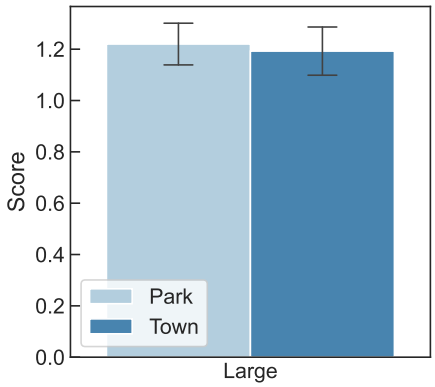

図 7　音量 large の実験結果

一方で，音量 normal，large の場合には有意な差が得られなかった．この原因として，被験者の数が十分ではなかった可能性が考えられる．今後，被験者数や音量パターンの増加など，より発展した調査が必要である．

4.2　主観的音量変化を考慮した情報提示

前述したように，視覚刺激によって主観的音量に変化が生じることが示唆された．従って，ヒアラブルデバイスでの情報提示時には，従来考慮されてきたユーザの周囲状況の音量のみならず，ユーザの視覚刺激も考慮した情報提示を行う必要がある．また，本研究では視覚刺激の影響のみを考慮するため，動画の音を削除していたが，実際の環境では視覚と聴覚に同時にその環境の情報が得られることになる．今後の課題として，ユーザの置かれた環境の視覚刺激及び聴覚刺激を複合的に考慮した最適な音声情報提示手法を確立する必要がある．

4.3　視覚刺激の種類

本稿では，視覚刺激として雑踏と公園を選択し，雑踏は視覚刺激が大きく，公園は視覚刺激が小さいと想定して実験を行った．しかし，視覚刺激のどのような要因が主観的音量に影響を与えたのかは明らかでない．例えば，同じ繁華街でも人通りが全くない場合は，主観的音量に与える影響が異なる可能性がある．視覚刺激のどのような要素が主観的音量の変化に与える影響が大きいのか，今後調査する必要がある．

4.4　刺激音の種類

本稿では，屋外環境でヒアラブルデバイスからナビゲーション等がある環境を想定し，刺激音として音声アナウンスを利用したが，刺激音が他の音の場合（純音や環境音）

には異なる結果が得られる可能性がある．また，同じアナウンスでもユーザの状況と合致する音声かどうかで主観的音量が変化する可能性は考えられる．例えば，駅の中でニュース音声と次の発車時刻の読み上げがあれば，後者の方がユーザにとっては重要と考えられ，主観的音量に差が生じる可能性がある．刺激音の種類や音声アナウンスの内容まで考慮した検討も今後必要だと考えられる．

5. まとめ

本研究では，ヒアラブルデバイスでの音声情報提示を想定し，視覚刺激と聴覚刺激が対応しない場合に，視覚刺激が主観的音量に与える影響を調査した．雑踏と公園の視覚刺激に対して主観的音量を調査したところ，音量 small の場合に，雑踏の視覚刺激よりも，公園の視覚刺激の方が主観的音量が有意に大きくなることが確認できた．また，得られた結果からヒアラブルデバイスにおける今後の音声情報提示手法について議論した．

謝辞

本研究の一部は，JST さきがけ JPMJPR2138，JPMJPR2039，および奈良先端科学技術大学院大学 CICP プロジェクトの助成を受けたものである．

参考文献

[1] 長谷川瑛一, 磯山直也, 酒田信親, 清川清. 動画の再生速度が視聴後の作業速度に与える影響. インタラクション 2021, pp. 41–48, 2021.

[2] Tomoyuki Shimizu, Kyosuke Futami, Tsutomu Terada, and Masahiko Tsukamoto. In-clock manipulator: information-presentation method for manipulating subjective time using wearable devices. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*, MUM '17, pp. 223–230, New York, NY, USA, November 2017. Association for Computing Machinery.

[3] 松田裕貴, 中村優吾, 諏訪博彦, 安本慶一. 人の知覚の集合知による参加型 IoT センサ調整プラットフォームの設計. マルチメディア, 分散, 協調とモバイル（DICOMO 2021）シンポジウム論文集, pp. 591–599, 2021.

[4] 矢高真一, 寺田努, 塚本昌彦. ウェアラブルコンピューティングのための主観的音量に基づく音声情報提示手法. 情報処理学会研究報告 Vol. 2011-UBI-29, No. 38, pp. 1–8, mar 2011.

[5] 安倍幸治, 小澤賢司, 鈴木陽一, 曽根敏夫. 視覚情報が環境音知覚に与える影響. 日本音響学会誌, Vol. 56, No. 12, pp. 793–804, 2000.

[6] 佐久間一平, 寺田努, 塚本昌彦. 視覚効果を用いた主観的音量の制御システムの設計と実装. エンタテインメントコンピューティングシンポジウム 2015 論文集, pp. 357–364, 2015.

[7] Yasunobu Tokunaga, Takane Terashima, and Ayumi Ishikawa. Influence of visual information on subjective evaluation of road traffic noise. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, Vol. 249, pp. 3607–3611, 2014.

[8] 藤本一寿, 坂田展甫, 谷田雅敏. 自動車騒音の不快感に関する現場実験と実験室実験の比較. 日本建築学会計画系論文報告集, Vol. 364, pp. 10–19, 1986.

[9] Malina Szychowska, Honorata Hafke-Dys, Anna Preis, Jędrzej Kociński, and Paweł Kleka. The influence of audio-visual interactions on the annoyance ratings for wind turbines. *Appl. Acoust.*, Vol. 129, pp. 190–203, January 2018.

[10] Zhang Bangjun, Shi Lili, and Di Guoqing. The influence of the visibility of the source on the subjective annoyance due to its noise. *Appl. Acoust.*, Vol. 64, No. 12, pp. 1205–1215, December 2003.

[11] 宮川雅充, 鈴木真一, 青野正二, 高木興一. 視覚情報が種々の環境音の印象に与える影響. 日本音響学会誌, Vol. 56, No. 6, pp. 427–436, 2000.

[12] 朝倉巧. 聴感実験における視覚情報の影響. 日本音響学会誌, Vol. 74, No. 12, pp. 655–662, 2018.

[13] Sylvie Charron, Marie‐ClaireBotte. Frequency selectivity in loudness adaptation and auditory fatigue. *The Journal of the Acoustical Society of America*, Vol. 83, No. 1, pp. 178–187, 1988.

[14] 立木孝. 聴覚の順応現象について. 耳鼻咽喉科臨床, Vol. 58, No. 11, pp. 741–751, 1965.

[15] 矢高真一, 田中宏平, 寺田努, 塚本昌彦, 西尾章治郎. ウェアラブルコンピューティングのための状況依存音声情報提示手法. 情報処理学会論文誌, Vol. 51, No. 12, pp. 2384–2395, 2010.