SEMULIN
Schlussbericht gem. Nr. 8.1 NKBF 98

Gefördert durch:

Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

Zuwendungsempfänger: Infineon Technologies AG, 81726 München	**Förderkennzeichen:** 19A20012D
Vorhabenbezeichnung: Verbundprojekt **SEMULIN**: Selbstunterstützende Multimodale Interaktion, Teilvorhaben Infineon Technologies AG: Ultraschallbasierende Gestenerkennung	
Laufzeit des Vorhabens: 01.11.2020 – 30.04.2024	
Berichtszeitraum: **01.11.2020 – 30.04.2024**	**Berichtsdatum:** 11.10.2024
Vorgelegt beim Projektträger: TÜV Rheinland Consulting GmbH Projektträger Bodengebundene Verkehrstechnologien Am Grauen Stein, 51105 Köln Ansprechpartner: Fr. Shilan Derakie, Hr. Philipp Wolf	
Vorgelegt durch: Simon Achatz, Dr. Christian Bretthauer, Elfi Fertl, Daniela Maier, Juan Mena Carrillo Telefon: +49 89 234 65292, Mail: daniela.maier@infineon.com	

Inhalt

1. Überblick Gesamtvorhaben ... 3
1.1. Ziele Gesamtvorhaben ... 3
1.2. Ergebnisse Gesamtvorhaben ... 3
2. Kurze Darstellung der Aufgabenstellung, der Voraussetzungen, der Planung und des Ablaufs des Teilvorhabens ... 4
2.1. Aufgabenstellung Teilvorhaben Infineon Technologies AG ... 4
2.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde ... 4
2.3. Planung und Ablauf des Vorhabens ... 5
2.4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde ... 5
3. Eingehende Darstellung der erzielten Ergebnisse und des voraussichtlichen Nutzens ... 6
3.1. AP1.a Intelligente Subsysteme ... 6
3.1.1. Task 1a.3 Teil I: Hardware zur Gestenerkennung ... 7
3.1.2. Task 1a.3 Teil II: Software zur Gestenerkennung ... 9
3.1.3. Task 1a.6 / Task 1a.7 Aufnahme und Aufbereitung von Blickdaten / Aufnahme und Aufbereitung der Sensordaten zur Entwicklung der Einzelkomponenten ... 12
3.2. AP1.b Systemarchitektur ... 13
3.2.1. Systemarchitektur und Anforderungsanalyse ... 13
3.2.2. Ausarbeitung und Anpassung der Systemarchitektur ... 14
3.2.3. Integration und Systemoptimierung ... 15
3.2.4. Finalisierung und Demonstratoraufbau ... 16
3.3. AP6 Demonstratoren ... 17
4. Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans ... 18
4.1. Während der Durchführung des Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen ... 20
4.2. Erfolgte oder geplante Veröffentlichungen des Ergebnisses ... 20
5. Abkürzungsverzeichnis ... 22
6. Abbildungsverzeichnis ... 23
7. Tabellenverzeichnis ... 23

1. Überblick Gesamtvorhaben

1.1. Ziele Gesamtvorhaben

Das übergeordnete Ziel des Projektes SEMULIN (Selbstunterstützende Multimodale Interaktion) bestand in der Entwicklung einer selbstunterstützenden, natürlichen und widerspruchsfreien Mensch-Maschine-Schnittstelle für automatisierte Fahrzeuge durch multimodale Ein- und Ausgabekonzepte.

Mit SEMULIN sollte eine selbstunterstützende, natürliche und widerspruchsfreie Kommunikation mittels multimodaler Ein- und Ausgabekonzepte ermöglicht werden. Die Natürlichkeit sollte durch die Unterstützung einer ganzheitlichen Multimodalität erreicht werden, die sich an der zwischenmenschlichen Kommunikation orientiert: Nicht eine ständige Verwendung von Aktivierungswörtern; sondern das Verständnis für den aktuellen Kontext und die Sinnzusammenhänge ermöglichen eine effektive und effiziente Kommunikation. Dafür muss ein Fahrzeug auf zahlreiche Sensoren (z.B. für Blickrichtung oder Sprache) zurückgreifen, diese Daten sofort semantisch interpretieren können (z.B. erkennen, ob die Personen miteinander oder eine Person mit dem Fahrzeug spricht) und in den aktuellen Kontext einordnen können.

Das Projekt konzentrierte sich auf Fahrszenarien des automatisierten Fahrens. Neben Fahrsituationen im automatisierten Fahrbetrieb bei höheren Automatisierungsgraden (SAE-Level 3-4/5) und Übergabeszenarien zwischen automatisiertem und nicht-automatisiertem Fahrbetrieb wurden auch Situationen des klassischen fahrergesteuerten Fahrens betrachtet.

1.2. Ergebnisse Gesamtvorhaben

Im Projektkonsortium, unter der Leitung von ElektroBit, arbeiteten Industrieunternehmen (ElektroBit, Infineon Technologies AG), Forschungsinstitute (Fraunhofer IIS, Universität Ulm) und Mittelständler (Audeering, eesy innovations, Blickshift) am gemeinsamen Ziel und den zugewiesenen Aufgaben. Über die Laufzeit wurde unter den Partnern eine zielgerichtete, offene und gute Zusammenarbeit aufrechterhalten. Die meisten der ehrgeizigen Projektziele wurden mit sorgfältiger Planung und fleißigem Einsatz erreicht; in der Zusammenarbeit wurden Lösungen erarbeitet, um technische, material- und beschaffungsbedingte Hindernisse und Risiken erfolgreich zu bewältigen. Auch die Zusammenarbeit mit dem Projektträger verlief stets freundlich und konstruktiv.

Die Teilvorhaben wurden aufgrund von Verzögerungen im Gesamtprojekt, die eine Zielerreichung im Rahmen der ursprünglich geplanten Projektlaufzeit nicht mehr ermöglichten, um sechs Monate bis zum 30. April 2024 verlängert. Nach Bewilligung der Projektverlängerung konnten die initialen Projektziele innerhalb der um sechs Monate erweiterten Projektlaufzeit erreicht werden. Die in der Vorhabenbeschreibung vorgesehenen Ergebnisse aus den Tasks sowie die Meilensteine wurden abgeschlossen. Auf die Ergebnisse des Gesamtvorhabens im Detail sei auf den gemeinsamen Abschlussbericht des Projekts verwiesen, diese werden dort beschrieben und vom Konsortialführer veröffentlicht.

2. Kurze Darstellung der Aufgabenstellung, der Voraussetzungen, der Planung und des Ablaufs des Teilvorhabens

2.1. Aufgabenstellung Teilvorhaben Infineon Technologies AG

Das Ziel von Infineon (ff. auch IFAG) war die Entwicklung einer ultraschallbasierten Gestenerkennung mit Silizium-Mikrofonen, sowie eines Audio-Entwicklungsbords zur Verarbeitung von Audiodaten. Diese Elektronikinfrastruktur sollte in der Lage sein, Audiodaten zusammen mit Ultraschall-Abstandsdaten zu verarbeiten, um Gestenerkennung im Fahrgastinnenraum zu ermöglichen; im Sinne einer ganzheitlichen Multimodalität, die sich an der zwischenmenschlichen Kommunikation orientiert. SEMULIN fokussierte auf Fahrszenarien des automatisierten Fahrens. Neben Fahrsituationen im automatisierten Fahrbetrieb in höheren Automatisierungsgraden (Level 3-4/5), und Übergabeszenarien zwischen automatisierten und nicht-automatisiertem Fahrbetrieb, sollten auch Situationen des klassischen fahrergesteuerten Fahrens betrachtet werden. IFAG leitete die Arbeitspakete AP1a „Intelligente Subsysteme“, AP1b „Systemarchitektur“ und AP6 „Demonstratoren“. Ein Arbeitsschwerpunkt lag im AP1a. Dessen Ziel war die Identifikation und Konzeption von intelligenten Teilsystemen. IFAG forschte hier, zusammen mit den Projektpartnern, an der Vorbearbeitung und Bereitstellung von aggregierten Sensordaten zur multimodalen Auswertung, um sie anschließend für ein Human-Machine-Interface (HMI) und Aktoren als Ausgabemodalitäten zu verwenden. Hier sollte die multimodale Nutzung von MEMS basierenden Silizium-Mikrofonen, neben der Gewinnung von Audiodaten auch zur Gestenerkennung, entwickelt werden. Es sollte ein Audio-Entwicklungsboard für den Anschluss von mehreren Mikrofonen entwickelt werden, um für Partner IIS ausreichend räumlich verteilte, zeitlich synchronisierte Audiodaten generieren zu können. In AP1b wurden sicherheitsrelevante Architekturen im Kontext HMI und automatisiertem Fahren erforscht, um die Spezifikation und Einrichtung einer Systemarchitektur in Form einer HMI-Plattform, deren Aufgabe die Fusionierung und Auswertung der Ein- und Ausgangssignale aus den Subsystemen ist, zu ermöglichen. Die Subsysteme (Mikrofone, Kameras, Sensoren etc.) sollten für die multimodale Verarbeitung mittels dieser Fusionsplattform Daten liefern. Im AP6 wurde die Integration der Systeme in die Demonstratoren und die geplanten Aufbauten bei den Projektpartnern koordiniert. Des Weiteren wurden im AP2 die Partner bei der Anbindung der intelligenten Sensoren unterstützt.

2.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Das Projekt SEMULIN wurde im Rahmen des Förderprogramms des BMWK „Neue Fahrzeug- und Systemtechnologien“ beantragt, mit Zuwendungsbescheid am 26.10.2020 bewilligt und durchgeführt. Es ist thematisch unmittelbar der Programmsäule “Automatisiertes Fahren” im Fachprogramm NFST zuzuordnen. Das Schwerpunktthema “Innovative Sensorik” aus dem Fachprogramm wurde durch die Verwendung intelligenter Sensoren, die bereits eine erste Stufe der Datenverarbeitung für die zu betrachtenden Eingabemodalitäten umsetzen, sowie durch die nachgelagerte Zusammenführung der Informationen mit Methoden der künstlichen Intelligenz und der Entwicklung einer hierfür angepassten Hardware- und Systemarchitektur, adressiert.

Für die Durchführung der Arbeiten zu SEMULIN waren bei der Infineon Technologies AG die notwendige Infrastruktur, die technische Ausrüstung und qualifizierte Ressourcen vorhanden. Die F&E-Aktivitäten zu SEMULIN wurden in den Geschäftsbereichen Automotive (ATV) und

Power und Sensor Systems (PSS) ausgeführt. PSS bietet eine breite Palette an Konnektivitäts-, RF- und Sensortechnologien. Hochpräzise Sensorlösungen bringen "menschliche" Sinne in IoT-Geräte und ermöglichen es ihnen, intuitiv auf ihre Umgebung zu reagieren. Dieses Knowhow stand im Rahmen der Zusammenarbeit innerhalb des Konzerns für die Anwendung im Projekt bereit. Das ATV-Portfolio umfasst Sensoren, Mikrocontroller, Hochleistungsspeicher für spezifische Anwendungen und Leistungshalbleiter auf Basis von Silizium und Siliziumkarbid sowie Komponenten für die Mensch-Maschine-Interaktion und die Fahrzeugkonnektivität. Der Bereich steuerte seine Expertise für die Entwicklung der angepassten Hardware- und Systemarchitektur bei.
Die Covid-19-Pandemie in den Jahren 2020 und 2021 hatte eine erhebliche Auswirkung auf die Projektarbeiten, da persönliche Arbeitstreffen für längere Zeit nicht möglich waren und sich die Lieferung wichtiger Materialien und Komponenten verzögerte. Dies erschwerte die Ausführung anstehender Aufgaben und die Anpassung angedachter Zeitpläne und -prozesse.

2.3. Planung und Ablauf des Vorhabens

Der Schlüssel zum Erfolg eines jeden Projektes ist eine Projektplanung, die die komplexen Aufgaben und Abhängigkeiten zwischen den Partnern effizient strukturiert. Die Abbildung zeigt das SEMULIN-Projektkonstrukt mit Entwicklungszyklen und Abhängigkeiten zwischen den Arbeitspaketen. Das Projekt SEMULIN wurde in acht Arbeitspakete aufgeteilt, die Arbeiten der einzelnen Partner im Gesamtkonsortium abgestimmt. Der Schwerpunkt der technischen Arbeiten der Infineon Technologies AG fand sich in den Arbeitspaketen (AP) 1.a, 1.b, 2 und 6. AP1a bildete die Ausgabeebene für die im AP1b verarbeiteten Modalitäten. Das AP1b beschäftigt sich außerdem mit der Systemarchitektur, das AP6 mit der Integration ins Gesamtsystem.

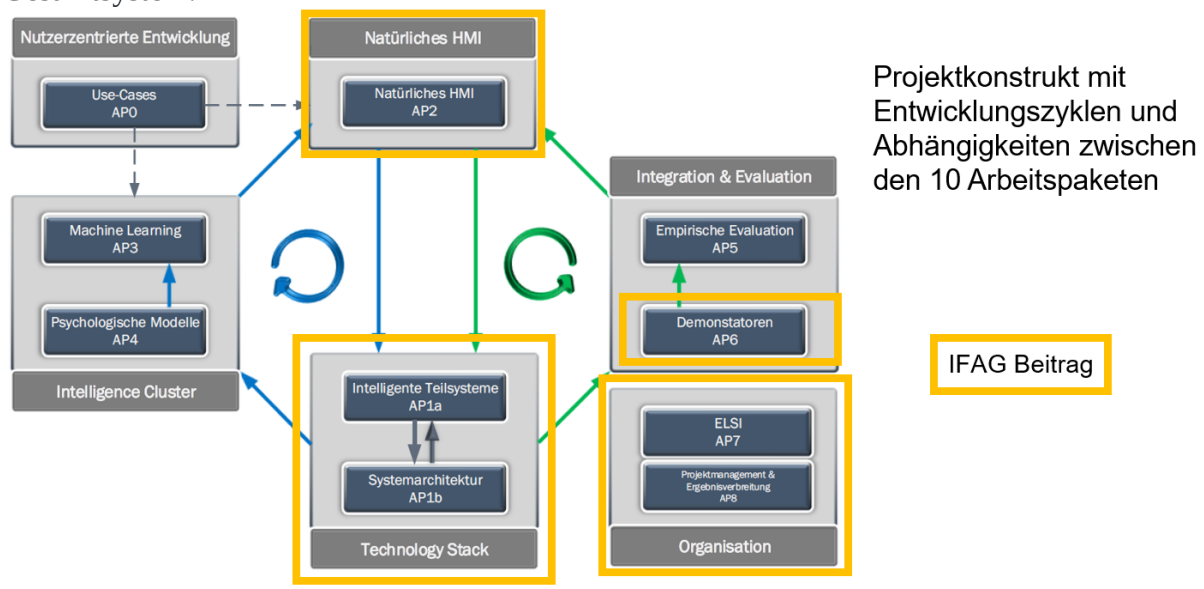

Abbildung 1: SEMULIN - Projektkonstrukt

2.4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Zum Zeitpunkt der Antragstellung beruhten Ansätze zur multimodalen Interaktion weitestgehend auf den Prinzipien der Parallelität und Redundanz. Hierdurch konnte beispielsweise eine einzelne Fahrzeugfunktion entweder durch Berührung, Sprache oder Gesten gesteuert werden, ohne die Möglichkeit, verschiedene Modalitäten zu kombinieren oder zwischen diesen zu wechseln. Dies führte zu widersprüchlichen Interaktionskontexten, und die

Bedienung des Fahrzeugs wurde in Tests ineffizient und als unnatürlich empfunden. Die Möglichkeit des Wechsels während einer Interaktionssequenz von gestenbasierter auf sprachbasierte Interaktion, ohne die Interaktionsabfolge neu zu starten, sobald der Nutzer eine Fahraufgabe mit seinen Händen übernehmen musste, fehlte. Weiterhin war es meist nicht möglich, verschiedene Modalitäten entsprechend ihrer Vorteile zu kombinieren. Die Fahrzeugindustrie versuchte, isolierte multimodale Interaktionen oder Erkennen auf Basis von verfügbaren Computerarchitekturen zu realisieren. Die Unterstützung mehrerer Modalitäten durch das Fahrzeug für explizite Eingabe und Ausgabe für natürliche Interaktion stellte somit einen wichtigen nächsten Schritt dar.
Das Förderprojekt ECSEL SILENSE, Laufzeit von 04/2017-03/2020, beschäftigte sich mit der Erforschung eines integrierten und ultraschallbasierten Konzepts zur berührungslosen Aktivierung und Steuerung von smarten Geräten oder Applikationen. Dabei wurden akustische Technologien untersucht und Konzepte zur Aktivierung und Steuerung von Geräten durch Gestenerkennung, Datenkommunikation sowie eine Indoor-Positionierung entwickelt. IFAG forscht hier an Ultraschall-Silizium-Mikrofonen für Sensor-Anwendungen mit smarter und innovativer Technologie.
Die Kerninnovation des Technologie-Stacks von SEMULIN bestand darin, dass die vorhandenen Sensoren durch, im Projekt entwickelte, softwaregestützte Verfahren in ihrem Funktionsumfang erweitert wurden. Die grundlegende Sensorhardware befand sich zu Projektstart in der Regel bereits auf einem sehr hohen Technologiereifegrad (TRL 8 oder 9). Um das Projektziel auf Basis der bestehenden Hardware-Lösungen erreichen zu können, musste jedoch die Sensorleistung mit neu zu entwickelnden, softwarebasierten Verfahren erweitert werden. Besonderes Augenmerk lag dabei auf den Ultraschallsensoren. Durch die Kombination von Audio-Sensorik und Ultraschallsensoren mit Hilfe einer Softwarelösung, die auf KI-Verfahren basiert, sollte der Technologiereifegrad von zu Beginn 2 bis 3 auf die Stufe 5 (Versuchsaufbau in Einsatzumgebung) am Projektende erhöht werden. Dies würde eine signifikante Weiterentwicklung darstellen, die die Integration und Nutzung von Ultraschallsensoren in zukünftigen Fahrzeugen erheblich verbessern sollte.

3. Eingehende Darstellung der erzielten Ergebnisse und des voraussichtlichen Nutzens

Die erzielten Ergebnisse und der voraussichtliche Nutzen, resultierend aus den Zielen und Arbeiten der Infineon Technologies AG, werden im Folgenden anhand der Arbeitspakete (AP) und Tasks dargestellt, in denen IFAG leitend bzw. überwiegend tätig war. Die dazugehörenden Meilensteine sind im AP dokumentiert.

3.1. AP1.a Intelligente Subsysteme

IFAG koordinierte das AP 1.a und damit die Zusammenführung der Arbeitsinhalte der einzelnen Partner, wie in der u.a. Übersicht angegeben.

Aufgabe	Beschreibung	
1a.1	Entwicklung von Ein-/Mehrkanaliger Quellentrennung für Sprachsignale	Fraunhofer IIS AM
1a.2	Entwicklung von Entwicklung von Echtzeit-Sprechererkennung	Fraunhofer IIS AM
1a.3	Entwicklung einer ultraschallbasierenden Gestenauswertung	Infineon Technologies
1a.4	Mimikerkennung	Fraunhofer IIS IS
1a.5	Sprechereigenschaftsanalyse	Audeering
1a.6	Aufnahme und Aufbereitung von Blickdaten	Unterauftrag IFAG / ElektroBit
1a.7	Aufnahme und Aufbereitung der Sensordaten zur Entwicklung der Einzelkomponenten	Unterauftrag IFAG / ElektroBit
1a.8	Entwicklung und Bereitstellung eines Human Centric steuerbaren Beleuchtungssystems	Eesy Innovation

Abbildung 2: AP1.a Tasks und Partner

3.1.1. Task 1a.3 Teil I: Hardware zur Gestenerkennung

Im Rahmen von AP1a, in der Task 1a.3, fanden die Arbeiten der Infineon Technologies AG zu Konzept, Design, Entwicklung und Inbetriebnahme des modularen Aufbaus für die ultraschallbasierte Gestenerkennung statt. MEMS-Mikrofone ermöglichen die Erfassung von Schallimpulsen und waren damit für den Aufbau eine entscheidende Komponente. Das Ziel der Arbeiten war ein Audio-und-Ultraschall-Entwicklungsboard für den Anschluss von mehreren Mikrofon-Ultraschall-Transceivern, bestehend aus einer MCU und mehreren MEMS – Mikrofonen, für ein stabiles Audio-Streaming zum Host. Dies wurde realisiert mit

1. der Hauptkomponente für die Gestensteuerung, dem „**Ultrahub Main Board**“, verantwortlich für die Verarbeitung der Daten und das Streaming zum Host
2. den damit zu verbindenden Sensorkomponenten, den „**Ultrahub Sensing Shields**“, mit je vier MEMS-Mikrophonen als Ultraschall- Transceivern.

Zusammenfassung der Funktionen des Aufbaus:

- Micro-USB-betriebenes Gerät
- Audio- und Ultraschalldaten-Streaming über die USB-Schnittstelle (USB-Audio 2.0)
- 192 kHz Abtastrate
- 24-Bit 4-Kanal-Audiodaten
- Modus-Schalter für DFU-Modus
- LED-Anzeigen für jeden der vier Ultraschall-Transceiver auf dem Sensorschild
- Serielle Schnittstelle zum Auslesen der Software-ID
- Serielle Schnittstelle zum Senden und Empfangen von Konfigurationsparametern
- EEPROM zum Speichern von Transceiver-, Board-ID- und Benutzerkonfigurationseinstellungen

Ultraschall Transceiver Mikrophone

Es wurde zunächst ein umfassender Entwicklungs- und Testprozess für eine große Anzahl von analogen und digitalen Ultraschall-Transceiver-Mikrophonen durchgeführt. Die wichtigsten Aspekte der Bewertung der Mikrofonleistung waren:

- Detaillierte Bewertung der Leistungsmerkmale jedes Mikrofons, wobei der Schwerpunkt auf Genauigkeit und Konsistenz lag
- Messung und Analyse des Frequenzgangs der Mikrofone, um sicherzustellen, dass sie die geforderten Spezifikationen erfüllen.

- Untersuchung der Richtcharakteristiken, um zu bewerten, wie gut die Mikrofone Schall aus verschiedenen Richtungen einfangen.

Sodann wurden Produktbeschreibungen mit den Leistungskennzahlen und Spezifikationen der getesteten Mikrofone erstellt, um die Ergebnisse zu dokumentieren.

Ultrahub Main Boards and Shields

Parallel zu den Mikrofontests wurde eine beträchtliche Anzahl (ca. 50) Ultrahub Main Boards mit den dazugehörigen analogen und digitalen Shields montiert und getestet. Die wichtigsten Aktivitäten waren:

- Zusammenbau und Tests: Aufbau und anschließendes Testen der Ultrahub-Hauptplatinen und ihrer Schilde, um den korrekten Betrieb und die Leistung zu überprüfen.
- Firmware- und Softwaretests: Überprüfung der Firmware (FW) und Software (SW) für die Hauptplatinen. Dazu gehörte auch das Testen der seriellen Schnittstellen, um sicherzustellen, dass sie korrekt funktionieren, wenn bis zu vier Geräte gleichzeitig mit einer Abtastrate von 192 kHz angeschlossen werden.

Um die ersten Ergebnisse zu dokumentieren und die weitere Entwicklung zu unterstützen, wurde eine erste Application Note (Klassifizierung: *Restricted*) erstellt und an die Projektpartner verteilt. Das Benutzerhandbuch „Infineon Ultrahub“ enthält Informationen zur Verwendung und Auswertung der analogen und digitalen MEMS-Ultraschall-Transceiver von IFAG mit Hilfe des Ultrahub-Evaluierungskits. Es macht den Nutzer mit dem Ultrahub-Evaluierungsboard vertraut und führt ihn durch die anfängliche Hard- und Software-Einrichtung sowie Messungen. Detaillierte Informationen über den Zugriff auf die serielle Schnittstelle und Protokolldefinitionen sind enthalten. Das Dokument gibt detaillierte Einblicke in die Testverfahren, Ergebnisse und Leistungskennzahlen sowohl der Ultraschall-Transceiver-Mikrofone als auch der Ultrahub-Hauptplatinen.

Von den Mikrofon-Shields wurden drei Varianten für den ElektroBit-Demonstrator und die Audio-Erfassung entwickelt und aufgebaut, um die am besten geeignete Anordnung für die ultraschallbasierte Gestenerkennung zu finden: Mit quadratisch-, Y-/sternförmig- und linear angeordneten MEMS-Mikrofonen (siehe Abbildungen). Die Anforderungen der Partner für den Aufbau wurden berücksichtigt. So wurde etwa der Abstand zwischen den Mikrofonen angepasst, um ein besseres ungestörtes Signal zu erlangen. Das „Ultrasonic-Hub Main Board“ sendet per USB-Schnittstelle die Amplitudendaten der vier Transceiver zur weiteren Verarbeitung an z.B. einen Laptop. Des weiteren können über serielle Kommunikation verschiedene Sende- und Empfangseinstellungen vorgenommen werden.

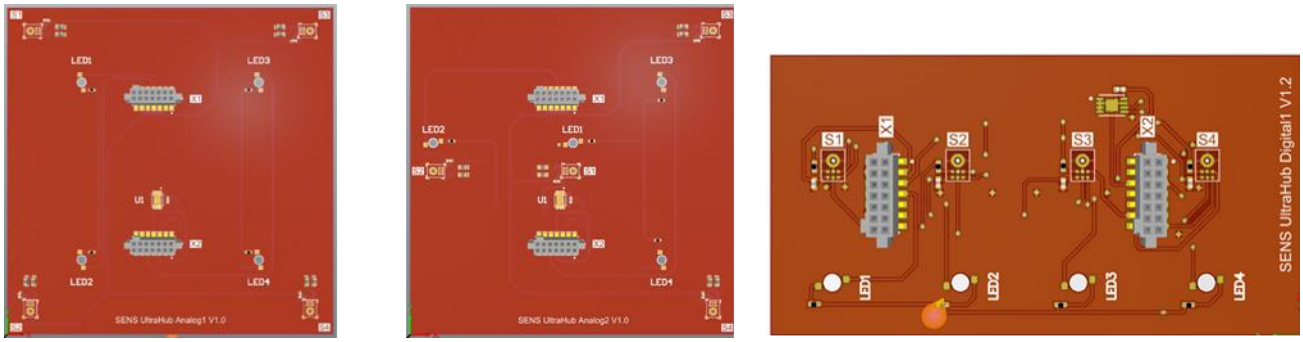

Abbildung 3: Ultrahub Sensing Shield Varianten

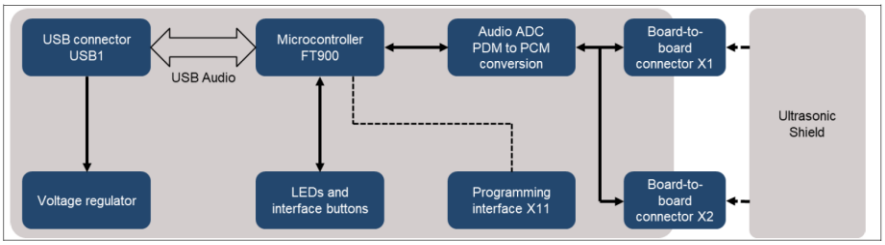

Figure 2 Ultrahub motherboard block diagram

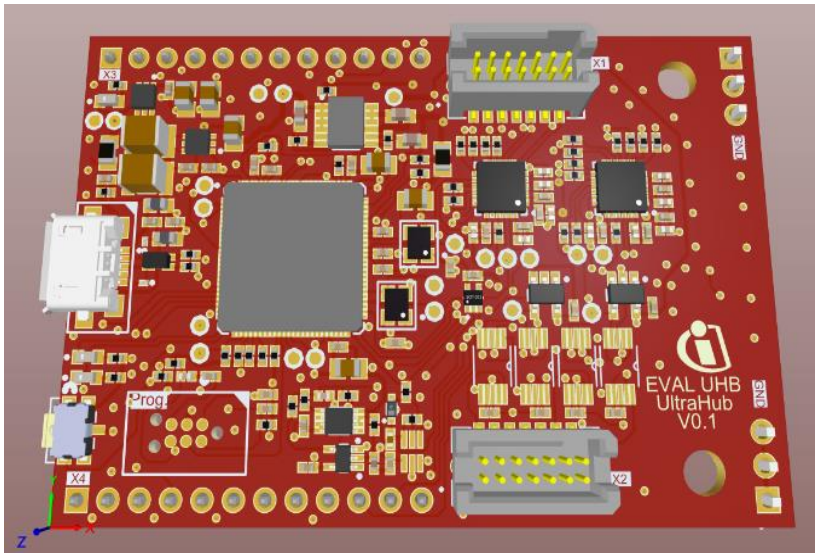

Abbildung 4: Ultrasonic-Hub Main Board

Die Hardware „Ultrasonic-Hub Main Board“ zur Ultraschalldatenverarbeitung wurde mit den verfügbaren „Ultrasonic-Hub Sensing Shield“-Versionen in Betrieb genommen, die Vorverarbeitungssoftware für die Daten entwickelt und verschiedene Ansätze getestet. Damit wurden erste Datenaufnahmen durchgeführt.

Die Projektpartner von ElektroBit, Audeering und Fraunhofer IIS wurden für Ihre weiterführenden Arbeiten mit der Hardware - Boards und Shields – mit entsprechenden Dokumentationen / User Manuals und Data Sheets durch IFAG ausgestattet. IFAG unterstützte die Partner bei der Inbetriebnahme der Hardware.

3.1.2. Task 1a.3 Teil II: Software zur Gestenerkennung

Neben der Hardware wurde eine Firmware entwickelt, die eine Verbindung mit dem

Audio/Ultraschall-Shield über eine serielle Schnittstelle ermöglichen sollte, um die Daten der verschiedenen Mikrophone auszulesen (Readout). Hierfür musste ein neues Konzept erarbeitet werden, da über die Standard-Schnittstelle ein Auslesen von bislang maximal zwei Mikrophonen (Stereo) möglich war. Mit der Softwareentwicklung wurde parallel zur Hardwareentwicklung gestartet, indem Synthetic Gesture Data und darauf basierende Algorithmen entwickelt wurden. (Siehe auch 4.2, Veröffentlichung „Object Positioning Algorithm Based on Multidimensional Scaling and Optimization for Synthetic Gesture Data Generation“).

Nach Fertigstellung der Hardware sowie der Inbetriebnahme der Komponenten „Ultrasonic-Hub Main Board“ und „Ultrasonic Sensor Shield“, erfolgte die Entwicklung der ersten Version der „Ultrasonic Main Board“ Firmware. Diese beinhaltet erste Vorverarbeitungsschritte inklusive der Bereitstellung der Amplitudendaten per USB.

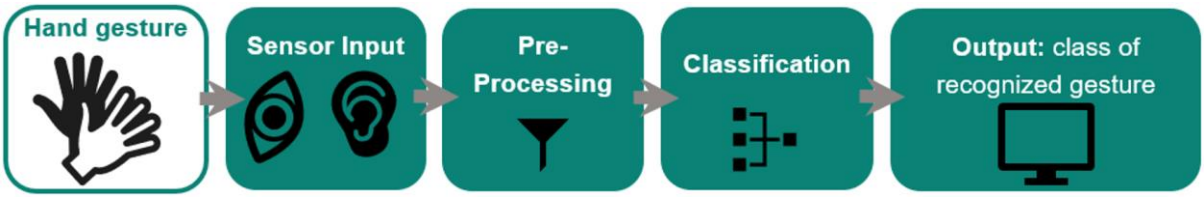

Abbildung 5: Ablaufdiagramm der Handgestenerkennung

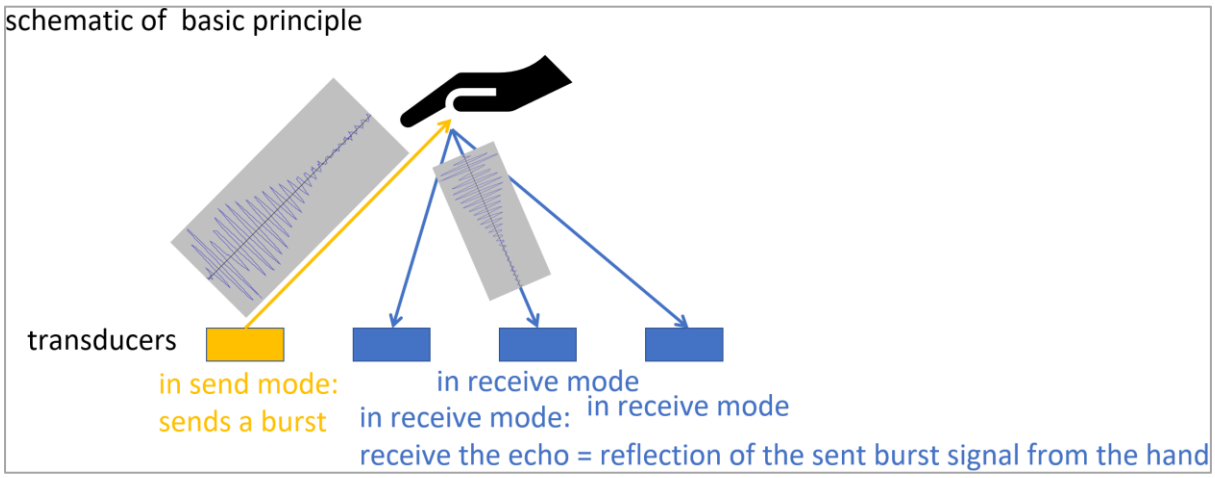

Abbildung 6: Basisprinzip der ultraschallbasierten Handgestenerkennung

Die Abbildung zeigt das Basisprinzip der ultraschallbasierten Handgestenerkennung. Angelehnt an Lidar bzw. Radar, wird ein Ultraschallpuls wiederholt gesendet und das zugehörige Echo empfangen. Die Ultrasonic-Hub Hardware ist dazu geeignet, jeden einzelnen Ultraschallwandler softwareseitig entweder als Sender und Receiver oder nur Receiver zu konfigurieren.

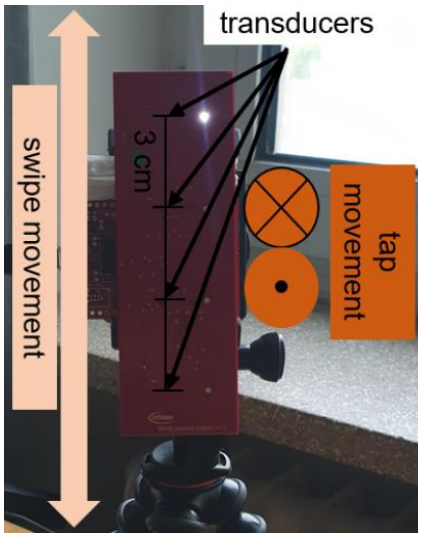

Abbildung 7: Ausführung der Handgesten

Die Ausführung der Handgesten erfolgt wie in Abbildung 7: Ausführung der Handgesten dargestellt. Es wurden drei Datensätze von Ultraschall-Handgesten aufgenommen, also ein Datensatz pro Sensing Shield. Diese Datensätze wurden in verschiedenen Formaten, mit unterschiedlichen Vorverarbeitungsschritten, abgespeichert. Das bedeutet, die Amplitudendaten, gefilterte Amplitudendaten, Spektrogramme, und Distanzdaten wurden erstellt. Vor- und Nachteile der verschiedenen Vorverarbeitungsschritte wurden verglichen. Die Software zur Datenaufnahme sowie die Firmware des Main Boards wurde für die Nutzung unter Linux angepasst. Auf Seiten der Firmware waren Anpassungen des Device-Namens und der USB-Kommunikation notwendig. Als Folge wird der Ultrasonic-Hub als CDC- und als Audio-Device korrekt erkannt.

Zwei Docker-Container, zur Datenaufnahme und Distanzberechnung (Container 1) und zur Klassifizierung auf Basis der Distanzdaten (Container 2), wurden generiert. Das Schema in Abbildung 8: Docker-Container Verarbeitungsfluss zeigt den Verarbeitungsfluss.

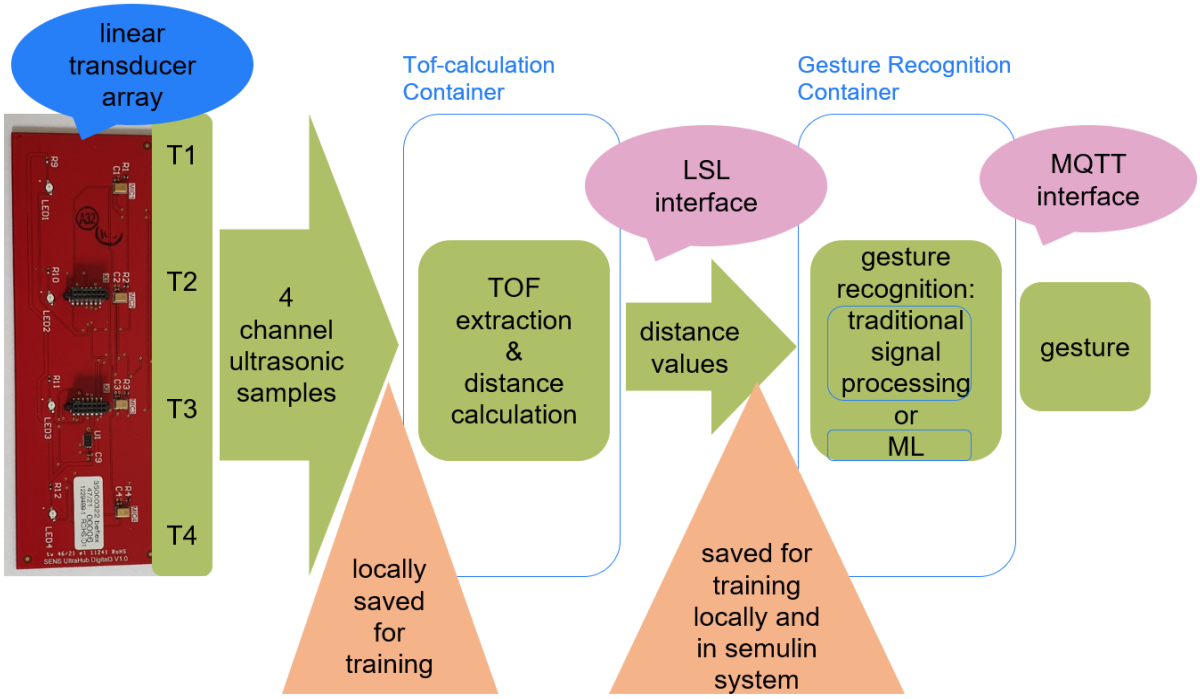

Abbildung 8: Docker-Container Verarbeitungsfluss

Container 1 greift auf die USB-Schnittstelle zum Ultrasonic-Hub zu und führt die Datenaufnahme und Vorverarbeitung, bestehend aus Rohdatenaufnahme, Bandpassfilter, Time Gain Compensation und Distanzberechnung, durch. Die Distanzdaten werden per LSL-Schnittstelle an das Gesamtsystem zur Verfügung gestellt. Container 2 greift auf die Distanzdaten, die per LSL zur Verfügung gestellt wurden, zu und führt die Klassifizierung von Gesten durch. Die erkannte Geste wird per MQTT im Gesamtsystem zur Verfügung gestellt.

Es wurden die Klassifizierungsstrukturen CNN, LSTM, GRU, ViT, CrossViT, Transformer, Random Forest und ohne ML auf dem aufgenommenen Datensatz und den unterschiedlichen Vorverarbeitungsstufen getestet. Die Tabelle zeigt die wichtigsten Ergebnisse.

Tabelle 1: Klassifizierungsstrukturen

Pre-processing	Processing	Accuracy on test set in %
TOF	Random forest	91
Spectrogram	CNN	95
No pre-processing	CNN	83
	LSTM	95
	Transformer	88

Mit dem linearen Sensing Shield können die Gesten „swipe up“, „swipe down“, „tap“, und „double tap“ unterschieden werden. Die beiden anderen Sensing Shields erlauben die Gesten “swipe up”, “swipe down”, “tap”, “double tap”, “swipe right”, and “swipe left”.

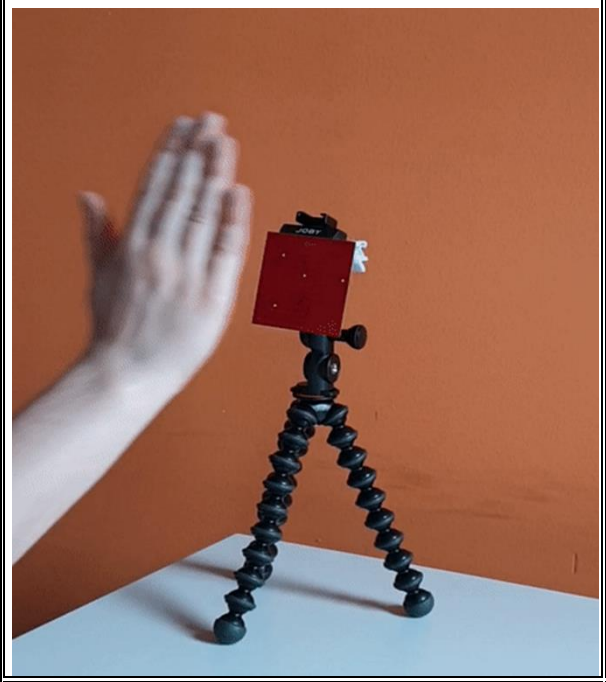

Abbildung 9: Gestenerkennung mit dem IFAG-Sensing Shield

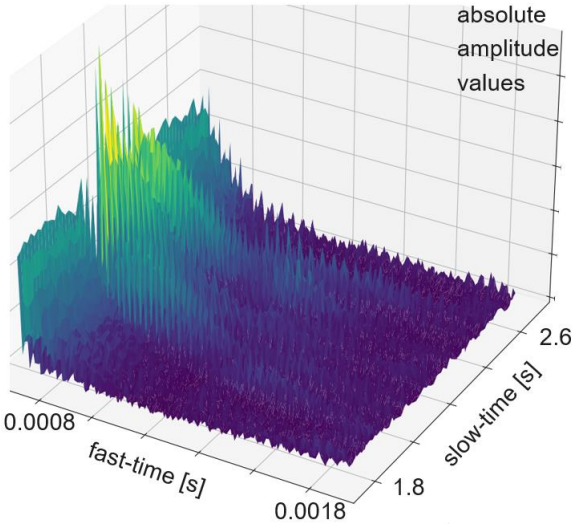

Abbildung 10: Rohdaten bei Datenaufnahme, Beispiel

Des Weiteren wurde ein weiteres Sensing Shield erstellt, das Beamforming mit acht Ultraschall-Wandlern erlaubt. Dieses kann entweder mit Oszilloskopen oder mit einem Aurix™ Mikrocontroller angesteuert und ausgelesen werden.

3.1.3. Task 1a.6 / Task 1a.7 Aufnahme und Aufbereitung von Blickdaten / Aufnahme und Aufbereitung der Sensordaten zur Entwicklung der Einzelkomponenten

Das im Unterauftrag eingebundene KMU Blickshift hat mit einem auf einer gekauften Blickrichtungserkennungslösung („Smart Eye Research“) basierenden Ansatz Eye-Tracking-

Daten mit angebunden. Für die Kombination der Blickdaten mit den anderen Modalitäten wurde die entwickelte Softwareplattform zur Aufnahme und Auswertung von Modalitäten um eine Aufnahme- und Abspielfunktion für alle anderen relevanten Sensorquellen erweitert. Eine weitere unterbeauftragte Aufgabe war die Bereitstellung einer Funktion zur Analyse und Identifikation für multimodale Zusammenhänge von Einzelmodalitäten. Die Ergebnisse der beiden Tasks sind detailliert im gemeinsamen SEMULIN-Schlussbericht dargestellt.

3.2. AP1.b Systemarchitektur

Dieser Abschnitt des Abschlussberichtes fasst die wesentlichen Fortschritte, Herausforderungen und Ergebnisse des Arbeitspaketes AP1b zusammen. Die Ziele des AP1b beinhalten folgende Aufgaben, die im weiteren Verlauf des Berichtes vertieft werden:

Abbildung 11: Ziele von AP1b

3.2.1. Systemarchitektur und Anforderungsanalyse

Zu Projektbeginn lag der Fokus auf der Entwicklung und Definition der Systemarchitektur des SEMULIN-Systems. Mehrere Workshops und Meetings wurden durchgeführt, um die Partner auf denselben Informationsstand zu bringen und die ersten Anforderungen an die Systemarchitektur zu definieren. Ein zentraler Workshop behandelte die Voranalyse und legte die Grundlage für die ersten Blockschaltbilder und Sensoranforderungen. Ein weiterer bedeutender Workshop konzentrierte sich auf "Safety & Security"-Anforderungen in automatisierten Fahrsituationen an die SEMULIN-Architektur. Einer der ersten Ansätze wurde in folgendem Blockschaltbild festgehalten. Auf der linken Seite die benötigte Sensorik, verarbeitet und fusioniert auf der Workstation „Fusion“, rechts die anschließende Visualisierung und Verbindung von Ausgabe-Modalitäten auf der Workstation „HMI“.

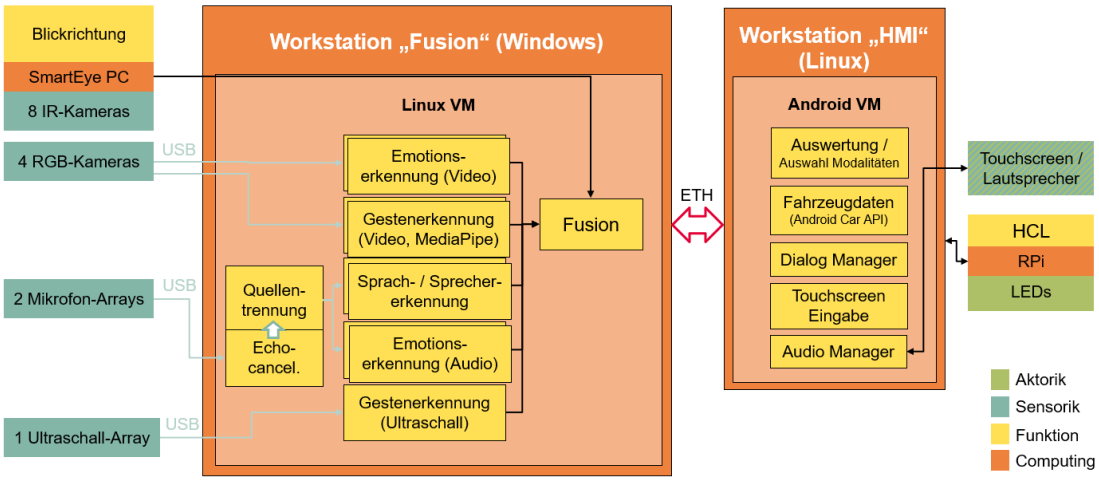

Abbildung 12: Erster Entwurf der Systemarchitektur

Die dabei gewonnenen Erkenntnisse führten zur Erstellung eines detaillierten Fragenkatalogs für die Schnittstellen zwischen Sensorik, Fusion Plattform und HMI Plattform. Dieser Katalog wurde an alle Partner mit Beiträgen verteilt und ausgewertet. Die Anforderungsanalyse wurde abgeschlossen und die ersten Architekturansätze unter Berücksichtigung von Kriterien wie Latenz, Performance und Skalierbarkeit erstellt und bewertet.

3.2.2. Ausarbeitung und Anpassung der Systemarchitektur

Das zweite Projektjahr stand im Zeichen der Ausarbeitung der System-, Hardware- und Softwarearchitektur. Aufgrund von Lieferschwierigkeiten bei bestimmten Hardwarekomponenten (z.B. Qualcomm™ KI-Beschleuniger) wurde ein Risikoanalyse durchgeführt und alternative Lösungen, wie die Nutzung einer Nvidia™ Evaluationsplattform oder die Verwendung eines Desktop PCs, erarbeitet.

Seitens der Softwarearchitektur wurde bereits die Verwendung einiger Architekturelemente beschlossen:

- MQTT als Kommunikationsprotokoll für Daten und Konfigurationen
- NTP für Zeit-Synchronisierung der einzelnen Teilsysteme
- Docker Umgebung und Containerisierung der einzelnen Softwaremodule

Zusammen mit einer weiteren Ausarbeitung der Hardwarearchitektur wurde ein weiterer Entwurf der Systemarchitektur festgelegt.

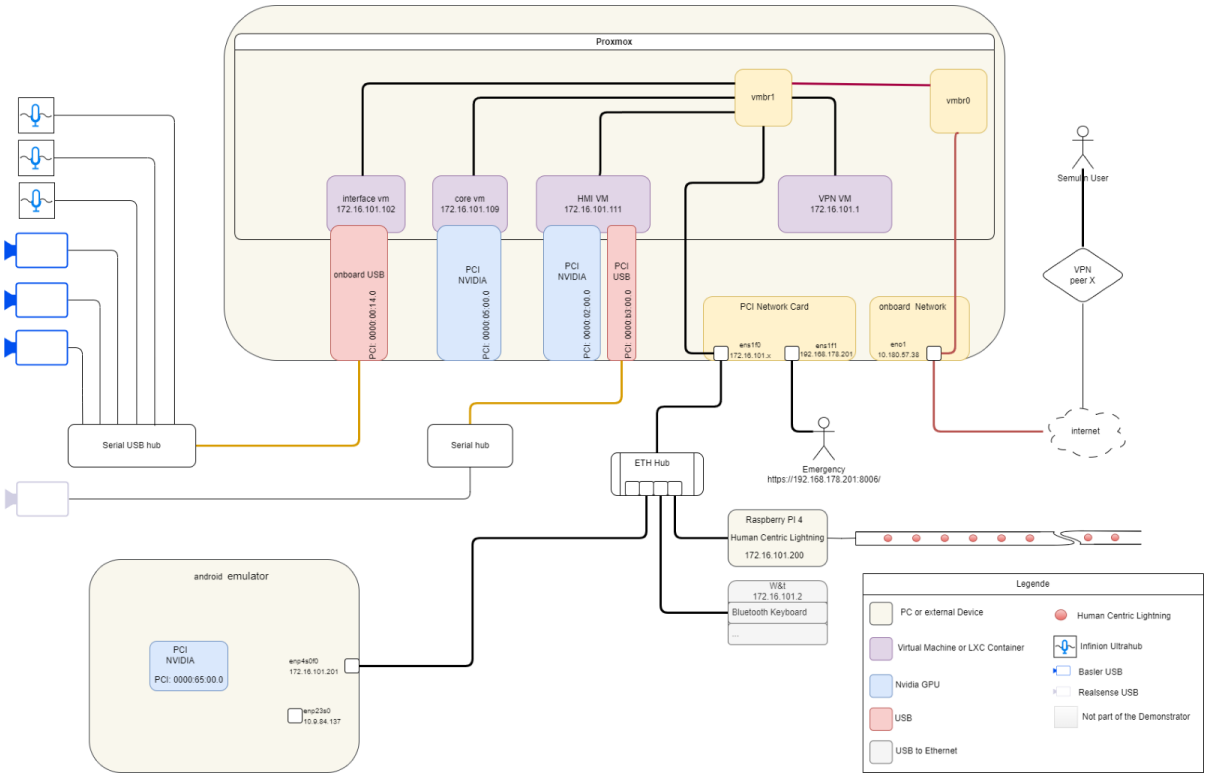

Abbildung 13: SEMULIN fortgeschrittener Entwurf der Systemarchitektur

Ein Präsenzworkshop führte zu einer weiteren Entscheidung zugunsten des "Lab Streaming Layer" als Software-Protokoll für das Audio Streaming und zur Implementierung der Echo Cancellation. Diese Maßnahmen waren entscheidend für die weitere Entwicklung der Systemarchitektur.

3.2.3. Integration und Systemoptimierung

Im dritten Jahr konzentrierten sich die Aktivitäten auf die Verfeinerung der Software- und Systemarchitektur sowie die Integration aller Teilsysteme. Wegen anhaltender Lieferengpässe wurde die Entwicklung auf einem zentralen PC-/Workstation-basierten System fortgesetzt.

Auf der Workstation wurden mehrere Virtuelle Maschinen (VM) erstellt, um die verschiedenen Anforderungen der Teilsystem abzudecken. Eine Windows-VM für die Blickshift Tools, eine Linux-VM für den Großteil der Algorithmik und Fusion, sowie spezielle VMs für die CARLA-Simulation und das HMI-System.

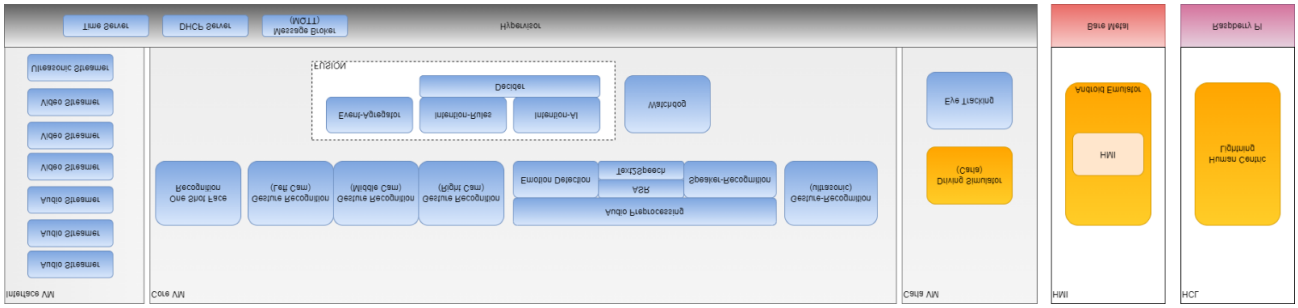

Abbildung 14: Finale Systemarchitektur

Eine Software-Integrationsumgebung wurde geschaffen, um Modultests zu ermöglichen und die Integration zu vereinfachen. Die erste erfolgreiche Datenaufnahme während der Probandenstudien zeigte Optimierungsbedarf, welcher durch ein Upgrade der Systemarchitektur adressiert wurde. Die Einführung einer Virtualisierungsebene in der Softwarearchitektur ermöglichte den flexiblen Austausch und die Erweiterung von Sensoren. Das gesamte SEMULIN-System konnte schließlich auf einem einzigen Rechner mit mehreren Linux-VMs abgebildet werden, was zu einem schlankeren und effizienteren Aufbau führte. Hiermit wurde das AP1b erfolgreich abgeschlossen und alle weiteren Interaktionen werden in AP6 Demonstratoren fortgeführt.

3.2.4. Finalisierung und Demonstratoraufbau

Die letzte Phase des Projektes beinhaltete hauptsächlich die Unterstützung und Koordination der weiteren Arbeitspakete. Der Aufbau des SEMULIN Demonstrators und die Durchführung der Probandenstudien an der Universität Ulm standen im Vordergrund. Während der Benutzung des SEMULIN-Demonstrators tauchten immer wieder technische Herausforderungen und Probleme auf, beispielsweise Synchronisierungsprobleme zwischen den Sensoren und USB-Port Zuordnungsprobleme, welche durch Modifikationen und die Einführung einer Virtualisierungsebene gelöst wurden.

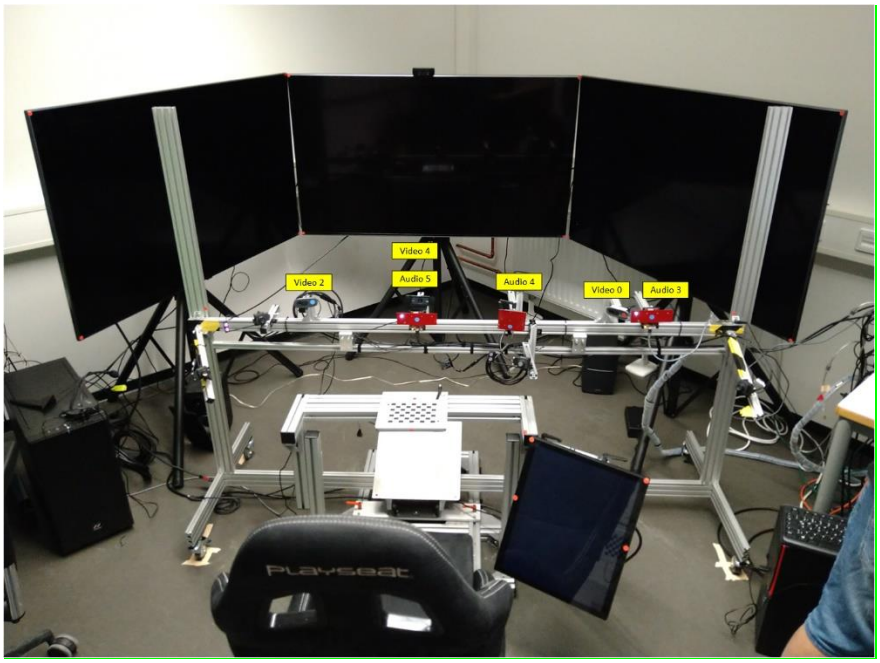

Abbildung 15: Probandenstudie

Die Auswertung und Überprüfung der Ergebnisse aus den ersten Probandenstudie brachte weitere kleine Probleme hervor. Diese wurden mit weiteren Modifikationen behoben. Die Einführung einer Virtualisierungsebene in der Softwarearchitektur ermöglichte den flexiblen Austausch und die Erweiterung von Sensoren. Das gesamte SEMULIN-System konnte schließlich auf einem einzigen Rechner mit mehreren Linux-VMs abgebildet werden, was zu einem schlankeren und effizienteren Aufbau führte. Hiermit wurde das AP1b erfolgreich abgeschlossen und alle weiteren Interaktionen werden in AP6 Demonstratoren fortgeführt.

3.3. AP6 Demonstratoren

Im Rahmen dieses Arbeitpakets wurde die Demonstration von multimodalen Interaktionen zwischen Nutzer und Fahrzeug realisiert. Durch IFAG wurde eine beispielhafte Umsetzung der Systemarchitektur und Anbindung von Sensoren für die multimodale Auswertung durchgeführt. Um frühzeitig Teilaspekte demonstrieren und verifizieren zu können, wurde ein mehrstufiges Konzept der Realisierung und Demonstration verfolgt.

Im Arbeitpaket AP1.a wurde die Hardware und Software des Infineon Audio und Ultrasound Boards für die Modalitäten Audio und Ultraschall-Gestenerkennung erforscht und entwickelt. Die ersten Tischdemonstratoren ermöglichten die frühzeitige Demonstration der Funktionsfähigkeit der Modalität Audio- und Ultraschall-basierte Gestenerkennung. Die entwickelte Hardware und Grundalgorithmen wurden den Partnern zur Verfügung gestellt, um die Entwicklung weiterer Modalitäten, wie z.B. Sprach- und Sprechererkennung, zu ermöglichen. Diese Aktivitäten wurden bereits in Kapitel 3.1.1. ausführlich beschrieben.

Während der Kompontentenentwicklung wurden in Rahmen von AP6 besondere Anforderungen aus Systemsicht frühzeitig erkannt und in die Entwicklung der Hardware und Software einbezogen. So zeigte sich, dass für die Audio-Modalität eine lineare Anordnung der Mikrofone von Vorteil war, während für Ultraschall eine Matrix-Anordnung bevorzugt wird. Daher wurden verschiedene Boards für Audio und Ultraschall konzipiert und entwickelt.

Im Arbeitpaket AP6 lag der Fokus auf der Integration der entwickelten Hardware in den Gesamt-SEMULIN-Demonstrator. Hierbei wurde die Integration der Hardware in den Simulator und anschließend in den SEMULIN-Demonstrator unter Berücksichtigung der spezifizierten Schnittstellen (LSL und MQTT) begleitet. Um eine reibungslose Integration zu gewährleisten, wurden Firmware-Updates notwendig, um das Zusammenspiel aller Komponenten im Demonstrator sicherzustellen. Die Abbildungen zeigen, an welchen Stellen das „Ultrasonic-Hub Main Board“ und „Ultrasonic-Hub Sensing Shield“ im Gesamtsystem verbaut wurden. In der ersten Abbildung sind die Boards noch frei zu sehen.

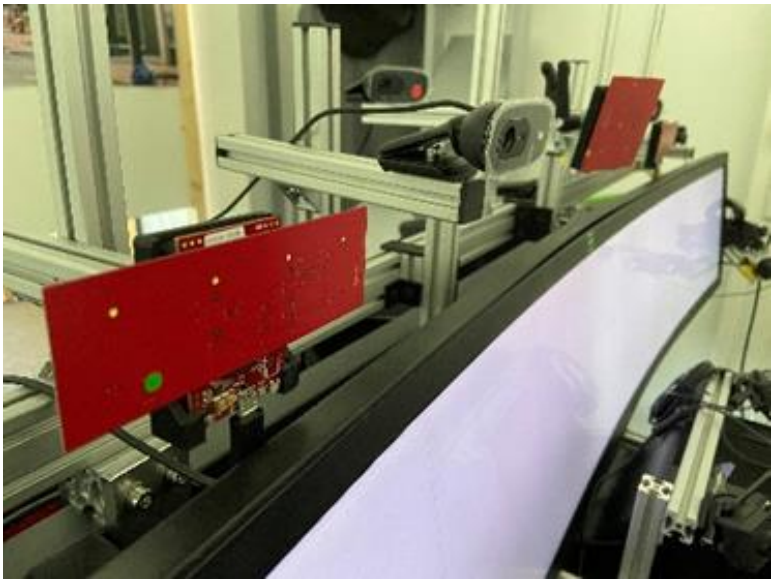

Abbildung 16: Ultrasonic-Hub im Gesamtsystem, frei

Abbildung 17: Ultrasonic-Hub im Gesamtsystem, nach Integration

In der zweiten Abbildung wurden die Boards hinter der Verkleidung integriert.

Neben den eigenen technischen Aktivitäten, v.a. in den AP1a und 1b, hat IFAG als AP6-Leitung die Koordination dieses Arbeitpakets durchgeführt. Unter anderen wurden regelmäßige Abstimmrunden mit den Projektpartnern organisiert, um die Zusammenarbeit und Interaktion bei der Integration in den Demonstrator im AP6 voranzutreiben, sowie die Schnittstellen zwischen den Partnern und andere APs zu berücksichtigen. Vor allem im dritten Jahr des Projekts fanden zahlreiche Vor-Ort-Termine statt, um auftretende Probleme direkt am Demonstrator, gemeinsam mit den Projektpartnern, zu lösen. Siehe auch Kapitel 3.2.4.

4. Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans

Die Anwendung der ultraschallbasierten Gestensteuerung im Automotive-Bereich wird auf Basis der Projektergebnisse aktuell intern evaluiert, wobei die Marktaussichten für Ultraschallanwendungen, speziell in der Fahrgastkabine, sehr vielversprechend sind. Der globale Markt für Ultraschallsensoren im Automobilbereich soll von 2020 bis 2027 mit einer jährlichen Wachstumsrate von 12,4% auf 4,4 Milliarden US-Dollar anwachsen.[1] Insbesondere für den Innenraum im Fahrzeug werden vermehrt Ultraschallsensoren diskutiert, um die Sicherheit und den Komfort der Insassen zu erhöhen, beispielsweise durch die Erkennung von Personen und Objekten im Fahrzeug. Namentlich die Regularien der Euro NCAP Child Presence Detection Initiative, die verhindern sollen, dass Kinder oder Hunde von Autofahrern versehentlich im Auto zurückgelassen werden, sind ein wesentlicher Treiber dieses Themengebiets. Infineon forscht daher kontinuierlich an In-Cabin-Monitoring-Systemen, hat eigene Produkte am Markt platziert und kann sich hier Synergieeffekte für die eigenen Systemlösungen erhoffen, wie etwa Radar- und TOF-Technologien.[2]

Für Infineon bilden die Ergebnisse des Projekts darüber hinaus die Grundlage für eine aktuelle sehr vielversprechende Kundeninteraktion im Bereich der Konsumgüter, speziell für den

[1] MarketsandMarkets, "Ultrasonic Sensor Market by Type (Proximity, Distance, Level, Temperature), Application (Parking Assist, Blind Spot Detection, Autonomous Vehicle), End-Use Industry (Automotive, Aerospace, Healthcare), and Region - Global Forecast to 2027"
[2] https://www.infineon.com/cms/en/applications/automotive/chassis-safety-and-adas/in-cabin-sensing/#!highlights

Bereich Augmented Reality (AR)- und Virtual Reality (VR)-Brillen sowie der True Wireless Stereo (TWS)-Gestensteuerung. Das AR-Marktsegment ist derzeit in vollem Aufschwung. Laut einer Studie von Grand View Research wird der globale AR/VR-Markt bis 2025 auf 143,3 Milliarden US-Dollar anwachsen, was einer jährlichen Wachstumsrate von 38,1% entspricht.3 Insbesondere das Thema eye-tracking mit Ultraschall in AR-Brillen ist aufgrund vielversprechender wissenschaftlicher Publikationen [4] [5] [6] [7] für die Brillen-Hersteller von großem Interesse. Die kleine Bauform, sowie der niedrige Leistungsbedarf und Unempfindlichkeit gegenüber Umgebungslicht sind die wesentlichen Argumente für eine solche Lösung. In Abbildung 18 ist schematisch dargestellt, wie der Einbau der Sensoren in einer AR-Brille erfolgen könnte. Derzeit werden mit Brillenherstellern bis zu 6 Ultraschall-Transceiver-Mikrophone pro Brille als realistisch diskutiert, was einen zweistelligen Millionenbedarf jährlich zur Folge haben würde und somit eine sehr vielversprechende Opportunität für Infineon darstellt.

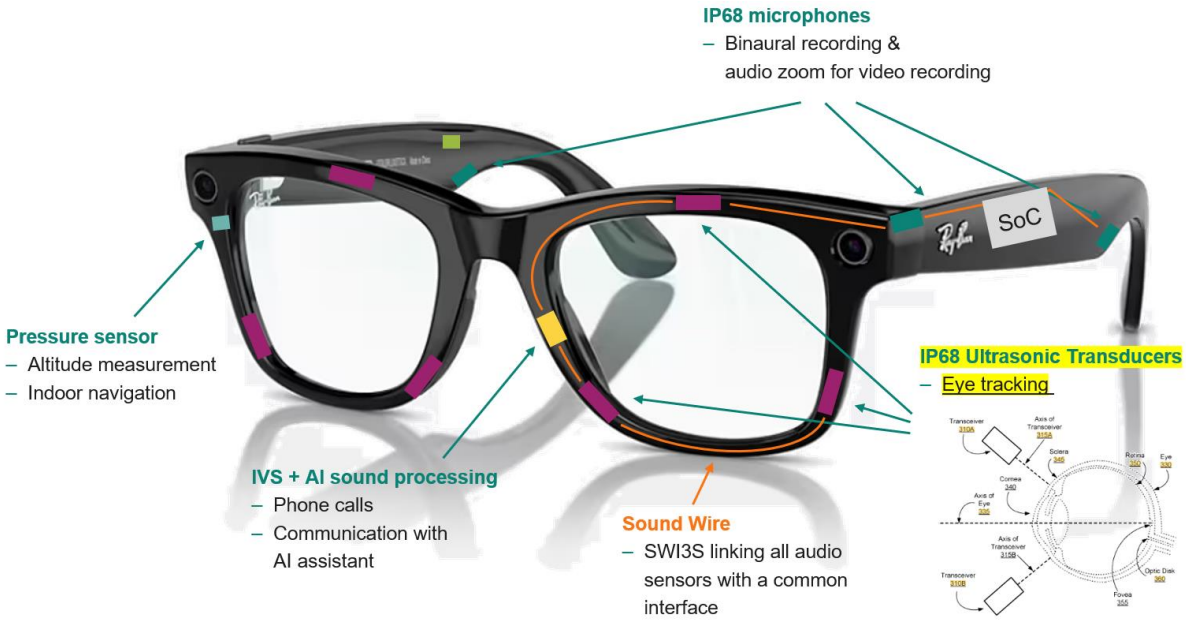

Abbildung 18: Opportunitäten für Infineons Consumer Sensorportfolio in AR-Brillen

Neben den AR-Brillen ist die steigende Nachfrage nach True-Wireless-Stereo (TWS) Kopfhörern mit Gestensteuerung ein weiterer wichtiger Aspekt. Laut einer Studie von MarketsandMarkets wird der globale TWS-Markt bis 2024 auf 10,3 Milliarden US-Dollar anwachsen, was einer jährlichen Wachstumsrate von 24,5% entspricht.[8] Diese Zahlen unterstreichen die Bedeutung einer frühzeitigen Positionierung in diesem Marktsegment. Die Projektergebnisse aus SEMULIN dienen der Anpassung spezifischer Anforderungen der

[3] Grand View Research, "Augmented Reality (AR) and Virtual Reality (VR) Market Size, Share & Trends Analysis Report by Component (Hardware, Software, Services), by Device (Head-Mounted Display, Gesture-Tracking Device), by Application, by Region, and Segment Forecasts, 2020 - 2025"
[4] 2023 IEEE International Ultrasonics Symposium (IUS) | 979-8-3503-4645-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/IUS51837.2023.10306398
[5] S. Sun et al., "Eye-Tracking Monitoring Based on PMUT Arrays," in Journal of Microelectromechanical Systems, vol. 31, no. 1, pp. 45-53, Feb. 2022, doi: 10.1109/JMEMS.2021.3125377
[6] Golard, A.; Talathi, S.S. Ultrasound for Gaze Estimation—A Modeling and Empirical Study. Sensors 2021, 21, 4502. https://doi.org/10.3390/s21134502
[7] GazeTrak: Exploring Acoustic-based Eye Tracking on a Glass Frame, https://arxiv.org/html/2402.14634v2
[8] MarketsandMarkets, "True Wireless Stereo (TWS) Earbuds Market by Type (Earbuds, Headphones), Application (Music, Calls, Gaming), and Region - Global Forecast to 2024"

Kunden und stärken somit Infineons Marktstellung in dem Segment. Insgesamt bieten die Projektergebnisse eine solide Grundlage für eine gute Positionierung in diesem Marktsegment, die Infineons Wettbewerbsfähigkeit stärken und langfristigen Erfolg sichern soll.

4.1. Während der Durchführung des Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Es liegen keine wesentlichen weiteren Erkenntnisse über bekannt gewordene Fortschritte auf dem Gebiet des Vorhabens bei anderen Stellen vor.

4.2. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Im Projektzeitraum wurden durch die Infineon Technologies AG im Zusammenhang mit dem Projekt Semulin die folgenden Arbeiten veröffentlicht:

- Gemeinsam von IFAG und der Universität Granada wurde ein Artikel mit dem Titel „Object Positioning Algorithm Based on Multidimensional Scaling and Optimization for Synthetic Gesture Data Generation (Borja Saez-Mingorance, Antonio Escobar-Molero, Javier Mendez-Gomez, Encarnacion Castillo-Morales and Diego P. Morales-Santo)“ veröffentlicht. Die Open-Access-Veröffentlichung erfolgte am 3.9.2021.
- Im Rahmen der Auswertung der Vorverarbeitung zu Distanzdaten versus Spektrogrammen wurde „COMPARISON OF APPROACHES FOR GESTURE RECOGNITION WITH MEMS ULTRASONIC TRANSDUCERS“ (Elfi Fertl, Maryna Shcherbak, Martin Krueger, Encarnación Castillo Morales, Manuel Pegalajar Cuéllar, Georg Stettinger) am 17. Januar 2024 von IEEE veröffentlicht.
- Eine rein ML-basierte Verarbeitung wurde in „End-to-End Ultrasonic Hand Gesture Recognition“ (Elfi Fertl, Do Dinh Tan Nguyen, Martin Krueger, Georg Stettinger, Rubén Padial-Allué, Encarnación Castillo, and Manuel P. Cuéllar) am 25 April 2024 veröffentlicht.

VERÖFFFENTLICHUNG / TITEL	DATUM	AUTOREN
Object Positioning Algorithm Based on Multidimensional Scaling and Optimization for Synthetic Gesture Data Generation https://www.mdpi.com/1424-8220/21/17/5923	03.09.2021	Borja Saez-Mignorance, Antonio Escobar-Molero, Javier Mendez-Gomez, Encarnacion Castillo-Morales, Diego P. Morales-Santos
Comparison of approaches for gesture recognition with MEMS Ultrasonic Transducers https://ieeexplore.ieee.org/document/10387761	28.-30.3. 2023 IEEE Explore, Brügge/BE	Elfi Fertl; Maryna Shcherbak; Martin Krüger; Encarnación Castillo Morales; Manuel Pegalajar Cuéllar; Georg Stettinger
End-to-End Ultrasonic Hand Gesture Recognition https://www.mdpi.com/1424-8220/24/9/2740	25.04.2024	Elfi Fertl, Do Dinh Tan Nguyen, Martin Krüger, Georg Stettinger, Rubén Padial-Allué, Encarnación Castillo Morales, Manuel Pegalajar Cuéllar

Tabelle 2: IFAG Veröffentlichungen in SEMULIN

Die wesentlichen Ergebnisse zu den Arbeiten an diesem Teilvorhaben werden im Herbst 2024 über die Technische Informationsbibliothek, Deutsche Forschungsberichte in Hannover veröffentlicht.

5. Abkürzungsverzeichnis

AR	Augmented Reality
CARLA	Open-Source Simulator für Forschung zum Autonomen Fahren
CNN	Convolutional Neural Network
CROSSVIT	Cross-Attention Multi-Scale Vision Transformer for Image Classification
DOCKER	Freie Software zur Isolierung von Anwendungen durch Containervirtualisierung
GRU	Gated Recurrent Unit
HMI	Human Machine Interface
LSL	Lab Streaming Layer, als Middleware-Protokoll für Virtualisierungsebene der Sensor Signale
LSTM	Long Short-Term Memory
MEMS	Micro-Electro-Mechanical System
ML	Machine Learning
MQTT	Offenes, zentrales Netzwerkprotokoll für Machine-to-Machine-Kommunikation
NCAP	New Car Assessment Programme
NTP	Network Time Protocol
Random Forest	Machine learning technique o. algorithm
TOF	Time of flight
Transformer	Artificial intelligence model
TWS	True Wireless Stereo
VIT	Vision Transformer
VM	Virtuelle Maschinen
VR	Virtual Reality

6. Abbildungsverzeichnis

Abbildung 1: SEMULIN - Projektkonstrukt 5
Abbildung 2: AP1.a Tasks und Partner 7
Abbildung 3: Ultrahub Sensing Shield Varianten 9
Abbildung 4: Ultrasonic-Hub Main Board 9
Abbildung 5: Ablaufdiagramm der Handgestenerkennung 10
Abbildung 6: Basisprinzip der ultraschallbasierten Handgestenerkennung 10
Abbildung 7: Ausführung der Handgesten 11
Abbildung 8: Docker-Container Verarbeitungsfluss 11
Abbildung 9: Gestenerkennung mit dem IFAG-Sensing Shield 12
Abbildung 10: Rohdaten bei Datenaufnahme, Beispiel 12
Abbildung 11: Ziele von AP1b 13
Abbildung 12: Erster Entwurf der Systemarchitektur 14
Abbildung 13: SEMULIN fortgeschrittener Entwurf der Systemarchitektur 15
Abbildung 14: Finale Systemarchitektur 15
Abbildung 15: Probandenstudie 16
Abbildung 16: Ultrasonic-Hub im Gesamtsystem, frei 17
Abbildung 17: Ultrasonic-Hub im Gesamtsystem, nach Integration 18
Abbildung 18: Opportunitäten für Infineons Consumer Sensorportfolio in AR-Brillen 19

7. Tabellenverzeichnis

Tabelle 1: Klassifizierungsstrukturen 12
Tabelle 2: IFAG Veröffentlichungen in SEMULIN 21

Berichtsblatt

1. ISBN oder ISSN **geplant**	2. Berichtsart (Schlussbericht oder Veröffentlichung) **Schlussbericht**
3. Titel **Verbundprojekt SEMULIN: Selbstunterstützende Multimodale Interaktion,** **Teilvorhaben Infineon Technologies AG: Ultraschallbasierende Gestenerkennung**	
4. Autoren [Namen, Vornamen] **Achatz, Simon; Dr. Bretthauer, Christian; Fertl, Elfi; Maier, Daniela; Mena Carrillo, Juan**	5. Abschlussdatum des Vorhabens **30.04.2024** 6. Veröffentlichungsdatum **11.10.2024** 7. Form der Publikation -
8. Durchführende Institutionen (Name, Adresse) **Infineon Technologies AG, Am Campeon 1-15, 85579 Neubiberg, DE**	9. Ber. Nr. Durchführende Institution - 10. Förderkennzeichen *) **19A20012D** 11. Seitenzahl **23**
12. Fördernde Institution (Name, Adresse) **Bundesministerium für** **Wirtschaft und Klimaschutz (BMWK)** **53107 Bonn**	13. Literaturangaben 14. Tabellen **2** 15. Abbildungen **18**
16. Zusätzliche Angaben -	
17. Vorgelegt bei (Titel, Ort, Datum) **TIB Technische Informationsbibliothek, Hannover, 11.10.2024**	
18. Kurzfassung Im kollaborativen SEMULIN-Projekt entwickelte die Infineon Technologies AG ein innovatives ultraschallbasiertes Gestenerkennungssystem, das MEMS-Mikrofone sowie ein Audio-Entwicklungsboard zur Verarbeitung von Audiodaten verwendet und so eine Gestenerkennung im Fahrgastraum ermöglicht. Im Mittelpunkt des Projekts stand ein ganzheitlicher Multimodalitätsansatz, der auf die zwischenmenschliche Kommunikation abzielt. Ein besonderer Schwerpunkt lag dabei auf automatisierten Fahrszenarien, einschließlich der Automatisierungsstufen 3-4/5 und Übergabesituationen zwischen automatisiertem und nicht-automatisiertem Fahren. Zusätzlich wurden auch klassische fahrergesteuerte Fahrsituationen betrachtet. Innerhalb des Projekts, mit sechs Projektpartnern, leitete IFAG drei Arbeitspakete, AP1.a "Intelligente Teilsysteme", AP1.b "Systemarchitektur" und AP6 "Demonstratoren". Der Schwerpunkt der Arbeit in AP1a lag auf der Identifizierung und dem Entwurf intelligenter Subsysteme sowie auf der Erforschung der Vorverarbeitung und Bereitstellung aggregierter Sensordaten für eine multimodale Auswertung, die letztendlich in einer Mensch-Maschine-Schnittstelle (HMI) und als Ausgabemodalitäten verwendet werden sollen. Dazu gehörte die Entwicklung von MEMS-basierten Siliziummikrofonen für den multimodalen Einsatz und die Erfassung von Audiodaten für die Gestenerkennung. Außerdem wurden in AP1b sicherheitsrelevante Architekturen erforscht, um die Spezifikation und Einrichtung einer Systemarchitektur für HMI und automatisiertes Fahren zu ermöglichen. In AP6 koordinierte IFAG die Integration der Systeme in die Demonstratoren und geplanten Aufbauten bei den Projektpartnern, während in AP2 die Partner bei der Anbindung intelligenter Sensoren unterstützt wurden.	
19. Schlagwörter Automatisiertes Fahren, Fahrassistenzsysteme, Gestenerkennung, HMI, Multimodal, MEMS, Mikrofon, Sensor, Ultraschall	
20. Verlag -	21. Preis -

Document Control Sheet

1. ISBN or ISSN **planned**	2. type of document (e.g. report, publication) **Final report**
3. title **Collaborative project SEMULIN: Self-supporting multimodal interaction, sub-project Infineon Technologies AG: Ultrasound-based gesture recognition**	
4. author(s) (family name, first name(s)) **Achatz, Simon; Dr. Bretthauer, Christian; Fertl, Elfi; Maier, Daniela; Mena Carrillo, Juan**	5. end of project **30.04.2024** 6. publication date **11.10.2024** 7. form of publication -
8. performing organization(s) (name, address) **Infineon Technologies AG, Am Campeon 1-15, 85579 Neubiberg, Germany**	9. originator's report no. - 10. reference no. **19A20012D** 11. no. of pages **23**
12. sponsoring agency (name, address) **Bundesministerium für Wirtschaft und Klimaschutz (BMWK)** **53107 Bonn**	13. no. of references 14. no. of tables **2** 15. no. of figures **18**
16. supplementary notes -	
17. presented at (title, place, date) **TIB Technische Informationsbibliothek, Hannover, Germany**, **11.10.2024**	
18. abstract In the SEMULIN project, Infineon Technologies AG aimed to develop an innovative ultrasound-based gesture recognition system. The system uses MEMS microphones and an audio development board to process audio data, thus enabling gesture recognition in the passenger compartment. The project focused on a holistic multimodality approach, aimed at interpersonal communication. A particular focus was put on automated driving scenarios, including automation levels 3-4/5 and handover situations between automated and non-automated driving. Classic driver-controlled driving situations were also considered. Within the project, with six project partners, IFAG led three work packages: WP1.a "Intelligent subsystems", WP1.b "System architecture" and WP6 "Demonstrators". The focus of the work in WP1a was on the identification and design of intelligent subsystems as well as research into the pre-processing and provision of aggregated sensor data for multimodal evaluation, which would ultimately be used in a human-machine interface (HMI) and as output modalities. This included the development of MEMS-based silicon microphones for multimodal use and the acquisition of audio data for gesture recognition. In addition, safety-relevant architectures were researched in WP1b to enable the specification and establishment of a system architecture for HMI and automated driving. In WP6, the integration of the systems into the demonstrators and planned setups at the project partners was coordinated, while in WP2 the partners were supported in connecting the intelligent sensors.	
19. keywords Automated driving, gesture recognition, HMI, MEMS, microphones, multimodality, sensors, ultrasound	
20. publisher -	21. price -