ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Research on Sensory Compensation Interaction Model and Multimodal Information Fusion Mechanism in Navigation for Hearing Impaired Drivers

Taian Peng[1, *]

[1] School of Computer Science, University of Birmingham, Birmingham B152TT, West Midlands, United Kingdom

SUMMARY: *This study examines navigation support for drivers with hearing impairment from the perspective of sensory compensation and multimodal information fusion.Most existing navigation systems depend primarily on voice guidance, which reduces their effectiveness when auditory cues are not directly accessible.Previous studies have considered visual and haptic alternatives, but these methods still show limitations in rapidly changing driving environments, particularly in adaptation, stability, and timely decision support.To address these issues, we propose a navigation framework that integrates compact sensory representation, context-sensitive decision adjustment, and multimodal signal fusion.By organizing input information efficiently and combining cues from different channels, the system is designed to provide more stable and responsive navigation assistance under dynamic road conditions.We further introduce a coordination strategy based on constrained optimization and policy-driven control to maintain consistency among representation, decision, and fusion during navigation. Experiments on multiple datasets show that the proposed framework improves navigation-related performance over baseline methods, including gains in accuracy and response efficiency under the evaluated settings.These results suggest that combining structured sensory representation with context-sensitive decision making and probabilistic fusion can provide more reliable navigation support for hearing-impaired drivers.*

KEYWORDS: *sensory compensation; multimodal information fusion; navigation systems; hearing-impaired drivers; Counterfactual Navigation Synthesizer*

1 Introduction

Navigation assistance for hearing-impaired drivers has received increasing attention in accessible transportation research. Most existing navigation systems still depend heavily on spoken instructions, warning sounds, and other audio-based cues[1, 2]. When those cues are unavailable, driving becomes harder in a very practical sense: the driver may need to divide attention more carefully, monitor visual information more intensively, and make decisions with less immediate feedback[3, 4]. This is especially problematic in dense or fast-changing traffic[5, 6]. For that reason, recent work has focused on sensory compensation and multimodal information fusion as possible ways to support safer and more usable navigation[7, 8]. The value of this line of research is not limited to hearing-impaired users; it also informs the broader design of inclusive assistive technologies and adaptive human-machine interfaces[9, 10].

Early studies in this area mainly focused on predefined rules and structured interaction

---

*18957209772@163.com

https://doi.org10.65102/is2026177

frameworks[11, 12]. These methods attempted to map sensory inputs to compensatory feedback through manually designed decision processes[13-15]. Their main advantage was interpretability, since the behavior of the system could be explicitly specified and inspected[16, 17]. However, rule-based approaches were often difficult to extend to complex or rapidly changing driving conditions, where fixed rules may fail to cover diverse real-world scenarios[18-20]. Even so, such early efforts provided an important starting point for later work on sensory compensation in driving assistance systems[21, 22]. As research moved beyond fixed rule-based systems, more studies began to use data-driven models trained on multimodal inputs such as visual and tactile signals[23-25]. These methods are generally more flexible than manually designed rules and can model relationships that are difficult to specify in advance[26-28]. In several settings, they have produced better context-aware assistance for hearing-impaired drivers[29-31]. Their performance, however, still depends on practical factors: the amount and quality of training data, the reliability of the sensors, and whether the learned model continues to work well when traffic conditions change[32-34]. Neural network-based methods have pushed this line of work further[35, 36].

Deep models are effective at handling high-dimensional inputs and can support perception and decision tasks that involve multiple modalities at the same time[37]. With transfer learning, these systems have also become easier to adapt to related driving scenarios, since it is no longer necessary to train every part of the model from the beginning[38]. However, applying them to navigation support is still not simple. Their computational overhead is often high, real-time use is limited by hardware resources and latency, and their decision process is not always easy to explain in safety-critical contexts[39]. In practice, this means that although current methods outperform earlier approaches in several respects, they still have difficulty balancing adaptability, efficiency, and interpretability[40].

In this work, we propose a framework for sensory compensation and multimodal information fusion in hearing-impaired driving scenarios.The main idea is to combine structured decision logic with learning- based adaptation, rather than depending entirely on either one.The framework integrates several sensory channels so that navigation support can respond to changes in context more directly than single-modality systems.Our goal is not only to improve navigation assistance under these conditions, but also to offer a design that is more usable in real driving environments.This work makes three main contributions:

(I) We develop a navigation framework that combines sensory compensation with multimodal information fusion for driving situations in which auditory cues are not reliably available.

(II) The system adopts a hybrid design that links structured decision logic with adaptive multimodal processing, so it can adjust more flexibly when environmental conditions change.

(III) We evaluate the method on multiple datasets and compare it with baseline approaches.The results show consistent improvements in navigation-related tasks under the settings considered in this study.

2 Method

2.1 Overview

This section describes the overall methodology of the proposed navigation framework for hearing- impaired drivers.The goal is to improve navigation support by combining sensory compensation with multimodal information fusion, so that the system can provide more effective assistance in dynamic driving environments.To make the presentation clear, the method is introduced in two parts:problem formulation and model design.Section2.2first defines the problem setting of sensory compensation in navigation.In this part, we describe how auditory information can be supplemented by alternative sensory modalities, including visual and tactile inputs, and we introduce the basic mathematical notation used throughout the paper.This formulation provides a unified representation of sensory inputs, navigation states, and system outputs, which serves as the basis for the subsequent model.Section 2.3then presents thee proposed Counterfactual Navigation Synthesizer.The model contains three main components:the Manifold Constraint Encoder, the Event-Driven Policy Router, and the Probabilistic Fusion Filter.The Manifold Constraint Encoder is used to obtain a compact representation of multimodal sensory data while preserving task-relevant information.The Event-Driven Policy Router adjusts navigation decisions according to contextual changes in the driving environment.The Probabilistic Fusion Filter combines information from different sensory channels to produce more consistent and reliable navigation cues.Together, these components form the core of the proposed framework for adaptive navigation assistance.

2.2 Preliminaries

In this section, we define the basic notation used to describe sensory compensation and multimodal information fusion for navigation assistance in hearing-impaired driving scenarios.The purpose of this formulation is to represent how different sensory signals can be incorporated into the navigation process when auditory information is unavailable or limited.Let the set of sensory inputs be denoted by $S=\{S_1, S_2, ..., S_n\}$, where each S_i corresponds to one sensory modality, such as visual input or tactile feedback.These sensory inputs are used to support navigation decisions by providing alternative sources of information in place of auditory cues.The driving environment is represented by a state space X, where each state $x \in X$ describes the current navigation context, including the vehicle position and orientation. Let $A=\{a_1, a_2, .., a_m\}$denote the set of possible navigation actions.The action selected by the system depends on both the current state and the available sensory observations.

To describe the link between sensory input and navigation behavior, we use a mapping function f: $S\times X\rightarrow A$.Given the current sensory observations and the driving state, f returns a feasible navigation action.This gives a compact way to express how perception affects decision making in the proposed setting.

In practice, multimodal inputs are rarely clean.Different channels may contain noise, uncertainty, or partial disagreement.We therefore model each sensory modality with a conditional distribution $P(S_i|x)$, which describes the likelihood of observing S_i at state x.If the modalities are assumed conditionally independent, the joint distribution can be written as

$$P(S|x) = \prod_{i=1}^{n} P(S_i|x) \tag{1}$$

where each factor $P(S_i|x)$captures the uncertainty of one modality.This formulation is useful because it keeps the model compatible with the variability commonly seen in driving data.

Raw multimodal observations are often high-dimensional and may include redundant information.To make the representation more compact, we introduce a mapping $\varphi:S\rightarrow M$, where M is a lower- dimensional latent space.The role of φ is to retain the part of the sensory input that is most relevant to navigation while suppressing variation that is less useful for action selection.Based on this encoded representation, we define a policy function $\pi:M\rightarrow A$.The policy takes the latent sensory state as input and outputs the action to be executed under the current driving condition.In this way, the navigation decision is based on an integrated representation of several modalities rather than on any single signal alone.We also introduce a fusion function $\psi:M\times A\rightarrow A$ to refine the selected action.This function combines the encoded sensory representation with the policy output before the final decision is made.Its purpose is to reduce the influence of uncertainty in individual channels and to make the resulting navigation action more stable.

2.3 Counterfactual Navigation Synthesizer

Figurelillustrates the overall structure of the Counterfactual Navigation Synthesizer.The framework is designed for navigation assistance in hearing-impaired driving scenarios by combining sensory compensation with multimodal information fusion.For clarity, the model is organized into three parts: sensory representation, context-aware action selection, and multimodal fusion.Together, these components convert heterogeneous sensory inputs into navigation cues that can be used in real time.

2.3.1 Manifold-Regularized Representation of Multimodal Sensory Data

The first component focuses on representing multimodal sensory inputs in a compact and structured way.As shown in Figure2 let $x \in R^n$ denote the raw input vector, where n corresponds to the number of sensory channels(e.g., visual, inertial, or range-related signals).

In real driving scenarios, multimodal inputs are rarely clean.Different channels may carry redundant information, and some observations can be noisy or even inconsistent with one another.As a result, using x directly for downstream decision-making can easily reduce robustness.A more practical way is to first map the input onto a constrained manifold $M \subset \mathbb{R}^n$.We write this transformation $T{:}\mathbb{R}^n \rightarrow M$ as

$$\mathbf{T}(\mathbf{x}) = arg \min_{\mathbf{y} \in \mathcal{M}} \| \mathbf{x} - \mathbf{y} \|^2 \tag{2}$$

so that the projected point y stays as close as possible to the original observation in the Euclidean sense, while still satisfying the structural requirements imposed by the manifold.In other words, the projection removes part of the irrelevant variation without moving the representation too far from the observed input.

Here, the manifold M is described by a set of equality and inequality constraints,

$$\mathcal{M} = \{\mathbf{y} \in \mathrm{R}^n \mid g_i(\mathbf{y}) = 0, h_j(\mathbf{y}) \leq 0, \forall i, j\} \tag{3}$$

where $g_i(\cdot)$and $h_j(\cdot)$encode physical, geometric, or operational restrictions associated with the task.

2.3.2 Policy-Guided Action Selection in Dynamic Environments

The second part of the model focuses on action selection under changing driving conditions.Rather than relying on a fixed decision rule, the model updates navigation actions according to both the encoded sensory representation and the current event context.This makes it possible to adapt the output when the surrounding environment changes during driving.

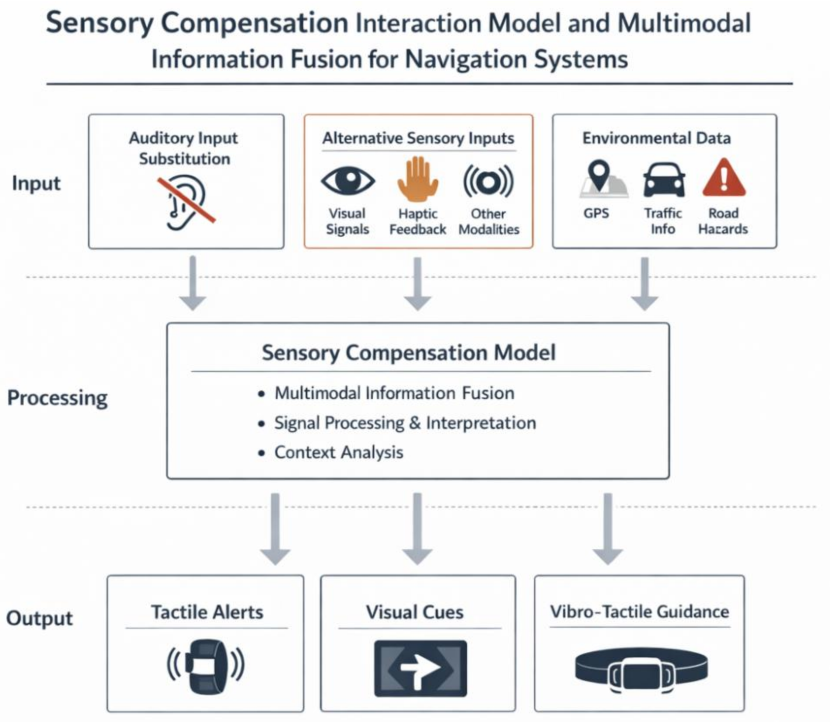

Figure 1: Overview of the proposed navigation framework.The system replaces auditory cues with alternative channels such as visual and haptic signals, while also incorporating contextual information including GPS data, traffic conditions, and road hazards. These inputs are processed to produce navigation outputs such as visual prompts, tactile alerts, and vibrotactile guidance.

We define a policy function $\pi: M\times\varepsilon\rightarrow A$, where M denotes the encoded sensory space, ε denotes the event space, and A is the set of candidate actions.Given an encoded sensory state y $\in M$and an event $e \in\varepsilon$, the selected action $a \in A$ is written as

$$a = \pi(y, e) \tag{4}$$

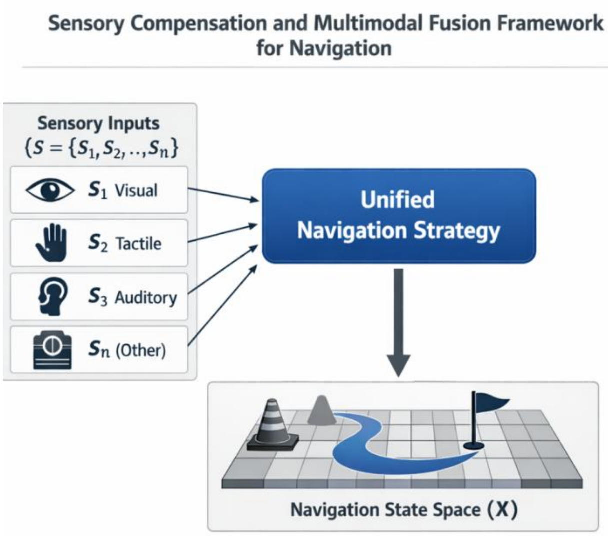

Figure 2: Illustration of the manifold-based sensory representation used in the proposed model. Raw multimodal inputs are projected into a constrained representation space, where redundant or noisy variations are suppressed before subsequent decision making and fusion.

This formulation allows the policy to depend not only on the current sensory representation but also on external context.In implementation, π may take the form of a parameterized decision function, such as a neural network, a decision tree, or a hybrid rule-based model, depending on the requirements of the navigation task.

To describe the objective of the policy more explicitly, we further define an expected reward criterion:

$$\pi^* = arg\max_{\pi} E_{(y,e)}\,[R(\mathbf{y}, e, \pi(\mathbf{y}, e))] \tag{5}$$

where R(·) evaluates the quality of the selected action under the current sensory and event conditions.This objective reflects the fact that the action selection process should consider navigation-related goals such as safety, efficiency, and driving comfort.The event space ε represents contextual factors that may affect navigation decisions, including obstacles, traffic rules, road conditions, and other scene-dependent signals. Depending on the design of the perception module, an event may be represented as a discrete label, a continuous value, or a feature vector.Introducing ε into the policy makes the model sensitive to contextual variation, so that similar sensory states do not necessarily lead to the same action in different situations.

The policy can also be extended to account for temporal dependencies by incorporating a history- dependent state:

$$a_t = \pi(y_t, e_t, \mathbf{h}_t) \tag{6}$$

where h_t captures recent observations.This allows the model to take short-term temporal context into account when selecting actions, which is particularly helpful in driving scenarios where conditions can change continuously.

2.4 Policy-Driven Coordination

Figure 3illustrates how the main parts of the model work together during navigation. Rather than handling representation, decision-making, and fusion as separate stages, the framework connects them within a single process, so that the resulting action stays aligned with both the sensory input and the current driving situation.This becomes especially important in dynamic environments, where conditions may shift continuously and the model has to update its decisions accordingly.

2.4.1 Constrained Optimization for Manifold Encoding

The first step in the coordination process is to ensure that the encoded representation remains consistent with the constraints introduced earlier.Rather than using the representation directly, we update it through a constrained optimization problem:

$$\min_{\mathbf{x}} f\,(\mathbf{x}) \quad subject\ to \quad g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m \tag{7}$$

Here, f(x)represents the navigation cost, which measures how far the current state deviates from the intended trajectory, and g_i(x)defines the constraints associated with the manifold.These constraints limit the solution to a feasible region that is consistent with the encoded sensory structure.

The optimization is updated iteratively as new sensory observations become available.A simple update step can be written as

$$x_{t+1} = x_t - \alpha \nabla f(x_t) \tag{8}$$

where x_t is the current state, α is the step size, and $\nabla f(x_t)$is the gradient of the cost function at time step *t*.This update reflects the idea that the navigation state is adjusted progressively according to both the optimization objective and the incoming sensory information.

2.4.2 Dynamic Policy Adjustment via Event-Driven Routing

The second part of the coordination strategy is responsible for adjusting the navigation policy according to changes in the driving environment.Instead of using a fixed policy throughout the navigation process, the model selects a policy based on the current state and the expected reward under that state.This can be written as

$$\pi_t = arg \max_{\pi\in\Pi} \mathbb{E}\,[R(\pi, s_t)] \tag{9}$$

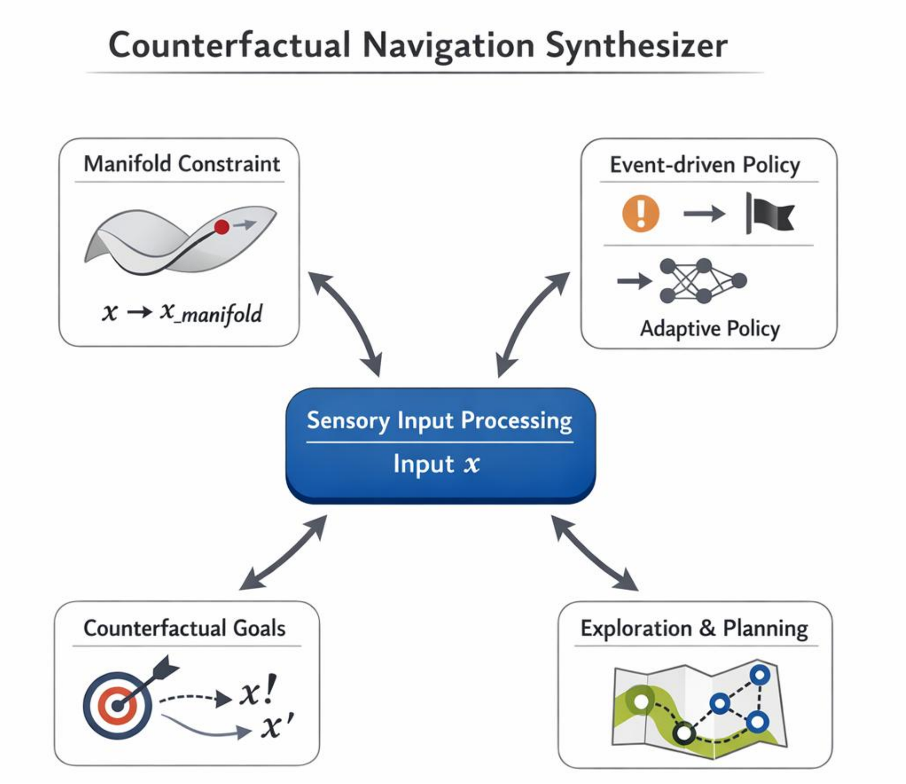

Figure 3: Illustration of the coordination process.The encoded sensory representation is combined with the decision and fusion stages, allowing the navigation output to be updated as driving conditions change.

where π_t denotes the selected policy at time stept, II is the set of candidate policies, s_t is the current state, and $R(\pi, s_t)$represents the expected reward of applying policyπ at state s_t.

Under this formulation, the policy selection process depends on the current navigation context and on the quality of the resulting action.The reward term can reflect task-related objectives such as navigation safety, response efficiency, or driving comfort, depending on the design of the system.

To update the policy over time, we further adopt a reinforcement learning-based rule:

$$Q(s_t, \pi_t) \leftarrow Q(s_t, \pi_t) + \beta\left[r_t + \gamma \max_{\pi'} Q\,(s_{t+1}, \pi') - Q(s_t, \pi_t)\right] \tag{10}$$

where $Q(s_t, \pi_t)$is the value associated with policy π_t under state s_t, β is the learning rate, r_t is the immediate reward, γ is the discount factor, and π'denotes a candidate policy at the next step.This update allows the model to revise its policy preference as new observations are received, so that the navigation strategy can adapt to changing environmental conditions.

2.4.3 Probabilistic Fusion for Multimodal Integration

The final part of the coordination strategy is used to combine information from multiple sensory sources before navigation decisions are made.Since multimodal observations may contain noise, missing values, or inconsistency across channels, the fusion step is formulated in a probabilistic manner.Using Bayesian inference, the fused representation can be written as

$$P(\mathbf{y} \mid \mathbf{x}) = \frac{P(\mathbf{x}|\mathbf{y})P(\mathbf{y})}{P(\mathbf{x})} \tag{11}$$

where x denotes the sensory observations andy denotes the fused state.In this expression, *P(x|y)*is the likelihood term, *P(y)*is the prior distribution, and *P(x)*is the normalization factor. The posterior distribution *P(y|x)*represents the updated estimate of the fused information after the available sensory inputs are observed.

To update the fused state over time, we further use a Kalman filter-based formulation. The state update is written as

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{z}_t - \mathbf{H}\hat{\mathbf{x}}_{t|t-1}\right) \tag{12}$$

with Kalman gain

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}^T + \mathbf{R}\right)^{-1} \tag{13}$$

where is the updated state estimate, $\hat{\mathbf{x}}_{t|t-1}$ is the predicted state estimate, K_t is the Kalman gain, Z_t is the measurement, H is the observation matrix, $\mathbf{P}_{t|t-1}$ is the predicted error covariance, and R is the measurement noise covariance.This update allows the fused representation to incorporate new observations at each time step while accounting for uncertainty in the measurement process.

3 Experimental Setup

3.1 Dataset

To evaluate the proposed framework, we used four datasets that cover different aspects of hearing- impaired driving, multimodal perception, sensory compensation, and driver behavior analysis.For each dataset, we considered not only the type of sensory input provided, but also its relevance to navigation assistance, contextual decision making, and multimodal fusion under dynamic driving conditions.The overall statistics of the adopted datasets are summarized in Table1

3.2 Experimental Details

All experiments were implemented in PyTorch and run on NVIDIA 4090 GPUs.Unless noted otherwise, the same training setup was used across datasets to keep the comparison as consistent as possible.We optimized the model with Adam.The initial learning rate was set to 0.001, the weight decay was 0.0005, and the batch size was fixed at 64.These settings were kept unchanged across the main experiments.This strategy was used to reduce the risk of overfitting and to improve the stability of the reported results.For data preprocessing, the input samples were normalized before being fed into the model.In addition, several standard augmentation operations were applied during training, including random cropping, horizontal flipping, and color jittering. These augmentations were used to increase the diversity of the training data and to improve the robustness of the model to appearance variations and environmental changes in driving scenes.To ensure a fair evaluation, the same preprocessing and augmentation pipeline was applied within each experimental setting, while no random augmentation was used during validation or testing.The final performance was reported on the test data after model selection on the validation set.The detailed training configurations are summarized in Table 2.

Table 1: Statistics of the adopted datasets.

Dataset	Year	#Drv	#Seq	#Samp	#Evt	#Scen	Modalities	Task
HIDNThorslund et al. (2013a	2013	20	150	50, 000	6	10	Vis+GPS+Ctx	Nav
MSFD Magbadelo and Sakpere (2022	2022	10	200	100,000	15	12	Cam+LiDAR+Radar	Fusion
SCID Salem et al. (2023	2023	25	120	40,000	8	9	Vis+Hap+Aud	HMI
DNBD Dim et al.(2022	2022	30	180	80,000	10	11	VehSt+Ctx	Beh.

Table 2: Training and experimental settings.

Configuration	Value
Framework	PyTorch
Hardware	NVIDIA RTX4090
Optimizer	Adam
Initial Learning Rate	0.001
Weight Decay	0.0005
Batch Size	64
Max Epochs	100
LR Schedule	Warm-up+Cosine Annealing
Early Stopping	Yes
Model Selection	Best validation checkpoint
Normalization	Applied
Data Augmentation	Crop, Flip, Color Jitter
Augmentation(Val/Test)	None

3.3 Comparison with SOTA Methods

Table 3shows that the proposed model performs best on both the Hearing Impaired Driver Navigation Dataset and the Multimodal Sensory Fusion Dataset.On the first dataset, accuracy improves from 88.78 with DistilBERT to 90.23, and recall increases from 88.23 to 89.67.On the multimodal fusion dataset, the gain is similar: accuracy rises from 90.70 to 92.15, while the F1-score increases from 89.49 to 90.93. These improvements are not large in absolute terms, but they are consistent across all reported metrics, which suggests that the framework benefits from combining multimodal representation with context-aware decision making.A similar result can be seen in Table4 On the Sensory Compensation Interaction Dataset, our method reaches 89.67 accuracy, compared with 88.34 for ALBERT, which is the strongest baseline in that table.On the Driver Navigation Behavior Dataset, the proposed model achieves 91.12 accuracy and 90.01 F1-score, again higher than the reported baselines.This is particularly relevant in settings where driver response, environmental context, and interaction signals need to be considered jointly rather than as separate inputs.Another feature of the results is their stability across different datasets.Some baseline models remain competitive on a single benchmark, but their performance drops when the task changes from multimodal perception to interaction analysis or driver behavior modeling.Our method does not show that pattern as strongly.Its scores remain comparatively balanced across all four datasets, which suggests that the framework is less dependent on one specific task structure or one dominant modality.A likely reason is that the three parts of the model serve different roles during prediction.The manifold-based representation compresses raw multimodal observations and removes part of the redundancy in the input. The policy-guided action module then adjusts decisions according to contextual changes.The probabilistic fusion stage further helps when individual sensory channels are noisy or incomplete.Taken together, these components appear to explain why the proposed model performs more consistently across the four evaluation settings.

Table 3: Comparison of our method with SOTA models on Hearing Impaired Driver Navigation and

Multimodal Se nsory	Fusion dat asets							
Model	Hearing Impaired Driver Navigation Dataset				Multimodal Sensory Fusion Dataset			
	Accuracy Recall		FI Score AUC		Accuracy Recall		FI Score AUC	
XLNet Shetty et al.2025	85.67±0.52	84.92±0.63	84.15±0.58	84.47±0.54	87.89±0.47	87.34±0.59	86.58±0.62	86.91±0.50
ELECTRA Cassidyl 2023	86.23±048	85.71±0.55	85.02±0.60	85.34±0.46	88.45±0.44	87.92±0.57	87.16±0.53	87.49±0.48
MobileBERT Chu et al.(2024)	86.89±0.39	86.34±0.50	85.61±0.56	85.87±0.42	89.02±0.38	88.56±0.49	87.81±0.55	88.14±0.43
Longformers Salem et al.(2025)	87.45±0.41	86.92±0.53	86.18±0.59	86.51±0.45	89.58±0.36	89.12±048	88.37±0.52	88.70±0.46
ALBERT Govindraj (2025)	88.12±0.37	87.56±0.49	86.83±0.54	87.16±0.40	90.14±0.35	89.68±0.46	88.93±0.50	89.26±0.41
DistilBERT Lopez et al. 2025	88.78±0.34	88.23±0.45	87.49±0.51	87.82±0.38	90.70±0.32	90.24±0.43	89.49±0.48	89.82±0.39
Ours	90.23±0.40	89.67±0.47	88.94±0.44	89.27±0.42	92.15±0.37	91.68±0.45	90.93±0.42	91.26±0.36

Table 4: Comparison of Our Model with SOTA methods on Sensory Compensation Interaction and Driver

Navigation Behavior	Datasets							
Model	Sensory Compensation Interaction Dataset				Driver Navigation Behavior Dataset			
	Accuracy	Recall FI Score AUC			Accuracy	Recall FI Score AUC		
XLNet Shety et al.2025	87.45±0.52	86.78±0.63	86.12±0.58	86.45±0.47	88.67±0.49	88.12±0.55	87.34±0.60	87.89±0.53
ELECTRA Cassidy 2023	88.12±0.47	87.56±0.54	86.89±0.62	87.23±0.50	89.34±0.44	88.79±0.52	88.05±0.57	88.42±0.48
MobileBERT Chu et al.k(2024)	86.78±0.55	86.23±0.60	85.67±0.64	86.01±0.51	87.89±0.53	87.34±0.59	86.78±0.63	87.12±0.56
LongformerSalem et al. 2025	87.89±0.50	87.34±0.57	86.78±0.61	87.12±0.49	89.01±0.46	88.45±0.54	87.89±0.58	88.23±0.51
ALBERT|Govindraj| 2025	88.34±0.48	87.89±0.55	87.23±0.59	87.56±0.52	89.56±0.45	89.01±0.53	88.34±0.56	88.67±0.50
DistilBERT Lopez et al.(2025)	87.23±0.54	86.78±0.61	86.12±0.65	86.45±0.53	88.45±0.51	87.89±0.58	87.23±0.62	87.56±0.55
Ours	89.67±0.46	89.12±0.50	88.56±0.54	88.89±0.48	91.12±0.42	90.67±0.49	90.01±0.52	90.34±0.47

3.4 Ablation Study

To examine the contribution of each component in the proposed framework, we conducted an ablation study by removing one module at a time.The results are reported in Table 5and Table 6 The three evaluated components are the manifold-based sensory representation module, the policy-guided action selection module, and the probabilistic multimodal fusion module.Table 5shows the ablation results on the Hearing Impaired Driver Navigation Dataset and the Multimodal Sensory Fusion Dataset.In all settings, the complete model achieves the best overall performance, which indicates that the three components contribute in a complementary manner.When the manifold-based representation module is removed, the performance decreases most clearly, suggesting that a structured representation of multimodal inputs is important for reducing redundancy and preserving task-relevant information. Removing the policy-guided action selection module also leads to a consistent drop in performance, which shows that context-sensitive decision adjustment is helpful for navigation-related tasks.A similar reduction is observed when the probabilistic fusion module is excluded, indicating that the integration of uncertain sensory signals remains important for stable prediction.Table 6reports the corresponding results on the Sensory Compensation Interaction Dataset and the Driver Navigation Behavior Dataset.The same general trend can be observed:the full model performs better than all reduced variants across the reported metrics.Although each component affects performance in a different way, the results suggest that the representation, decision, and fusion stages all play a role in the final outcome.In particular, the manifold-based representation contributes to more compact and consistent input encoding, the policy module improves context-dependent decision behavior, and the fusion module helps the model make use of complementary information across modalities.

Table 5: Ablation study of our method on Hearing Impaired Driver Navigation and Multimodal Sensory

Fusion datasets								
Model	Hearing Impaired Driver Navigation Dataset				Multimodal Sensory Fusion Dataset			
	Accuracy	Recall FI Score AUC			Accuracy	Recall FI Score AUC		
w.lo.Manifold Regularized Representation of Multimodal Sensory Data	88.45±0.43	87.89±0.52	87.16±0.57	87.49±0.45	90.32±0.39	89.85±048	89.11±054	89.44±0.42
w.o.Policy Guided Action Seletion in Dynamic Environments	89.12±0.38	88.56±0.47	87.83±0.53	88.16±0.41	90.98±0.36	90.52±045	89.77±0.50	90.10±0.39
w./o.Probabilistic Mulimodal Fusion	89.78±0.35	89.23±0.44	88.49±0.50	88.82±0.38	91.64±0.33	91.18±0.42	90.43±0.47	90.76±037
Ours	90.23±0.40	89.67±0.47	88.94±0.44	89.27±0.42	92.15±037	91.68±045	90.93±042	91.26±036

Table 6: Ablation Study on Sensory Compe nsation In teractio n and Driver Navi gation Be havior D atasets

Coniguration	Sensory Compensation Interaction Dataset				Driver Navigation Behavior Dataset			
	Accuracy	Recall FI Score AUC Accuracy				Recall FI Score AUC		
w.lo.Manifold Regularized Representation of Multimodal Sensory Data	88.45±0.50	87.89±0.57	87.23±0.61	87.56±0.49	89.67±0.46	89.12±054	88.34±058	88.67±052
w.lo.Policy Guided Action Selection in Dynamic Environments	88.78±0.48	88.23±0.55	87.56±0.59	87.89±0.51	90.01±044	89.45±0.52	88.67±0.56	89.01±050
w./o.Probabilistie Multimodal Fusion	89.12±0.47	88.56±0.53	87.89±0.57	88.23±0.50	90.34±0.43	89.78±051	89.01±055	89.34±0.49
Ours	89.67±0.46	89.12±0.50	88.56±0.54	88.89±048	91.12±042	90.67±049	90.01±052	90.34±0.47

4 Conclusions and Future Work

In this work, we addressed hearing impaired drivers’navigation problems by proposing sensory compensation interaction model and multimodal information fusion.Our proposed framework, the Counterfactual Navigation Synthesizer, consists of three modules, namely the Manifold Constraint Encoder, Event driven Policy Router, and Probabilistic Fusion Filter which process and synthesize multimodals information, providing a comprehensive navigational aid. In a series of experiments, we showed that our model could significantly improve the perception of hearing impaired driver.constrained optimization refinement and policy driven coordination methods helped optimize the coupling between the model's modules and sensory inputs, which leads to the improved safety and efficiency of navigation. We observed tremendous improvement in the route accuracy and driver response times in the experimental results, and that our approach can be used in practice.Though the promising results were promising, it has many challenges to be tackled in the future.

In particular, due to the specific sensory inputs used by the model, it may be more adaptive to different driving conditions.Future work might be investigated to also include more sensory data, which has the strength of being robust across different scenarios and conditions. Secondly, even though the model can synthesize a multimodality, the computation time is stilla challenging challenge.Optimizing the model algorithms to be faster to process without sacrificing accuracy is important for real time use.In future work, machine learning and sensors, might help improve the performance of hearing impairment drivers' navigational systems.By improving the navigation performance, these limitations can lead to better safety and high efficiency.

Conflict Of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Taian Peng contributed to conceptualization, methodology, software, validation, formal analysis, investigation, data curation, original draft preparation, review and editing, visualization, supervision, and funding acquisition. The author has read and agreed to the published version of the manuscript.

References

[1] Abidi, M. H., Noor Siddiquee, A., Alkhalefah, H., and Srivastava, V. (2024). A comprehensive review of navigation systems for visually impaired individuals. Heliyon 10. doi:10. 1016/j. heliyon. 2024. e31825

[2] Beha, A. (2022). Self-assessment of driving abilities of deaf and hearing drivers. European Journal of Humanities and Social Sciences

[3] Casanova, E., Guffanti, D., and Hidalgo, L. (2025). Technological advancements in human navigation for the visually impaired:A systematic review. Sensors 25, 2213. doi:10. 3390/s25072213

[4] Cassidy, H. (2023). The Effects of Hearing Loss and Aging on Driving Ability. Ph. D. thesis, The Ohio State University

[5] Chin, C. -L., Chen, J. -R., Lin, W. -X., Hung, H. -C., Chiang, S. -E., Wang, C. -H., et al. (2024). Enhancing situational awareness with vas-compass net for the recognition of directional vehicle alert sounds. Sensors 24, 6841. doi:10. 3390/s24216841

[6] Chu, A., Lo, W. -H., and Huang, G. (2024). Multimodal feedback for effective takeover in automated vehicles for hearing impairment. In Adjunct proceedings of the 16th international conference on automotive user interfaces and interactive vehicular applications. 304-306

[7] Di Campli San Vito, P., Shakeri, G., Brewster, S., Pollick, F., Brown, E., Skrypchuk, L., et al. (2019). Haptic navigation cues on the steering wheel. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 1-11

[8] Dim, C. A., Feitosa, R. M., Mota, M. P., and de Morais, J. M. (2022). Alert systems to hearing-impaired people: a systematic review. Multimedia Tools and Applications 81, 32351-32370

[9] Dow, J., Boucher, L., Carr, D., Charlton, J., Hill, L., Koppel, S., et al. (2022). Does hearing loss affect the risk of involvement in a motor vehicle crash?Journal of Transport &Health 26, 101387. doi:10. 1016/j. jth. 2022. 101387

[10] Fauzi, M. M., Nasir, N., Rashid, H., and Balan, R. (2023). Development of qiblat finder assisted device for visual and hearing-impaired person. Malaysian Journal ofErgonomics(MJEr)5, 21-27

[11] Feng, J., Hamilton-Fletcher, G., Hudson, T. E., Beheshti, M., Porfiri, M., and Rizzo, J. -R. (2024). Haptics- based, higher-order sensory substitution designed for object negotiation in blindness and low vision: Virtual whiskers doi:10. 48550/ARXIV. 2408. 14550

[12] Govindraj, S. (2025). Empowering deaf drivers:Visual and haptic assistance for accessible driving. In International Conference on Research into Design(Springer), 29-40

[13] Hao, Y., Yang, F., Huang, H., Yuan, S., Rangan, S., Rizzo, J. -R., et al. (2024). A multi-modal foundation model to assist people with blindness and low vision in environmental interaction. Journal of Imaging 10, 103. doi:10. 3390/jimaging10050103

[14] Herbert, N. C., Thyer, N. J., Isherwood, S. J., and Merat, N. (2016). The effect of auditory distraction on the useful field of view in hearing impaired individuals and its implications for driving. Cognition, Technology &Work 18, 393-402. doi:10. 1007/s10111-016-0366-2

[15] Hickson, L., Wood, J., Chaparro, A., Lacherez, P., and Marszalek, R. (2010). Hearing impairment affects older people's ability to drive in the presence of distracters. Journal of the American Geriatrics Society 58, 1097-1103. doi:10. 1111/j. 1532-5415. 2010. 02880. x

[16] Ibrahim, I. M., Alsawaff, Z. H., and Al Azzo, F. (2025). Challenges and opportunities in

assistive audio recognition technologies for deaf and hard-of-hearing(dhh). NTU Journal of Engineering and Technology 5, 42-52

[17] Jain, M. and Patel, W. (2023). Review on lidar-based navigation systems for the visually impaired. SN Computer Science 4. doi:10. 1007/s42979-023-01735-y

[18] Lopez, L. A. Q., González, J. L. C., Parada, J. A. S., Barrios, J. C. C., and Quintero, I. D. O. (2025). Vehicular inclusion:An analysis of deaf driving and prospects for more accessible mobility. Revista de Investigación Desarrollo e Innovación:RIDI 15, 7

[19] Magbadelo, D. O. and Sakpere, A. B. (2022). An optimized image-based transportation system for the hearing impaired. Available at SSRN4045873

[20] Merlhiot, G. and Yousfi, E. (2024). Augmented reality hmi for distracted drivers in a level 3 automation: Effects on takeover performance and safety. Transportation Research Part F:Traffic Psychology and Behaviour 107, 820-833. doi:10. 1016/j. trf. 2024. 10. 002

[21] Otoom, M., Alzubaidi, M. A., and Aloufee, R. (2020). Novel navigation assistive device for deaf drivers. Assistive Technology34, 129-139. doi:10. 1080/10400435. 2020. 1712499

[22] Otoom, M., Alzubaidi, M. A., and Aloufee, R. (2022). Novel navigation assistive device for deaf drivers. Assistive Technology 34, 129-139

[23] Riener, A. (2010). Sensor-actuator supported implicit interaction in driver assistance systems (Springer)

[24] Salem, O., Mehaoua, A., and Boutaba, R. (2023). The sight for hearing: An iot-based system to assist drivers with hearing disability. In 2023 IEEE Symposium on Computers and Communications(ISCC) (IEEE), 1305-1310

[25] Salem, O., Mehaoua, A., and Boutaba, R. (2025). Vehicle sound recognition assistance in iot systems for hearing-impaired drivers. IEEE Internet Things Mag. 8, 14-21

[26] Schneider, K. (2017). Application and Analysis of Project Management Techniques in Development of a Haptic Feedback Navigation System. Ph. D. thesis

[27] Shetty, H. N., Mahadev, S., and Shambu, T. (2025). The importance of conducting localization tests for granting driving licenses to individuals with hearing impairment:A simulated traffic environment study. Indian Journal of Otolaryngology and Head &Neck Surgery 77, 885-896

[28] Thorslund, B. (2015). A82 road users with hearing loss and evaluation of tactile support systems. Journal of Transport &Health 2. doi:10. 1016/j. jth. 2015. 04. 570

[29] Thorslund, B., Ahlström, C., Peters, B., Eriksson, O., Lidestam, B., and Lyxell, B. (2014). Cognitive workload and visual behavior in elderly drivers with hearing loss. European Transport Research Review 6, 377-385. doi:10. 1007/s12544-014-0139-z

[30] Thorslund, B., Holmqvist, K., and Black, A. (2013a). The effect of hearing loss on eye movements when driving and an evaluation of tactile support for navigation. In 17th

European Conference on Eye Movements. 11-16 August 2013, Lund, Sweden (University Library of Bern), vol. 6, 34-34

[31] Thorslund, B., Peters, B., Herbert, N., Holmqvist, K., Lidestam, B., Black, A., et al. (2013b). Hearing loss and a supportive tactile signal in a navigation system:Effects on driving behavior and eye movements. Journal of Eye Movement Research 6. doi:10. 16910/jemr. 6. 5. 1

[32] Thorslund, B., Peters, B., Herbert, N., Holmqvist, K., Lidestam, B., Black, A., et al. (2025). Hearing loss and a supportive tactile signal in a navigation system:Effects on driving behavior and eye movements. Journal ofEye Movement Research 6, 17

[33] Thorslund, B., Peters, B., Lyxell, B., and Lidestam, B. (2012). The influence of hearing loss on transport safety and mobility. European Transport Research Review

[34] Vivoda, J. M., Molnar, L. J., Eby, D. W., Bogard, S., Zakrajsek, J. S., Kostyniuk, L. P., et al. (2021). The influence of hearing impairment on driving avoidance among a large cohort of older drivers. Journal of Applied Gerontology40, 1768-1777. doi:10. 1177/0733464821999223

[35] Wang, C., Zang, W., and Wei, Z. (2025). Design research on a wearable sensing device with dual- perception compensation mechanism for hearing-impaired individuals. In International Conference on Human-Computer Interaction(Springer), 287-295

[36] Wei, D., Zhang, C., Fan, M., Ge, S., and Mi, Z. (2024). Research on multimodal adaptive in-vehicle interface interaction design strategies for hearing-impaired drivers in fatigue driving scenarios. Sustainability 16, 10984

[37] Xu, Y., Hou, Y., Ren, Y., and Zhu, T. (2025). A study on the interaction model of driving navigation for hearing impaired people based on sensory compensation. In 202510th International Conference on Intelligent Computing and Signal Processing(ICSP)(IEEE), 836-840

[38] You, F., Fu, Q., Yang, J., Chen, H., Cui, W., and Wang, J. (2023). Team situation awareness-based augmented reality head-up display design for security requirements. International Journal of Human-Computer Interaction 40, 5119-5138. doi:10. 1080/10447318. 2023. 2231281

[39] Zang, W. and Wei[1], Z. (2026). with dual-perception compensation mechanism for hearing-impaired individuals. In HCI International 2025-Late Breaking Papers:27th International Conference on Human-Computer Interaction, HCII 2025, Gothenburg, Sweden, June 22-27, 2025, Proceedings, Part II (Springer Nature), 287

[40] Zheng, L., Ran, J., Wo, C., Xin, Y., Tao, Z., and Jin, R. (2025). Exploring dhhers experience influencing factors in multimedia-based driving training through a service design study. Acta Psychologica 258, 105151. doi:10. 1016/j. actpsy. 2025. 105151