



HOCHSCHULE FÜR
TECHNIK UND WIRTSCHAFT
DRESDEN
UNIVERSITY OF APPLIED SCIENCES

Entwicklung eines kamerabasierten Evaluierungsverfahren für Fahrerassistenzsysteme mit autonomen Notbremseingriff

Masterarbeit

zur Erlangung des Abschlusses

Master of Science

im Studiengang

Angewandte Informationstechnologien

eingereicht von

Patrick Richter

1. Gutachter: Herr Prof. Dr. Kai Bruns

2. Gutachter: Herr Prof. Dr. Toralf Trautmann

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Zielstellung	2
1.3	Eingrenzung	3
2	Theoretische Grundlagen.....	4
2.1	Der Notbremseingriff in der Fahrzeugtechnik	4
2.2	Physikalische Hintergründe	5
2.2.1	Projektion optischer Systeme.....	5
2.2.2	Abbildungsfehler optischer Systeme	7
2.2.3	Reflexion von Schallwellen	8
2.3	Praxisrelevante Sensorik.....	11
2.3.1	Kameras.....	11
2.3.2	Lidar.....	17
2.3.3	Radar	20
3	Technologischer Kontext.....	26
3.1	Problemanalyse.....	26
3.2	Bewertungskriterien.....	27
3.3	Verwandte Projekte	28
3.4	Bekannte Verfahren zur Distanzbestimmung in der Computer Vision	30
3.4.1	Marker Recognition.....	33
3.4.2	Stereo Vision.....	34
3.5	Auswahl geeigneter Ansätze	35
4	Konzeption geeigneter Verfahren	38
4.1	Distanzmessung mit einer Kamera: Einsatz der Monovision	38
4.2	Einführung von normierten Markern	41
4.2.1	Gradient-based Pattern Finding Algorithmus (GPF)	44

4.2.2	Feature-based Pattern Finding Algorithmus (FPF)	46
4.2.3	Gegenüberstellung der theoretischen Vor- und Nachteile beider Verfahren	49
4.3	Zeitliche Synchronisation von Videodaten	52
4.3.1	Visuell	52
4.3.2	Audio.....	53
4.4	Abgrenzung	61
4.5	Annahmen	62
5	Implementierung	64
5.1	Anordnung der Prüftechnik.....	64
5.2	Testdatensätze	66
5.3	Verwendete Bibliotheken	68
5.4	Algorithmischer Aufbau aus Sicht der Nutzereingaben	69
5.5	Ausgewählte Details der Implementierung	71
5.5.1	Sub-Frame Audio-Synchronisation	71
5.5.2	Beschleunigung der Erfassung der relevanten Messdaten	73
5.5.3	Auswahl der relevanten Messdaten	74
5.5.4	Interpolation und Glättung der Messergebnisse.....	75
5.6	Vorstellung der Testergebnisse	77
5.6.1	Gegenüberstellung der Genauigkeit der implementierten Verfahren ...	77
5.6.2	Bewertung der Notbremsfunktion.....	79
5.6.3	Mögliche Fehlerquellen und ihre Ursachen.....	81
6	Zusammenfassung und Ausblick	83
7	Abbildungsverzeichnis	87
8	Literaturverzeichnis	89

Abkürzungsverzeichnis

ABS	Anti-Blockier-System
CAN	Controller Area Network
TÜV	Technischer Überwachungsverein
CV	(Open) Computer Vision
FPS	Frames per Second; Bilder pro Sekunde
ROI	Region of Interest
Laser	Light emulation by stimulated emission of radiation
Radar	Radio Detection and Ranging
Lidar	Light Detection and Ranging

Glossar

Kombi-Instrument	Vgl. Tachometer; Wird häufig zur Anzeige von Geschwindigkeit, Drehzahl und Signalen genutzt.
Bit	Ein digitaler Wert, der "0" oder "1" annehmen kann.
Byte	Acht Bit ergeben einen Byte.
OpenCV	"Open Computer Vision": Funktionsbibliothek, die viele Algorithmen zur Bildverarbeitung und -analyse bereit hält.
ROI	Eine ROI ist ein kleinerer Teil einer Abbildung, der die für ein angestrebtes Ziel enthaltenen Objekte beinhaltet. In der Regel werden damit Algorithmen auf den wesentlichen Bildbereich beschränkt und damit beschleunigt.
Winkelposition	Ausgehend von einem Ursprung (Punkt) im Raum, einer gegebenen Richtung (Winkel) und Entfernung kann ein Endpunkt beschrieben werden (vgl. Vektoren), der Winkelposition genannt wird.
Stereokamera	Zwei in gleiche Richtung "blickende" Kameras, vorzugsweise mit gleichen Eigenschaften, die im Raum versetzt zueinander stehen. Deren Bildfeld muss sich jedoch in bestimmter Entfernung überdecken. Die räumliche Translation ist abhängig von der zu erfassenden Minimalentfernung.
Marker	Eindeutig zu identifizierende, künstliche Markierung, welche durch Bildanalyse gut auffindbar und beschreibbar ist. Ein natürlicher Marker heißt "Landmark".

1 Einleitung

Fahrerassistenzsysteme bilden, im Zusammenhang mit der computergestützten Verarbeitung von Sensordaten, einen der bedeutendsten und am schnellsten wachsenden Forschungsbereich der modernen Fahrzeugtechnik. Vor allem im Gebiet der Fahrzeugsicherheit wurden seit der Nutzung verschiedener einfacher Sensoren enorme Erfolge zur Unfallprävention und Insassenschutz erzielt, sodass bekannte Systeme wie das Anti-Blockier-System (ABS) [1] und Airbags ein Standard für moderne Fahrzeuge geworden sind. Auch aus Sicht des autonomen Fahrens konnten durch den Einsatz komplexerer Sensoren und leistungsstarker Computer beispielsweise Einpark-Assistenten realisiert werden, die das Fahrzeug autonom in Parklücken positionieren können.

Die Kombination beider Aspekte, bestehend aus Sensorik und der computergestützten Verarbeitung derer Daten, ermöglichte die Entwicklung moderner Fahrerassistenzsysteme. Dazu zählen aktive Fahrspurassistenten, Abstandsregelungs-Assistenten und Notbremsassistenten.

1.1 Motivation

Fahrerassistenzsysteme dienen der Sicherheit und können je nach System sogar das Fahrzeug autonom kontrollieren. Die vom Fahrzeug verwendeten Sensoren zur Wahrnehmung der Umwelt sind allerdings nicht dauerhaft zuverlässig und können Daten außerhalb ihrer Spezifikation an das Fahrzeug übermitteln, wodurch ein falsches Verhalten der Fahrerassistenzsysteme ausgelöst werden kann. Dieser Effekt kann vor allem bei der Abstandsmessung durch Radarsensoren beobachtet werden. Eine interne Beobachtung hat gezeigt, dass Radarsensoren durch Alterung zunehmend der Spezifikation abweichende Daten an das Fahrzeug übermitteln können, wodurch aktive Fahrerassistenzsysteme, wie der Notbremsassistent, verzögert, respektive falsch reagieren können. Jedoch ist dies noch nicht wissenschaftlich belegt und sollte deshalb kritisch betrachtet werden.

Durch den Aspekt des autonomen, aktiven Eingriffs in das Fahrzeug und des daraus entstehenden Gefahrenpotentials bei Fehlfunktion des Systems, müssen die Systeme regelmäßigen Prüfungen durch die in Deutschland akzeptierten Prüforganisationen für Fahrzeuge (TÜV, Dekra) unterzogen werden. Jedoch gibt es bis dato keine standardisierten, zuverlässigen Prüfverfahren zur Sicherstellung der vom Hersteller vorgesehenen Funktion dieser Fahrerassistenzsysteme, sodass die Prüfungen der Systeme noch immer durch die persönliche, subjektive Einschätzung des jeweiligen Prüfers durchgeführt werden.

Meine Motivation ist es die Möglichkeiten der automatisierten Prüfung von Notbremsassistenten zu betrachten. Gefordert ist ein System, was mit allen Fahrzeugtypen kompatibel und schnell durchführbar ist, sowie zuverlässige Ergebnisse ausgibt. Es gilt dabei besonders zu beachten, dass den Notbremsassistenten verschiedener Fahrzeuge eine Vielzahl sich unterscheidender Prinzipien und Sensoren zugrunde liegen, was die Auswahl geeigneter Messtechnik und Algorithmen einschränkt.

1.2 Zielstellung

Ziel der Masterarbeit ist es die Möglichkeiten der Prüfung von Notbremsassistenten verschiedener Fahrzeuge zu ergründen und ein geeignetes Verfahren zur Evaluierung der korrekten Funktion zu entwickeln. Es werden mithilfe der Funktionsbibliothek OpenCV verschiedene Ansätze zur Distanzgewinnung mit nur einer Kamera untersucht. Der gewählte Ansatz soll die Anbringung der Prüftechnik im theoretischen, späteren Prüfbetrieb so einfach wie möglich gestalten, die Anschaffungskosten des Systems, gegenüber einer Stereokamera¹ (zwei Kameras zur Distanzmessung), müssen gering sein, und Ressourcen-schonend sowie mit wenig Berechnungsaufwand auszuwerten sein, was auf eine entsprechend anzustrebende geringe Berechnungszeit schließen lässt.

Betrachtet werden Distanz-Indikatoren in Form von speziellen, angebrachten Markern. Es ist darauf zu achten, dass das System eine möglichst hohe

¹ Zwei, meist achsparallele, Kameras, welche die gleiche Szene zeitgleich beobachten und an das Stereo-Sehen des Menschen angelehnt sind. Dadurch werden Distanzmessungen ermöglicht.

Kompatibilität zur Verwendung verschiedener Kameras aufweist, die im Spektralbereich der menschlichen Wahrnehmung arbeiten. Abschließend soll mithilfe des entwickelten Verfahrens eine Aussage über die Zuverlässigkeit und Güte des Notbremsassistentensystems durch eine fachkundige Person (z.B. Prüfer) getroffen und die Messergebnisse grafisch dargestellt werden können.

1.3 Eingrenzung

Das Fahrzeug bewegt sich im Rahmen dieses Projektes auf einem Rollenprüfstand, wobei eine reale Fahrsituation simuliert wird. Dabei wird mithilfe eines Schienensystems und einer Prüfmaske dem zu prüfenden Fahrzeug ein sich näherndes, vorausfahrendes Fahrzeug simuliert. Die Reaktionen des zu prüfenden Fahrzeugs wird visuell sowie akustisch über eine Kamera im Innenraum aufgezeichnet, welche auf das Kombi-Instrument (vgl. Tachometer) gerichtet ist. Somit können Warnsignale, die für den Fahrer bestimmt sind, aufgenommen und ausgewertet werden. An der Frontscheibe des Fahrzeuges ist eine weitere Kamera befestigt, welche die theoretische Fahrtrichtung des zu prüfenden Fahrzeugs und die Prüfmaske aufnimmt.

2 Theoretische Grundlagen

Das nachfolgenden Kapitel beschäftigt sich theoretischen Grundlagen auf die sich die Masterarbeit stützt und die zum allgemeinen Verständnis der darauf folgenden Kapitel dienen.

2.1 Der Notbremseingriff in der Fahrzeugtechnik

Ein zentraler Begriff stellt in dieser Arbeit der Notbremsassistent dar (angelehnt an [2] S. 595f.). Ein Notbremsassistent ist ein Fahrerassistenzsystem, welches vorausschauend auf eine entstehende Gefahrensituation reagieren kann. Die Reaktion unterteilt sich in drei verschiedene Reaktionsarten (Anlehnung an [3]). Die erste Reaktionsart ist eine einfache Auffahrwarnung, welche den Fahrer meist hauptsächlich akustisch auf einen drohenden Auffahrunfall hinweist. Die zweite Reaktionsart ist die Bremskraftunterstützung und bedient sich der Auffahrwarnung und einem Bremsassistenten. Sobald das Fahrzeug eine Auffahrwarnung auslöst, und der Fahrer plötzlich vom Gas geht und daraufhin in kurzer Zeit die Bremse betätigt, wird sofort die volle Bremskraft ausgelöst. Die letzte Reaktionsart ist die autonome Notbremsung, bei der an die Auffahrwarnung erweiterte Algorithmen darüber entscheiden, ob unabhängig vom Fahrer eine Notbremsung eingeleitet wird.

Um Überschneidungen in den Begrifflichkeiten zu vermeiden wird in dieser Arbeit für alle Reaktionsarten von einem Notbremsassistenten, bzw. Notbremssystem gesprochen.

Die Bewertungskriterien zur Überprüfung der korrekten Funktion eines Notbremseingriffs sind im Abschnitt 3.2 beschrieben.

2.2 Physikalische Hintergründe

Neben der Klärung der Begrifflichkeiten stellt vor allem die Beschreibung wichtiger physikalischer Hintergründe, auf denen alle nachfolgenden Verfahren basieren, eine wichtige Grundlage zum Verständnis von Fahrerassistenzsystemen und den Möglichkeiten ihrer Prüfung dar. Das ist bereits durch den Fakt begründet, dass die heutige Fahrzeugentwicklung, und nicht zuletzt moderne Fahrerassistenzsysteme, nicht mehr nur von Maschinenbau geprägt sind, sondern auf viele andere Bereiche übergreifen, wie z.B. die Informatik und die direkt mit ihr, der Datenerfassung, sowie -verarbeitung, im Zusammenhang stehenden physikalischen Hintergründe.

2.2.1 Projektion optischer Systeme²

Das Funktionsprinzip aller modernen, optischen Kameras ist identisch, unabhängig von ihrer Empfindlichkeit für bestimmte Wellenlängen. Durch eine Kamera können Bildinformationen von einer Szene gewonnen werden. Dabei gehen Informationen verloren, da die dreidimensionale Szene auf einer nur zweidimensionalen Bildebene abgebildet wird. Diesen Vorgang nennt man Projektion. Sie kann mithilfe eines vereinfachten Kameramodells mathematisch beschrieben werden. Die nachfolgende Beschreibung ist an die Quelle [4 S. 153f.] angelehnt.

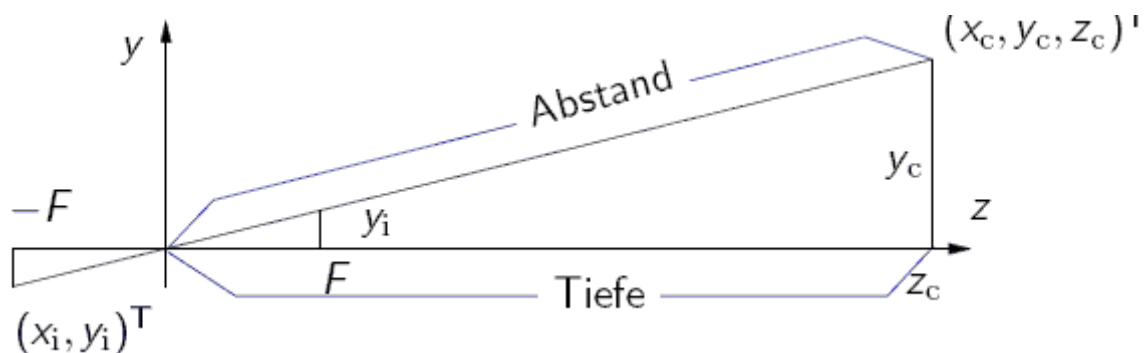


Abbildung 1: Lochkameramodell (Quelle: Mindpicnic [5], Alternative Quelle: [6])

Das Lochkameramodell (Abbildung 1) ist das einfachste der Kameramodelle und ähnelt gleichzeitig dem Strahlengang innerhalb einer realen Kamera. Das abzubildende Objekt befindet sich in dem Modell in dem Bereich positiver z -

² Entnommen aus eigener Arbeit: [9]

Koordinaten. Die zweidimensionale Abbildungsebene (vgl. Bildsensor) befindet sich bei $z = -F$.

Ein dreidimensionales Objekt reflektiert eintreffende Lichtstrahlen innerhalb einer Szene. In der Annahme, dass ein Lichtstrahl vom Punkt C mit den Raumkoordinaten $\mathbf{C} = (x_c, y_c, z_c)'$ durch die Lochblende der Modellkamera im Punkt $\mathbf{O} = (0, 0, 0)$ tritt, würde der Punkt C den Punkt $\mathbf{T}_c = (x_i, y_i)^T$ auf der Bildebene treffen.

Die Koordinaten des resultierenden Punktes \mathbf{T}_c sind direkt proportional abhängig zur Brennweite der Lochkamera. Würde die Brennweite F verdoppelt werden $F' = F * 2$, so verdoppelt sich auch das zweidimensionale Abbildungsmaß $\mathbf{T}_c' = (x_i * 2, y_i * 2)^T$.

Die Entfernung, bzw. der Abstand, des Objektes verhält sich ähnlich. Er steht im umgekehrt proportionalen Verhältnis zum Abbildungsmaß mit den Abbildungskordinaten x_i und y_i . Verdoppelt sich also die Entfernung $\mathbf{C}' = (x_c, y_c, z_c * 2)^T$ des identischen Objektes zur Lochblende des Kameramodells, so halbiert sich das Abbildungsmaß $\mathbf{T}_c' = (x_i/2, y_i/2)^T$.

Letztendlich ist auch das Abbildungsmaß von den Koordinaten x_c und y_c abhängig und umgekehrt. Es ist ein direkt proportionales Verhältnis und sagt aus, dass eine Verdopplung der x_c und y_c Koordinaten $\mathbf{C}' = (x_c * 2, y_c * 2, z_c)'$ auch eine Verdopplung des Abbildungsmaß $\mathbf{T}_c' = (x_i * 2, y_i * 2)^T$ zur Folge hat (vgl. Verdopplung der Größe des Objektes).

Daraus ergeben sich folgende Verhältnisse für das gesamte Projektionsmodell.

- **direkt proportionales Verhältnis**

$$F \sim x_i \quad F \sim y_i \quad x_c \sim x_i \quad y_c \sim x_i$$

- **umgekehrt proportionales Verhältnis**

$$z_c \sim \frac{1}{x_i} \quad z_c \sim \frac{1}{y_i}$$

2.2.2 Abbildungsfehler optischer Systeme³

Das Abbildungssystem Kamera ist in der Realität nicht vollständig mit dem Lochkameramodell identisch. Reale Kameras weisen je nach ihrem optischen Strahlengang verschiedene Abbildungsfehler auf, die im folgenden Abschnitt erklärt und an die Quellen [7] und [8] angelehnt sind.

Der Strahlengang wird maßgeblich durch die optischen Elemente der Kamera bestimmt. Man unterscheidet hauptsächlich zwischen Linsen- und Spiegelsystemen. In der Regel treten Abbildungsfehler vom Zentrum zum Rand des Bildfeldes stärker auf. Die für diese Arbeit bedeutendsten Abbildungsfehler sind nachfolgend aufgelistet.

- Chromatische Aberration
- Sphärische Aberration
- Verzeichnung

Linsensysteme weisen oft einen Farbfehler auf, welcher auch als *chromatische Aberration* bezeichnet wird. Dadurch werden unterschiedliche Wellenlängen unterschiedlich stark gebrochen, sodass das Linsensystem keinen einheitlichen Brennpunkt für alle Wellenlängen aufweist. Im Grunde heißt das, dass rote, grüne und blaue Lichtstrahlen in verschiedenen Fokuspunkten gebündelt werden, was zu einem unscharfen und teils farbverschobenen Bild führen kann.

Ein verwandter Abbildungsfehler ist die *sphärische Aberration*. Sie kann in Linsen- und Spiegelsystemen entstehen, wenn Lichtstrahlen eines punktuellen Ursprungs nach dem Durchgang durch das Linsensystem nicht die gleiche Schnittweite haben. Dadurch entstehen mehrere Brennpunkte für die gleiche Wellenlänge (entspricht Farbe). Das führt zu einem weichen und unscharf wirkenden Bild.

Beide Arten von Aberration kann man mit speziellen Linsenkombinationen und besonderen Linsen- oder Spiegelformen vermeiden.

Ein Problem aller realen Abbildungssysteme ist jedoch die Verzeichnung (s. Abbildung 2). Sie ist je nach Aufbau des Linsen- oder Spiegelsystems unterschiedlich stark ausgeprägt. Es wird zwischen kissenförmiger und tonnenförmiger Verzeichnung unterschieden.

³ Entnommen aus eigener Arbeit: [9]

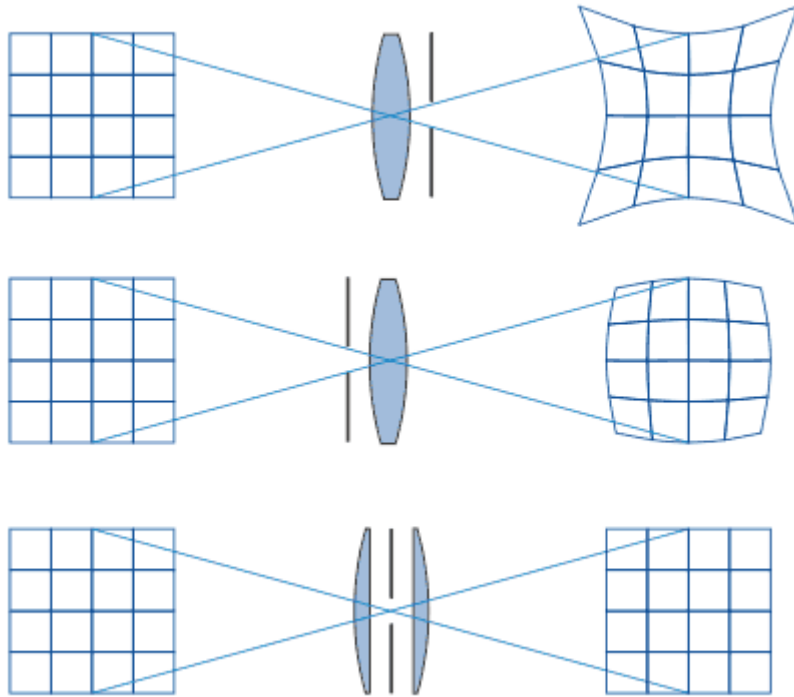


Abbildung 2: Optische Verzeichnung (von oben nach unten: kissenförmig, tonnenförmig, keine) (Quelle: Olypedia - Olympus-Wiki [9])

Die Verzeichnung hat eine besondere Bedeutung für die Computer Vision. Sie hat den Effekt, dass Geraden des realen Raum als gebogene Linien in der projektiven Abbildung dargestellt werden. Dadurch ist zum Beispiel das Finden paralleler Linien innerhalb zweidimensionaler Bilder nur noch eingeschränkt möglich. Es ist daher in vielen Anwendungsfällen von Computer Vision üblich die Verzeichnungen durch das Analysieren der projektiven Abbildung auszugleichen. Der Vorgang steht meist im Zusammenhang mit der Kamerakalibrierung⁴ und wird vor allem beim Arbeiten mit zwei Kameras als Stereokamerapaar, oder bei der Linearisierung relativer Koordinaten der Abbildung gegenüber der absoluten Koordinaten angewendet.

2.2.3 Reflexion von Schallwellen⁵

Ähnlich der Reflektion von Licht an optisch reflektierenden Oberflächen (vgl. Reflektion an Spiegel-Objektiv; diffuse Reflektion an weißen Oberflächen) entstehen

⁴ Zum Ausgleich von Bildverzerrungen oder zur Erfassung der intrinsischen und extrinsischen Kameraparameter

⁵ Anlehnung an Quelle [12]

beim Auftreffen von Schallwellen auf Objekte reflektierte Schallanteile. Bei der Reflektion einer Schallwelle wird, wie auch beim Licht, nicht die gesamte enthaltene Energie reflektiert. Je nach akustischen Materialeigenschaften der Objekte (Resonanzfrequenz, Oberflächenhärte, etc.) werden Frequenzen (Energie) der eintreffenden Schallwelle absorbiert, in kinetische und anschließend in thermische Energie umgewandelt. Der Grad der Absorption und die Absorptionsfrequenzen sind abhängig von den Materialeigenschaften des Objekts. Im Allgemeinen gilt, dass ein Objekt die sogenannte Eigenfrequenz, also die Frequenz bei der das Objekt selbst am besten zum Schwingen angeregt werden kann (Resonanzfrequenz), am besten dämpft. So können einfache Stofftücher und Vorhänge zum Beispiel vor allem hohe akustische Frequenzen dämpfen, wobei der Tieftonanteil (Bass) völlig unberührt bleibt, weil sie eine recht hohe Eigenfrequenz haben. Ein Gegenteiliges Beispiel ist ein Körper mit hoher Masse, niedriger Eigenfrequenz, aber harter Außenwand, eine sogenannte Bassfalle. Diese reflektiert durch die hohe Oberflächenhärte hohe Schallfrequenzen fast vollständig, während tiefe Schallfrequenzen in den Körper eindringen und im Inneren gedämpft werden können. Ein alltagstaugliches Beispiel für Bassfallen sind Polstermöbel, wie Sofas, die jedoch durch ihre weiche Oberfläche auch mittlere und hohe Frequenzanteile dämpfen. Eine Dämpfung des Tieftonbereichs ist somit eher selten durch normale Möblierung zu realisieren.

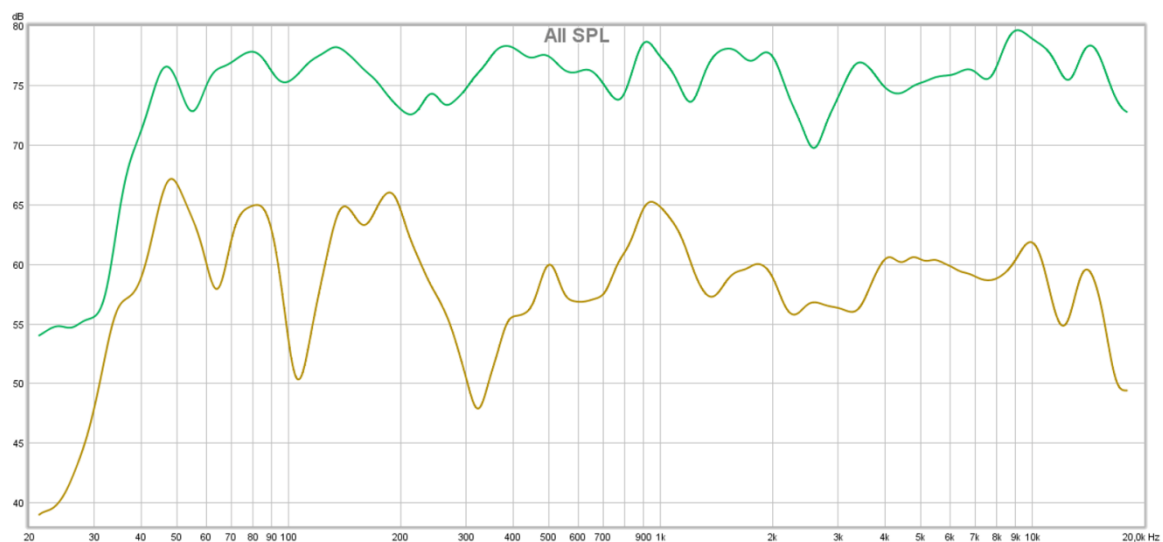


Abbildung 3: Beeinflussung des Frequenzganges eines neutralen 2-Wege-Lautsprechers durch die Raumakustik (Messung in unterschiedlicher Entfernung frontal vor Lautsprecher | Grün: 1m | Gelb: 3m)
(Quelle: eigenes Projekt, Software: Room EQ Wizard v5.14)

Alle Möbel, Stoffe, Pflanzen, Gemälde, und ähnliche Objekte, beeinflussen die Raumakustik und verändern den diffusen Schallanteil, also den an Objekten reflektierten Schall. Das führt dazu, dass ein Schall-Aufnahmegerät, wie z.B. ein Mikrofon, immer den Direktschall und den teilweise gedämpften Diffusschall aufnimmt, wodurch eine Mischung verschiedener Schallanteile zeitlich verzögert und in ihrem Frequenzgang verändert am Aufnahmegerät ankommen. Durch die zeitliche Verzögerung entsteht außerdem, beim Auftreffen von Schallwellen gleicher Frequenzen, Interferenzen mit Veränderung der Amplituden (respektive Lautstärke) einzelner Frequenzen an einem bestimmten Ort. Für alle akustisch normalen Räume, also Räume ohne gleiche akustische Dämpfungseigenschaften über das gesamte hörbare Frequenzspektrum nach DIN 18041 (Mai 2004), gilt deshalb, dass es zu Schallreflektionen in Form von Diffusschall kommt, der mit allen Schallanteilen im Raum Interferenzen ausbildet und somit den wahrgenommenen Frequenzgang in dessen Amplitude verfälscht. Trivial beschrieben bedeutet dies, dass Schallwellen über das gesamte hörbare Frequenzspektrum, die ideal und in gleicher Lautstärke von einer Schallquelle ausgesendet werden, mit ungleicher Lautstärke beim Hörer ankommen, wie es Abbildung 3 zeigt. In der Darstellung kann man deutlich sehen, dass der mit einem Mikrofon aufgenommene Frequenzgang einen Meter vor dem Lautsprecher (Grün) deutlich neutraler (linearer) ist als in drei Metern Entfernung (Gelb). In größerer Entfernung und je nach Ort im Raum kommt es zu deutlichen Interferenzen bestimmter Frequenzen und somit zu Abschwächungen (z.B. 64Hz, 106Hz, 322Hz) und Verstärkungen (z.B. 48Hz, 81Hz, 142Hz, 186Hz, 940Hz) der Schallwellen dieser Frequenzen. Gemessen wurde in einem Raum mit 2,60m Deckenhöhe, 3,45m Breite und 4,70m Länge. Weitere Details zum Raum können der Quelle [10] entnommen werden.

Der Einfluss des Direkt- und Diffusschallanteils der Raumakustik wird im Abschnitt 4.3.2 auf Seite 53 von besonderer Bedeutung sein.

2.3 Praxisrelevante Sensorik

In Fahrzeugen werden verschiedene Sensoren eingesetzt, die auf den grundlegenden physikalischen Gegebenheiten basieren, die in Absatz 2.2 vorgestellt wurden. Derartige Sensoren werden in Fahrzeugen eingesetzt, um die Umwelt wahrzunehmen und die gewonnenen Daten in die Fahrzeugsicherheit- und Assistenzsysteme einfließen zu lassen. Die eingehenden Daten werden mithilfe von Mikrocontrollern und Computersystemen verarbeitet, sodass dem Fahrer entsprechende Informationen gegeben werden, oder autonome Eingriffe in die Fahrt durch Sicherheitssysteme erfolgen können. Für den gegebenen Anwendungsfall werden Sensoren eingesetzt, welche die Umgebung vor dem Fahrzeug in Fahrtrichtung in ausreichender Reichweite (Distanz) wahrnehmen können. Vor allem die Anforderung an eine ausreichende Sensorreichweite führt dazu, dass einige Sensorsysteme nur eingeschränkt für unterschiedliche Fahrsituationen einsetzbar sind. Die verschiedenen Fahrsituationen sind dabei von variierenden Umwelteinflüssen, wie Licht, Wetter und Verschmutzung, und der fahrdynamischen Situation des Fahrzeugs, wie Kurvenfahrten, abhängig.

2.3.1 Kameras⁶

Optische Kamerasysteme stellen unter den eingesetzten Systemen den komplexesten Ansatz dar. Klassische optische, digitale Kameras besitzen im wesentlichen einen Bildsensor und ein Objektiv, welche die Aufnahme der Umgebung beeinflussen. Dabei fallen Lichtstrahlen durch das Objektiv⁷ auf den Bildsensor, wo sie sich im besten Fall fokussieren und ein physikalisches Pixel⁸ treffen auf dem sie Helligkeitsinformationen (Reflektionsverhalten des abgebildeten Objektes) hinterlassen.

Die Gesamtmenge aller Pixel des Bildsensors und deren physikalische Anordnung erzeugt ein zweidimensionales Bild der realen Szene. Der Bildsensor kann daraufhin ausgelesen und die Daten der einzelnen Pixel in digitaler Form auf einem digitalen

⁶ Angelehnt [13]

⁷ Zusammensetzung aus Linsen-, oder Spiegelementen, um eine Fokussierung des diffusen Lichts der Umgebung zu erhalten.

⁸ einheitlich abgegrenzte Fläche auf Sensoroberfläche, die zur Aufnahme und Speicherung von elektrischer Ladung durch den inneren photoelektrischen Effekt dient.

Speichermedium festgehalten werden, wie auf Festplatten, SD-Karten, oder dem Arbeitsspeicher eines Computers.

2.3.1.1 Innerer photoelektrischer Effekt⁹

Die Grundlage zum Verständnis für das Funktionsprinzip moderner Kameras ist der Aufbau von Bildsensoren und deren Nutzung des inneren photoelektrischen Effekts.

Bildsensoren bestehen aus einem Halbleiter-Material, wie z.B. Silizium, auf dem einzelne einheitlich große Flächen, die Pixel, strukturell abgegrenzt sind (s. Abbildung 4). Idealerweise besteht zwischen den einzelnen Pixeln keine elektrische Verbindung, sodass jedes einzelne Pixel elektrische Ladungen aufnehmen und speichern kann. Durch das Eintreten eines Lichtquants (Photon) mit ausreichender Energie (abhängig von der Wellenlänge) in das p- oder n-dotierte Halbleitermaterial des Pixels wird die Energie dieses Lichtquants auf ein Elektron des Halbleitermaterials übertragen. Es entsteht ein negativ geladenes, freies Elektron und ein positiv geladenes Defektelektron (Loch), das sogenannte Elektronen-Loch-Paar.

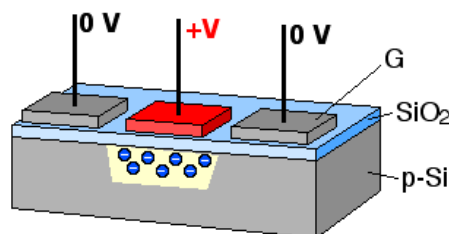


Abbildung 4: Schematischer Aufbau eines Pixel auf einem CCD-Bildsensor (Quelle [11])

Aus physikalischer Sicht bezeichnet man das Elektron bei p-dotiertem Halbleitermaterial als Minoritätsladungsträger und das Defektelektron als Majoritätsladungsträger. Durch das Anlegen einer Spannung mithilfe einer Elektrode, an der Oberfläche des Halbleitermaterials, kann eine Art der beiden Ladungsträger in dem so entstandenen *Potentialtopf* gesammelt werden. Die Majoritätsladungsträger werden stets in das Innere des Halbleitermaterials abgeleitet, wogegen sich die Minoritätsladungsträger im Potentialtopf sammeln. Typischerweise bestehen Bildsensoren aus p-dotiertem Silizium, was bedeutet, dass die freien Elektronen

⁹ angelehnt an [14] S. 983-1000 und [15]

(Minoritätsladungsträger) im Potentialtopf gespeichert und ausgelesen werden. Sie stellen das Nutzsignal dar.

Die dadurch über eine bestimmte Zeit (Belichtungszeit) entstandene elektrische Ladung kann nun anhand von signalverarbeitender Elektronik ausgelesen und gemessen werden (analog, stetig). Durch einen Analog-Digital-Wandler (AD-Wandler) werden abschließend aus dem analogen Signal digitale Werte erzeugt (diskret). Typische Wertebereiche sind 8 Bit (256 Helligkeitsstufen) und 16 Bit (65535 Helligkeitsstufen) je Pixel.

2.3.1.2 Laufzeitproblem

Die Daten (Menge der Helligkeitswerte aller Pixel), die von einer Kamera gewonnen werden, besitzen einen enormen Informationsgehalt, sind aber nicht ohne weiteres eindeutig auszuwerten. Der hohe Informationsgehalt eines Kamerabildes ist durch die typische große Anzahl von Pixeln gegeben, von denen durch die *intrinsischen Kameraparameter*¹⁰ zwar bekannt ist aus welcher Richtung (gesehen vom Brennpunkt) sie Licht aufnehmen, jedoch sind die Informationen über die Entfernungen der Reflektionsquelle der Lichtstrahlen, und damit der abgebildeten Objekte, verloren gegangen. Diese Zusammenhänge müssen also durch die computergestützte Auswertung der digitalen Bildinformationen zurückgewonnen, oder geschätzt werden. In der Fahrzeugtechnik bedient man sich oft Verfahren der Computer Vision, um Informationen aus dem Bildmaterial zu extrahieren, wie z.B. die Segmentierung¹¹ von interessanten Objekten (z.B. Fahrzeugen) gegenüber der Szene, Entfernung zu solchen Objekten (Rückgewinnung von Tiefeninformationen), oder der Klassifizierung von Objekten (z.B. Fußgänger-, Fahrspur-, Fahrzeugerkennung).

Beim Einsatz von Verfahren zur Bildauswertung in Fahrzeugen muss jedoch immer die Berechnungszeit der Verfahren beachtet werden. Durch die Komplexität von zweidimensionalen Bildern ist zum Teil mit einem hohen rechnerischem Aufwand und damit hohen Laufzeiten zu rechnen, was mitunter dazu führt, dass leistungsstarke

¹⁰ innere Parameter eines Kamerasystems, wie Brennweite und physikalische Pixelgröße, aus denen sich Informationen über den Strahlenverlauf des aufgenommenen Lichts im Raum, relativ zur Kamera, ausgehend von jedem einzelnen Pixel, ableiten lassen.

¹¹ Hervorhebung/Abgrenzung einer Gruppe von zusammengehörigen Pixeln in einem digitalen Bild als Objekt. Der Ausschnitt dieser Pixelgruppe aus dem gesamten Bild wird Segment genannt.

Rechner in die Fahrzeuge verbaut werden müssen. Auch dann muss die mögliche lange Antwortzeit (Laufzeit von Signalaufnahme bis Ergebnis vorliegt), durch den hohen Berechnungsaufwand der Auswertung eines in Echtzeit aufgenommenen Kamerabildes, beachtet werden. **Lange Antwortzeiten resultieren in einer niedrigen effektiven Bildwiederholrate und einer zeitlich verzögerten Reaktionsfähigkeit des Fahrassistenzsystems.** Ist die Antwortzeit sehr lang (z.B. eine Sekunde), also die Bildwiederholrate sehr niedrig (hier: ein FPS¹²), dann kann die maximale Geschwindigkeit bei der das Fahrassistenzsystem eingesetzt werden kann nur so hoch sein, wie eine gewünschte Reaktion nach der Antwortzeit noch problemlos ausgeführt werden kann, wie z.B. ein Notbremseingriff beim gegebenen Notbremsassistent. **In realen Echtzeit-Einsatzszenarien ist also die Komplexität von eingesetzten Algorithmen immer ein Kompromiss aus Genauigkeit und Laufzeit.**

2.3.1.3 Physikalische Grenzen und Beschränkungen¹³

Über die zeitliche Verzögerung hinaus müssen außerdem die physikalischen Beschränkungen eines optischen Kamerasystems beachtet werden. So unterliegt das optische System einer Fehleranfälligkeit, bzw. funktionalen Einschränkung, durch Verschmutzungen und Gegenlicht¹⁴, was das aufgenommene Bild verschlechtern kann.

Verschmutzungen erzeugen Obstruktionen¹⁵, die den Lichteinfall auf dem Bildsensor flächenbezogen reduzieren. Die Reduktion des Lichteinfalls und die Größe der betroffenen Fläche auf dem Bildsensor ist durch die Position der Verschmutzung im Lichtweg¹⁶ und deren Beschaffenheit bestimmt.

Gegenlicht beeinflusst ebenfalls Flächen auf dem Bildsensor, also Bildbereiche im digitalen Bild. Als Gegenlicht bezeichnet man Lichtquellen, welche direkt oder indirekt in den Strahlengang des optischen Systems (Kamera) strahlen und darin an Bauteilen des optischen Systems diffus gestreut werden, oder direkt und fokussiert auf den Bildsensor treffen. Fokussierte Gegenlichter sind meist nahezu punktförmig

¹² *Frames per Second*; Bilder pro Sekunde

¹³ [2] S. 546ff. und S. 466f.

¹⁴ Eine Lichtquelle, welche direkt in den Strahlengang (Objektiv) strahlt und dessen Licht auf der Bildebene (Bildsensor) gestreut oder fokussiert auftritt und Überstrahlungen erzeugen kann. Im Allgemeinen wird dadurch Kontrast in betroffenen Bildbereichen verloren.

¹⁵ Verdeckungen, Abdeckung eines optischen Elements (teilweise, vollständig)

¹⁶ Weg eines Lichtstrahls vom Ort der Aussendung (Objekt) durch das Objektiv der Kamera bis zum Eintreffen auf dem Bildsensor

und überstrahlt, was bedeutet, dass mehr Licht ankommt als das Pixel physikalisch gesehen Ladung durch eintreffende Photonen aufnehmen kann. Aus technischer Sicht ist das Pixel in diesem Zustand als gesättigt zu bezeichnen. Alle zusätzlich eintreffenden Photonen werden im resultierenden Bild ignoriert, wodurch Informationen verloren gehen. Wird das analoge Signal eines Pixels des Bildsensors nun durch den Analog-Digital-Wandler der Kamera in ein digitales Bild umgewandelt, so kann das Signal noch zusätzlich verstärkt werden. Wird das Signal nicht verstärkt, so gelten gesättigte Pixel als digitales Maximum des Wertebereichs (z.B. ein Wertebereich von 8 Bit entsprechen 256 Helligkeitsstufen). Wird jedoch das analoge Nutzsignal noch zusätzlich verstärkt, so können bereits nicht-gesättigte Pixel die Ladung eines nicht-verstärkten gesättigten Pixels übersteigen. Alle diese Pixel werden somit bei der Umwandlung in ein digitales Signal als digitales Maximum ausgegeben, wodurch abermals Informationen verloren gehen (Clipping¹⁷). Im Allgemeinen werden durch Gegenlicht auch eintreffende Lichtstrahlen durch das optische System aufgrund von Reflektionen, Brechung und Beugung von Lichtstrahlen gestreut, also beispielsweise örtlich neben der nahezu punktförmigen Abbildung der Gegenlichtquelle abgebildet. Die Streuung erzeugt so auch bei anderen, dadurch betroffenen Pixeln eine Aufhellung und senkt den Kontrast durch das Hinzufügen dieses optischen Störsignals.

2.3.1.4 Rauschen als Störsignal¹⁸

Weitere Störungen des Bildes werden durch das Signalrauschen erzeugt. Rauschen ist im Allgemeinen eine Überlagerung des Nutzsignals (freie Elektronen) mit dem Rauschen als Störsignal. Das Signalrauschen unterteilt sich in zwei verschiedene Rauschanteile, welche sich in ihrem Entstehungsort und Auswirkungen auf das Bild unterscheiden. Der erste und bei modernen Bildsensoren am stärksten auftretende Rauschanteil ist das *Dunkelrauschen*, welches sich wiederum in *thermisches Rauschen* und *Ausleserauschen* unterteilt. *Thermisches Rauschen* entsteht durch den Einfluss von Wärmestrahlung (fernes Infrarot) und führt dazu, dass in der Halbleiterschicht der Pixel des Bildsensors freie Elektronen entstehen, wie auch durch Photonen des sichtbaren Lichts. Das *Ausleserauschen* entsteht in den

¹⁷ Das Abschneiden oder Kappen von Signalen oder Zahlen durch eine gegebene physikalische oder technische Begrenzung (Minimum und Maximum)

¹⁸ angelehnt [16]

signalverarbeitenden Bauteilen beim Auslesen der Pixel. Durch strukturelle Unterschiede und thermische Einflüsse auf die signalverarbeitenden Bauteile wird auf das Signal der einzelnen Pixel ein weiteres Störsignal in Form des Ausleserauschens aufaddiert.

Der zweite und weniger hervortretende Rauschanteil ist das *Photonenrauschen* und spielt in den meisten Anwendungsfällen eine untergeordnete Rolle. Das Photonenrauschen entsteht durch eine nicht konstante Anzahl ankommender Lichtquanten in einem konstanten Zeitraum (Schwankung der Photonenzahl). Dies ist aber nur in lichtschwachen Anwendungsszenarien von Bedeutung, wie z.B. der Astrofotografie, bei denen mitunter über mehrere Stunden Belichtungszeit nur relativ wenige Photonen von einem lichtschwachen Himmelsobjekt aufgenommen werden können. Nicht zuletzt ist jedes Pixel eines idealen Bildsensors ein Photonenzähler, da bekannt ist mit welcher Wahrscheinlichkeit (Quanteneffizienz) und welchen Wellenlängen (Transmissionskurve) ein einzelnes Elektron mit definierter Ladung in der Halbleiterschicht eines Pixels freigesetzt wird.

2.3.1.5 Monovision und Stereovision¹⁹

Beim Einsatz von Kameras in Fahrzeugen wird versucht möglichst viele und vereinfachte Informationen aus den komplexen Bildern zu berechnen. Die Daten sollen im Falle des Notbremsassistenten Informationen, wie die Position und den Abstand zum vorausfahrenden Fahrzeug, enthalten, um berechnen oder schätzen zu können ob und mit welcher Stärke ein Brems Eingriff erfolgen muss.

Oft werden dabei zwei verschiedene Strategien verfolgt. Eine Möglichkeit den Abstand zu vorausfahrenden Fahrzeugen zu schätzen ist der klassische Einsatz einer Kamera (Monovision) an der Fahrzeugfront, die den vorausfahrenden Verkehr beobachtet. Dabei wird versucht Fahrzeuge im digitalen Bild zu finden und die Umrisse des Fahrzeuges herauszustellen. Gelingt dies, kann der Abstand zum gefundenen, vorausfahrenden Fahrzeug grob geschätzt werden. Je nach verwendetem Verfahren kann diese Schätzung recht genau erfolgen. Oft ist dafür eine angelernte Datenbank mit vielen Vergleichsfahrzeugen hinterlegt, in der viele detaillierte Informationen zu Fahrzeugformen und Dimensionen (z.B. Breite) gängiger Modelle hinterlegt sind. So kann beispielweise in der Regel zwischen einem vergleichsweise schmalen PKW

¹⁹ [16] S. 415ff. "Triangulation"

und einem breiteren LKW unterschieden werden. Die dabei abgebildete Größe des Fahrzeugs auf dem Kamerabild wäre bei breiten Fahrzeugen entsprechend größer, obwohl der Abstand der gleiche ist.

Ein genauerer Ansatz ist die Verwendung von zwei Kameras, die als Stereokamerapaar angeordnet sind und ebenso den vorausfahrenden Verkehr beobachten. Das Prinzip ist an das menschliche Sehen mit zwei Augen angelehnt und nutzt die Parallaxe²⁰ zum Bestimmen von Entfernungen. Die Stereovision bedient sich dabei ebenso dem Vorgehen der Monovision, sodass die Position von vorausfahrenden Fahrzeugen im Kamerabild mithilfe nur eines Kamerabildes grob bestimmt werden. Die so gewonnenen eingegrenzten Bildregionen werden nun mithilfe des zweiten Kamerabildes mit der Information der Parallaxe erweitert, sodass die Distanz genauer berechnet werden kann. Der Nachteil der Stereovision gegenüber der Monovision ist allerdings der vergleichsweise höhere Berechnungsaufwand und somit die höhere Laufzeit (s. 2.3.1.2). Je nach geforderter Genauigkeit und Assistenzsystem muss also beachtet werden ob genauere Distanzen nötig sind und die erhöhte Laufzeit akzeptabel ist.

2.3.2 Lidar²¹

Ein Lidar ist die Abkürzung für *Light Detection and Ranging* und bezeichnet den Einsatz eines Lasersystems, um genaue Distanzen zu Objekten mithilfe von Licht zu bestimmen. Häufig ist das Lasersystem dadurch gekennzeichnet, dass nur eine bestimmte Anzahl an Winkelpositionen, also Distanzen in bestimmten Winkelabständen, erfasst werden können, wobei die Erfassung meist nur in der horizontalen Ebene erfolgt. Ein Lidar ist außerdem durch eine bestimmte maximale Aufnahmefrequenz gekennzeichnet, mit der die Szene erfasst werden kann (vgl. Bildwiederholrate). Nicht selten liegt die Bildwiederholrate unterhalb der einer optischen Videokamera.

²⁰ Die Parallaxe ist die scheinbare Veränderung der Position eines Objektes, wenn der Beobachter selbst seine Position ändert (hier: Wechsel eines Auges zum anderen, bzw. Kameras)

²¹ angelehnt an Quelle [17] S. 172ff. "Lidarsensorik"

2.3.2.1 Technischer Aufbau

Der Lidar besteht aus der technischen Sicht, wie sie in der Abbildung 5 zu sehen ist, aus einer einfachen Laserdiode (*Laser Diode*), einem Umlenkspiegel (*Mirror*) und einer Fotodiode (*Photo Diode*).

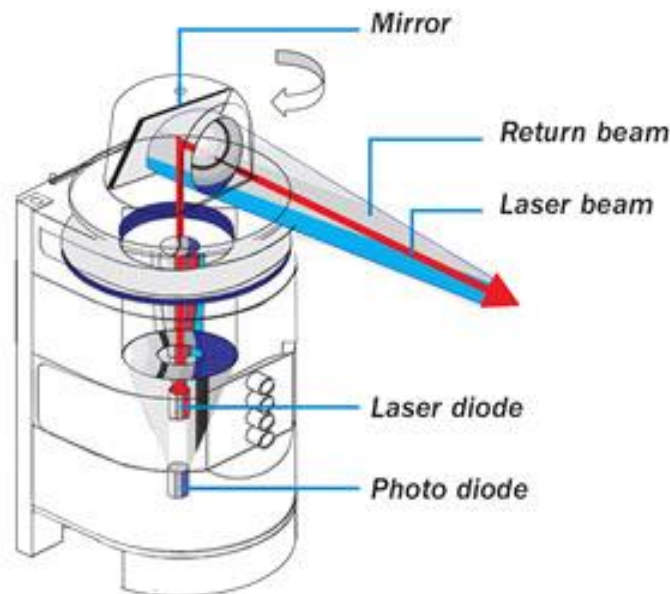


Abbildung 5: Aufbau eines Lidar-Sensors (Angelehnt an Quelle: SICK, Inc.)

Der durch die Laserdiode emittierte Laserstrahl wird an dem Umlenkspiegel um 90° abgelenkt und tritt in den Raum. Dort wird er von einem Objekt in bestimmter Entfernung (diffus) reflektiert. Die reflektierten, abgeschwächten Lichtstrahlen (*Return Beam*) werden nun am Lidar wieder aufgenommen und mit einer Fotodiode detektiert. Die Zeit, die seit dem Aussenden des Lichts durch die Laserdiode und dem Empfangen durch die Fotodiode vergangen ist, wird nun zur Berechnung des vom Licht zurückgelegten Weges benutzt. Das System des Lidar muss dabei sehr kleine Zeitunterschiede im Pikosekundenbereich (Billionstel Sekunden) auflösen können, um eine Genauigkeit der Distanzbestimmung von einigen Millimetern bis Zentimetern zu erreichen. Ein zeitliches Auflösungsvermögen von 33 Pikosekunden entspricht einer möglichen Genauigkeit von ca. einem Zentimeter bei der Distanzbestimmung (mit Lichtgeschwindigkeit $c = 300000 \frac{\text{km}}{\text{s}}$).

2.3.2.2 Funktionsprinzip

Zur Erfassung mehrerer Richtungen in horizontaler Ebene dreht sich der Umlenkspiegel um seine Horizontalachse, sodass eine theoretische Erfassung der Entfernungen um 360° in horizontaler Ebene möglich ist, wenn es die konstruktiven Merkmale des Gehäuses und die Montage am Fahrzeug erlauben. Je nach Drehgeschwindigkeit des Umlenkspiegels dauert die Erfassung (Erfassungszeit) der Distanzen der gesamten Umgebung in horizontaler Ebene solange wie der Spiegel für eine Umdrehung benötigt. Zwar verringert die Einschränkung des Erfassungswinkels von 360° auf beispielsweise 180° die theoretische Erfassungszeit eines Bildes auf die Hälfte einer ganzen Umdrehung, jedoch muss der Spiegel weiterhin den Rest der Drehung vollziehen, um wieder am Startwinkel der Erfassung anzukommen. Damit verbleibt die Erfassungszeit von 180° konstant mit der von 360° Erfassungswinkel. Somit ist lediglich die Drehgeschwindigkeit des Umlenkspiegels das ausschlaggebende mechanische Kriterium für die Wiederholrate²².

Neben dem Einfluss der Mechanik auf die Wiederholrate haben vor allem die rechentechnischen elektronischen Komponenten einen entscheidenden Einfluss auf die maximal mögliche Wiederholrate und das Winkelauflösungsvermögen. Umso schneller (zeitlich) diese Komponenten die Daten der Distanzmessung eines Winkels verarbeiten können, desto kürzer kann die nicht-nutzbare Totzeit bis zur nächsten Distanzmessung eines anderen, darauffolgenden Winkels in Drehrichtung sein. Eine kurze Totzeit ermöglicht entweder höhere Wiederholraten, oder ein höheres räumliches Winkelauflösungsvermögen der Messungen. Ein langsamer drehender Umlenkspiegel würde somit eine wesentlich höhere räumliche Winkelauflösung ermöglichen, jedoch in einer geringeren Wiederholrate resultieren.

2.3.2.3 Einschränkungen

Durch alle vorgestellten Parameter wird deutlich, dass der Betrieb eines Lidar immer ein Kompromiss aus Wiederholrate, räumlichen Winkelauflösungsvermögen und aus wirtschaftlicher Sicht vor allem aus den entstehenden Kosten ist, da das System durch die Ausnutzung der Lichtgeschwindigkeit eine sehr hohe zeitliche Auflösung und somit exakt arbeitende und teure Komponenten benötigt.

²² Anzahl an Erfassungen der gesamten Szene pro Sekunde (vgl. Bilder pro Sekunde, FPS in der Videotechnik)

Darüber hinaus ist das Messprinzip mit Licht einigen materialbedingten Fehlern unterworfen. Sehr stark lichtabsorbierende Oberflächen und hochgradig transparente Werkstoffe verfälschen die Distanzmessung, sodass die eigentliche Zielstellung, die Distanz zum räumlich nächsten Festkörper in Richtung des Lichtstrahls zu messen, nicht für alle Materialien erfüllt werden kann. Außerdem sind Lidar, ebenso wie optische Kamerasysteme, wetterbedingten Einschränkungen unterworfen. Sie reagieren zwar weniger empfindlich auf Gegenlicht, jedoch werden sie im Besonderen durch Regen- und Schneefall eingeschränkt.

2.3.3 Radar²³

Die Radar-Technologie (*Radio Detection And Ranging*) ist heute bei Fahrzeugen mit Notbremsassistent weit verbreitet. Seinen Ursprung hat diese Technologie in der Militärtechnik zur Bestimmung von Distanzen zu Flugkörpern. Die Automobilindustrie nutzt die Technologie auf ähnliche Weise. Durch einen entsprechenden Radar-Sensor, wie er in der Abbildung 6 vom Unternehmen Hella beispielhaft zu sehen ist, können vorausfahrende Fahrzeuge detektiert und der Abstand zu ihnen ermittelt werden. Der Sensor funktioniert im wesentlichen durch das Aussenden von gebündelten elektromagnetischen Wellen (Radiowellen) in Form eines Impulses, der Reflektion dieser an einem Objekt und der Aufnahme der Impulsantwort (reflektierte Radiowellen). Daraufhin werden die gewonnen Messergebnisse, wie die Laufzeit und Frequenz der Impulsantwort, ausgewertet und so Distanzen und Winkelpositionen zu vorausfahrenden Fahrzeugen, Passanten, oder allgemeinen Hindernissen berechnet.

2.3.3.1 Technischer Aufbau

Aus technischer Sicht erfolgt das Aussenden der Radarwellen durch das im Radar-Sensor eingebaute *Radar Front-End*, welches aus einer Platte von mehreren Radioelementen und deren integrierten Verdrahtung besteht. Der Aufbau erinnert zum Großteil an einen Bildsensor mit niedriger Auflösung (Bildsensor: mehrere Megapixel; Front-End: typischerweise 9-128 Elemente).

²³ Angelehnt an Quelle [17] S. 123ff. "Radarsensorik"

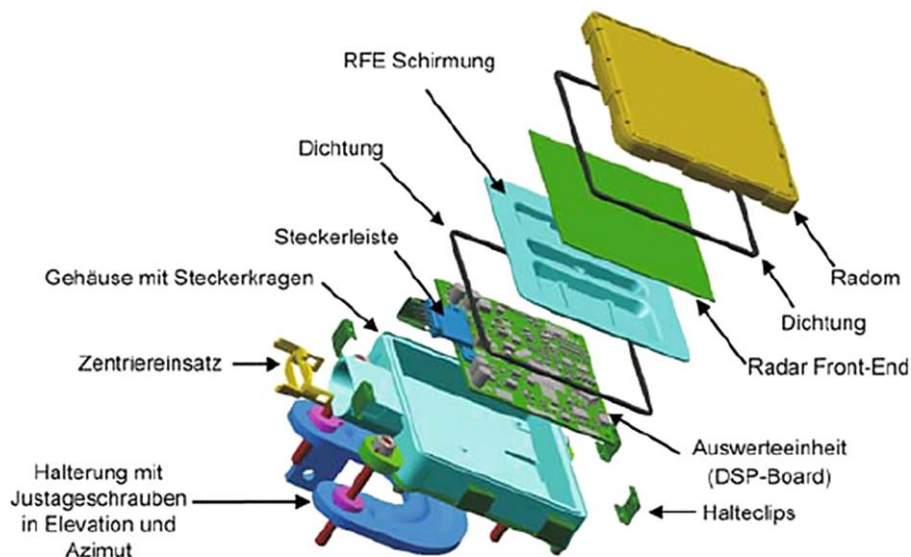


Abbildung 6: Explosionsdarstellung des Hella-ACC-Radars (Quelle: Hella)

2.3.3.2 Funktionsprinzip

Jedes Radarelement sendet eine elektromagnetische Welle aus, welche sich mit Lichtgeschwindigkeit $c \approx 300.000 \frac{\text{km}}{\text{s}}$ fortbewegt. Die Radarwellen verlassen den Sensor gebündelt als Primärstrahl. Die Art der Bündelung wird durch die Bauform der Antenne (Front-End) vorgegeben. Die angestrebte Bündelung ist beim Einsatz in der Fahrzeugtechnik hoch, um ein möglichst hohes laterales und vertikales Auflösungsvermögen zu erreichen. So besitzt die Strahlungskeule²⁴ eine vergleichsweise niedrige vertikale und eine höhere horizontale Ausbreitung. Die niedrige vertikale (vgl. flache) Ausbreitung unterbindet die Reflektion der Radarstrahlen auf der Fahrbahn, oder Deckenreflektionen in potentiellen Tunnelfahrten, und somit längere, verfälschte Signalwege zu einem Objekt.

Am Zielobjekt angekommen wird die elektromagnetische Welle mit unterschiedlicher Intensität und verschiedenen Abstrahlrichtungen reflektiert, man spricht vom Reflexionsvermögen, dem Radarquerschnitt σ , eines Radarzieles. Der Radarquerschnitt wird beeinflusst vom Material, Form und Fläche des Objekts. Im Straßenverkehr relevante Objekte wie ein LKW besitzt beispielsweise einen Radarquerschnitt von $\sigma \approx 1.000 \text{ m}^2$, ein PKW $\sigma \approx 100 \text{ m}^2$, ein Motorrad $\sigma \approx 10 \text{ m}^2$ und ein Fußgänger $\sigma \approx 1 \text{ m}^2$. Das ist jedoch der Idealfall, der von der

²⁴ Charakteristische Form des Raumes in dem sich die Radarstrahlen ausbreiten.

Fahrzeugrückseite ausgeht. Ein großer Einflussfaktor für das Reflexionsvermögen ist außerdem die Geometrie der Flächen des Fahrzeugs, die in die Richtung des Radars zeigen. Flächen, die direkt auf das Abstrahlzentrum des Radars zeigen weisen das beste Reflexionsvermögen auf. Ist die gerade Fläche jedoch im Raum verdreht, bzw. geneigt, so werden die Radarstrahlen nicht mehr zum Radar zurückreflektiert, wodurch das Objekt nicht detektiert werden könnte. Ein vertikal um 45° gedrehter, quer stehender LKW bietet beispielsweise nur noch an den Kanten des Aufbaus und an vereinzelten Stellen des Fahrgestells Flächen, die für den Radar sichtbar sind.

Beim Aussenden der Radarwellen muss das Signal jedoch moduliert werden, um der elektromagnetischen Welle Informationen zur Laufzeitmessung und Wiedererkennung der fahrzeugeigenen Radarwellen mitzugeben. Vor allem letzter Punkt ist technisch einfach umsetzbar, jedoch von enormer Bedeutung, wenn das System fehlerfrei neben anderen Fahrzeugen mit simultan arbeitenden Radaranlagen eingesetzt werden soll.

2.3.3.3 Bestimmung von Informationen aus dem Radarsignal

Die Informationen zur Laufzeitmessung werden in modernen Radaranlagen hauptsächlich mit zwei Ansätzen realisiert. Der erste Ansatz ist das *Puls-Doppler*-Prinzip und sendet die Radarwelle in kurzen Impulsen mit einer definierten Länge des Impulses und Periodendauer aus. Zwischen zwei Impulsen, also innerhalb einer Periodendauer, kann nun eine Laufzeitmessung erfolgen. Dabei wird die Zeit bestimmt, die nach dem Empfangen des reflektierten Signals bis zum Aussenden vergangen ist. Der zurückgelegte Weg der Radarwellen wird mithilfe der Lichtgeschwindigkeit c bestimmt. Das Prinzip des Verfahrens ist der Laufzeitmessung der Lichtstrahlen (ebenso elektromagnetische Wellen) des Lidar identisch und kann dem Kapitel 2.3.2 entnommen werden. Die Impulslänge des Signals bestimmt die Ortsauflösung des Radars. Ist die Impulslänge zu lang können die reflektierten Signale von zwei Objekten, mit nur geringfügig anderen Distanzen zum Radar, nicht mehr eindeutig trennbar sein und so für den Radar zu einem einzigen Objekt werden. Optimal ist eine möglichst kurze Impulslänge, jedoch reduziert das die Möglichkeiten dem Impuls ein eindeutiges Wiedererkennungssignal mitzugeben.

Zusätzlich zur Laufzeitmessung wird nach dem Empfangen die Veränderung der Frequenz des Signals auf Grundlage des Doppler-Effekts analysiert, welcher besagt,

dass die Frequenz des reflektierten Signals nicht gleich der Frequenz des ausgesendeten Signals ist, wenn sich das Zielobjekt relativ zum Empfänger (hier: Radar) hin oder weg bewegt. Das physikalische Prinzip für dieses Verhalten wurde bereits im Jahr 1842 von Christian Doppler beschrieben und ist im allgemeinen für alle Arten von Wellen (wie Schall und Elektromagnetismus) gültig. Durch die relative Bewegung eines Zielfahrzeugs vor dem Radar mit gleicher Geschwindigkeit würde sich die Frequenz des Radarimpulses also nicht verändern. Bewegt sich das Zielfahrzeug hingegen auf den Radar zu, dann wird die Frequenz erhöht. Analog dazu wird die Frequenz niedriger, wenn sich das Zielfahrzeug vom Radar weg bewegt.

Der zweite Ansatz ist das FMCW-Prinzip (*Frequency Modulated Continuous Wave*), dessen auffälligster Unterschied zum *Puls-Doppler-Prinzip* das Aussenden einer fortlaufenden veränderlichen Frequenzfolge, anstelle von Impulsen, ist. Ähnlich dem Puls-Doppler-Prinzip erfolgt beim FMCW-Prinzip die Distanzmessung über die Zeit, die bis zum Empfangen eines ausgesendeten Signals vergangen ist. Darüber hinaus wird im gleichen Zug die Veränderung der Frequenz (*Doppler Shift*) betrachtet, wodurch sich ebenso auf die Differenzfrequenz, ähnlich dem *Puls-Doppler-Prinzip*, schließen lässt. Der Vorteil des FMCW- gegenüber dem *Puls-Doppler-Prinzip* ist die Einfachheit der Signalauswertung beider Informationen aus einer Mischung durch Signalmultiplikation von Ausgangs- und Empfangssignal über die Zeit. Dem so entstandenen Differenzsignal kann die Frequenzabweichung, sowie die zeitliche Länge des Signals entnommen werden.

2.3.3.4 Bestimmung der Winkelposition von Objekten

Die dritte Information, die beim Einsatz eines Radar gewonnen werden muss ist der Winkel bei dem ein Objekt erkannt wurde. Ohne diese Information kann nicht unterschieden werden ob mehrere kleine, oder ein größeres Objekt detektiert wurde. Insbesondere bei Fahrsituationen im Stadtverkehr ist die Information über die Winkelposition, also die Richtung (Winkel) und Distanz, eines Objektes von enormer Bedeutung, da sonst Objekte auf einer benachbarten Fahrbahn oder am Straßenrand ebenso als Hindernis erkannt werden könnten. Diese *false-positives*, also als gefährlich eingestufte ungefährliche Objekte, könnten Notbremseingriffe und

somit Gefahrensituationen auslösen, in denen nachfolgende Fahrzeuge ungewollt involviert werden könnten (z.B. Auffahrunfälle).

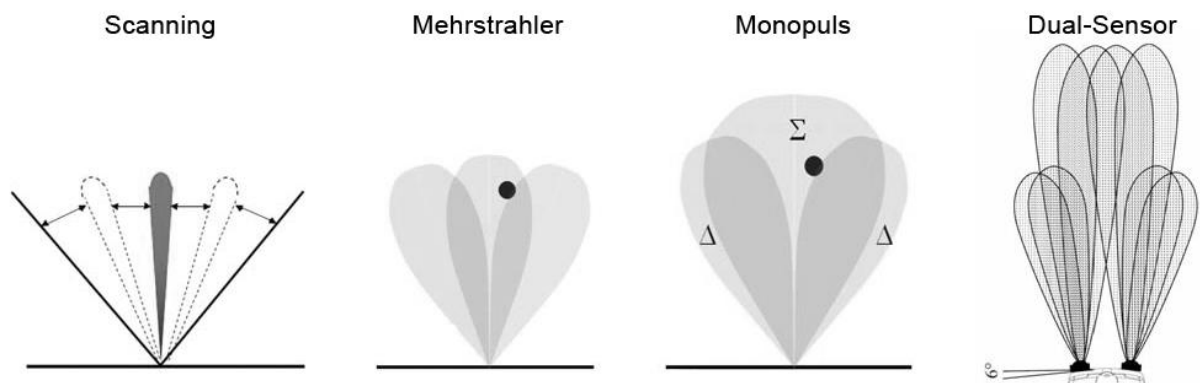


Abbildung 7: verschiedene Ansätze zur Bestimmung der Winkelposition bei Radarsystemen (Angelehnt an Bildquellen: [2] S. 141-144)

Die Winkelposition wird bei den meisten Radarsystemen für Fahrzeuge durch drei Ansätze gewonnen. Der erste Ansatz, das mechanische *Scanning* (s. Abbildung 7), verfolgt ähnlich dem Lidar (s. 2.3.2) einem mechanischen Ansatz zur Erfassung der Winkelposition. Dabei wird der Radar schnell in verschiedene Winkel geschwenkt (z.B. durch vollständige Drehung der Radarantenne oder oszillierende Bewegungen) und so mehrere Radarkeulen erfasst. Der Nachteil dieser Methode ist, dass sich die Zeit zur vollständigen Erfassung der Szene (Erfassungszeit) über alle zu erfassenden Winkel erstreckt, sich also um ein vielfaches gegenüber einer einfachen Distanzmessung verlängert.

Den zweiten Ansatz stellt der *Mehrstrahler* (s. Abbildung 7) dar, bei dem mehrere Antennen innerhalb eines Radar genutzt werden, die in verschiedene Richtungen abstrahlen und so ebenso mehrere Radarkeulen erfassen. Der Vorteil ist die weiterhin kurze Erfassungszeit eines einfachen Radarsensors, jedoch erhöht sich die technische Herausforderung einen solchen Radar in entsprechend kleiner, für Fahrzeuge geeigneter Größe, zu bauen, sodass viele Radarsensoren mit diesem Ansatz mehr Platz in der Fahrzeugfront benötigen.

Der dritte Ansatz, der *Monopuls*, nutzt die im Signal entstehende Phasenverschiebung, die man erhält, wenn zwei Empfangsantennen, die das reflektierte Empfangssignal aufnehmen, technisch so realisiert sind, dass sie hauptsächlich Empfangssignale einer bestimmten Richtung aufnehmen und darüber

hinaus in ihrer räumlichen Position um einige Zentimeter voneinander getrennt sind. Dadurch nimmt eine Antenne bei seitlich positionierten Objekten ein stärkeres Empfangssignal auf, als die andere. Außerdem erhält die weiter vom Objekt abgewandte, bzw. weiter weg liegende, Antenne das Signal zeitlich später, sodass es zu der genannten Phasenverschiebung in der Signalverarbeitung, gegenüber der dem Objekt zugewandten Antenne, kommt. Die Radarkeule wird dabei in der Regel ausschließlich durch eine zentrale Sendeantenne erzeugt, wie in der genannten obigen Abbildung zu sehen.

Ein vierter und relativ neuer Ansatz ist das *Dual-Sensor-Konzept*, welches die Erfassung der Winkelposition durch zwei parallel zueinander abstrahlende Radarsensoren, in Stereo-Anordnung mit einem Abstand von mehreren Zentimetern bis Metern zueinander, nutzt. Die so entstehende Disparität erlaubt es die Position von Objekten in Abstrahlrichtung sehr genau zu bestimmen. Die beiden Radarsensoren sind dabei spiegelsymmetrisch angeordnet, sodass die von ihnen erzeugten Radarkeulen insbesondere den mittleren Bereich in Fahrtrichtung abdecken und die schwächeren Nebenkeulen den nahen Randbereich neben der eigenen Fahrspur (s. Abbildung 7). Vor allem die zentralen und signalstarken Hauptkeulen der beiden Radarsensoren überlappen sich, sodass die Robustheit der Distanzmessung des kritischen Bereichs innerhalb der Fahrspur gegenüber einem einzelnen Radarsensor verbessert werden kann. Die Winkelbestimmung kann nun durch die bekannte Abstrahlrichtung aller Radarkeulen, oder durch die Laufzeitunterschiede und die somit erfolgende Phasenverschiebung im Signal, wie schon im dritten Ansatz, erfolgen. Der Vorteil gegenüber den beiden vorherigen Ansätzen besteht darin, dass die Erfassungszeit wie auch bei einzelnen Radarsensoren kurz ist, da die Messung beider Radarsensoren simultan erfolgen kann. Außerdem kann in der Regel auf zusätzliche Nahbereichssensoren verzichtet werden, da die signalschwächeren Nebenkeulen diesen Bereich abdecken können. Es müssen allerdings entsprechende Positionen in der Front des Fahrzeuges gefunden werden, an denen beide Radarsysteme montiert werden können. Ein weiterer Nachteil sind die höheren Kosten für zwei Radarsensoren gegenüber den beiden Ansätzen mit nur einem Sensor.

3 Technologischer Kontext

In dem folgenden Kapitel werden die geeigneten Möglichkeiten und Algorithmen erarbeitet, um ein geeignetes Verfahren bestimmen zu können, welches die Zielstellung hinreichend erfüllen kann. Dabei wird insbesondere der gegebene Problemfall selbst genauer analysiert und daraufhin Anforderungen an ein geeignetes Verfahren formuliert. Ebenso werden die angewendeten Verfahren verwandter Projekte untersucht, um festzustellen, welche für die eigene Arbeit geeignet sind. Auf dieser Basis erfolgt die Betrachtung bekannter Algorithmen der Computer Vision. Um die theoretischen Grundlagen erarbeiten zu können, muss eine Problemanalyse aufgestellt werden, welche die vorhandene Situation benennt und auf deren Basis eine Formulierung von Anforderungen gegeben ist.

3.1 Problemanalyse

In dem gegebenen Problemfall handelt es sich um eine durchzuführende Kontrollmessung, um festzustellen, ob das gegebene Notbremsassistentensystem die modellspezifische Reaktion reproduzierbar auslöst. Für die Durchführung steht ein PKW der Marke VW, Modell Passat CC (Baujahr 2012), ein PKW der Marke BMW, Modell i3 (Baujahr 2013), sowie ein Rollenprüfstand MFP 3000 von MAHA, zur Verfügung. Der Rollenprüfstand ermöglicht es eine Fahrt auf gerader Strecke und leichten Kurven mit Geschwindigkeiten bis zu 80 km/h zu simulieren. Im Idealfall "bemerkt" der PKW nicht, dass er nur auf einem Rollenprüfstand betrieben wird und lässt das Prüfen aller Fahrerassistenzsysteme zu. Zusätzlich wird der Rollenprüfstand mit einer Schienenkonstruktion ausgestattet, die es erlaubt eine Prüfmaske auf das Auto zuzubewegen und so Fahrerassistenzsysteme mit Abstandsregelung zu prüfen. Diesen Systemen lässt sich auch der zu prüfende Notbremsassistent zuordnen.

Da das System für alle Fahrzeugmodelle einsetzbar sein soll, ist das Auslesen relevanter Daten bezüglich des Notbremsassistentensystems vom CAN-Bus²⁵ ausgeschlossen. Die CAN-Bus-Daten sind modell- und herstellerspezifisch, wodurch

²⁵ *Controller Area Network; Serielles Bussystem; Datenschnittstelle zur Kommunikation zwischen versch. Steuergeräten im Fahrzeug*

eine Beschaffung etwaiger Daten zum vollständigen Auslesen des CAN-Bus mit Rücksprachen und Kosten gegenüber dem Fahrzeughersteller verbunden wäre. Ziel ist es eine Lösung außerhalb der modellspezifischen Schnittstellen zu finden, welche unabhängig vom Fahrzeug operiert, Signale vom Fahrzeug an den Fahrer jedoch aufnehmen kann.

Geeignet dafür ist ein externes Kamerasystem, welches in die Lage versetzt werden muss, Abstände möglichst genau und mit hoher zeitlich kontinuierlicher Frequenz zu messen, um Abstandsmessungen unabhängig von den fahrzeugeigenen Abstandsmessungen durchzuführen. Nach dem Abgleich der durchgeführten Abstandsmessung und der Reaktion des Fahrzeugs kann eine Aussage darüber getroffen werden, ob das Fahrzeug zu den richtigen Zeitpunkten, respektive Abständen, die korrekten Reaktionen ausgeführt hat.

3.2 Bewertungskriterien

Um die Messergebnisse einer Prüfung des Notbremsassistenten bewerten zu können müssen entsprechende Kriterien aufgestellt werden, an dem sich ein theoretischer Prüfer orientieren kann.

Als optimal kann die Kenntnis über die im Fahrzeug eingesetzten Algorithmen, bzw. deren konkreten Reaktionen im Zusammenhang zu absoluten Entfernungen und Relativgeschwindigkeiten, betrachtet werden, bei denen das Notbremsassistentensystem ausgelöst wird. Da die Fahrzeughersteller jedoch derartige Daten nicht offenlegen müssen alternative Ansätze gefunden werden, die eine herstellerübergreifende Prüfung des Notbremsassistentensystems erlauben.

Eine Möglichkeit die nicht offengelegte Daten über das Verhalten des Fahrzeugs unter bestimmten Bedingungen zu ermitteln wäre das Anlegen einer globalen Datenbank auf der Grundlage der Messung einer großen Anzahl verschiedener Fahrzeuge. Durch die so gewonnene große Anzahl von Referenzmessungen über einen langen Zeitraum können Durchschnittswerte für jedes Fahrzeugmodell berechnet und sich an diesen orientiert werden, um die korrekte Funktionalität der Notbremsassistenten einzuschätzen zu können. Die Langzeitbeobachtung durch das

Anlegen der erwähnten Datenbank ermöglicht es darüber hinaus Veränderungen, bzw. Abweichungen, der Reaktionen verschiedener Fahrzeuge über die Zeit festzustellen und an Hersteller weiterzugeben. Damit kann für die allgemeine Prüfung der Notbremsassistenten verschiedener Fahrzeuge auf konkrete Messergebnisse zurückgegriffen werden, die als eindeutiges Bewertungskriterium dienen.

Im Rahmen dieser Arbeit kann jedoch, aufgrund der kleinen Testmenge von Fahrzeugen und Messungen, nicht auf entsprechende Erfahrung in Form einer Datenbank zurückgegriffen werden. Dennoch ist die Bewertung der korrekten Funktion des Notbremsassistenten in eingeschränktem Umfang möglich. Da die Langzeitbeobachtung entfällt, und somit keine mögliche "Alterung" der fahrzeugeigenen Messungen und Reaktionen festgestellt werden kann, kann lediglich die Streuung der Messergebnisse innerhalb eines kurzen Zeitraums als Bewertungskriterium herangezogen werden. Dabei ist anzustreben die Streuung durch mehrere Messdurchläufe, mit mehrmaligen Auslösungen des Notbremsassistenten, zu messen, was im Rahmen dieser Arbeit als hinreichendes Bewertungskriterium dienen soll.

3.3 Verwandte Projekte

Es existieren einige Projekte, die sich dem maschinellen Sehen bedienen, um vorausfahrende Fahrzeuge zu erkennen. Das Interesse ist groß, da der Konkurrenzkampf in der Automobilindustrie hoch ist und neue innovative Systeme für einen besseren Absatz eines Fahrzeugmodells sorgen. Darüber hinaus werden Systeme, welche die Sicherheit der Fahrzeuginsassen erhöhen in der Regel in der Öffentlichkeit als sehr positiv wahrgenommen. In diesem Abschnitt werden dem Thema dieser Arbeit entsprechend einige passende Projekte, bzw. bekannte Ansätze, grob vorgestellt.

Die meisten bewährten Ansätze auf Basis von Radar- und Lidar-Sensoren sind als klassische Ansätze zu verstehen. In dieser Arbeit geht es jedoch u.a. um die Detektion und Abstandsmessung zu vorausfahrenden Fahrzeugen mithilfe einer Kamera. Die nachfolgend vorgestellten verwandten Projekte sind zwei sehr bekannte

Beispiele für die Integration von Kamerasensorik um ein Fahrzeug teil- oder vollständig autonom auf die Umgebung reagieren zu lassen.

Das System zur Abstandsbestimmung zu vorausfahrenden Fahrzeugen des, im Rahmen dieser Arbeit untersuchten, BMW i3 nutzt die entwickelte Soft- und Hardware eines unabhängigen Unternehmens, die in ähnlicher Form u.a. auch schon in LKWs von Volvo eingesetzt wurde (Volvo CWAB) [12]. Das System wurde von dem Unternehmen *Mobileye* entwickelt und in Form einer eigenständigen Hardware, mit Schnittstelle zum Datensystem des PKW, als Mikroprozessor *EyeQ* und *EyeQ²* vermarktet [13]. Das Unternehmen bietet u.a. Softwarefunktionalitäten wie Spur-, Fahrzeug-, Fußgänger- und Verkehrsschilderkennung an (s. Abbildung 8). Das System nutzt zur Beobachtung des vorausfahrenden Verkehrs eine Kamera und wertet den aufgenommenen Bildbereich mit Software-Algorithmen aus. Große Teile der dafür verwendeten Prinzipien sind bekannt, jedoch legt *Mobileye* die verwendeten Algorithmen und Parameter zur Anpassung selbstverständlich nicht offen.

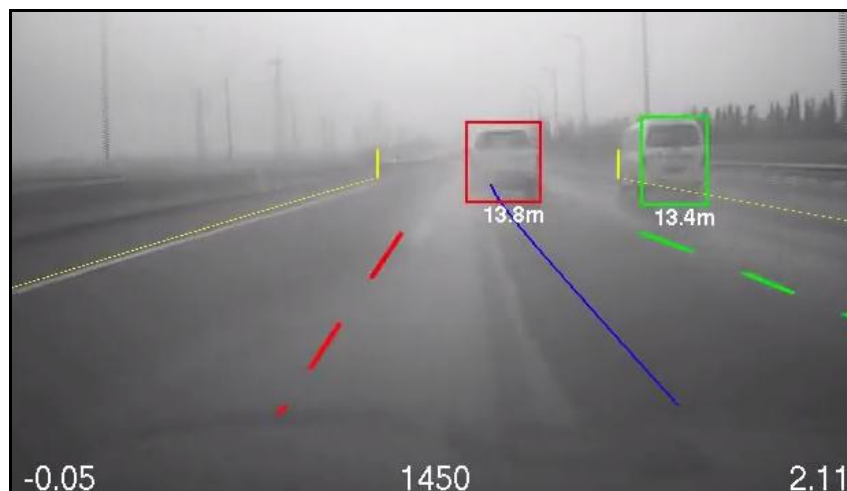


Abbildung 8: Fahrzeug- und Spurerkennung von Mobileye (2005) (Quelle: Einzelbild aus Video von <http://www.mobileye.com/technology/applications/vehicle-detection/forward-collision-warning/>)

Ein anderes Projekt ist die DARPA-Challenge [14], welche unregelmäßig von einer Außenstelle des US-amerikanischen Verteidigungsministerium organisiert wird. DARPA ist die *Defense Advanced Research Projects Agency*, welche teils militärische Ziele verfolgt und mit den organisierten "Challenges" Industrie und Forschung weltweit zur Teilnahme an Herausforderungen anregt. Von besonderem Interesse für diese Arbeit ist die DARPA Urban Challenge, die im Jahr 2007

ausgetragen wurde. Die Aufgabe der teilnehmenden Teams war es ein Fahrzeug vollständig autonom durch einen 93 km langen Stadt-Parcours zu bewegen und eine bestimmte vorgegebene Strecke mit Hindernissen zu absolvieren und das Ziel zu erreichen. Die teilnehmenden Teams statteten die Fahrzeuge, neben anderer Sensorik, auch mit Kameras aus, deren Bildmaterial in Echtzeit ausgewertet wurde.

Neben den beiden sehr bekannten Beispielen gibt es noch eine Vielzahl von Arbeiten der Grundlagenforschung zur Erkennung von Fahrzeugen [15] in bewegten und unbewegten Kamerabildern, wie z.B. mithilfe von *Knowledge-Based Methods* (Symmetrie, Farbe, Schatten, Ecken, Kanten, Textur und Fahrzeugbeleuchtung), *Stereo-Vision-Based Methods* (*Disparity Map* und *Inverse Perspective Mapping*) und *Motion-Based Methods*. Woraufhin meist eine Absicherung der gefundenen, möglichen Fahrzeuge durch *Template-Based Methods* und *Appearance Methods* erfolgt.

3.4 Bekannte Verfahren zur Distanzbestimmung in der Computer Vision

Im Rahmen der Computer Vision wird häufig versucht Informationen über eine reale dreidimensionale Szene aus der zweidimensionalen Abbildung dieser ursprünglichen Szene zurückzugewinnen. Das ist nötig, da durch die Projektion der realen Szene Größen (vgl. Informationen) transformiert und in einer Abbildung, mit weniger Informationsgehalt, resultieren. Man spricht in diesem Zusammenhang von *Linearer Transformation* (Anlehnung an Quelle [4] und [16]) und beschreibt den Informationsgehalt verschiedener Transformationsgruppen. **Bereits der Fakt, dass die durch Projektion erfasste reale Szene von drei Raumdimensionen auf zwei Dimensionen reduziert wird impliziert den Informationsverlust.**

Die gewünschten Informationen können aus der Abbildung, bis hin zur vollständigen Rekonstruktion aller Informationen der realen Szene, durch die Analyse des gegebenen Bildmaterials erfolgen. Dafür werden Invarianten²⁶ durch bekannte

²⁶ unveränderliche Größen

Größen eingeführt und somit Freiheitsgrade²⁷ beseitigt. Ein typisches Beispiel für die Einführung von Invarianten ist die Kenntnis über eine absolute Größe eines realen Objektes, welches in der Abbildung wiedergefunden werden kann. Will man die Abbildung eines Objektes zum Beispiel von der projektiven Gruppe in die metrische Gruppe überführen, so können Teilverhältnisse zusammen mit relativen Entfernungen eingeführt werden.

Gruppe	Freiheitsgrade	Invarianten
Projektiv	8	Doppelverhältnis, Kollinearität, Inzidenz
Affin	6	Vorherige, Parallelität Teilverhältnisse
Metrisch	4	Vorherige, Winkel Relative Entfernungen
Euklidisch	3	Vorherige, Absolute Entfernungen

Tabelle 1: Lineare Transformationen (vgl. Quelle: Thomas Nindel, Überdeckungsfreie Fassadentexturen aus mehreren Nahbereichsfotografien, S. 17)

Im vorliegenden Beispiel der Abstandserfassung mithilfe einer Kamera reicht es aus das relative Größenverhältnis eines abgebildeten markanten Objektes im Zusammenhang mit der relativen Entfernung zu kennen. Dieser Fall der Informationsrückgewinnung wird im nachfolgenden Abschnitt 3.4.1 weiter verfolgt und zu einer praktikablen Strategie zur Distanzmessung von vorausfahrenden Fahrzeugen weiterentwickelt. Darüber hinaus nutzt auch das darauf folgende Verfahren, die Stereo Vision im Abschnitt 0, einen ähnlichen Ansatz und führt bekannte absolute Größen (Invarianten) in Form der extrinsischen und intrinsischen Kameraparameter ein, um Informationen der metrischen und euklidischen Gruppe zurückzugewinnen.

Alles in Allem ist für die Erfassung der Distanzen zu einem vorausfahrenden Fahrzeug, respektive Prüfmaske, lediglich die metrische Gruppe notwendig. Die Distanzmessung erfolgt immer relativ zum Fahrzeug, wodurch ein

²⁷ unbekannte Größen

Fahrzeugkoordinatensystem mit Ursprung am fahrzeugeigenen Sensor zur Distanzmessung aufgespannt wird. Das Fahrzeugkoordinatensystem entspricht wiederum dem Koordinatensystem der Kamera, mit der die Distanzmessung vorgenommen wird. Lediglich der Ursprung der beiden Koordinatensysteme ist verschoben (Translation), weshalb die ermittelte relative Distanz mithilfe der Kamera mit der Messung der Positionsdivergenz zwischen beiden distanzmessenden Systemen ausgeglichen werden muss.

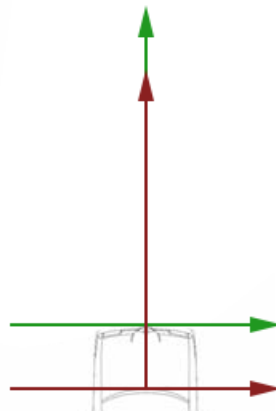


Abbildung 9: Translation des Kamerakoordinatensystem (Rot) in das Fahrzeugkoordinatensystem (Grün) ausgehend vom im Fahrzeug eingesetzten Sensor zur Distanzmessung vorausfahrender Fahrzeuge
[eigene Darstellung]

In der Abbildung 9 wurde eine solche Translation der Koordinatensysteme beispielhaft dargestellt, wobei die Kamera (im Ursprung des roten Koordinatensystems) auf der Windschutzscheibe angebracht wurde und das Fahrzeug über einen Radarsensor an der Front des Fahrzeugs (im Ursprung des grünen Koordinatensystems) verfügt. Die Translation der Koordinatensysteme erfolgt nach der Berechnung der Distanz ausgehend von der Kamera (Kamerakoordinatensystem), also am Ende des zu entwickelnden Verfahrens, durch triviale Addition der vorzeichenbehafteten Positionsdivergenz, bzw. -distanz.

3.4.1 Marker Recognition²⁸

Die Marker Recognition (vgl. "*Marker Detection and Tracking*") ist ein Verfahren in der Computer Vision um definierte und optisch gut zu unterscheidende Marker in einem Bild zu finden und verschiedene Eigenschaften, die den Markern entnommen werden können, auszuwerten. Dazu zählen relative Eigenschaften wie Translation, Rotation, und Größe der Marker im zweidimensionalen Bild.

Um Marker für die Distanzbestimmung zu nutzen, ist es nötig die relative Größe eines Markers, oder die relativen Abstände mehrere Marker zueinander zu bestimmen. Dabei muss die absolute Größe des Markers, oder der absolute Abstand der Marker zueinander bekannt sein. Aufgrund der bekannten absoluten Größen im Verhältnis zu den relativen Größen, kann eine zuverlässige Aussage über die absoluten Distanzen gemacht werden.

Ist beispielsweise bekannt, dass ein Referenzmarker 10 Zentimeter absolute Größe besitzt und er bei einer gemessenen Distanz von einem Meter eine relative Größe von 100 Pixel ausweist, so kann vorhergesagt werden, dass ein Marker mit den gleichen Eigenschaften des Referenzmarkers und nur 50 Pixel relative Größe, trotzdem 10 Zentimeter absolute Größe besitzt, da er baugleich zum Referenzmarker ist, und doppelt so weit entfernt sein muss wie der Referenzmarker, da seine relative Abbildung im zweidimensionalen Bild nur halb so groß ist. Es ergibt sich also ein einfaches umgekehrt proportionales Verhältnis f der Entfernungen des Referenzmarkers d_{referenz} zu allen anderen Markern mit gleichen Eigenschaften anhand der bestimmten Pixelbreiten des Referenzmarkers p_{referenz} und des gemessenen Markers p_{relativ} .

$$f = \frac{1}{p_{\text{relativ}}}$$

$$d_{\text{Marker}} = f \times p_{\text{referenz}} \times d_{\text{referenz}}$$

Die Abbildung wird durch das Einführen der absoluten Entfernung zu einem Marker, oder der Gruppe von Markern, in die metrische Gruppe überführt, sodass eine Aussage über die absoluten Entfernungen und relativen Positionen in Abhängigkeit von der Kameraposition im Raum (vgl. reale Szene) getroffen werden kann.

²⁸ angelehnt an Quelle [19]

3.4.2 Stereo Vision²⁹

Im Gegensatz zu der Distanzbestimmung durch Marker nutzt man bei der Stereo Vision die Disparität zweier Kameras ([17] S. 415ff.).

Beide Kameras können zueinander beliebig ausgerichtet sein. Die überschneidenden Gesichtsfelder ergeben einen gemeinsam sichtbaren Bereich, der für die Ermittlung von Invarianten von Bedeutung ist.

Für die Bestimmung von Distanzen ist hauptsächlich der Einsatz von zwei identischen, achsparallelen Kameras mit definiertem Abstand von Bedeutung. Aufgrund dessen wird im nachfolgenden Abschnitt vorrangig der Einsatz von zwei Kameras als zueinander fixiertes Stereokamerapaar betrachtet.

Mit der Anwendung eines gerichteten Stereokamerapaar versucht man in der Regel annähernd das Prinzip des natürlichen Stereosehens nachzuahmen (vgl. menschliches Augenpaar). Das heißt, dass die Kameras in einem definierten Abstand zueinander fixiert sind und gemeinsam in eine definierte Richtung "blicken". Im Fall des menschlichen Sehens kreuzen sich die optischen Achsen der Kameras (vgl. Augen) und sind auf ein Objekt im Raum gerichtet, sodass sie auf einen gemeinsamen Punkt fixiert sind. Durch die unterschiedlichen Winkelausrichtung beider Augen kann die Entfernung zum fixierten Objekt approximiert werden, wobei der fixierte Punkt bei der Überlagerung der Bilder beider Augen theoretisch an den gleichen Bildkoordinaten abgebildet werden. Umso weiter entfernt ein fixiertes Objekt ist, desto mehr nähern sich die optischen Achsen der Augen einer parallelen Ausrichtung an. Dieses Verhältnis wird Parallaxe genannt.

Im Gegensatz zum menschlichen Auge nutzen die meisten Stereokameras das Prinzip der parallelen Ausrichtung der optischen Achsen, sodass sich die optischen Achsen im Unendlichen schneiden, wodurch unendlich weit entfernte Objekte in beiden Kameras auf den gleichen Bildkoordinaten abgebildet werden. Die Abbildung naher Objekte wird jedoch maßgeblich durch eine scheinbare Verschiebung zwischen beiden Kamerabildern bestimmt. Der Effekt der Parallaxe tritt ein, indem ein Objekt in beiden Bildern auf unterschiedlichen Bildkoordinaten abgebildet wird, woraus sich bestimmbare Winkelabweichungen ergeben.

²⁹ angelehnt an Quelle [20] S. 415ff.

Durch die Einführung von intrinsischen³⁰ und extrinsischen³¹ Kameraparametern ist es möglich die Abbildung von der projektiven Gruppe in die metrische oder gar euklidische Gruppe zu überführen. Die Voraussetzung für die Überführung in die metrische Gruppe ist das Einbeziehen der intrinsischen Kameraparameter, wodurch Aussagen über die relativen Positionen von Objekten, ausgehend von der Kamera, ermöglicht werden. Sollen absolute Positionen bestimmt werden ist es über die intrinsischen Kameraparameter hinaus nötig auch die extrinsischen einzubeziehen, wodurch auch die Lage der Stereokamera im Raum bekannt ist und so durch die bereits bekannten relativen Entfernungen (zur Stereokamera), sowie relativer Position (durch intrinsische Kameraparameter mit Brennweite, Sensorgröße) die absolute Positionen bestimmt werden können.

3.5 Auswahl geeigneter Ansätze

Für die Konzeption von Verfahren, welche die Erfassung der Distanzen von einem zu testenden Fahrzeug zu einem Hindernis und der damit verbundenen Auslösung des Notbremsassistenten ermöglichen, müssen geeignete Ansätze gefunden und daraufhin angepasst werden. Dazu werden die in diesem Kapitel vorgestellten grundlegenden Ansätze der Computer Vision kritisch betrachtet, sodass geeignete Lösungen konzipiert werden können, die möglichst viele Kriterien der Zielstellung (s. 1.2 und 3.1) erfüllen. Die Stereovision wird bereits in der Zielstellung aufgrund betriebsnaher Kriterien ausgeschlossen, soll jedoch in dieses Kapitel zur Auswahl geeigneter Ansätze als möglicherweise geeigneter Ansatz einfließen.

Ein wichtiger Punkt der Zielstellung ist es möglichst robuste Ergebnisse mit genauen Distanzmessungen zu erhalten. Das ist durch beide vorgestellten Ansätze realisierbar und verlangt lediglich einer genauen Kalibrierung, um die metrische oder euklidische Gruppe erreichen zu können, wobei die metrische Gruppe in jedem Fall für eine Distanzmessung und Beurteilung der korrekten Funktion des Notbremsassistenten ausreichend ist.

³⁰ inneren Kameraparameter: Brennweite, Sensorgröße (Pixelskalierung in x- und y-Richtung), Bildmittelpunkte (x-Wert und y-Wert)

³¹ äußeren Kameraparameter: Translation und Rotation der Stereokamera im Raum

Die Kalibrierung bei der *Marker Recognition* kann so mit dem einmaligen Messen der absoluten Distanz zu dem Marker, bzw. den Markern, erfolgen, wobei die relative Größe der Marker in der Abbildung bei diesem absoluten Abstand ebenso festgehalten wird. Die Kalibrierung kann bei gleichbleibenden Markern und Markieranordnung beibehalten werden.

Bei der Stereo Vision wiederum erfolgt eine Kalibrierung durch die Erfassung der intrinsischen, und bei Bedarf zusätzlich der extrinsischen, Kameraparameter. Diese Kalibrierungsdaten können entweder ermittelt werden oder bekannt sein, wie z.B. intrinsische Daten über die Brennweite und Sensorgröße oder extrinsische wie die Pose³² der Stereokamera im Raum. Die Ermittlung der intrinsischen Kameraparameter kann mithilfe der optischen Kalibrierung mit einem Schachbrettmuster und entsprechendem Algorithmus zur Auswertung der gesammelten Bilddaten erfolgen. Die extrinsischen Kameraparameter, also die Pose der Kamera, können als Translation des relativen Kamerakoordinatensystem in absolute Weltkoordinaten genutzt werden und nur durch bekannte absolute Größen ermittelt werden. D.h. sie werden entweder gemessen, oder anhand vorher gemessener Größen (z.B. feste, bekannte, definierte Marker im realen Raum) aus dem Bildmaterial heraus bestimmt. Die intrinsischen Kameraparameter umfassen jedoch bereits alle relevanten Daten zur Bestimmung der Distanzen zu einem vorausfahrenden Fahrzeug (respektive Prüfmaske), weshalb eine Überführung in die euklidische Gruppe nicht nötig ist.

Ein weiterer Punkt der Zielstellung umfasst die geringe Berechnungszeit und Ressourcen-Schonung für das im Prüfbetrieb eingesetzte Verfahren. Theoretisch ist die Stereovision der Monovision dabei unterlegen, da die doppelte Anzahl Pixel (zwei Bilder statt ein Bild) ausgewertet werden muss. Die robuste Berechnung der Tiefeninformationen bei der Stereovision ist außerdem an entsprechend robuste Verfahren geknüpft, wozu klassische Verfahren, wie das *Block Matching* (BM) und *Semi Global Block Matching* (SGBM) [17] nur bedingt in der Lage sind, da sie zu Bildartefakten und der Zusammenfassung von Bildregionen mit ähnlicher Distanz, und damit sprunghafter Veränderung der ermittelten Tiefeninformationen, neigen. Robustere Verfahren basieren auf *Features*³³ und ermitteln die Distanz durch die

³² Translation und Rotation

³³ markante Merkmale in einem Bild, welche gut beschreibbar sind und wiedergefunden werden können; s. Quelle [12] S. 321ff.

Betrachtung der Verschiebung einiger markanter Punkte (*Landmarks*) zwischen den beiden Stereobildern.

Abgesehen von den theoretischen Nachteilen der Stereovision kann die Berechnungszeit durch Beschränkung auf eindeutig beschreibbare Bildregionen (Indikatoren wie Marker, s. 4.2) gering gehalten werden, sodass nicht mehr das gesamte Bild der Stereokamera ausgewertet werden muss. Analog ist dies natürlich auch auf die Monovision anwendbar, sodass auch dort der Vorteil der deutlich geringeren Berechnungszeit genutzt werden kann.

4 Konzeption geeigneter Verfahren

Im nachfolgenden Kapitel werden einige bekannte Verfahren kritisch betrachtet und untersucht, ob Anpassungen beziehungsweise Erweiterungen erfolgen können, um geeignetere Verfahren für die eigene Arbeit zu konzipieren. Im späteren Verlauf dieses Kapitels werden eigene Verfahren schließlich selbstständig konzipiert und im Kapitel 5 in einen Software-Prototypen überführt.

4.1 Distanzmessung mit einer Kamera: Einsatz der Monovision

Die Distanzmessung erfolgt aus technischer Sicht traditionell mit dafür konzipierten Messgeräten. Solche Messgeräte basieren vor allem auf den im Kapitel 2.3 erwähnten Technologien basierend auf Radar-, Laser-, und klassischer Kameratechnik. Vor allem die Distanzmessung mit Kameras erfolgt meist mit zwei Kameras, die wie das menschliche Augenpaar gerichtet in eine Richtung schauen und sich die Disparität zu Nutze machen, um absolute (real räumliche) Distanzen zu Objekten zu berechnen.

Über diese klassischen Ansätze hinaus gibt es noch zwei Möglichkeiten mit nur einer Kamera absolute Distanzen zu bestimmen. Der erste Ansatz nutzt das **Verhältnis von mindestens einer absoluten zu einer relativen Größe**. Beide veränderliche Größen müssen dabei jedoch ein **Proportionalitätsverhältnis zueinander aufweisen**, also immer im selben Verhältnis zueinander stehen. Ein anwendungsbezogenes Beispiel dafür ist das Verhältnis der relativen, abgebildeten Breite der Rückseite eines vorausfahrenden PKW zu der absoluten, realen Distanz zur Rückseite des vorausfahrenden Fahrzeugs. Das so gebildete Verhältnis kann mithilfe der Mathematik und Projektion auf den **Strahlensatz** übertragen werden, der im Fall der Projektion aussagt, dass sich die zwei veränderlichen Größen (absolute Distanz, relative Abbildungsgröße) umgekehrt proportional zueinander verhalten (siehe Kapitel 3.4.1).

Misst man nun einmalig, zu einem beliebigen Zeitpunkt, die absolute reale Distanz bei einer beliebigen relativen Abbildungsgröße des Objektes, dann kann man diesen

Referenzwert nutzen, um die absolute, reale Distanz für alle anderen Abbildungsgrößen zu berechnen. Die halbe relative Abbildungsgröße ausgehend vom Referenzwert würde somit die doppelte absolute Distanz ergeben. Durch den aufgezeigten Ansatz kann somit mit nur geringem algorithmischen Laufzeitaufwand die absolute Distanz trivial bestimmt werden. Für diesen Ansatz sind keinerlei Kenntnisse über das Kamerasystem als solches nötig, sodass jederzeit ein anderes Kameramodell eingesetzt werden kann, bei denen sich intrinsischen, kameraspezifischen Parameter, wie Brennweite und Sensorgröße, unterscheiden (Kamerakalibrierung). Es muss lediglich erneut einmalig ein Referenzwert für dieses Kameramodell bestimmt werden, der als Vergleichswert für alle nachfolgenden Messungen gilt (Referenzkalibrierung).

Ein anderer Ansatz nutzt eben diese **intrinsischen Kameraparameter, um die absolute Distanz aus der relativen Abbildungsgröße und einer zweiten absoluten Größe zu berechnen**. Die zweite absolute Größe ist in diesem Anwendungsfall die **reale, absolute Breite der Fahrzeugrückseite** zu der die Distanz bestimmt wird. Die Funktionsweise erklärt sich erneut am einfachsten durch die Projektion (s. Kapitel 2.2.1). Die intrinsischen Kameraparameter sind bei Beachtung des Strahlensatzes der Projektion optischer Systeme beide konstanten Größen auf Seite der Kamera in Form von Sensorgröße und Brennweite. Ist nun die absolute Breite der Fahrzeugrückseite bestimmt worden (Referenzkalibrierung) und vergleicht diese anschließend mit ihrer relativen Abbildung auf dem Kamerasensor, so ist nur noch die absolute Distanz als unbekannte Größe vorhanden. Durch die bekannten intrinsischen Kameraparameter kann die relative Abbildungsbreite der Fahrzeugrückseite nun ins Verhältnis gestellt und die absolute Distanz ebenso trivial berechnet werden. Lediglich die Bestimmung der intrinsischen Kameraparameter ist nicht trivial und zeigt sich in den meisten Anwendungsfällen als aufwändig und zeitraubend.

Bei beiden vorgestellten Ansätzen müsste außerdem theoretisch, neben der Kamerakalibrierung beim Wechsel des Kameramodells, außerdem mit jeder Veränderung des vorausfahrenden Fahrzeugtyps erneut eine Referenzkalibrierung durchgeführt werden, um die Fahrzeugparameter zu erfassen. **Das erste Verfahren benötigt als Referenzwert die absolute Distanz und das zweite Verfahren benötigt die absolute Breite, bzw. Größe der Fahrzeugrückseite.** Für dieses Problem kann jedoch

auf einen einfachen Lösungsansatz zurückgegriffen werden, welcher im nächsten Kapitel (s. 4.2, S. 41) vorgestellt wird.

Der Fokus dieser Arbeit liegt auf einem möglichst robusten Verfahren mit geringer algorithmischer Laufzeit (vgl. schneller Algorithmus) und geringem Kalibrierungsaufwand.

In direkten Vergleich zwischen der Stereovision und Monovision sind durch beide vorgestellten Ansätze enorme Laufzeitvorteile und eine robustere Bestimmung Distanz zu der Prüfmaske möglich. Unter Beachtung aller Gesichtspunkte ist der vielversprechendste Ansatz demzufolge der Einsatz von Monovision mit einmaliger Bestimmung der realen, absoluten Distanz zur Fahrzeugrückseite.

Um das Notbremsassistentensystem eines Fahrzeugs zu testen, muss jedoch auch eine zuverlässige, robuste Erkennung der Fahrzeugrückseite und eine anschließende möglichst genaue Bestimmung der absoluten Distanz gewährleistet sein. Außerdem soll im Zweifelsfall ausschließlich die korrekte Fahrzeugrückseite des vorausfahrenden Testfahrzeugs erkannt werden, auch wenn in Videoaufnahmen mehrere Fahrzeugrückseiten sichtbar sind. Den Problemen kann am besten durch die Verwendung von Markern begegnet werden. Marker eignen sich besonders gut, um das vorausfahrende Testfahrzeug zu markieren und die Erkennung der Fahrzeugrückseite auf das korrekte Fahrzeug zu beschränken. Darüber hinaus bieten Marker, durch die definierte Verwendung einer bestimmten Form und Anordnung, die Erkennung der Fahrzeugrückseite robuster, sowie die Berechnung der Distanz genauer umzusetzen, als es durch eine beliebige, unscharf abgegrenzte Fahrzeugrückseite möglich wäre. Der Vorteil des Einsatz von Markern wurde bereits im Kapitel 3.4.1 ausführlich aufgeführt.

4.2 Einführung von normierten Markern

Für die Bestimmung der Abstände wird jedoch im Gegensatz zu dem genannten Kapitel nicht die Größe der Marker beachtet, sondern mehrere, definiert angeordnete Marker verwendet. Dabei wird stets versucht die Mittelpunkte der Marker (fortan allgemein als Position bezeichnet) so genau wie möglich zu bestimmen. Als Anordnung können zum Beispiel zwei Marker vertikal oder horizontal angebracht und die relative Distanz zwischen den beiden Markern bestimmt werden. Das hat den Vorteil, dass größere relative Pixelabstände in der digitalen Abbildung (vgl. Einzelbild einer Videoaufnahme) zwischen den absoluten, physikalischen Zentren der Marker erreicht werden können, als bei der Messung der Projektionsgröße eines scheinbar kleineren einzelnen Marker. Größere relative Pixelabstände resultieren wiederum in einer genaueren Bestimmung der absoluten Distanz, da das Auflösungsvermögen der digitalen Abbildung besser ausgenutzt wird. Es muss lediglich darauf geachtet werden, dass die Anordnung der Marker auf der Fahrzeugrückseite so gewählt wird, dass die Marker beim kleinstmöglichen Abstand noch in das Sichtfeld der Kamera passen. Das kann erreicht werden, wenn entweder die Anordnung der Marker entsprechend verändert, oder die Brennweite des Kameraobjektivs so gewählt wird, dass die gesamte Anordnung der Marker das Bild sinnvoll ausfüllen.

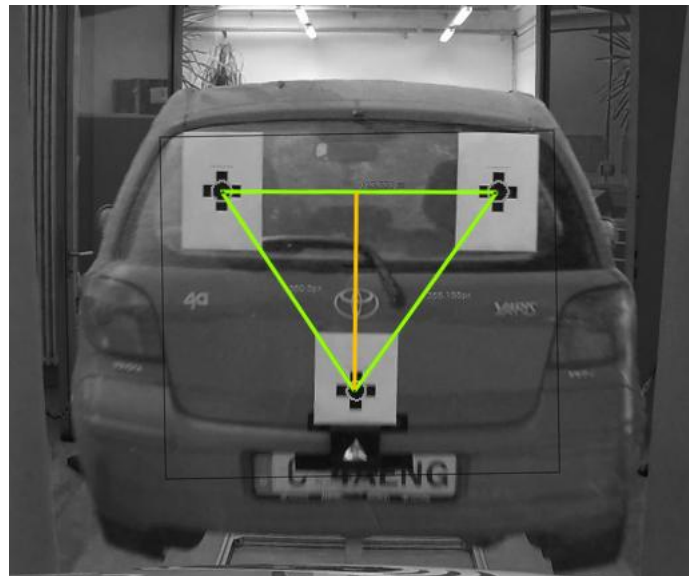


Abbildung 10: Dreieckige Anordnung von Markern auf einer Prümmaske [eigene Darstellung]

Eine sinnvolle Anordnung der Marker auf der Fahrzeugrückseite ist ein gleichschenkeliges Dreieck. Diese Sonderform des Dreiecks bietet bei entsprechender

Ausrichtung zwei diagonal verlaufende Schenkel und eine horizontale Kante, wie in Abbildung 10 zu sehen. Die horizontale Kante verläuft bei dieser Ausrichtung entlang den ebenso horizontal angeordneten, physikalischen Pixel (Def.: lichtempfindliche Siliziumflächen auf Fotosensor). Das führt bei idealer horizontaler Ausrichtung des Dreiecks und Kamera dazu, dass die horizontale Kante des Dreiecks keine vertikale Ausdehnung auf dem Fotosensor besitzt, was bedeutet, dass die physikalischen Zentren der Marker theoretisch nur eine ganzzahlige relative, abgebildete Distanz aufweisen können. Die erreichbare Genauigkeit wäre so durch die Auflösung beschränkt. Dem wirkt der Einsatz von den zu dieser horizontalen Kante diagonal verlaufenden Schenkel des Dreiecks entgegen. Der diagonale Verlauf nutzt nicht nur die horizontale Dimension des Fotosensors, sondern auch die vertikale Dimension. So sind nicht-ganzzahlige, rationale Abstände durch die Abbildung bestimmbar, was die Genauigkeit, durch die bessere Ausnutzung der Auflösung des Fotosensors, signifikant erhöht.

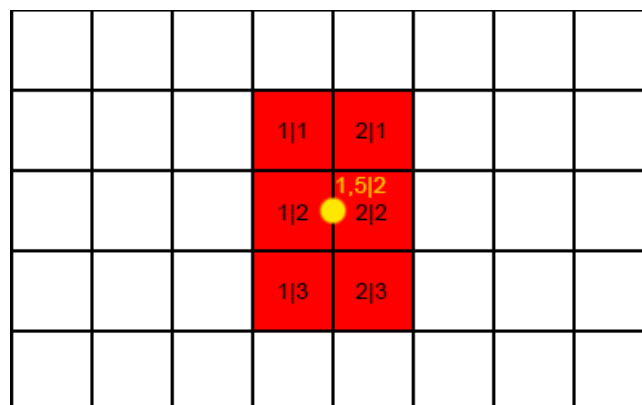


Abbildung 11: Berechnung der Mittelpunktkoordinate (Gelb) eines abgebildeten Objekts (Rot) mit Subpixelgenauigkeit [eigene Darstellung]

Im Allgemeinen kann durch die Bestimmung von Abständen zwischen den ganzzahligen Koordinaten der drei Eckpunkte des gleichschenkligen Dreiecks bereits eine Erhöhung der Genauigkeit auf $\frac{1}{3}$ Pixel erreicht werden. Man spricht in diesem Zusammenhang von Subpixelgenauigkeit, also einer Berechnung von Positionen und Längen zwischen zwei ganzzahligen Pixeln. Eine anschauliche Erklärung dafür ist die Erfassung von scharf abgegrenzten Objekten auf dem Fotosensor, dessen Abbildung sich auf mehrere benachbarte Pixel verteilt (siehe Abbildung 11). Ein solches Objekt ist zwar nur auf Pixeln mit ganzzahligen Koordinaten abgebildet, jedoch erhält man keine ganzzahligen, sondern rationale Koordinaten, wenn man

den Mittelpunkt (Gelb) aller zu dem Objekt gehörigen Pixel (Rot) über das arithmetische Mittel bestimmt. Das Ergebnis für die Breitenkoordinaten (X-Koordinaten) ist in diesem Beispiel rational, das Ergebnis für die Höhenkoordinaten (Y-Koordinaten) ist ebenso rational, obwohl erneut ein ganzzahliges Ergebnis vorzuliegen scheint. Die einfachste Begründung dafür ist die theoretisch erreichbare Genauigkeit bei der Mittelung von drei ganzen Zahlen (drei Höhenpixel in Reihe) über das arithmetische Mittel. Das Ergebnis kann theoretisch in $\frac{1}{3}$ -Schritten variieren, also als Bruch dargestellt werden. Die so erreichte Genauigkeit beträgt $\pm \frac{1}{n}$ Pixel, wobei n die Anzahl der zur Berechnung herangezogenen ganzzahligen Pixelkoordinaten ist.

Die durch die drei Seiten der Dreieckanordnung erreichte Subpixelgenauigkeit wird nachfolgend als Grund-Subpixelgenauigkeit bezeichnet, da in späteren Abschnitten weitere Subpixel-Ansätze zur konsekutiven Erhöhung der nun erreichten Genauigkeit vorgestellt werden.

Zusätzlich kann mithilfe der Breiten (horizontale Anordnung) und der Höhen (diagonale Anordnung der Schenkel des Dreiecks) festgestellt werden, ob das Dreieck räumlich verdreht abgebildet wird und somit die Kamera bzw. die Prüfmaste nicht parallel zueinander ausgerichtet sind. Entsprechende Korrekturrechnungen können damit erfolgen, falls sie benötigt werden.

4.2.1 Gradient-based Pattern Finding Algorithmus (GPF)

Beim GPF wird versucht den normierten Marker im gesamten Bild mithilfe von Helligkeitsverläufen (Gradienten) zu finden und dessen genaue Position zu bestimmen.

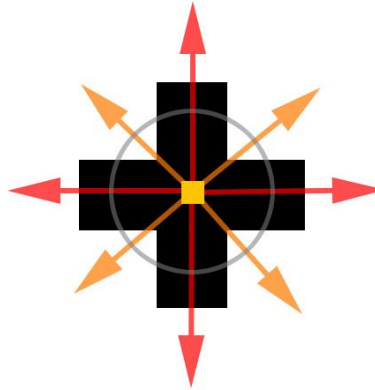


Abbildung 12: Gradient-based Pattern Finding, mit den Diagonalen (Orange), Achsparallelen (Rot) und einheitlicher Gradientenlänge (Grau) [eigene Darstellung]

Dabei werden zunächst mögliche Zentren der abgebildeten Marker gesucht, indem dunkle, entsättigte Pixel auf benachbarte hellere, entsättigte Regionen (vgl. Diagonalen des Kreuzmusters) geprüft werden (Diagonalenprüfung). Dabei müssen ausgehend vom möglichen zentralen Pixel alle benachbarten diagonalen Gradienten die gleiche Länge aufweisen, also der Abstand zwischen möglichem zentralen Pixeln und den Bildkoordinaten bei denen der Helligkeitszuwachs (Gradient) groß genug war, um als helle Region zu gelten.

Ist die Gradientenprüfung erfolgreich gewesen, müssen nachfolgend auch die Horizontalen und Vertikalen geprüft werden (Achsparallelenprüfung). Ausgehend vom ermittelten möglichen Zentrum, müssen sie ähnliche Pixelhelligkeiten und Sättigung aufweisen wie das ermittelte mögliche zentrale Pixel. Dafür werden nur die Pixel geprüft, die um die einheitliche Länge der Diagonalenprüfung vom Zentrum entfernt sind (siehe Abbildung 12).

Ist die Diagonalen- und Achsparallelenprüfung erfolgreich erhält man in der Regel mehrere mögliche Zentren innerhalb eines im Bildfeld abgebildeten Markers (vergleichbar mit den ersten beiden der drei abgebildeten Marker in der Abbildung 13).

Diese möglichen Zentren werden nun zusammengefasst, wenn sie direkt benachbart sind. Eine direkte Nachbarschaft ist definiert durch alle Pixel welche die Kanten oder Ecken eines zentralen Pixels berühren. Es werden folglich alle möglichen Zentren als zusammengehörig markiert, die sich innerhalb dieser direkten Nachbarschaftsbeziehung zueinander befinden (siehe Abbildung 13).

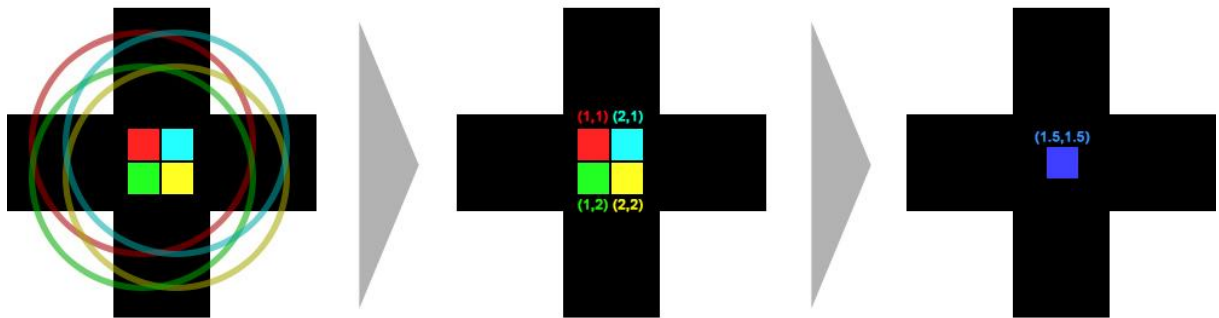


Abbildung 13: Bestimmung von möglichen Zentren (Rot, Türkis, Grün, Gelb) und anschließende Mittelung zum echten Zentrum mit Subpixelkoordinaten (Blau) [eigene Darstellung]

Für die Zusammenfassung und damit Eliminierung aller markierten zusammengehörigen Zentren wird die Position des nun ermittelten echten Zentrums über das arithmetische Mittel subpixelgenau berechnet. Mit zunehmender Anzahl der zur Mittelung herangezogenen zusammengehörigen Zentren steigt die mögliche Subpixelgenauigkeit nach folgender Formel an.

$$[\text{Subpixelgenauigkeit}] = \frac{1}{[\text{Anzahl d. Pixel die zur Mittelung herangezogen werden}]}$$

So weisen beispielsweise zwei horizontal nebeneinander liegende Pixel (X-Bildkoordinate) eine Subpixelgenauigkeit von $\frac{1}{2}$ Pixel auf. Sind es vier horizontal nebeneinander liegende Pixel, beträgt die Subpixelgenauigkeit bereits $\frac{1}{4}$ Pixel. Analog gilt dies für die vertikale Bildachse.

In der Abbildung 13 ist die Mittelung des echten Zentrums aus mehreren ermittelten möglichen Zentren und ihren einheitlichen Gradientenlängen dargestellt und verdeutlicht das Prinzip der Subpixelgenauigkeit.

4.2.2 Feature-based Pattern Finding Algorithmus (FPF)

Ähnlich dem GPF nutzt das FPF die Helligkeitsverläufe des kreuzförmigen, normierten Markers, durch den Einsatz von FAST Features (Quelle [18] und [19]). FAST ist ein besonders schnelles und effektives Verfahren, um Kanten, bzw. Ecken von abgebildeten Objekten im Bildfeld zu finden.

Der Algorithmus vergleicht die Intensität (Helligkeitswert) eines zu untersuchenden, zentralen Pixels und dessen kreisförmige Umgebung bestehend aus 16 Pixeln (siehe Abbildung 14: Prinzip von FAST Features) mithilfe eines einfachen Vergleichs der Intensitäten (siehe Formel).

$$S_{p \rightarrow x} = \begin{cases} d, & I_x \leq I_p - t & \text{(darker)} \\ s, & I_p - t < I_x < I_p + t & \text{(similar)} \\ b, & I_p + t \leq I_x & \text{(brighter)} \end{cases}$$

Formel 1: Entscheidungsbaum FAST

Die Intensität I_x jedes Pixels innerhalb des umgebenden Kreis wird mit der Intensität I_p des zentralen Pixels verglichen und einem Entscheidungsbaum (engl. *Decision Tree*) übergeben, welcher beschreibt, ob das Pixel eine ähnliche s (*Similar*), geringere d (*Darker*), oder höhere b (*Brighter*) Intensität als das zentrale Pixel I_p aufweist. Für den Entscheidungsbaum wird außerdem ein Schwellwert t (*Threshold*) verwendet, welcher den Einfluss von Bildrauschen reduzieren soll und so eine robustere Zuordnung der Zustände d, s, b erlaubt. Das Ergebnis des FAST Algorithmus sind drei Zähler für jedes zentrale Bildpixel I_p , welche die Anzahl der dunkleren, gleichen und helleren Umgebungspixel I_x enthalten.

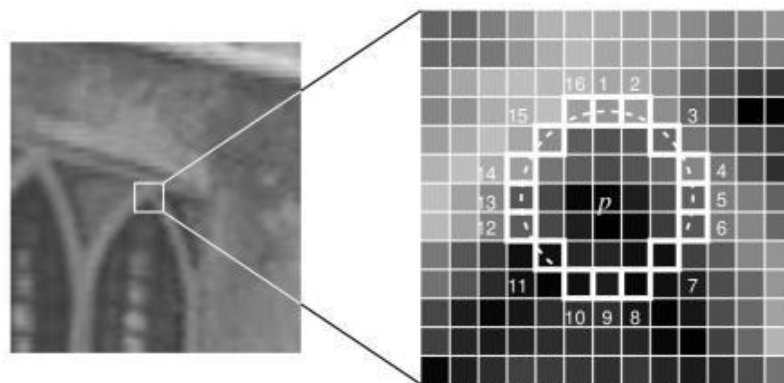


Abbildung 14: Prinzip von FAST Features (Quelle: [20] und [19])

Anhand der Zähler kann eine Aussage darüber getroffen werden ob es sich um eine Kante, bzw. Ecke handelt. Eine gerade Kante hat typischerweise acht Pixel der gleichen Klassifikation (50% der 16-Pixel-Umgebung) und eine Ecke zwölf Pixel (75% der 16-Pixel-Umgebung).

Im Fall der eingesetzten normierten Marker kann durch die Klassifikation und ihre Zähler bestimmt werden, ob es sich um eine außenliegende (hellere Pixel überwiegen) oder eine innenliegende Ecke (dunklere Pixel überwiegen) handelt, wie in Abbildung 15 verdeutlicht.

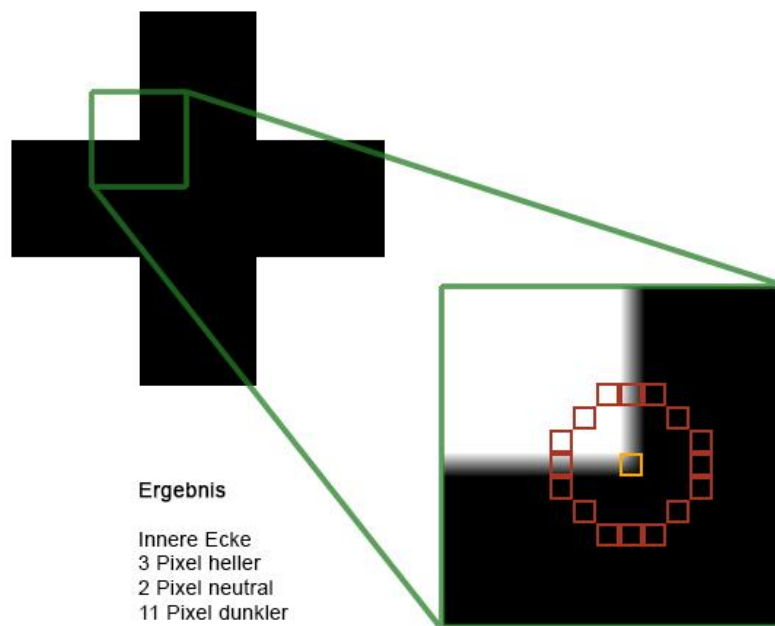


Abbildung 15: Beispielhafte Anwendung von FAST auf den normierten Marker während der Laufzeit
 [eigene Darstellung]

Die nun geschaffene Basis aus FAST Features und die dadurch ermöglichte eindeutige Bestimmbarkeit von außen- und innenliegenden Ecken kann nun dazu genutzt werden, die normierten Marker anhand der charakteristischen Anordnung ihrer Ecken im Bildfeld zu finden.

Der eingesetzte normierte Marker ist spiegelsymmetrisch entlang beider Bildachsen ausgehend von seinem Zentrum (s. Abbildung 16). Das bedeutet folglich, dass die inneren Ecken i (Orange) des Markers im gleichen Verhältnis zueinander stehen, wie die Ecken eines gleichseitigen Vierecks. Der inneren Ecke i_1 liegen folglich zwei innere Ecken i_2 und i_3 am nächsten,

welche den gleichen Abstand zur ursprünglichen inneren Ecke i_1 besitzen (Rot).

$$d_{12} = d_{13}$$

Die ursprüngliche innere Ecke i_1 besitzt darüber hinaus den gleichen Abstand zur vierten inneren Ecke i_4 , wie die beiden inneren Ecken i_2 und i_3 zueinander (Gelb).

$$d_{14} = d_{23}$$

Nachdem die Verhältnisse der inneren Ecken verifiziert wurden kann ein Mittelpunkt aus der Position der vier inneren Ecken bestimmt werden. Der so erzeugte Mittelpunkt ist bereits subpixelgenau mit einer Subpixelgenauigkeit von $\frac{1}{4}$ Pixel (vgl. S. 45 Abbildung 13). Die ermittelte Subpixelgenauigkeit wird nun durch das Einbeziehen der zum normierten Marker zugehörigen äußeren Ecken erhöht. Das Auffinden und Verifizieren der äußeren Ecken geschieht mit ähnlichen Methoden, wie sie schon beim Bestimmen der inneren Ecken Anwendung fanden. Die zur inneren Ecke i_1 gehörigen äußeren Ecken a_1 und a_8 weisen den gleichen Abstand $d_{i_1 a_1}$ und $d_{i_1 a_8}$ zu i_1 auf und deren diagonalen Abstand $d_{a_1 a_8}$ ist gleich dem durch die inneren Ecken bestimmten Abstand d_{14} und d_{23} , sodass folgende Verhältnisse gelten.

$$d_{i_1 a_1} = d_{i_1 a_8} \text{ und } d_{a_1 a_8} = d_{14} = d_{23}$$

Das Vorgehen ist analog für die Bestimmung der anderen äußeren Ecken.

Für die Erhöhung der Genauigkeit werden weitere Mittelpunkte aus den äußeren Punkten a_1, a_2, a_5, a_6 und a_3, a_4, a_7, a_8 erzeugt, welche nah am vorher aus den inneren Ecken erzeugten Mittelpunkt liegen. Die so erhaltenen drei Mittelpunkte haben die gleiche Subpixelgenauigkeit, da sie jeweils aus vier Eckpunkten erzeugt wurden. Die Bildung des einfachen arithmetischen Mittels aus den drei Mittelpunkten mit einer Subpixelgenauigkeit von je $\frac{1}{4}$ ergibt einen Mittelpunkt mit einer theoretischen Gesamt-Subpixelgenauigkeit von $\frac{1}{12}$ Pixel.

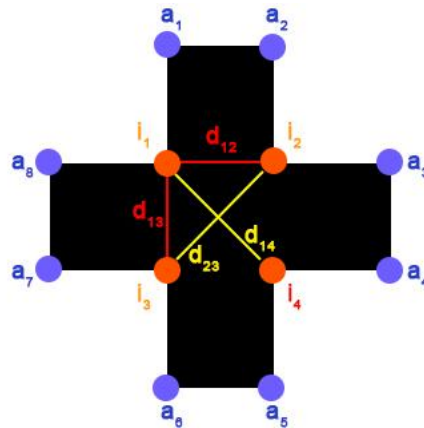


Abbildung 16: Logische Identifizierung von Markern mithilfe von FAST Features (FPF Algorithmus)
[eigene Darstellung]

Die vergleichsweise höhere Subpixelgenauigkeit des *Feature-based Pattern Finding* gegenüber dem *Gradient-based Pattern Finding*, resultiert theoretisch in einer genaueren Abstandsbestimmung der Prüfmaste, was im Kapitel 5.6 "Vorstellung der Testergebnisse" nachgewiesen wird.

4.2.3 Gegenüberstellung der theoretischen Vor- und Nachteile beider Verfahren

Die theoretische, höhere Subpixelgenauigkeit des FPF gegenüber dem GPF ist ein großer Vorteil, der bei der Verwendung des FPF für akkuratere Ergebnisse genutzt werden kann. Jedoch werden die Vorteile theoretisch verfahrensbedingt durch die gleichzeitige Einführung einiger Nachteile begleitet. Dadurch sollten in jedem Fall Vorüberlegungen getroffen werden welches der beiden Verfahren für bestimmte Situationen das geeignetere ist.

Dieser Abschnitt beschäftigt sich folglich mit der Gegenüberstellung der theoretischen Vor- und Nachteile und soll die Grenzen beider Verfahren aufzeigen. Dafür wurden hauptsächlich drei Kriterien betrachtet, die im konkreten Anwendungsfall direkt auf die Verfahren Auswirkungen haben und nicht-trivial zu beseitigen sind. Damit sind sie von großer Bedeutung für die Robustheit der Verfahren. Die zwei beeinflussenden Parameter sind die Auflösung des Gesamtsystems (Auflösung d. Bildsensor, Brennweite, Güte der Optik, Unschärfen)

und das Rauschen. Beide Parameter stehen im direkten Zusammenhang mit der Kamera-Hardware und bestimmen die Qualität des aufgenommenen Bildmaterials.

Ein wesentlicher Unterschied beider Verfahren ist das Prinzip mit dem der normierte Marker identifiziert wird. Während beim GPF eine relativ grobe Analyse des Helligkeitsverlauf ausgehend von einem angenommenen Zentrum erfolgt, wird beim FPF nach (scharf abgebildeten) Ecken mithilfe von FAST gesucht. Die grobe Analyse des GPF ist, durch das "Abtasten" der Helligkeitsverläufe (Gradienten) vorgegebener Richtungen in verschiedenen Pixelabständen ausgehend von einem Zentrum, unempfindlicher gegenüber geringer Auflösung als das FPF. Beim FPF wird durch FAST gezielt ein enger Bereich um ein angenommenes Eck-Pixel abgetastet. Die Größe des abgetasteten Bereichs wird dabei nie verändert, sodass eine statische Abtastung für variierend groß abgebildete Marker erfolgt. Um eine Ecke der Marker erfolgreich identifizieren zu können muss also der Abstand von einem Eck-Pixel zum nächsten mindestens drei Pixel betragen, da das nächste Eck-Pixel sonst innerhalb des abgetasteten Radius von FAST liegt und die eigentliche Ecke nicht als solche detektiert werden kann. Die Gesamtbreite der Abbildung des Markers darf bei FPF somit 9 Pixel keinesfalls unterschreiten. Darüber hinaus bestimmt FAST verfahrensbedingt bei unscharfen Abbildungen vermehrt ungenaue Klassifikationen der umgebenden Pixel. So kann es bei unscharfen Abbildungen vorkommen, dass viele Pixel häufig als "neutral" klassifiziert werden, was die Erkennung von Eck-Pixeln erschwert.

Das GPF hingegen kann auch in sehr kleinen Pixelabständen den Marker vollständig erkennen, wenn die durch begrenzte Kameraauflösung und Unschärfe resultierende geringere Auflösung der abgebildeten Marker ausreichend Gradienten aufweist. Theoretisch kann beim GPF ein Gradient in einem Pixelabstand von nur 1,5 Pixel erkannt werden, was mit der Beachtung aller Richtungen in einer minimalen Gesamtbreite des abgebildeten Marker von 3 Pixeln resultiert. Zusammengefasst muss die gesamte Auflösung des abgebildeten Marker beim GPF lediglich so hoch sein, dass der abgebildete Marker differenziert genug abgebildet ist um die hellen Diagonalen und dunklen vertikalen, sowie horizontalen Gradienten zu identifizieren.

Zusätzlich zu den auflösungsbegrenzenden Faktoren spielt das Bildrauschen eine entscheidende Rolle für die Robustheit der beiden Verfahren. Mit zunehmendem Bildrauschen zeigt sich theoretisch erneut das FPF als das sensiblere der beiden

Verfahren. Höheres Bildrauschen wirkt beim FPF auf den bereits beschriebenen statischen Abtastungsradius. Die Klassifizierung von dunkleren, helleren und neutralen Pixeln kann durch Rauschen stark beeinflusst werden und somit auch wieder die korrekte Detektion einer Ecke verhindern. Das GPF hingegen mittelt immer zwei benachbarte Pixel in Richtung des Gradienten zur Bestimmung der Helligkeitsverläufe ausgehend vom angenommenen Zentrum des abgebildeten Markers. Durch die Mittelung der beiden Pixel wird ein Mittelwert gebildet und das Rauschen so ebenso gemittelt. Dadurch wird der Einfluss des Rauschen reduziert und fällt beim GPF gegenüber dem FPF verfahrensbedingt geringer aus.

Anschließend zu den Nachteilen können nun die in den Abschnitten 4.2.1 und 4.2.2 genannten Vorteile beider Verfahren ergänzt und eine Übersicht über die Eignung beider Verfahren für bestimmte Ausgangssituationen aufgestellt werden. Die Übersicht nutzt dabei die drei Abstufungen "gering", "mittel" und "hoch" zur Einschätzung der theoretischen Vor- und Nachteile beider Verfahren.

	GPF	FPF
Genauigkeit	– mittel	+ hoch
Benötigte Auflösung	+ gering	– hoch
Rauschempfindlichkeit	+ mittel	– hoch
Berechnungsaufwand	+ gering	– mittel

Tabelle 2: Übersicht der Vor- und Nachteile von GPF und FPF

Zusammenfassend ist das FPF also dann zu bevorzugen, wenn das Kamerasystem nur wenigen Störeinflüssen unterliegt und über eine hohe Auflösung verfügt, da es die genaueren Ergebnisse liefert. Jedoch geht die höhere Genauigkeit in Gegenüberstellung zum GPF mit einer ebenso höheren Berechnungszeit einher. Das GPF sollte somit dann eingesetzt werden, wenn die Genauigkeit der Ergebnisse nicht sehr hoch sein muss und die Ausgangsbedingungen durch hohes Bildrauschen und eine geringere Auflösung schlechter sind.

4.3 Zeitliche Synchronisation von Videodaten

Beim angestrebten Einsatz von zwei Kameras, zur zeitgleichen Erfassung der Prüfmaske und den Signalen des Fahrzeugs, ist es nötig beide voneinander unabhängig aufgenommene Videodaten zeitlich zu synchronisieren, sodass die Signale des Fahrzeuges zeitlich mit den ermittelten Abständen zur Prüfmaske übereinstimmen. Dadurch soll eine korrekte Auswertung der Daten sichergestellt werden.

4.3.1 Visuell

Bei der visuellen Synchronisation werden Bewegungen von Objekten innerhalb der Szene mithilfe von *Motion Tracking*³⁴ in den Aufnahmen beider Kameras verfolgt.

Beim *Motion Tracking* handelt es sich um ein Verfahren der Computer Vision, bei dem Bewegungen innerhalb von Videoaufnahmen mithilfe von Features verfolgt werden. Es werden so für jedes Frame markante Features bestimmt und mit dem nachfolgenden Frame verglichen (*Feature Matching*). Stimmen Features in diesem Vergleich überein wird deren Änderung der Position (Translation) durch einen Vektor repräsentiert und gespeichert. Das robuste Wiederfinden von vorher gefundenen Features stellt dabei eine der größten Herausforderungen des Motion Trackings dar. Wurden für alle aufgenommenen Frames der beiden Kameras entsprechende *Motion-Tracking*-Daten gefunden, werden nun charakteristische Bewegungen im zeitlichen Verlauf markiert und abschließend die Gesamtheit der Markierungen einer Kamera so lange in negative und positive Richtung der Zeitachse verschoben, bis die Anzahl übereinstimmender Markierungen maximal ist.

Besonders gute Referenzereignisse sind plötzliche Richtungswechsel bzw. anfangende und endende Bewegungen. Durch die so erhaltenen Daten, über die abgebildeten Bewegungen gleicher Objekte innerhalb der Szene, kann der zeitliche Versatz (vgl. *Offset*) bestimmt werden. Der Abgleich der so über die Zeit beobachteten Bewegungskurven kann mit der Abbildung 19 der Audio-

³⁴ Bewegungsverfolgung in einer Videoaufnahme von markanten Bildteilen, bzw. Objekten mithilfe von Tracking-Algorithmen auf Basis von z.B. Features

Synchronisation verglichen werden, da sich die Bestimmung des Offset beider Verfahren ähnelt.

Die zeitliche Auflösung $t_{\text{Auflösung}}$ (vgl. Genauigkeit) des *Offset* ist umgekehrt proportional zu der Bildwiederholrate f_{Video} (vgl. engl. *Frames Per Second*, FPS).

$$t_{\text{Auflösung}} = \frac{1}{f_{\text{Video}}}$$

Entspricht die Bildwiederholrate der Aufnahmen beider Kameras beispielsweise 30 FPS, beträgt die maximal erreichbare zeitliche Auflösung $\frac{1}{30}$ Sekunden.

4.3.2 Audio

Im Gegensatz zur visuellen Synchronisation von Videodaten (s. 4.3.1) nutzt die im folgenden konzipierte akustische Synchronisation (vgl. Audio-Synchronisation) eine mit den Videoaufnahmen zusammen aufgenommene Tonspur. Die Tonspuren enthalten theoretisch sich ähnelnde akustische Abschnitte, die gegeneinander um das Offset t_{offset} zeitlich verschoben in jeder der beiden Tonaufnahmen aufgetreten sind. Ein zur Synchronisation geeigneter akustischer Abschnitt ist ein mehrere Millisekunden langes Audio-Sample, welches eindeutig zu beschreiben ist und über die gesamte zeitliche Länge der Tonaufnahme vorzugsweise einmalig auftritt. Ein eindeutig zu beschreibendes Audio-Sample hat wiederum einen charakteristischen Amplitudenverlauf, was beispielhaft in der Abbildung 17 dargestellt wird.

Die in der Abbildung dargestellten zwei digitalen Tonspuren sind eine zeitdiskrete, komplexe Zusammensetzung einer bestimmten Anzahl von einzelnen Auslenkungen (Amplitude) pro Zeiteinheit (Abtastrate; engl. *Sampling Rate*, vgl. *Samples Per Second*). Das Auftreten von Amplitudenänderungen über die Zeit stellt die akustische Frequenz dar. Im Beispiel wurde eine Abtastrate f_{abtast} von 44100 Hz verwendet, was 44100 Amplitudenänderungen pro Sekunde entspricht. Diese Abtastrate ist ein gängiger Wert für Tonaufnahmen, was sich physiologisch durch das menschliche Gehör und dem Nyquist-Shannon-Abtasttheorem begründet.

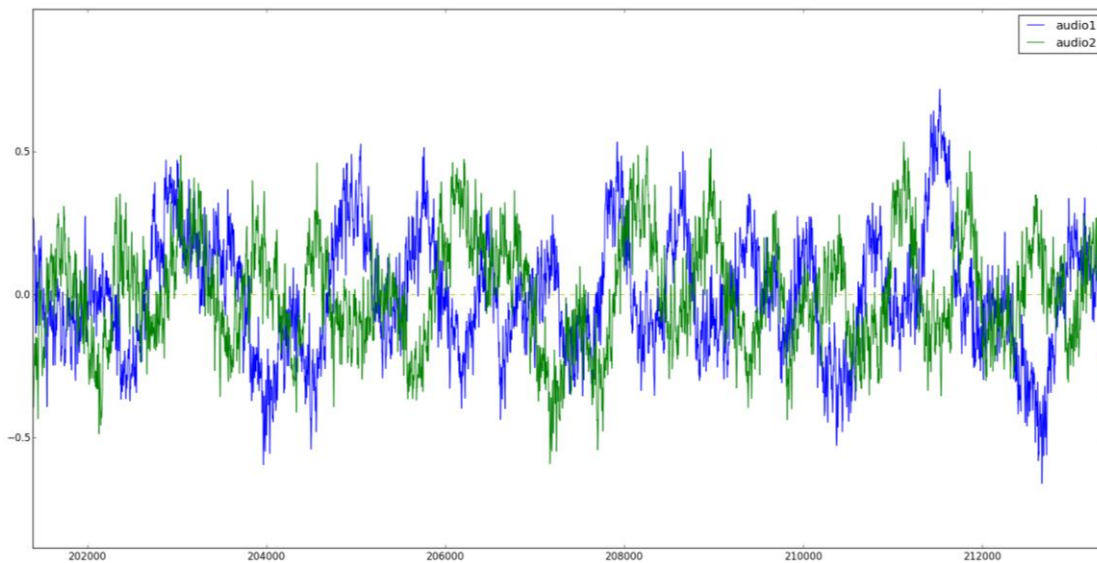


Abbildung 17: Typische Audiospur mit gleichem, jedoch zeitlich verschobenem, Frequenz- und Amplitudengang (Offset; grüne Tonspur zeitlich später) [eigene Darstellung]

Das Abtasttheorem nach Nyquist und Shannon besagt, dass die Abtastrate für eine korrekte Rekonstruktion aus dem zeitdiskreten, digitalen Signal mindestens doppelt so groß sein muss wie die maximal akustisch wiederzugebende, obere Grenzfrequenz f_{max} . Dadurch ergibt sich eine obere Grenzfrequenz von $f_{max} = 22050 \text{ Hz}$ bei der doppelten Abtastrate $f_{abtast} = 44100 \text{ Hz}$.

$$f_{max} = 22050 \text{ Hz}$$

$$f_{abtast} \geq 2 \times f_{max}$$

$$f_{abtast} \geq 44100 \text{ Hz}$$

Formel 2: Nyquist-Shannon-Abtasttheorem

Die Abtastrate von 44100 Hz und die obere, korrekt rekonstruierbare, Grenzfrequenz von 22050 Hz erfüllt hinreichend den von Menschen wahrnehmbaren Frequenzbereich von 20 – 20000 Hz. In der Regel ist die Wahrnehmbarkeit von hohen Frequenzen im Ultraschallbereich ($> 16 \text{ kHz}$) vom Alter des Hörers abhängig. Laut einiger Untersuchungen [21] können die meisten Menschen ab dem 30. Lebensjahr bereits Frequenzen oberhalb von 17 kHz nur noch eingeschränkt wahrnehmen, bereits Frequenzen oberhalb von 8 kHz erfahren ab dem 30. Lebensjahr eine Beeinflussung. Aus diesen Gesichtspunkten wäre eine Abtastrate von 40000 Hz für das Gehör des Menschen mehr als ausreichend und würde je nach Zustand des Hörorgans und Alter des Hörers sogar Reserven bieten. Die um

4100 Hz höhere Abtastrate gegenüber diesem physiologisch sinnvollem Wert begründet sich durch nicht-ideale technische Bauteile. Ein Audiogerät benötigt für die Signalwiedergabe zur Minimierung von akustischen Aliasing-Effekten einen Tiefpassfilter, der jedoch mit realen Bauteilen nicht steil, bzw. abrupt (ideal), bei 20kHz trennen kann. Ein Tiefpass weist immer eine sogenannte Flankensteilheit auf. Durch den Puffer von 4100 Hz kann sichergestellt werden, dass auch mit einem nicht sehr steil trennenden (und kostengünstigen) Tiefpass eine ungedämpfte Wiedergabe bei 20 kHz möglich ist. So legten sich die Unternehmen Sony und Philips im Jahr 1979 auf diese nur scheinbar willkürlichen 44100 Hz fest.

Die aufgenommenen, komplexen Tonspuren können so relativ trivial auf ihre Auslenkungen (Amplituden) und Frequenzgänge hin verglichen und das Offset ermittelt werden. Geht man jedoch davon aus, dass die Tonspuren mitunter mehrere Sekunden, in beide Richtungen der Zeitachse versetzt sein können und es 44100 (Amplituden-)Werte pro Sekunde gibt, die mit jedem Amplitudenwert der anderen Tonspur verglichen werden, dann wird deutlich, dass sich die Laufzeit einer solchen Vergleichsrechnung exponentiell verhalten wird. Bereits am Beispiel von einer maximal tolerierten Zeitverschiebung Δt_{max} von einer Sekunde erhält man so eine hohe Anzahl von durchzuführenden Vergleichsrechnungen $n_{vergleich\ he} \cong 2 \text{ Milliarden}$.

$$n_{vergleich\ he} = 2 \times (\Delta t_{max} \times f_{abtast})^2$$

$$n_{vergleich\ he} = 2 \times (1s \times 44100Hz)^2 = 1.944.810.000$$

Zwei Milliarden Vergleichsoperationen sind für moderne Rechner mit geringer Laufzeit lösbar, geht man jedoch davon aus, dass die Abtastrate der Tonaufnahme höher sein kann, typischerweise bis zu 96000 Hz, und die Zeitverschiebung Δt_{max} größer als eine Sekunde ist, dann erhöht sich dieser Wert exponentiell. Es ist daher nötig den vorliegenden komplexen Amplitudenverlauf der beiden Tonaufnahmen zu vereinfachen und somit leichter beschreib- und vergleichbar zu machen.

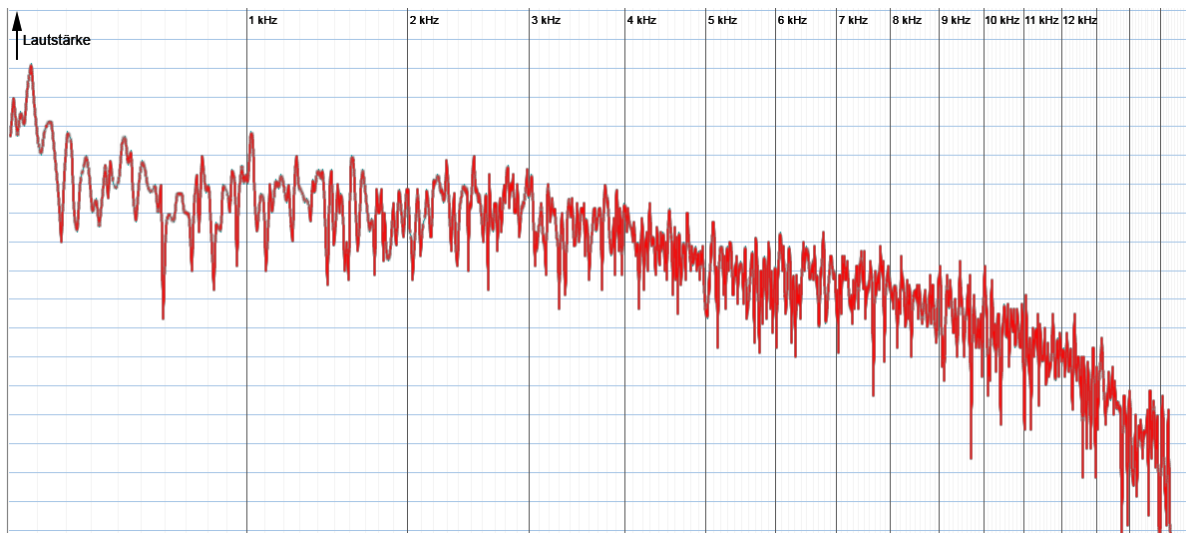


Abbildung 18: Frequenzanalyse der Tonaufnahme eines Prüfstandlaufes des VW Passat CC mit Dieselmotor, des Laborbereichs Fahrzeugtechnik der HTW Dresden, auf dem Prüfstand MFP 3000 von Maha (Darstellung wurde mit der Software WavePad Audio-Editor erzeugt) [eigene Darstellung]

Für eine einfachere Beschreib- und Vergleichbarkeit werden auffallende und leicht wiederzuerkennende Eigenschaften der Amplitudenverläufe in Form von *Features* (vgl. visuelle *Features*; *FAST Features* Kapitel 4.2.2) bestimmt und die Anzahl der Amplitudenwerte (Samples) deutlich reduziert. Als geeignete Eigenschaften für das Erzeugen von *Features* sind charakteristische Auslenkungen des Amplitudenverlaufs, wie z.B. nach einer Änderung des Anstiegs (Wechsel von steigenden zu fallenden Werten, und umgekehrt), oder einem Durchgang des Amplitudenverlaufs durch die Nulllinie (Nulldurchgang; vgl. Wechsel von positiven Amplitudenwerten zu negativen, und umgekehrt). Soll jedoch ein Feature nach jedem Nulldurchgang bestimmt werden ist es sinnvoll die in zeitlich kurzen Abständen auftretenden hohen Frequenzen herauszufiltern, da hohe Frequenzen in kurzen zeitlichen Abständen viele Nulldurchgänge auslösen können und so die Anzahl von Samples (Anzahl *Samples* entsprechen Anzahl *Features*) nicht wesentlich reduzieren. Für eine Filterung der hohen Frequenzen spricht darüber hinaus auch das Einsatzszenario auf einem Rollenprüfstand auf dem ein Fahrzeug (typischerweise mit Verbrennungsmotor) betrieben, und eine reale Fahrt simuliert wird. Die bei einer realen Fahrt auftretenden Geräusche sind vor allem durch den Motor des Fahrzeugs und den Abrollgeräuschen der Reifen bestimmt. Anhand der vorhergehenden Abbildung 18 ist zu erkennen, dass die akustische Frequenz während der simulierten Fahrt eines PKW mit Verbrennungsmotor auf dem Prüfstand nur geringfügig den oberen Frequenzbereich des menschlichen Gehörs und der

Tonaufnahmen ausnutzt. Somit ist eine Fokussierung auf den Tief- und Mitteltonbereich sinnvoll und eine Filterung des Hochtonbereichs für das gegebene Szenario geeignet.

Nach der Filterung hoher Frequenzen wird für die gesamte Tonaufnahme zwischen zwei Nulldurchgängen die größte bzw. kleinste Auslenkung, ein Extrempunkt, gesucht. Der Extrempunkt stellt ein *Feature* der Tonaufnahme dar und enthält in Nachbarschaft mit anderen Features die folgenden Informationen.

- minimale oder maximale Amplitude (Auslenkung)
- Zeit des Auftretens in der gesamten Tonaufnahme
- zeitlicher Abstand zu benachbarten Features

Zur Wiederfindung der *Features* in der jeweils anderen Tonaufnahme können nun diese Eigenschaften verglichen und die zeitliche Verschiebung "Offset" t_{offset} bestimmt werden. In der Abbildung 19 ist das Wiederfinden von Features von zwei sich ähnelnden, aber zeitlich verschobenen Tonspuren dargestellt. Die für die Bildung von Features genutzten Quell-Tonspuren der Abbildung 19 entsprechen den bereits in der Abbildung 17 gezeigten Tonspuren mit vollem Frequenzspektrum.

Für das Wiederfinden von gleichen Features, und den damit einhergehenden Vergleich der Eigenschaften der ermittelten Features, sollte jedoch mehr als nur ein Feature verwendet werden, da ein Feature allein nur wenige Eigenschaften besitzt und so mehrfach in einer mehrere Sekunden langen Tonspur auftreten kann. Dafür können die Beziehungen zwischen benachbarten Features, bzw. die Folge von Features mit so vorhersagbaren Eigenschaften, mit einbezogen werden. Umso mehr Features mit einbezogen werden, desto eindeutiger kann eine Folge von Features (*Feature-Sequenz*) in der anderen Tonspur wiedergefunden werden. Die rot markierte *Feature-Sequenz* zeigt eine Zusammensetzung aus jeweils drei *Features*, die violette Sequenz setzt sich aus sechs Features zusammen. Für beide Sequenzlängen n_{features} (Anzahl von Features pro Sequenz) ist das gesuchte Offset t_{offset} hinreichend bestimmbar.

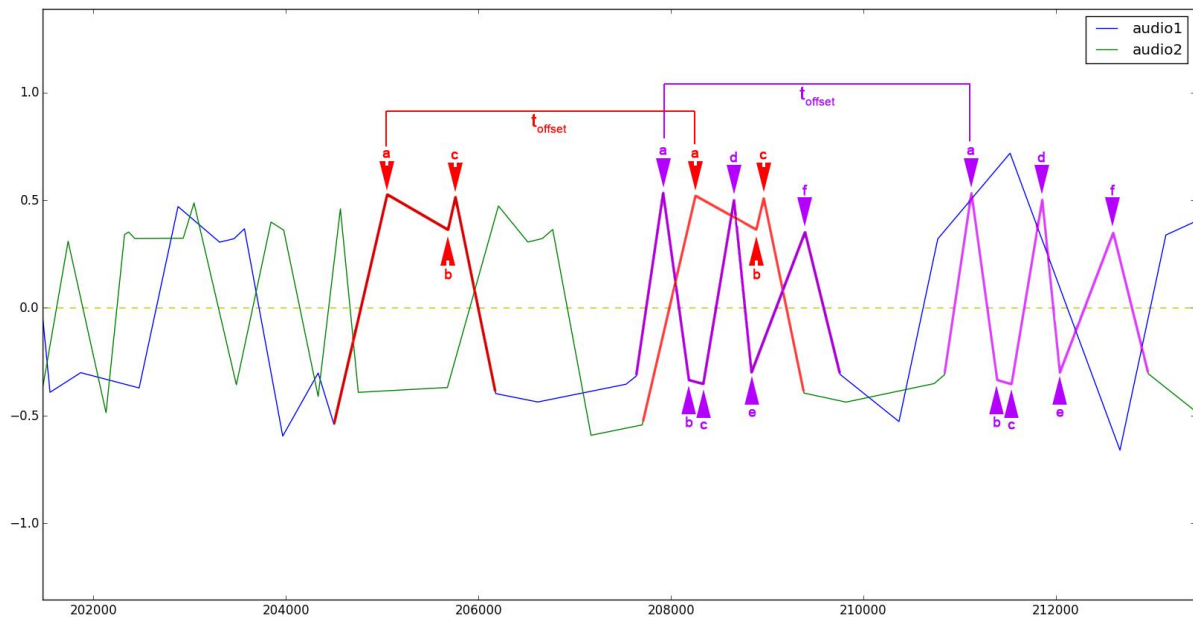


Abbildung 19: Audio Synchronisation - Bestimmung des Offset sich ähnelnder akustischer Abschnitte
[eigene Darstellung]

Bisher wurde von idealen Bedingungen ausgegangen und mögliche Störeinflüsse (Störgrößen), welche die Signale (Tonaufnahmen) verfälschen, nicht beachtet. Jedoch sind die elektronischen Komponenten, die für die Tonaufnahme verantwortlich sind, nicht ideal und unterliegen Störeinflüssen, die sich als akustisches Rauschen in den Aufnahmen niederschlagen (vgl. Rauschabstand³⁵). Hauptsächlich erzeugen nicht-ideale elektronische Bauteile bei der Signalverarbeitung entsprechende Störungen. Das analoge Eingangssignal von einem Mikrofon muss durch zwei wesentliche elektronische Komponenten fließen, durch die es eine Störung erfahren kann. Die erste Komponente ist ein Vorverstärker, bei der das schwache elektrische Signal vom Mikrofon verstärkt und damit dessen Leistung erhöht wird. In Zusammenhang mit der Akustik spricht man in Zusammenhang mit der Leistung eines Signals auch oft von der Lautstärke. Wurde die Lautstärke des Eingangssignals erhöht gelangt es zum Digital-Analog-Wandler (DA-Wandler), wo es im Takt einer vorgegebenen Abtastrate in ein zeitlich diskretes, digitales Signal umgewandelt wird.

Das Rauschen durch elektronische Bauteile kann den Amplitudenverlauf stark beeinflussen und so das Ermitteln des Offset t_{offset} erschweren oder mit steigendem

³⁵ Rauschabstand, auch Störabstand oder Signal-Rausch-Verhältnis, bezeichnet die technische Qualität eines Nutzsignals (hier: Audio), welches von einem Rauschsignal überlagert ist. Es ist definiert als das Verhältnis vom mittleren Nutzsignal zum mittleren Rauschsignal.

Rauschpegel³⁶ gar unmöglich machen. Das Rauschen beeinflusst in der Regel einen breiten unspezifischen Frequenzbereich und äußert sich in unregelmäßigen, nicht vorhersagbaren Schwankungen der Amplitude des Audiosignals. Dadurch muss darauf geachtet werden möglichst hochwertige und rauscharme elektronische Bauteile für die Tonaufnahmen zu verwenden. Das Auftreten von Rauschen unterschützt weitergehend das Glätten des Signals durch das Filtern von hohen Frequenzen.

Eine andere Störgröße ist die Raumakustik, die ihren Ursprung in der Reflexion von Schallwellen hat (s. 2.2.3). In dem vorliegenden Szenario werden zwei getrennte Kameras mit je einer eigenen Verarbeitungs-Pipeline³⁷ für das Audio-Eingangssignal in Form von elektronischen Bauteilen verwendet. Jede der beiden eingesetzten Kameras besitzt neben dieser Signal-Pipeline ein integriertes Mikrofon, welches die akustischen Signale (Schalldruck) in elektrische Energie umwandelt (Eingangssignal). Durch die Bauform der Gehäuse, und vor allem der unterschiedlichen räumlichen Montage der Kameras im Testszenario, nehmen die integrierten Mikrofone unterschiedliche Schalldruckpegel über die Zeit auf. Die größten zu nennenden Einflüsse sind Laufzeitunterschiede aufgrund des räumlichen Weges zwischen beiden Kameras und der Ausbreitungsgeschwindigkeit von Schallwellen, die unterschiedlich starke Dämpfung verschiedener Frequenzen und Schallreflektionen. Die Einflüsse durch diese Größen sind in der Audiotechnik auch als Raummoden bekannt. Jeder Raum, in dem sich Schallwellen ausbreiten, besitzt bestimmte Eigenschaften, welche die Schallwellen verschiedener Frequenzen unterschiedlich stark reflektieren oder absorbieren. Es kommt zu Schallüberlagerungen von zurückgeworfenen Schallwellen, sodass gewisse Frequenzen ortsabhängig verstärkt oder ausgelöscht werden können. Vor allem eine Garage, wie sie im gegebenen Szenario vorzufinden ist, reflektiert durch große kahle Betonflächen Schallwellen aller Frequenzen besonders stark. Dadurch kommt es zu ungewollten Halleffekten, die kurz auftretende charakteristische Geräusche im Hall verschwimmen lassen können. So nimmt eine beispielsweise auf der Motorhaube montierte Kamera zwar gleiche charakteristische Frequenzänderungen wahr, wie eine Kamera, die beispielsweise in einem akustisch gedämpften Fahrzeuginnenraum

³⁶ Der Pegel bezeichnet in der Audiotechnik die Lautstärke (oft in dB angegeben).

³⁷ (Vgl. Verarbeitungs-Kette); Von einer Pipeline wird in der Informatik gesprochen, wenn Daten eine feste Abfolge von Einzelprozessen durchlaufen müssen, um ein bestimmtes Endergebnis zu erhalten.

montiert ist, jedoch wird der Amplitudenverlauf bestimmter Frequenzen abweichend aufgenommen (verzerrt). Das Wiederfinden von Features ist so nur bedingt möglich.

Eine Strategie um den Einfluss der Raummoden auf zwei räumlich getrennte Mikrofone aus dem Weg zu gehen, ist der Einsatz eines einzelnen, gemeinsamen Mikrofons für beide Kameras. Im Fall der eingesetzten zwei Kameras mit eigener Mikrofon-Vorverstärker-Elektronik bedeutet dies, dass Umgebungsgeräusche durch eine einheitliche und im Raum definiert positionierte Aufnahmequelle erfasst werden, was bedeutet, dass an der Mikrofon-Vorverstärker-Elektronik das gleiche analoge Audiosignal anliegt und verarbeitet wird. Die einzige nun vorhandene Störquelle ist das eingebrachte Rauschsignal durch die signalverarbeitende Elektronik der jeweiligen Kamera, sodass sich die nach der Verarbeitung getrennt gespeicherten digitalen Audiosignale ähneln, jedoch nicht identisch sind. Das Problem der Bestimmung und Wiederfindung von Features in zwei durch Rauschsignale gestörte Audioaufnahmen besteht also weiterhin und muss bei der Entwicklung des Algorithmus zur Audio-Synchronisation beachtet werden.

Dem Algorithmus sind jedoch natürliche Grenzen gesetzt. Bei einem sehr geringen Nutzsignal und sehr hohem Rauschsignal (schlechtes Signal-Rausch-Verhältnis) können die Audiospuren nicht zuverlässig synchronisiert werden. Es ist daher durch den Anwender bereits bei der Auswahl der Kameras, bzw. deren elektronischen Bauteile, auf möglichst hochwertige Komponenten zu achten.

4.4 Abgrenzung

Auf Basis der Problemanalyse (s. 3.1) und der Evaluierung der vorgestellten Verfahren können nun die folgenden Abgrenzungen formuliert werden:

Im Rahmen der Implementierung vorgestellter Ansätze soll ein Prototyp entstehen, der auf einem x86-Windows-Betriebssystem mit entsprechend vorhandenen Laufzeitkomponenten und Treibern lauffähig ist. Der Prototyp wird mit einer anschaulichen grafischen Benutzeroberfläche (GUI) ausgestattet, sodass Eingaben in das System durch klassische Eingabemethoden (Maus und Tastatur) ermöglicht werden. Die GUI soll auf mehreren Dialogen basieren, welche den Nutzer die Bedienung verglichen mit einem Konsolenprogramm vereinfachen und ausreichend Hilfestellung und Hinweise in geeigneter Form vermitteln.

Die gesamte Messung erfolgt ausschließlich auf einem dafür vorgesehenen Prüfstand (MFP 3000) in Verbindung mit einem Schienensystem auf dem eine präparierte Prüfmaske bewegt und so eine Situation simuliert werden kann, in der ein Notbremseingriff erforderlich ist.

Betrachtet wird ausschließlich die Monovision, also der Einsatz einer Kamera, zur Distanzmessung. Desweiteren wird eine Kamera für die Beobachtung des Kombi-Instruments (Tachometer) benutzt, um Signale des Notbremsassistenten an den Fahrer aufzuzeichnen und später entsprechend auszuwerten.

Bedingt durch die in diesem Kapitel vorgestellten und auf den Anwendungsfall spezialisierten Ansätze wird das GPF und das FPF zur Distanzmessung, sowie die Audio-Synchronisation aufgrund ihrer höheren Genauigkeit gegenüber der visuellen Synchronisation, implementiert.

Für die Kalibrierung ist, im Rahmen des zu entwickelnden Prototypen, nur eine Referenzmessung mithilfe des Kamerabildes der distanzmessenden Kamera und eines Laser-Entfernungsmesser (Hersteller: Bosch) vorgesehen. Die Werte für diese Kalibrierung werden in eine entsprechende Eingabemaske manuell eingetragen und temporär, oder bei Bedarf dauerhaft, gespeichert. Die Rektifizierung des Bildmaterials (Ausgleich der Bildverzerrungen, s. 2.2.2 S. 7) wird in den Prototypen nicht implementiert, da die gegebene Verzerrung des kritischen Bildbereiches

unkritisch ist (ROI³⁸; sehr zentral, fern vom Bildrand) und die Genauigkeit der Ergebnisse zur Demonstration des Prinzips der konzipierten Verfahren ausreichend ist.

Die abschließend berechneten Ergebnisse sollen mithilfe einer geeigneten grafischen Bibliothek anschaulich dargestellt werden, sodass ein entsprechender Prüfer des Fahrzeugs die Ergebnisse eindeutig zuordnen kann. Der Prototyp soll keine automatische Beurteilung der Messergebnisse beinhalten, da eine eindeutige Beurteilung nur durch konkrete Vorgaben seitens des Herstellers erfolgen kann. Dem Notbremsassistenten kann daher eine korrekte Funktion zugesprochen werden, wenn das System nach mehrfacher Durchführung der Tests konstante Ergebnisse aufweist, oder der Prüfer über die nötige Kenntnis, bzw. Erfahrung, verfügt, um eine entsprechende Einschätzung anhand der ermittelten konkreten Messergebnisse abgeben zu können. Die Beurteilung der Messergebnisse unterliegt im Rahmen des Prototyps ausschließlich dem Prüfer des Fahrzeugs.

4.5 Annahmen

Abschließend sollen im Rahmen der Konzeption nun Annahmen formuliert werden auf deren Basis die Implementierung der gewählten Verfahren erfolgen soll.

Die erste Annahme bezieht sich auf den letzten Absatz der Abgrenzung. Demnach sollte ein Prüfer qualifizierte Aussage über korrekte Funktionalität des Notbremsassistenten, anhand der Distanzmessung und aufgenommenen (Warn-) Signalen des Fahrzeugs an den Fahrer, treffen können.

Darüber hinaus ist davon auszugehen, dass die Fahrzeugmaske vom Fahrzeug stets als ein reales, vorausfahrendes Fahrzeug erkannt wird. Wenn dies nicht gegeben ist, muss eine entsprechende Anpassung der Prüfmaske durch qualifizierte Personen erfolgen.

Es wird angenommen, dass während aller Messfahrten auf dem Prüfstand gleichbleibende Lichtverhältnisse vorherrschen, sodass die Umweltbedingungen für

³⁸ ROI: "Region of Interest" (Beschränkung der Berechnungen auf einen kleineren/interessanten Bildbereich des gegebenen Bildes, da nur dort relevante Ereignisse auftreten)

das optische System (eingesetzte Kameras) bei jeder Messung identisch sind. Daher wird von einer Umgebung mit kontrollierbaren Lichtbedingungen ausgegangen (wie z.B. eine Prüfhalle).

Die Audio-Video-Synchronisation beider eingesetzten Kameras, für die Innenraum- sowie Distanzmessung, wird als gleichbleibend angenommen. Das bedeutet, dass mit jeder Auslösung der Aufnahme eines Videos die Verschiebung von Audio und Video konstant ist. Außerdem wird angenommen, dass die erwähnte Verschiebung innerhalb der Serie eines Kameramodells identisch ist.

Zuletzt wird die Annahme getroffen, dass das Ausgangsmaterial (Videoaufnahmen) unwesentlich verzerrt ist, oder bei deutlicher Verzerrung bereits entsprechend entzerrt wurde. Es ist daher von der Benutzung von rektifiziertem, kalibriertem Ausgangsmaterial im Rahmen des Prototypen auszugehen.

5 Implementierung

In diesem Kapitel werden die Besonderheiten implementierter Algorithmen beschrieben und geeignete Testdatensätze vorgestellt. Abschließend werden die Testergebnisse der einzelnen implementierten Verfahren in Bezug zu ihrer Robustheit, bzw. Verwendbarkeit für fortführende Analysen, bewertet. Die ausgewählten Verfahren wurden entsprechend ihrer Beschreibung (s. 3.4), bzw. deren Konzeption (s. Kapitel 4) umgesetzt und können im Anhang der Arbeit gefunden werden.

5.1 Anordnung der Prüftechnik

Ein wichtiger Bestandteil der Implementierung ist die Aufnahme der Testdatensätze, die mithilfe der vorhandenen Prüftechnik angefertigt werden. Die nachfolgende Abbildung zeigt die schematische Anordnung aller relevanten Bestandteile während der Prüfung, wie sie auch für der Aufnahme der Testdatensätze gegeben war.

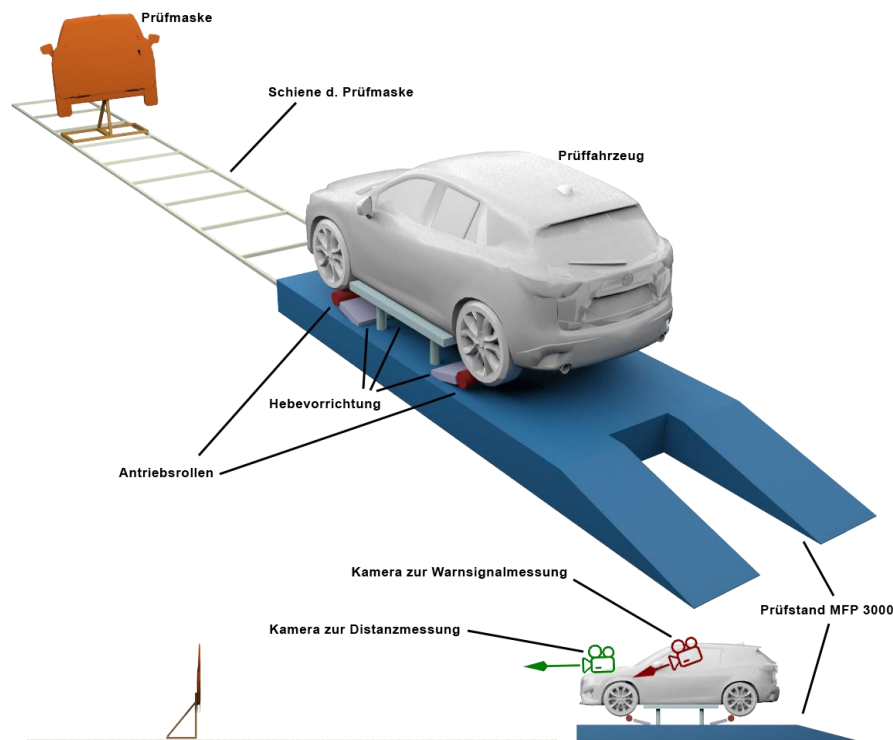


Abbildung 20: Anordnung der Prüftechnik [eigene Darstellung]

Der Begriff "Prüftechnik" ist die Gesamtheit aller am Prüfprozess beteiligten funktionalen Bauteile, bzw. Geräte, was folgende einschließt:

- Prüfstand MAHA MFP 3000 (Blau)
 - Hebevorrichtung (helles Blau, Lila)
 - Antriebsrollen (Rot)
 - Schiene der Prüfmaske (helles Grau)
 - bewegliche Prüfmaske (Braun)
- zu prüfendes Fahrzeug (Grau)
- Kamerasystem
 - Kamera zur Warnsignalmessung (Rot)
 - Kamera zur Distanzmessung (Grün)

Für die Aufnahme der Testdatensätze wurde so vorgegangen, wie es auch ein normaler Prüfbetrieb vorsehen würde. Zuerst wurde das Fahrzeug auf den Prüfstand gefahren und für das Anheben vorbereitet. Anschließend wurde das Fahrzeug mithilfe der zugehörigen Steuereinheit des Prüfstandes im Hebemodus³⁹ an den dafür vorgesehenen Punkten, welche am Unterboden zu finden sind, angehoben. Dadurch wurde das Chassis des Fahrzeugs fixiert und die Radaufhängung des Fahrzeugs entlastet, wobei die Räder selbst nur wenige Zentimeter (i.d.R. 5-15cm) über dem Prüfstand kontaktlos positioniert waren. Anschließend wurde an der Steuereinheit in den Prüfstandmodus gewechselt, wodurch die Kontrolle über die Antriebsrollen und deren eigene Hebevorrichtung ermöglicht wurde. Die Hebebühne kann in diesem Modus nicht bewegt werden. Die Antriebsrollen der Vorder- und Hinterachse wurden daraufhin mit entsprechender Funktion der Steuereinheit des Prüfstandes automatisch angehoben und gegen die Räder des Fahrzeugs gedrückt. Nachdem ein Gegendruck von 25kg pro Rad aufgebaut wurde, werden die Antriebsrollen für den Prüfbetrieb freigegeben. Mit Aktivierung der Achssynchronisation an der Steuereinheit kann das Fahrzeug auf dem Prüfstand regulär⁴⁰ betrieben werden, wodurch eine "normale" Fahrt simuliert und die Assistenzsysteme getestet werden können. Für den Test des Notbremsassistenten wurde zusätzlich die Fahrzeugmaske durch eine manuelle, externe Steuerung auf das auf dem Prüfstand "fahrende" Fahrzeug, mithilfe eines im Schlitten der

³⁹ Der Prüfstand besitzt einen Hebemodus in dem die Hebebühne bedient werden kann und einen Prüfstandmodus in dem die Antriebsrollen kontrolliert werden können.

⁴⁰ vgl. Fahrt/Betrieb des Fahrzeug auf der Straße

Fahrzeugmaske eingebauten Motors, zubewegt. Für die Prüfung eines Fahrzeugs auf dem Prüfstand ist lediglich zu beachten, dass die vom Hersteller definierten Begrenzungen des Prüfstandes, wie maximaler Lenkwinkel, Gewicht und Geschwindigkeit, nicht überschritten werden.

Nach erfolgter Aufnahme des Prüfungsszenarios durch die beiden Kameras wurden die gewonnen Aufnahmen auf einen Computer übertragen auf dem die prototypisch entwickelte Software zur Auswertung der Datensätze installiert war. Für die Auswertung wurde ein Notebook mit *Intel Core i7 4710MQ* Prozessor, unter der Nutzung eines Prozessorkerns mit 3,3 bis 3,5 GHz (schwankend) und 16 GB RAM mit 1600 MHz eingesetzt.

5.2 Testdatensätze

Die Testdatensätze wurden anschließend mit den zwei zur Verfügung stehenden Fahrzeugen aufgenommen. Das "erste Fahrzeug" (Fahrzeug A) ist ein VW Passat CC (Baujahr 2012) mit Radarsensor und das "zweite Fahrzeug" (Fahrzeug B) ein BMW i3 (Baujahr 2013) mit Kamerasystem zur Fahrzeug-eigenen Abstandsmessung. Die Testdatensätze umfassen selbst angefertigte Aufnahmen mit einer *GoPro Hero 4*, welche auf der Windschutzscheibe, und einer *GoPro Hero 3+*, welche am Himmel⁴¹ im Innenraum des Fahrzeug, montiert war.

Die Aufnahmen des VW Passat CC (Testdatensatz A) zeichnen sich durch den Einsatz des Radarsensor aus, für den die Fahrzeugmaske aus Schaumstoff, durch die Wellenlänge von Radarwellen, nicht sichtbar ist. Deshalb wurde an der Fahrzeugmaske ein Dreieck-Reflektor aus Metall angebracht, der für die verlustarme und direkte Reflektion von Radarwellen optimal geformt ist. Dadurch konnte die Fahrzeugmaske, obwohl für die elektromagnetischen Wellen des Radarsensor transparent, trotzdem vom Fahrzeug wahrgenommen werden.

Die Aufnahmen des BMW i3 (Testdatensatz B) erfolgten analog zu denen des VW Passat CC. Da es sich bei dem distanzmessenden Sensor dieses Fahrzeugs um ein Kamerasystem handelt, war lediglich die optische Erscheinung der Fahrzeugmaske

⁴¹ Dach des Fahrzeug im Innenraum

im sichtbaren Licht von Bedeutung. Die Fahrzeugmaske ist einem realen Fahrzeug nachempfunden, sodass auf Kameras basierte Sensorsysteme die Fahrzeugmaske im Idealfall als Fahrzeug erkennen und ein entsprechendes Verhalten des Fahrzeugs ausgelöst wird. Aufgrund von einer zu kleinen Testmenge von Prüfstandläufen fließen die Testdatensätze des BMW i3 jedoch nicht in die nachfolgenden Betrachtungen ein.



Abbildung 21: Test des ACC des BMW i3 mit Auswertung durch den entwickelten Prototypen [eigene Darstellung]

Feldversuche mit dem BMW i3 und einem entsprechend präparierten vorausfahrenden Fahrzeug haben jedoch gezeigt, dass das Fahrerassistenzsystem des i3 auf Basis von *MobileEye* (s. 3.3: Verwandte Projekte) bei entsprechenden Witterungsbedingungen gut funktioniert. Für diese Versuche wurde versucht das Fahrzeug mit dem fahrzeugeigenen ACC-System vollständig autonom bremsen und beschleunigen zu lassen, ohne Eingriff durch den Fahrer. Eine Videosequenz, welche diese Feldversuche zusammen mit der Auswertung der Messdaten durch den entwickelten Prototypen zeigt, liegt dieser Arbeit auf dem Datenträger bei (vgl. Abbildung 21).

5.3 Verwendete Bibliotheken

Für die Implementierung der gewählten Verfahren wurde hauptsächlich die Funktionsbibliothek *OpenCV*, in Verbindung mit der Programmiersprache C++ und *matplotlib* Funktionsbibliothek von *Python* genutzt, die mithilfe eines Python-Wrappers an das C++-Programm angebunden wurden. Zur Gestaltung und Darstellung der grafischen Oberflächen (Dialoge) kam Qt in der Version 5.5 zum Einsatz.

Der größte Teil der Implementierung stützt sich auf die Funktionsbibliothek *OpenCV*, welche unter Open-Source-Lizenz steht und sich im Wesentlichen auf die Computer Vision und die Bildverarbeitung konzentriert. Einer der Hauptaugenmerke von *OpenCV* ist es, dem Anwender eine einfach zu verwendende und gleichzeitig mächtige Computer-Vision-Funktionsbibliothek zu bieten, die je nach den gestellten Anforderungen angepasst werden kann. Sie bietet mehr als 500 Funktionen, die viele Bereiche der Computer Vision abdecken. In dieser Arbeit wurden jedoch nur wenige der bekannten Algorithmen genutzt, wohingegen viele der Standardfunktionalitäten zur Bildberechnung zur Anwendung kamen. Die meisten Algorithmen waren in der Funktionsbibliothek nicht implementiert und konnten selbstständig mithilfe grundlegender *OpenCV*-Funktionen realisiert werden. Demzufolge konnten einfache Bibliotheksfunktionen genutzt und im Rahmen dieser Arbeit um einige komplexe Algorithmen zur Bildanalyse erweitert werden.

Desweiteren wurde auf Standardbibliotheken von C++ zurückgegriffen, welche durch die Installation der Entwicklungsumgebung *Microsoft Visual Studio 2010* verfügbar waren. Letztendlich wurde zur Visualisierung der Ergebnisse noch die weitreichend bekannte grafische Bibliothek *matplotlib* von *Python* mithilfe eines Python-Wrapper in das C++-Programm eingebunden, da nur wenige anschauliche Funktionsbibliotheken zur Darstellung von grafischen Diagrammen (engl. *Plots*) für die Programmiersprache C++ existieren. Darüber hinaus zeichnet sich *matplotlib* durch die einfache Handhabung und die Kompaktheit der Befehlsketten zur Ausgabe von Diagrammen aus.

Für die Erstellung und Darstellung der grafischen Benutzeroberfläche (GUI) kam Qt zum Einsatz. Insbesondere wurde die Unterstützung für mehrere gleichzeitig geöffnete Dialoge genutzt, um die teils komplexen Nutzeingaben sinnvoll und

anschaulich zu verteilen. Ein weiteres besonderes *Feature* von Qt, welches ebenso verstärkt im Rahmen der Implementierung zum Einsatz kam, ist ein asynchrones, objektorientiertes Kommunikationskonzept, was auf *Signals & Slots* basiert. Dieses Konzept ermöglicht es prozessunabhängig Daten zwischen mehreren GUI-Objekten auszutauschen, ohne dass die Ausführung des Programms blockiert wird. Dadurch wurde es ermöglicht mehrere Berechnungen im Hintergrund auszuführen (z.B. Audio-Synchronisierung), während die grafische Benutzeroberfläche noch vollständig nutzbar bleibt (Form der Parallelisierung).

5.4 Algorithmischer Aufbau aus Sicht der Nutzereingaben

Der vollständige algorithmische Aufbau des entwickelten Prototypen erfolgte mit den genannten Bibliotheken und ist aus Sicht der Nutzereingaben an die schematischen Darstellung in der Abbildung 22 (siehe unten) angelehnt. Die Berechnungsschritte (zwischen zwei schwarzen Balken), welche in wichtigen, dauerhaft zu speichernden Daten münden (vollständig mit Großbuchstaben bezeichnete Aktionen), wurden in der Abbildung aufgrund ihrer Komplexität nicht vollständig abgebildet. In jedem dieser Schritte erfolgte die Berechnung der jeweiligen Daten durch die im Kapitel 4 (Konzeption geeigneter Verfahren) vorgestellten Verfahren, die prototypisch implementiert wurden. Häufig wird die Berechnung mithilfe der konzipierten Verfahren in der Abbildung als "ausführen" bezeichnet.

Ein Nutzer des Prototypen besitzt an jedem der schwarzen Balken die Möglichkeit auf das Menü (engl. *Menu*) der GUI zuzugreifen und jedem der mit dem schwarzen Balken (Menübalken) Aktionen anzuführen. Die gewählte Folge der Aktionen von oben nach unten im Diagramm ist die ideale, vorgesehene Reihenfolge nach welcher der Prototyp bedient werden sollte.

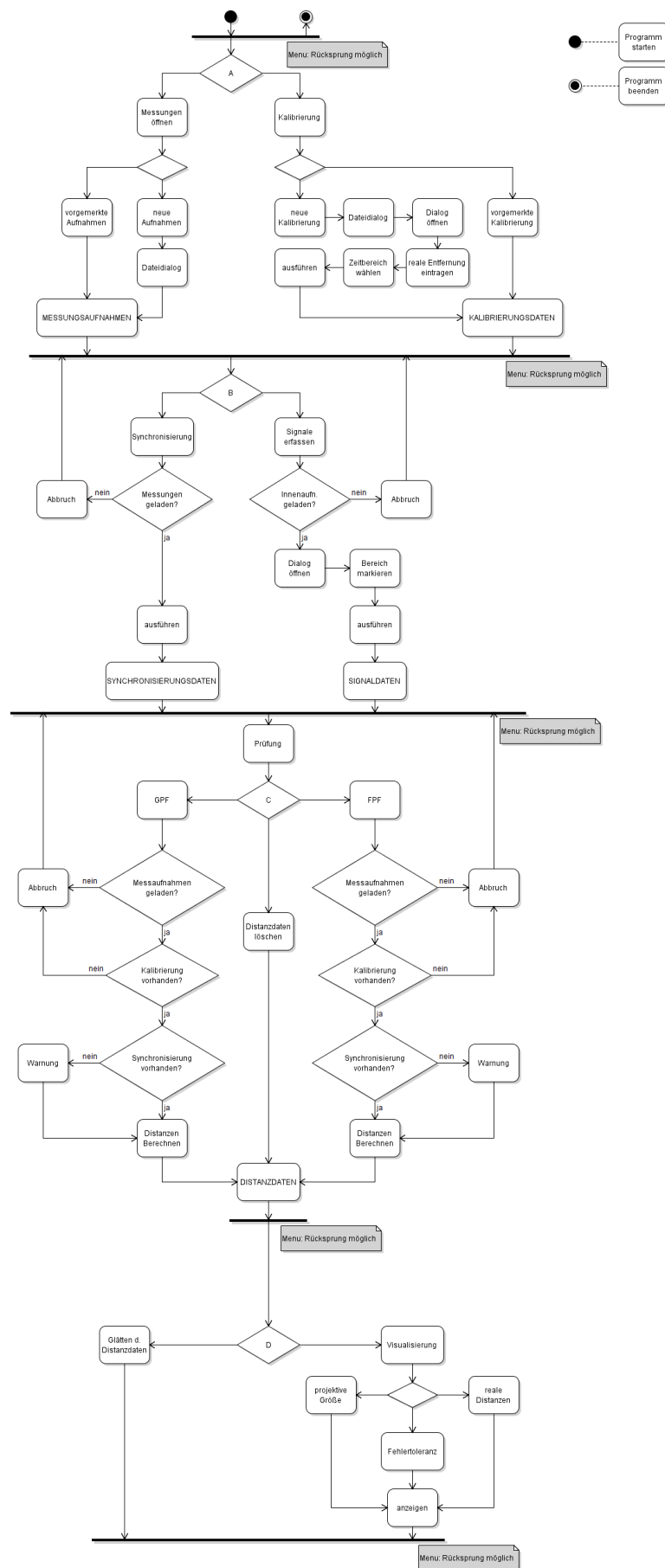


Abbildung 22: Algorithmischer Ablaufplan nach GUI-Funktionen des Prototypen (angelehnt an UML-Strukturdiagramm) [eigene Darstellung erstellt mit Violet UML Editor]

Abweichungen von dieser Reihenfolge, z.B. durch das Überspringen eines Schrittes, können dazu führen, dass benötigte Daten zur Berechnung fehlen. Diese Fälle werden durch entsprechende Prüfungen abgefangen, welche sicherstellen, dass alle benötigten Daten für diesen Berechnungsschritt vorhanden sind. Sollte solch eine Prüfung zum Abbruch des Berechnungsschrittes führen wird eine entsprechende Warnung ausgegeben, die auf den Grund des Abbruchs hinweist.

5.5 Ausgewählte Details der Implementierung

Im folgenden Unterkapitel werden ausgewählte Details der implementierten Verfahren ausgewählt und beschrieben. Es handelt sich dabei vor allem um Besonderheiten der Algorithmen, sowie die Beschreibung zusätzlicher in die Verfahren integrierter Ideen, die über die Konzeption hinaus gehen und erst während der Implementierung konzipiert wurden. Zumeist handelt es sich um kleinere algorithmische Schritte, die nötig waren, um die Daten für die Weiterverarbeitung oder Darstellung aufzubereiten. Die grundlegenden Verfahren des Kapitel 4, welche in diesem Abschnitt nicht aufgeführt sind, wurden wie beschrieben prototypisch implementiert und entsprechen überwiegend ihrer Konzeption.

5.5.1 Sub-Frame Audio-Synchronisation

Die Implementierung der Audio-Synchronisation folgt größtenteils der Konzeption wie sie im Absatz 4.3.2 beschrieben ist. Bei der Umsetzung ist jedoch deutlich geworden, dass das Finden von Übereinstimmungen der *Audio-Features* eine besondere Herausforderung mit vielen Lösungsstrategien darstellt. Aufgrund der gestellten Anforderung an ein Verfahren (respektive Implementierung), welches robust ist und nur eine geringe Laufzeit (Berechnungsaufwand) aufweist, wurde versucht eine möglichst eindeutige und schnelle Möglichkeit zur Bestimmung von übereinstimmenden Features zu entwickeln.

Das *Matching* der Features wurde durch den Vergleich weniger Feature-Eigenschaften zwischen beiden Tonaufnahmen vergleichsweise trivial implementiert.

Das Kriterium für eine Übereinstimmung (*Match*) war lediglich die gleiche Änderung der Richtung der Amplitudendifferenz von einem Feature zum nachfolgenden. Dieses Kriterium wurde für eine Reihe von Features in Form einer Sequenz (Vergleichssequenz) mit einer bestimmten Anzahl Features n (Sequenzlänge) wiederholt. Im entwickelten Prototypen hat sich durch Tests mit mehreren Testdatensätzen die Sequenzlänge $n = 8$ als eine zuverlässige Größe herausgestellt. Für jede Vergleichssequenz wurde daraufhin eine Zeitabweichung t_{offset} berechnet und gespeichert, wenn alle Features den Kriterien entsprachen. Wurde keine passende Vergleichssequenz gefunden, wurde diese Sequenz übersprungen und mit der nächsten fortgefahren. Die Zeitabweichung ergibt sich aus der Differenz der ursprünglichen zeitlichen Position des ersten Feature beider Vergleichssequenzen.

Aufgrund dessen, dass die eingehenden Tonaufnahmen sehr monoton sind (Motorengeräusche) und sich an manchen Stellen ähneln, kommt es darüber hinaus nicht selten vor, dass einer Vergleichssequenz mehrere Matches zugeordnet werden können. Um dem entgegenzuwirken, wird zusätzlich zu jedem Match die Summe der Amplituden-Abweichungen der aufeinanderfolgenden Features gespeichert. Damit werden nur die beiden übereinstimmenden Vergleichssequenzen mit der geringsten Amplitudenabweichung akzeptiert und deren Zeitabweichung t_{offset} gespeichert. Für diese Amplitudendifferenzberechnung müssen jedoch vorher beide Tonaufnahmen in ihrer Lautstärke mit einem geeigneten Standardverfahren normalisiert werden, da sonst die gesuchten übereinstimmenden Vergleichssequenzen immer eine entsprechend hohe Amplitudendifferenz aufweisen können.

Durch das, verglichen mit allen verfügbaren Eigenschaften der Audio-Features, triviale Matching entstehen relativ viele mögliche Zeitabweichungen für eine Vielzahl von Vergleichssequenzen. Über alle ermittelten Zeitverschiebungen gibt es einen großen Anteil von Ausreißern, die in Form von Streuung auftreten. Werden alle Zeitabweichungen geordnet und betrachtet die Verteilung so kann eine Normalverteilung der Werte festgestellt werden. Um die reale Abweichung zu ermitteln, können nun folglich entsprechende statistische Berechnungen auf diese Verteilung angewandt werden. Im Rahmen der Implementierung und dem Testen mit den vorhandenen Testdatensätzen hat sich jedoch herausgestellt, dass eine zuverlässige und einfache Methode die Filterung der geordneten Werte nach

Häufigkeiten ist, da die gesuchte echte Zeitabweichung überwiegend mehrfach und Ausreißer oft nur einmalig auftraten. So wurden alle Werte der Liste gelöscht, die keine Nachbarn besaßen, welche innerhalb eines schmalen Toleranzbereichs (Toleranzzeit t_{toleranz}) lagen. So konnte eine große Anzahl von Ausreißern eliminiert und abschließend die echte Zeitabweichung durch die Anwendung eines einfachen Medians, auf die nun sehr große Menge der Werte mit echter Zeitabweichung und wenigen Ausreißern, berechnet werden.

5.5.2 Beschleunigung der Erfassung der relevanten Messdaten

Die Suche von Messdaten, welche der dreieckigen Anordnung der normierten Marker auf der Prüfmaske (siehe Abbildung 23) entsprechen, hat je nach angewandtem Verfahren das Ziel spezielle Merkmale innerhalb der Videoaufnahme zu suchen. In der Regel wird der gesamte zur Verfügung stehende Bildbereich nach diesen Merkmalen abgesucht, sodass sich die Berechnungszeit abhängig von der Bildgröße erhöht. Daher ist es sinnvoll eine Vorauswahl der Bildbereiche (ROI) zu treffen in denen entsprechende Marker und deren Merkmale zu finden sind und die Berechnungszeit somit zu reduzieren (schwarzes Viereck in der Abbildung) und nicht relevante Bereiche damit zu ignorieren (Grau gestreifter Bereich).

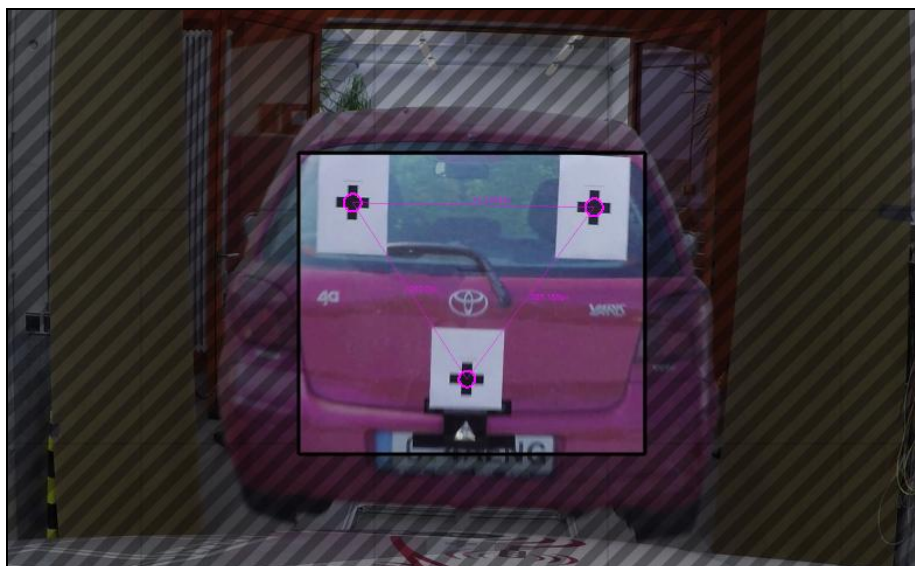


Abbildung 23: ROI (Region of Interest; schwarz) zur Einschränkung des zur Berechnung herangezogenen Bildbereichs [eigene Darstellung]

Im vorliegenden Fall wird das zu prüfende Fahrzeug auf einem Rollenprüfstand betrieben, auf dem es über die Hebevorrichtung fixiert ist und mit Kameras versehen ist, deren Ausrichtung im Weltkoordinatensystem kaum Änderungen unterliegt. Der erfasste Bereich der realen Szene ist quasi unveränderlich. Darüber hinaus werden bei beiden implementierten verfahren Marker auf einer Prüfmaske gesucht, die sich auf einem Schienensystem bewegen, welches wiederum nur Bewegungen der Prüfmaske direkt auf das zu prüfende Fahrzeug zu und von diesem Fahrzeug weg erlaubt.

Zusammengefasst verändert sich also der Bildbereich in dem die Marker erscheinen kaum, sodass der zu untersuchenden Bildbereich einschränken werden kann. Während der Implementierung hat sich herausgestellt, dass die Einschränkung des Bildbereichs, also die Bestimmung eines ROI, durch den Schritt der Kalibrierung sinnvoll erscheint. Dabei wird die Größe des ROI mithilfe der scheinbaren Größe der zur Kalibrierung genutzten relativen Größen der Dreiecke, über dem durch den Nutzer manuell bestimmten Zeitbereich, bestimmt. Die Größe des ROI wird abschließend mit einem Faktor multipliziert, der das ROI so vergrößert, dass die Marker vollständig hinein passen.

5.5.3 Auswahl der relevanten Messdaten

Bei der nun beschleunigten Erfassung der Dreiecke (Teil der Messdaten), werden, bedingt durch das breite Ansprechen der implementierten Verfahren auf Kreuzformen im Bildbereich, Dreiecke gefunden, welche von der ursprünglichen Anordnung der Marker auf der Prüfmaske abweichen.

Dadurch muss eine Auswahl der relevanten Dreiecke erfolgen, die alle Ausreißer eliminieren können. Durch die Anwendung einer einfachen Häufigkeitsanalyse kann bereits ein großer Teil der Ausreißer gefunden werden. Die häufigsten Größenverhältnisse (*Aspect Ratio*) der Dreiecke sind in der Regel die gesuchten Dreiecke auf der Prüfmaske. Das kann mit den bekannten Daten vom Schritt der Kalibrierung, und der damit verbundenen Kenntnis über die Anordnung der Dreiecke auf der Prüfmaske, verbunden werden. Es können die Eigenschaften der zur Kalibrierung herangezogenen Dreiecke genutzt werden, um die erfassten Messdaten

abzusichern und zu säubern. In der Regel können so Ausreißer vollständig eliminiert und anschließend die erfassten Messdaten über *Plots* dargestellt werden.

5.5.4 Interpolation und Glättung der Messergebnisse⁴²

Die gesäuberten Messdaten sind direkt nach der Anwendung des GPF oder FPF nicht in jedem Fall für die Darstellung in einem Diagramm (*Plot*) geeignet. Durch Bildstörungen, wie (Farb-)Rauschen oder Reflektionen auf den Markern, können nicht in jedem Frame der Videoaufnahmen zuverlässig alle Marker und somit nicht immer die Dreieckanordnung auf der Prüfmaske erkannt werden. Dadurch ist eine akkurate Distanzbestimmung nicht möglich. Die so entstehenden Lücken in den Messdaten kann durch Interpolation ausgeglichen werden, was eine Abschätzung der vorherrschenden Distanzen zu beliebigen Zeitpunkten zwischen zwei erfassten Messungen ermöglicht.

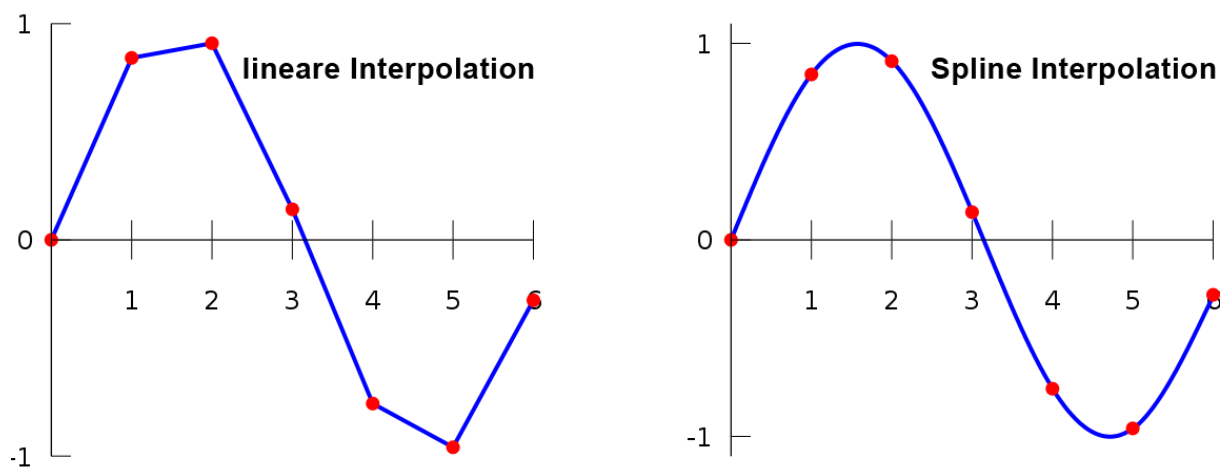


Abbildung 24: verschiedene Interpolationsmethoden (Angelehnt an Darstellungen von [22])

Die einfachste Form der Interpolation ist die lineare Interpolation, also eine einfache Verbindung von zwei Distanzmessungen zu entsprechenden Zeitpunkten. Dies führt bei gleichbleibenden, geradlinigen Bewegungen zu korrekten Ergebnissen. Sollte die Prüfmaske jedoch eine beschleunigte Bewegung ausführen, kann keine zuverlässige Aussage, bzw. Schätzung, mehr über die Distanz zu einem nicht messbaren Zeitpunkt gegeben werden (siehe Abbildung 24). Mithilfe fortgeschrittener Interpolationsverfahren, wie der *Spline Interpolation*, können auch Beschleunigungen

⁴² Anlehnung an [28]

approximiert und Distanzen zu beliebigen Zeitpunkten zwischen zwei vorhandenen Messzeitpunkten abgeschätzt werden.

Im Rahmen der Entwicklung des Prototypen wurde eine Form der Interpolation gewählt, die der *Spline Interpolation* ähnelt. Dabei wird anstelle fehlender Messungen zum Zeitpunkt realer Frames der Videoaufnahme eine lineare Interpolation angewendet und weiter verarbeitet. Aus einem realen Messpunkt mit interpoliertem Nachbarn in positiver Zeitrichtung und eben diesem interpolierten Messpunkt wird nun das arithmetische Mittel gebildet und somit eine Form der Glättung vollzogen. Der so entstehende Mittelwert ersetzt den linear interpolierten Wert und wird temporär als "realer Wert" für mögliche nachfolgende interpolierte Messpunkte genutzt. Nach diesem Prinzip wird mit allen interpolierten Messpunkten verfahren. Durch die gerichtete Glättung (entlang der Zeitachse) wird, ausgehend von realen Messpunkten, die Kurvencharakteristik so verändert, dass sich Berge und Täler leicht in die positive Zeitrichtung der Glättung verschieben. Um dem entgegenzuwirken, wird ausgehend vom zeitlich letzten realen Messpunkt in negative Zeitrichtung analog vorgegangen. Die Kurve wird dadurch entgegengesetzt zur vorherigen Glättungsrichtung in negative Zeitrichtung verlagert und die Effekte der beiden Glättungen gleichen sich aus.

Diese einfache Form der Interpolation führt bei mehrfacher Ausführung zu einem deutlich glatteren Verlauf und lässt zuverlässigere Aussagen über die Distanz zu beliebigen Zeitpunkten zwischen zwei realen Messpunkten zu, als es durch die lineare Interpolation möglich wäre.

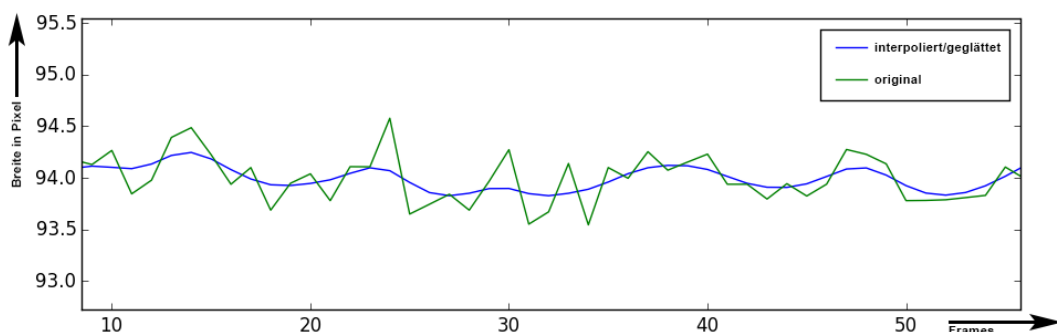


Abbildung 25: Interpolation und Glättung der Messdaten [eigene Darstellung]

In der Abbildung 25 wurde diese Form der Interpolation und Glättung angewendet. Die Kurven in der Abbildung stellen die relative Breite der dreieckigen Anordnung der

normierten Marker auf der stillstehenden Prüfmaske dar. Die grüne Kurve zeigt deutlich die Ungenauigkeiten der Messung der relativen Größe des Dreiecks (hier: Breite in Pixel) um einen Wert von $\pm 0,5$ Pixel, durch beschriebene Störeinflüsse. Mithilfe der implementierten Interpolation und Glättung konnten Lücken geschlossen und die Ungenauigkeit verringert werden.

5.6 Vorstellung der Testergebnisse

Ziel der Arbeit war die Durchführung von Messungen mithilfe des entwickelten Prototypen und entsprechenden Testdatensätzen durch die Simulation einer Gefahrensituation mithilfe der zur Verfügung stehenden Prüftechnik (s. 5.1). Für die Messungen auf dem Prüfstand wurde hauptsächlich das Fahrzeug A, der VW Passat CC, eingesetzt und insgesamt 16 Messungen durchgeführt. Jedoch zeigte das Fahrzeug A in drei von sechzehn Fällen keine Reaktion auf die entgegenkommende Prüfmaske, sodass diese Messungen nicht nutzbar waren. Letztendlich standen dreizehn erfolgreiche Durchgänge zur Verfügung, die im Prototypen ausgewertet werden konnten, wobei wiederum drei der dreizehn Durchgänge eine unzureichend späte Reaktion auf die simulierte Gefahrensituation zeigten.

5.6.1 Gegenüberstellung der Genauigkeit der implementierten Verfahren

Zur Berechnung der absoluten Entfernungen der Testdatensätze wurden die eigenständig entwickelten und im Kapitel 4 beschriebenen Verfahren, das *Gradient-based Pattern Finding* und das *Feature-based Pattern Finding*, genutzt. Dennoch unterliegen beide Verfahren unterschiedlicher Zuverlässigkeiten bei der Erkennung, und schwankender Genauigkeit bei der Bestimmung, der Mittelpunkte der normierten Marker, wie es bereits im Abschnitt 4.2.3 "*Gegenüberstellung der theoretischen Vor- und Nachteile* beider Verfahren" beschrieben wurde. Nach der Implementierung der Verfahren und die darauf folgende Nutzung des Prototypen mit den Testdatensätzen sind die theoretischen Vorüberlegungen bestätigt worden, was in den nachfolgenden beiden Abbildungen beispielhaft dargestellt wurde.

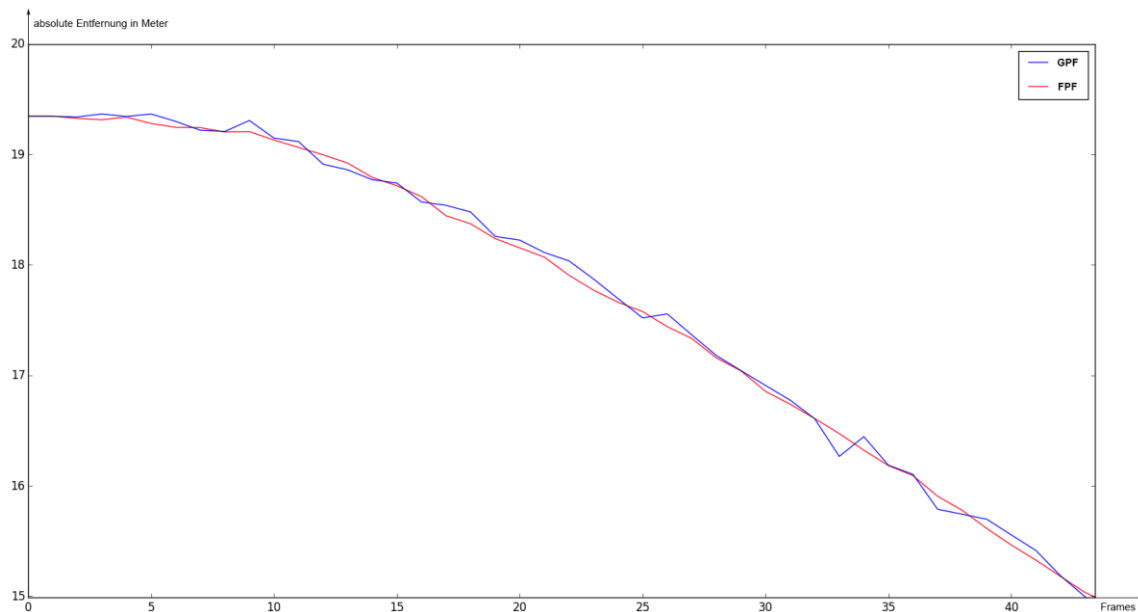


Abbildung 26: Gegenüberstellung der ermittelten abs. Entfernungen von GPF u. FPF [eigene Darstellung]

Beide Abbildungen wurden dem gleichen Testdatensatz ("run3") entnommen und stellen das GPF dem FPF direkt gegenüber. Dabei zeigte sich, dass das FPF eine wesentlich genauere Entfernungsbestimmung erlaubt als das GPF, was sich auch in einer niedrigeren Streuung der Fehlertoleranz in Abbildung 27 verdeutlicht. Die höhere Genauigkeit des FPF kommt, wie bereits in den theoretischen Vorüberlegungen erläutert, durch die größere Anzahl von Markermerkmalen zur Mittelpunktberechnung zustande.



Abbildung 27: Gegenüberstellung der Fehlertoleranz von GPF und FPF [eigene Darstellung]

Trotz der gezeigten Vorteile des FPF ist die Zuverlässigkeit des FPF niedriger als die des GPF. Das GPF nutzt lediglich den Helligkeitsverlauf der Marker ausgehend von der Mitte der normierten Marker um potentielle Marker zu bestimmen. Sollten Details der Marker durch Störeinflüsse, Bewegungsunschärfe, oder der begrenzten Auflösung des eingesetzten Kamerasystems undeutlich abgebildet werden, so besitzt das GPF eine höhere Erkennungsrate als das FPF.

<pre>Start processing GPF. triangles interpolated: 16 (8% of all frames) Success: Markers processed. (Gradient Pattern) Start processing FPF. triangles interpolated: 77 (42% of all frames) Success: Markers processed. (Feature Pattern)</pre>
--

Abbildung 28: Zuverlässigkeit des GPF und FPF [eigene Darstellung]

Die Begründung darin liegt, wie im Abschnitt 4.2.3 vermutet, an der sensiblen Suche des FPF nach Eckpunkten innerhalb der Marker. Sollten die Marker durch genannte Einflüsse nicht scharf und in ausreichender Auflösung abgebildet werden, so können nicht alle vier der benötigten inneren Ecken in der Abbildung der Marker gefunden werden. Dieses Problem zeigt sich in allen vorliegenden Testdatensätzen, was in der Abbildung 28 beispielhaft mithilfe des Testdatensatz "run3" gezeigt wird, welcher bereits in den vorherigen beiden Abbildungen genutzt wurde. Zusammenfassend wird damit deutlich, dass die Interpolation der fehlenden Messdaten von großer Bedeutung ist (s. 5.5.4).

5.6.2 Bewertung der Notbremsfunktion

Bereits bei der Aufnahme der sechzehn Durchläufe zeigte sich, dass die Notbremsfunktion des Fahrzeugs A unregelmäßig auslöste, sodass das vorausfahrende Fahrzeug, in Form der Prüfmaske, bei nahezu gleichbleibender Geschwindigkeit manchmal nicht oder während der Annäherung nicht durchgehend erkannt wurde. Die unregelmäßige Erkennung der Prüfmaske führte dazu, dass der Notbremsassistent zum Teil nicht, oder nur sehr spät ausgelöst wurde. Diese Probleme können auch den Messergebnissen entnommen werden. In der nachfolgenden Abbildung 29 ist zu erkennen, dass die erstmalige Erkennung der Gefahrensituation (Warnstufe) durch die entgegenkommende Prüfmaske (Rot),

sowie die Auflösung der Warnung (Grün) zum Teil unregelmäßig über die Kurve der ermittelten absoluten Entfernungen verteilt ist.

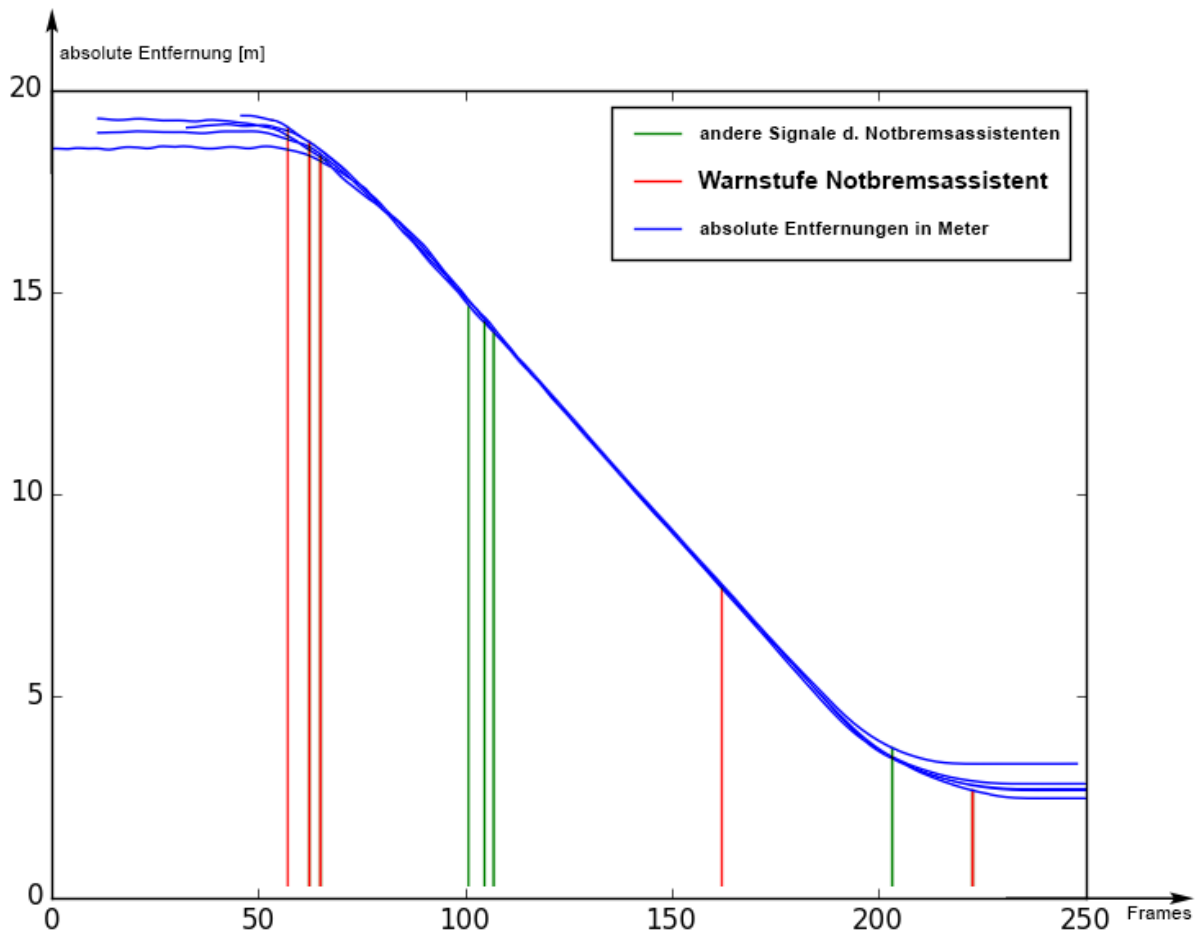


Abbildung 29: Gesamtergebnis der Reaktion des Notbremsassistenten nach fünf Durchläufen (VW Passat CC) [eigene Darstellung]

Für die gezeigte beispielhafte Abbildung wurden fünf aufeinanderfolgende der sechzehn Durchläufe genutzt. Die gewählten Durchläufe spiegeln die Gesamtheit aller Durchläufe ausreichend wieder und sind zur Veranschaulichung der Bewertung der Notbremsfunktion von Fahrzeugen gut geeignet. Im vorliegenden Beispiel wurde darauf geachtet, dass alle Durchläufe in ihrer Bewegung synchronisiert sind, d.h. die Maske fängt etwa zum gleichen Zeitpunkt an sich zu bewegen. Die Start- und Endentfernung der Prüfmaske variiert in geringem Maße zwischen den Durchläufen. Dem Beispiel kann nun nach den definierten Bewertungskriterien (s. 3.2) entnommen werden, dass das Fahrzeug in drei von fünf Durchläufen die Notbremsfunktion in nahezu gleichen Entfernungen (zwischen 18,5 und 19 Meter) auslöst (Warnstufe), welche in etwa gleichen Zeitabständen (erste drei grüne Markierungen) wieder aufgelöst werden (zwischen 14,5 und 15 Meter). In zwei von drei Fällen wurde die

Notbremsfunktion des Fahrzeugs jedoch zu spät ausgelöst (bei ca. 7,5 und 3 Meter). In einer realen Gefahrensituation hätte dies zu einer unzureichenden Auslösung der Notbremsfunktion geführt und somit wäre die Aufprallenergie eines potentiellen Unfalls nicht ausreichend reduziert worden.

Nach diesem Beispiel könnte ein Prüfer zu dem Schluss kommen, dass das Fahrzeug aus dem vorgestellten Beispiel die Prüfung des Notbremsassistenten nicht, oder nur mit Einschränkungen, bestanden hätte, sodass eine konkrete Überprüfung der Funktionalität in einer Werkstatt erfolgen und das Fahrzeug dem Prüfer erneut vorgestellt werden müsste. Mit dem in diesem Beispiel vorgestellten Prinzip können nun entsprechende Prüfungen beliebiger Fahrzeuge erfolgen, um zumindest die Streuung und Fehleranfälligkeit des Notbremsassistenten des Fahrzeugs zu prüfen. Die Prüfung der langfristigen Veränderung (z.B. durch Alterung) der Reaktionen des Notbremsassistenten kann wie in Abschnitt 3.2 beschrieben nur unter Zuhilfenahme einer globalen Datenbank erfolgen.

5.6.3 Mögliche Fehlerquellen und ihre Ursachen

Bei den Testläufen mit dem Fahrzeug A (VW Passat CC) hat sich gezeigt, dass bei drei der sechszehn Testläufe die Prüfmaske nicht detektiert wurde. Bei weiteren drei Durchläufen erfolgte die Reaktion des Fahrzeug A sehr spät, sodass die Verwicklung des Fahrzeug A in einen Unfall nicht ausgeschlossen werden kann. Diese sechs von sechszehn Fälle, also $\frac{6}{16} = 37,5\%$ aller Fälle, sind in der Praxis als problematisch zu betrachten. Jedoch muss davon ausgegangen werden, dass die Testläufe im gegebenen Szenario gewissen Fehlerquellen unterliegen, welche die fahrzeugeigene Sensorik gestört und so das Messergebnis verfälscht haben könnte.

Der Versuchsaufbau (s. 5.1: Anordnung der Prüftechnik) der Testläufe war in einer kleineren Werkstatthalle aufgebaut. Das zu prüfende Fahrzeug war dabei in Richtung des Schienensystems ausgerichtet, welches sich selbst entlang einer tunnelartig aufgebauten Raumstruktur erstreckte, die außerdem von zwei Torbögen unterbrochen ist. Dadurch kann davon ausgegangen werden, dass die Fahrzeugsensorik von diesem "Prüftunnel" zum Teil gestört worden sein könnte. Die Begrenzungsflächen des Prüftunnel bestanden dabei zum größten Teil aus massiven

Beton, oder metallischen Flächen. Darüber hinaus verwendet das Fahrzeug A einen Radarsensor zur Abstandsbestimmung zu vorausfahrenden Fahrzeugen, welcher in "Tunneln" durch Mehrfachreflektionen der Radarwellen an den Begrenzungsflächen (Tunnelwänden) störanfällig ist (s. 2.3.3: Radar). Gerade die Begrenzungsflächen des gegebenen Prüftunnels bestanden aus Materialien, welche Radarwellen besonders gut reflektieren, sodass eine nicht zu vernachlässigende Beeinflussung durch diese Störgröße nicht zu vernachlässigen ist. Zusammenfassend ist es daher notwendig einen späteren Prüfbetrieb in einem entsprechend geeigneterem Umfeld auszuführen, um möglichst viele Störeinflüsse zu umgehen.

6 Zusammenfassung und Ausblick

Anhand der vorgestellten Verfahren und den prototypischen Implementierungen konnte gezeigt werden, dass geeignete Lösungen auf Basis der Computer Vision und dem Einsatz von nur einer Kamera zur Distanzmessung entwickelt werden können. Das Ergebnis ist ein Prototyp, welcher die gestellten Anforderungen erfüllt und zwei eigenständig entwickelte Verfahren in einer Software vereint, deren Aufbau an die potentielle Integration in das Tagesgeschäft von Prüforganisationen angelehnt ist.

Bei der Ausarbeitung der Arbeit wurden einige grundlegende Ansätze der Computer Vision aufgegriffen und verändert, um sie entsprechend der Problemstellung anzupassen. Dadurch konnten vor allem aus den Ansätzen der Monovision geeignete Verfahren entwickelt werden. Dazu zählt das *Gradient-based Pattern Finding* und das *Feature-based Pattern Finding*, welche auf das Wiederfinden eines normierten, rotationssymmetrischen Markers spezialisiert sind. Die so entstandene robuste Bestimmung der Mittelpunkte der Marker konnte für die Berechnung von absoluten Entfernungen mit hoher Genauigkeit genutzt werden.

Neben den für die Problemstellung gegebenen Lösungsvorschlägen existieren noch eine Vielzahl weiterer möglicher Herangehensweisen, sodass fortführende Untersuchungen in Betracht gezogen werden können.

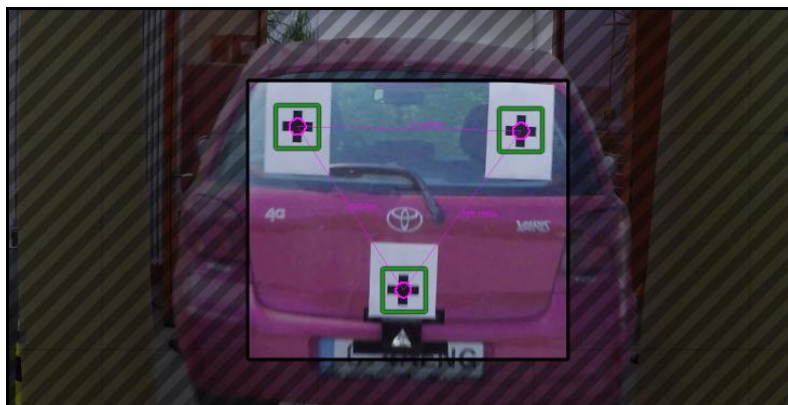


Abbildung 30: Vorauswahl von kleineren ROIs für das FPF mithilfe des GPF [eigene Darstellung]

Eine triviale Möglichkeit die Berechnungszeit des implementierten FPF zu reduzieren und dennoch die hohe Genauigkeit beizubehalten, ist die Kombination des GPF mit dem FPF. Das GPF kann, als das Verfahren mit geringerer Berechnungszeit, zur Vorauswahl der Marker im Bildbereich genutzt werden. Anschließend kann das FPF

in den ermittelten kleineren Bildbereichen angewendet werden (s. Abbildung 30). Im Vergleich mit der beschriebenen ROI (s. 5.5.2) kann so die Berechnungszeit des FPF-Teilschrittes deutlich reduziert werden. In der Summe liegt die so erzielte Berechnungszeit des kombinierten Verfahrens folglich zwischen der des GPF und FPF.

Die beiden Verfahren können darüber hinaus noch dahingehend verbessert werden, verschiedene andere Marker zu erkennen und auszuwerten. Ein relativ naheliegender Ansatz wäre die Implementierung der Unterstützung von spiegel- und rotationssymmetrische Marker, die vom Nutzer manuell eingegeben und angelernt werden können. Das FPF würde dabei weiterhin Marker benötigen, welche charakteristische Ecken aufweisen. Das GPF hingegen würde auch mit Markern funktionieren, welche charakteristische Helligkeitsverläufe (Gradienten) in verschiedenen Richtungen ausgehend vom Zentrum des Markers besitzen (s. Abbildung 31).

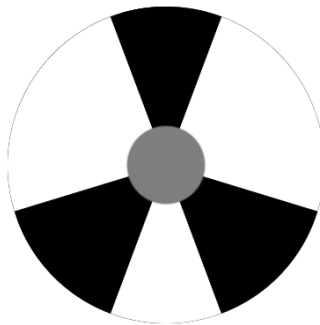


Abbildung 31: beispielhafte Abbildung eines geeigneten Marker zum Anlernen für das GPF [eigene Darstellung]

Wie bereits in einem vorhergehenden Absatz dieses Abschnitts beschrieben, ist die Vorauswahl von Bereichen in Form einer ROI sinnvoll, um die Berechnungen auf einen Bereich zu beschränken in dem sich annähernd ausschließlich zur Berechnung benötigte Abbildungen von Objekten befinden. So kann die Erkennung von Fahrzeugrückseiten (mithilfe *Haar Classifier*, siehe beispielhafte Abbildung 32) zur groben Abschätzung der Distanzen zu vorausfahrenden Fahrzeugen, oder zur Vorauswahl für die Definition einer ROI, dienen. Die so ermittelte ROI kann für zukünftige, geeignete Algorithmen verwendet werden, welche, anhand der Abbildung der Fahrzeugrückseite, Entfernungen genauer einschätzen können. Zum Beispiel kann dies durch eine hinterlegte Datenbank erfolgen, welche die Erscheinung

bekannter Fahrzeuge enthält. Die so erkannten Fahrzeugrückseiten (Bildung von ROIs) können anschließend beispielsweise mit einer Stereokamera gezielt analysiert werden um exakte Entfernungen zu den erkannten Fahrzeugen zu ermitteln. Vor allem der Ansatz mit der Stereokamera ist ein vielversprechender Ansatz, der weiter verfolgt werden sollte, wenn die Entfernung von beliebigen vorausfahrenden Fahrzeuge im realen Straßenverkehr ermittelt werden soll.

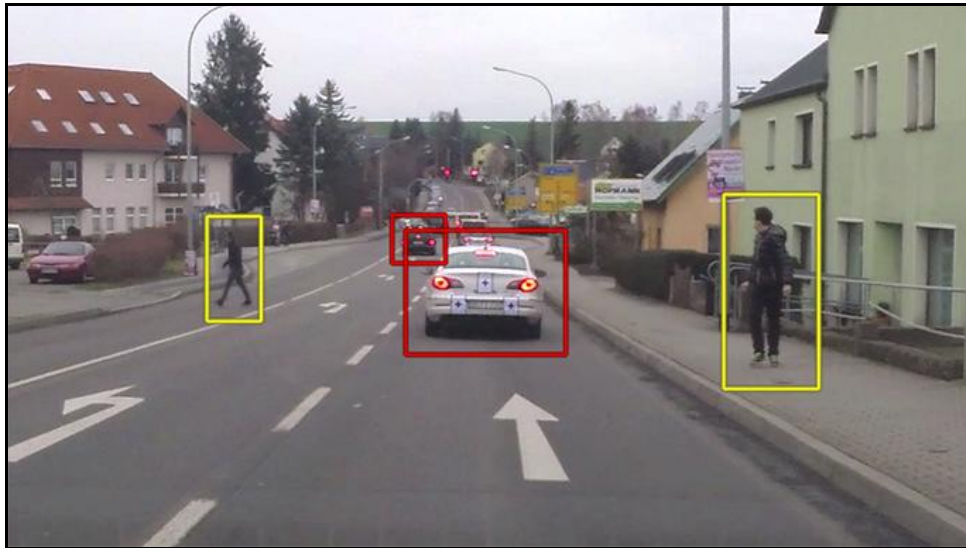


Abbildung 32: Mögliche Anwendung eines *Haar Classifier* zur Erkennung von Fahrzeugen (Rot) und Personen (Gelb) [eigene Darstellung]

Abschließend kann auch eine Verbesserung des Ansatz zur der Auswahl der relevanten Messdaten angestrebt werden. Bei mehreren erkannten Dreiecken im Frame einer Messungsaufnahme kann neben implementierter Auswahl zusätzlich eine weitere Auswahl durch Vergleich des Dreieckmittelpunkte benachbarter Frames erfolgen. Die Positionen der Mittelpunkte müssen sich bei zugehörigen Dreiecken über die Zeit harmonisch bewegen und dürfen sich nicht sprunghaft ändern.

7 Abbildungsverzeichnis

Abbildung 1: Lochkameramodell (Quelle: Mindpicnic [5], Alternative Quelle: [6])	5
Abbildung 2: Optische Verzeichnung (von oben nach unten: kissenförmig, tonnenförmig, keine) (Quelle: Olypedia - Olympus-Wiki [9])	8
Abbildung 3: Beeinflussung des Frequenzganges eines neutralen 2-Wege-Lautsprechers durch die Raumakustik (Messung in unterschiedlicher Entfernung frontal vor Lautsprecher Grün: 1m Gelb: 3m) (Quelle: eigenes Projekt, Software: Room EQ Wizard v5.14)	9
Abbildung 4: Schematischer Aufbau eines Pixel auf einem CCD-Bildsensor (Quelle [11])	12
Abbildung 5: Aufbau eines Lidar-Sensors (Angelehnt an Quelle: SICK, Inc.)	18
Abbildung 6: Explosionsdarstellung des Hella-ACC-Radars (Quelle: Hella)	21
Abbildung 7: verschiedene Ansätze zur Bestimmung der Winkelposition bei Radarsystemen (Angelehnt an Bildquellen: [2] S. 141-144)	24
Abbildung 8: Fahrzeug- und Spurerkennung von Mobileye (2005) (Quelle: Einzelbild aus Video von http://www.mobileye.com/technology/applications/vehicle-detection/forward-collision-warning/)	29
Abbildung 9: Translation des Kamerakoordinatensystem (Rot) in das Fahrzeugkoordinatensystem (Grün) ausgehend vom im Fahrzeug eingesetzten Sensor zur Distanzmessung vorausfahrender Fahrzeuge [eigene Darstellung]	32
Abbildung 10: Dreieckige Anordnung von Markern auf einer Prüfmaske [eigene Darstellung]	41
Abbildung 11: Berechnung der Mittelpunktkoordinate (Gelb) eines abgebildeten Objekts (Rot) mit Subpixelgenauigkeit [eigene Darstellung]	42
Abbildung 12: Gradient-based Pattern Finding, mit den Diagonalen (Orange), Achsparallelen (Rot) und einheitlicher Gradientenlänge (Grau) [eigene Darstellung]	44
Abbildung 13: Bestimmung von möglichen Zentren (Rot, Türkis, Grün, Gelb) und anschließende Mittelung zum echten Zentrum mit Subpixelkoordinaten (Blau) [eigene Darstellung]	45
Abbildung 14: Prinzip von FAST Features (Quelle: [20] und [19])	46
Abbildung 15: Beispielhafte Anwendung von FAST auf den normierten Marker während der Laufzeit [eigene Darstellung]	47
Abbildung 16: Logische Identifizierung von Markern mithilfe von FAST Features (FPF Algorithmus) [eigene Darstellung]	49

Abbildung 17: Typische Audiospur mit gleichem, jedoch zeitlich verschobenem, Frequenz- und Amplitudengang (Offset; grüne Tonspur zeitlich später) [eigene Darstellung]	54
Abbildung 18: Frequenzanalyse der Tonaufnahme eines Prüfstandlaufes des VW Passat CC mit Dieselmotor, des Laborbereichs Fahrzeugtechnik der HTW Dresden, auf dem Prüfstand MFP 3000 von Maha (Darstellung wurde mit der Software WavePad Audio-Editor erzeugt) [eigene Darstellung].....	56
Abbildung 19: Audio Synchronisation - Bestimmung des Offset sich ähnelnder akustischer Abschnitte [eigene Darstellung].....	58
Abbildung 20: Anordnung der Prüftechnik [eigene Darstellung]	64
Abbildung 21: Test des ACC des BMW i3 mit Auswertung durch den entwickelten Prototypen [eigene Darstellung]	67
Abbildung 22: Algorithmischer Ablaufplan nach GUI-Funktionen des Prototypen (angelehnt an UML-Strukturdiagramm) [eigene Darstellung erstellt mit Violet UML Editor]	70
Abbildung 23: ROI (Region of Interest; schwarz) zur Einschränkung des zur Berechnung herangezogenen Bildbereichs [eigene Darstellung]	73
Abbildung 24: verschiedene Interpolationsmethoden (Angelehnt an Darstellungen von [22]).....	75
Abbildung 25: Interpolation und Glättung der Messdaten [eigene Darstellung]	76
Abbildung 26: Gegenüberstellung der ermittelten abs. Entfernungen von GPF u. FPF [eigene Darstellung].....	78
Abbildung 27: Gegenüberstellung der Fehlertoleranz von GPF und FPF [eigene Darstellung]	78
Abbildung 28: Zuverlässigkeit des GPF und FPF [eigene Darstellung]	79
Abbildung 29: Gesamtergebnis der Reaktion des Notbremsassistenten nach fünf Durchläufen (VW Passat CC) [eigene Darstellung]	80
Abbildung 30: Vorauswahl von kleineren ROIs für das FPF mithilfe des GPF [eigene Darstellung]	83
Abbildung 31: beispielhafte Abbildung eines geeigneten Marker zum Anlernen für das GPF [eigene Darstellung].....	84
Abbildung 32: Mögliche Anwendung eines <i>Haar Classifier</i> zur Erkennung von Fahrzeugen (Rot) und Personen (Gelb) [eigene Darstellung].....	85

8 Literaturverzeichnis

1. **LEIBER, H, CZINCZEL, A und ANLAUF, J.** *ANTIBLOCKIERSYSTEM (ABS) FÜR PERSONENKRAFTWAGEN*. s.l. : Bosch, 1980.
2. **(Hrsg.), Hermann Winner | Stephan Hakuli | Gabriele Wolf.** *Handbuch Fahrerassistenzsysteme*. Wiesbaden : Vieweg+Teubner | GWV Fachverlage GmbH, 2009.
3. Wikipedia. *Wikipedia*. [Online] Wikipedia, 2016. [Zitat vom: 25. März 2016.] <https://de.wikipedia.org/wiki/Notbremsassistent>.
4. **Zisserman, Richard Hartley and Andrew.** *Multiple View Geometry in computer vision (Second Edition)*. Cambridge : Cambridge University Press, 2003.
5. Mindpicnic. [Online] 2013. [Zitat vom: 21. 09 2013.] <https://www.mindpicnic.de/>.
6. **Richter, Patrick.** *Möglichkeiten und Grenzen der Erkennung von Hindernissen für ein Wasserfahrzeug unter Verwendung von Computer Vision*. Dresden : s.n., 2013.
7. **Thöniß, Thomas.** *Abbildungsfehler und Abbildungsleistung*. Göttingen : s.n., 2004.
8. Abbildungsfehler. *Wikipedia*. [Online] [Zitat vom: 22. August 2013.] <http://de.wikipedia.org/wiki/Abbildungsfehler>.
9. Olypedia - Olympus-Wiki. [Online] 2013. [Zitat vom: 16. 09 2013.] http://olypedia.de/static/images/d/d2/Verzeichnung_Blende_Wikimedia.png.
10. **Richter, Patrick.** *Eigene akustische Raumanalyse (Software gestellt von: Sonatech)*. 2016.
11. Wikimedia. [Online] 2014. [Zitat vom: 20. Mai 2014.] http://upload.wikimedia.org/wikipedia/commons/6/66/CCD_charge_transfer_animation.gif.
12. Volvo Collision Warning with Auto Brake (CWAB). *Volvo Cars*. [Online] [Zitat vom: 20. August 2013.] https://www.media.volvocars.com/media/articles/html/12210_3_2.aspx.

13. Mobileye. [Online] Mobileye, 2016. [Zitat vom: 25. März 2016.] <http://www.mobileye.com/technology/applications/vehicle-detection/>.
14. **Martin Buehler, Karl Iagnemma, Sanjiv Singh (Eds.).** *The DARPA Urban Challenge - Autonomous Vehicles in City Traffic*. Heidelberg : Springer, 2009.
15. **Zehang Sun, Member, IEEE, George Bebis, Member, IEEE, and Ronald Miller.** On-Road Vehicle Detection: A Review. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL. 28, NO. 5. s.l. : IEEE, 2006.
16. **Nindel, Thomas.** *Überdeckungsfreie Fassadentexturen aus mehreren Nahbereichsfotografien*. Dresden : s.n., 2009.
17. **Kaehler, Gary Bradski & Adrian.** *Learning OpenCV - Computer Vision with the OpenCV Library*. s.l. : O'REILLY, 2008.
18. **Miroslav Trajković, Mark Hedley.** Fast Corner Detector. *Image and Vision Computing* 16 (1998) 75-87. 1998, 16.
19. opencv documentation. *opencv documentation*. [Online] 2016. [Zitat vom: 27. März 2016.] http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_fast/py_fast.html.
20. **Edward Rosten.** edwardrosten. *edwardrosten*. [Online] 2016. [Zitat vom: 27. März 2016.] <http://www.edwardrosten.com/work/corner.png>.
21. **Gerhard Hesse, Armin Laubert.** *Hörminderung im Alter - Ausprägung und Lokalisation*. 2005.
22. Wikipedia. *Wikipedia*. [Online] 2016. [Zitat vom: 27. März 2016.] <https://en.wikipedia.org/wiki/Interpolation>.
23. **Borucki, Hans.** *Einführung in die Akustik*. 1980.
24. Wikipedia. *Wikipedia*. [Online] Wikipedia, 2016. [Zitat vom: 25. März 2016.] <https://de.wikipedia.org/wiki/Digitalkamera>.
25. *Ueber den Einfluss des ultravioletten Lichtes auf die electrische Entladung.* **Hertz, Heinrich.** 267, 1887, Bd. *Annalen der Physik* (267).

26. **Herrmann, Klaus H.** *Der Photoeffekt – Grundlagen der Strahlungsmessung.* Braunschweig : Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, 1994.
27. **Wikibooks.** *Digitale bildgebende Verfahren.* s.l. : Wikipedia, 2012.
28. **Billinghurst, Hirokazu Kato and Mark.** *Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System.* Hiroshima & Washington : Hiroshima City University & University of Washington.