

音声入力を用いた自転車利用者による危険箇所記録方式の提案と評価

佐藤 光 【コミュニケーション & コラボレーション研究室】

Proposal and Evaluation of a Voice Input Method for Recording Hazardous Locations by Bicycle Users

SATO, Ko 【Communication and Collaboration Lab.】

1 はじめに

日本では自転車が広く利用される一方、交通事故が社会課題となっている [1]. 主な危険要因として、静的な「狭い道路」、路上駐車などの動的な「障害物」、歩行者の「密集」が挙げられるが、既存のナビゲーションアプリでは、利用者が感じる主観的な危険情報をリアルタイムに共有する仕組みが不十分である。従来の自転車安全支援システムは、加速度センサー等の客観的データに依存しており、恐怖や不安といった主観的危険を捉えることが困難であった。そこで本研究では、利用者が危険を感じた際に「狭い」「障害」「いっぱい」のキーワードを音声入力し、危険箇所を記録・共有する方式を提案し、その有効性を評価する。

2 提案方式

本提案方式は3つの設計原則に基づく。第一に、ハンズフリー操作による安全性の確保である [2]. 第二に、主観的危険認知の即時記録である。危険を感じた瞬間に発話することで、記憶バイアスを排除する。第三に、低認知負荷な入力インターフェースである。事前定義された短いキーワードを発話するのみで記録が完了する設計とした。

3 実験

室内シミュレーション環境にて、固定自転車を用いた走行実験を実施した。被験者は一人称視点の走行映像 (11.5km/h) に合わせてペダルを漕ぎ、襟元のマイクを用いて、危険認知時に指定されたキーワードを発話した。音声処理には Vosk エンジンを使用した。実験は、昼間環境 (被験者 30 名) と、その後同一被験者 15 名による夜間環境の 2 条件で実施した。夜間条件では、街灯を 30m 間隔で配置し、道路交通規則に準拠した前照灯 (前方 10m の視認性を確保) を点灯させた。いずれの条件においても、3 種類の危険場面 ((1) 狭路, (2) 障害物, (3) 密集) 図 1 を各 3 回、計 9 試行ランダムに提示した。提案方式の 3 つの設計原則の有効性を検証するため、以下の評価指標を設定した。原則 1 および原則 3 に対しては、走行環境下での操作性と認知負荷を測定した。原則 2 に対しては、危険認知から発話までのタイ

ミングを測定した。また、原則 3 に対しては音声認識性能も評価した。



図 1 危険箇所と照度条件. 上段：昼間環境, 下段：夜間環境. 左から狭路, 障害物, 密集.

4 結果と考察

4.1 音声認識性能

提案方式を用いて危険箇所が正しく記録された割合 (記録成功率) は 84.1% (302 件中 254 件) であった。一方、環境音による誤作動等を除外した、走行環境下での音声認識の正確性は 93.0% (273 件中 254 件) と高い水準を示した。

キーワード別の詳細を表 1 に示す。「いっぱい」が F1-score 0.967 と最も安定していた一方、「障害」は Precision 1.000 を記録したものの、「狭い」への誤分類 (7 件) により Recall が低下した。記録失敗 (全 48 件) の内訳は、環境音の誤認識が 58.3% (28 件) と過半数を占めた。なお、場面別の成功率 (狭路 84.5%, 障害物 80.0%, 密集 87.9%) に大きな乖離は見られなかった。

「障害」の誤分類は、「狭い」との音韻的類似性が原因であると推察される。失敗の主因である環境音については、VAD (Voice Activity Detection) の導入によりノイズを除去することで、記録成功率を認識の正確性と同等水準 (約 94%) まで向上できる可能性が示唆された。また、場面間で成功率に偏りがないことから、提案方式は多様な交通状況下において一定の堅牢性を

表1 キーワード別の音声認識性能

キーワード	Precision	Recall	F1-score	Support
狭い	0.870	0.946	0.906	92
障害	1.000	0.899	0.947	89
いっぱい	0.989	0.946	0.967	92

有すると考えられる。

4.2 キーワードの受容性と認知負荷

キーワードの「話しやすさ」には有意差が認められ ($p=0.001$)、「狭い」(平均 4.53) が最も高く評価された。形容詞である「狭い」は日常会話で頻繁に用いられるため心理的抵抗が低い一方、「障害」は語彙として硬く、口語的表現としての違和感が指摘された。一方で、3条件間で認知負荷の有意差は認められず、提案方式の汎用性が示された。これは、日常の運転における発話行為の慣れや、予測可能なシミュレータ環境が負荷を抑制したためと考えられる。また、全キーワードで心理的抵抗に昼夜の有意差は認められなかった。暗所においても「他者から見られている可能性」という潜在的な社会的要因が心理的ハードルに影響しており、発話抵抗感も物理的な明るさよりも、利用場面における社会的な文脈に依存することが示唆された。

4.3 発話タイミングと認知プロセス

危険箇所通過時刻と発話時刻の差分(以下、先行時間)を指標とし、環境条件および危険種別が認知から発話までのプロセスに与える影響について分析した。図2は昼夜環境における発話タイミング(先行時間)の比較を示す。昼環境($n=15$)と夜環境($n=15$)の発話タイミングについて、対応のあるt検定を用いて比較した所、3つの危険箇所すべてにおいて、昼間環境と夜間環境の発話タイミングに統計的有意差は認められなかった(狭路 $p=.1149$, 障害物 $p=.8753$, 密集 $p=.1084$)。これは本実験の照明条件により、危険箇所の検出に必要な視認性が昼環境と同程度に維持されたと考えられる。このことは実環境においても街灯と自転車ライトが適切に整備されていれば、提案方式が夜間時においても昼間同様に機能する可能性を示唆している。また、主観評価では狭路および密集において夜環境で認知が遅れた ($p < .05$) が、客観的な発話タイミングには昼環境との有意差が生じていない。これは、視認性の低下により危険認知時点は遅延するものの、認知後は速やかに音声入力プロセスへ移行可能であり、時間的余裕が維持されたと推察される。

昼環境における先行時間について、危険種別(障害物, 狭路, 密集)を要因としたフリードマン検定を行った結果、平均時間は障害物 3.23 秒, 狭路 3.05 秒, 密集 3.23 秒であり、統計的有意差は認められなかった ($p=.121$)。一方で、主観評価「どのくらい前から危険だと感じていたか」についてフリードマン検定およびボンフェローニ

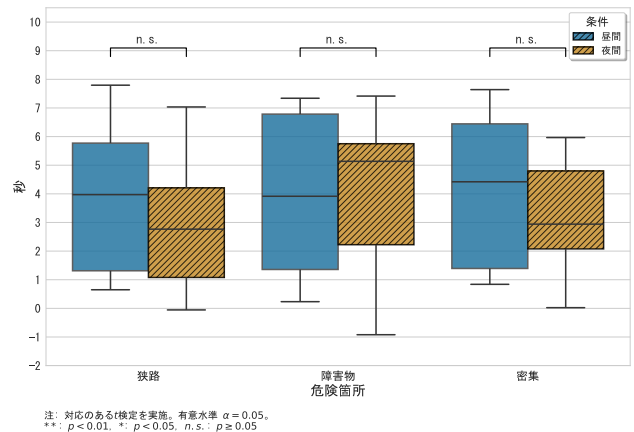


図2 発話タイミングの昼夜比較(箱ひげ図)

法による多重比較を行ったところ、障害物・密集の方が狭路より有意に早期に認知されていた ($p < 0.01$)。これは主観的に早く危険を認知しても、歩行者や路上駐車のような動的で予測困難な対象に対しては、具体的な行動計画や発話の確信度などのプロセスを経るため、発話のタイミングに差が出なかったと推察される。また、被験者間の個人差の寄与したことも考えられる。危険察知の閾値やリスク許容度が個々に異なるため、これが発話タイミングに影響し、平均値での有意差を打ち消した一因と考えられる。

5 まとめ

本研究では、自転車走行中の危険箇所を音声入力で記録する方式を提案した。実験の結果、本方式は走行時の環境下においても、運転操作を阻害しない低い認知負荷で運用可能であることが確認された。特に、入力キーワードに日常語を用いることで利用者の心理的抵抗が低減され、受容性が高まることが明らかとなった。また、視認性の低下により主観的な危険認知が遅れる状況下であっても、発話による記録操作自体は遅延なく遂行可能であることが示された。以上より、本提案方式は走行時の安全性や環境条件(昼夜)に左右されず、従来困難であった主観的な危険情報の即時記録を実現可能であると結論付けられる。

参考文献

- [1] 国土交通省道路局, 警察庁交通局. 安全で快適な自転車利用環境創出ガイドライン, 6 2025.
- [2] Pawel W. Wozniak, Lex Dekker, Francisco Kiss, Ella Velner, Andrea Kuijt, and Stella F. Donker. Brotate and tribike: Designing smartphone control for cycling. MobileHCI '20, New York, NY, USA, 2020. Association for Computing Machinery.