

Institut für Visualisierung und Interaktive Systeme
Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3231

Fast-Forward Video Visualization

Videovisualisierungen für den Schnellvorlauf

Kuno Kurzhals

Studiengang: Informatik

Prüfer: Prof. Dr. Daniel Weiskopf

Betreuer: Dipl.-Inf. Markus Höferlin

begonnen am: 11. August 2011

beendet am: 10. Februar 2012

CR-Klassifikation: I.3.3, I.3.m, H.5.1

Inhaltsverzeichnis

1. Einleitung	5
1.1. Allgemeine Motivation	5
1.2. Aufgabenstellung	6
1.3. Gliederung	7
2. Related Work	9
2.1. Methoden des Video-Browsings	9
2.1.1. Statische Repräsentationen	10
2.1.2. Dynamische Repräsentationen	10
2.2. Studien zum Thema Video-Browsing	11
3. Fast-Forward Visualisierungen	13
3.1. Prinzip der Implementierungen	14
3.2. Frameskip	16
3.2.1. Implementierung	16
3.2.2. Vorteile der Frameskip-Methode	16
3.3. Blending	17
3.3.1. Prinzip des biologisch motivierten Blendings	17
3.3.2. Implementierung	18
3.3.3. Vorteile der Blending-Methode	19
3.4. Differenzen	19
3.4.1. Implementierung	19
Hintergrundentfernung	20
Objektmasken verrechnen	20
3.4.2. Vorteile der Differenzen-Methode	23
3.5. Tracking	23
3.5.1. Implementierung	24
Tracking Daten	24
Pfeile und Schweife	25
3.5.2. Vorteile der Tracking-Methode	27

3.6. Videoplayer mit Eingabefunktion	28
4. Geschwindigkeitsvisualisierungen	29
4.1. Tachometer	29
4.2. Rahmen	31
4.3. Störstreifen	32
5. Studie	35
5.1. Hypothesen und Fragen	35
5.1.1. Erkennung	36
5.1.2. Bewegungsverfolgung	36
5.1.3. Geschwindigkeitseindruck	37
5.2. Aufgabenstellungen	37
5.2.1. Variablen	38
5.2.2. Erste Aufgabe - Fremdobjekte erkennen	39
5.2.3. Zweite Aufgabe - Bewegungsverfolgung vergleichen	41
5.2.4. Dritte Aufgabe - Geschwindigkeitsvisualisierungen vergleichen	41
5.3. Durchführung	43
5.3.1. Pilotstudie	43
5.3.2. Benutzerstudie	44
Umgebungsbedingungen	44
Teilnehmer	44
Studienablauf	45
5.4. Ergebnisse	48
5.4.1. Auswertung der ersten Aufgabe	48
F-Scores	48
Signifikanztests	50
Auswertung der Fragebögen	52
Freie Kommentare	56
5.4.2. Auswertung der zweiten Aufgabe	57
5.4.3. Auswertung der dritten Aufgabe	59
Freie Kommentare	60
6. Probleme und Verbesserungen	63
6.1. Fast-Forward Visualisierungen	63
6.2. Geschwindigkeitsvisualisierungen	66
6.3. Probleme der Studie	68
7. Zusammenfassung und Ausblick	69
A. Tabellen und Histogramme	73
Literaturverzeichnis	81

Einleitung

1.1. Allgemeine Motivation

Die Verwendung von CCTV (Closed Circuit Television Cameras) ist inzwischen ein gängiges Mittel zur Prävention von Verbrechen und zur Sicherheitsüberwachung bestimmter Orte[GS05]. In der Vergangenheit wurden dafür häufig Videobänder verwendet, auf denen Aufnahmen im Zeitraffer festgehalten wurden.[Bro95]. Bei zukünftigen Projekten kann man davon ausgehen, dass digitale Aufnahmen die analogen ersetzen werden. Beispiele für CCTV Projekte finden sich bei Gill und Spriggs[GS05]. Die verwendete Anzahl an Kameras variiert dabei von 8 (Project Borough) bis zu 646 (Project London Hawkeye) Stück. Dies soll nur einen kleinen Einblick über die möglichen Größenordnungen solcher Projekte gewähren. Würde jede dieser Kameras zur Aufzeichnung verwendet werden, so wird einem klar, dass sich dabei sehr viel Videomaterial ansammeln kann. Bei Bedarf bieten intelligente Überwachungssysteme der zweiten und dritten Generation Möglichkeiten für eine automatisierte Analyse des Videomaterials mittels Computer Vision Technologie[VV05]. Geht man davon aus, dass es noch einige Zeit braucht, bis vollautomatisierte Systeme robust genug sind, um eine Detektion von gesuchten Inhalten zu gewährleisten, so benötigt man immer noch Menschen, die sich bei Bedarf Teile des Videomaterials ansehen. Da das Überwachungspersonal nicht dauerhaft das Bild jeder Kamera beobachten kann (Beispiel: Nur 15,9% der Zeit innerhalb von 20 Stunden wurden Kamerabilder in Einkaufszentren beobachtet[HF04]), ist es durchaus möglich, dass ein bestimmtes Ereignis während eines Zeitpunktes geschieht, in dem niemand die Bilder der Kamera live überwacht. Kann man den möglichen Zeitraum des Ereignisses einschränken, so ist das zwar hilfreich, aber es können dennoch einige Stunden an Videomaterial übrigbleiben. Dieses Material in Echtzeit zu betrachten wäre sehr langwierig und anstrengend. Vernünftigerweise wird man das Video beschleunigen, wobei im Regelfall ab einer gewissen Beschleunigung Bilder übersprungen werden, wenn das Video digital vorliegt.

Ein Problem beim Überspringen der Bilder ist die Verstärkung der sogenannten „Change Blindness“ [SL97][SA05]. Dies ist ein Effekt der zur Folge hat, dass Menschen unter bestimmten Umständen nicht in der Lage sind, selbst große Veränderungen innerhalb einer Szene zu erkennen. Als bekanntes Beispiel dient das Experiment von Simons und Levin[SL98]. Dabei wurden Passanten angesprochen und nach dem Weg zu einem Gebäude gefragt. Während der Unterhaltung wurde die Sichtlinie zwischen den Personen unterbrochen, indem eine Tür zwischen ihnen vorbeigetragen wurde. Während der Unterbrechung wurde die Person, die nach dem Weg gefragt hatte, durch eine andere Person ersetzt. Nur die Hälfte der Versuchspersonen bemerkte diesen Unterschied.

Das von Rensink et al.[ROC97] „Flicker Paradigm“ beschreibt ein Experiment, in dem den Versuchsteilnehmern ein Bild gezeigt wurde, das sukzessiv von grauen Feldern unterbrochen wurde. An einer bestimmten Stelle wurde dann eine Änderung in dem Bild vorgenommen, die erkannt und beschrieben werden sollte. Dabei stellte sich heraus, dass es durch die Unterbrechungen wesentlich länger brauchte, die Änderungen im Bild wahrzunehmen, als ohne.

In der Praxis spielt dieser Effekt bei der Unterbrechung von Bewegung, durch das Überspringen von Bildern eine wichtige Rolle. So kann beispielsweise kriminelles Verhalten bei Zeitrafferaufnahmen wesentlich schlechter wahrgenommen werden, als bei einer normalen Abspielgeschwindigkeit[SBC07].

In dieser Arbeit werden Visualisierungen vorgestellt, die diese übersprungenen Bildinformationen in die sichtbaren Bilder während der Beschleunigung wieder einfügen. In einer anschließenden Benutzerstudie soll dann herausgefunden werden, wie sich diese Visualisierungen in der Effektivität der Erkennung von Objekten, der Verfolgung von Bewegungen und ihrer allgemeinen Wirkung auf die Benutzer unterscheiden.

1.2. Aufgabenstellung

Nachdem die Motivation geklärt wurde, weshalb ein beschleunigtes Abspielen von Überwachungsvideos erwünscht ist, soll nun die Aufgabenstellung dieser Arbeit näher erläutert werden.

Diese lässt sich in 3 Hauptpunkte unterteilen:

1. **Entwicklung und Umsetzung von Methoden:** Um Videos zu beschleunigen, existieren bereits vorhandene Methoden. Eine Untersuchung der Methoden bezüglich ihrer Verwendbarkeit im Rahmen einer Fast-Forward Visualisierung dient dabei als Ausgangspunkt. Die Entwicklung und Implementierung eigener Methoden folgt darauf. Hierbei soll darauf geachtet werden, dass visuelle Informationen durch die Methoden erhalten bleiben, und dem Benutzer in unterstützender Weise präsentiert werden.

2. **Umsetzung von Methoden zur Visualisierung von wechselnden Geschwindigkeiten:** Durch den Einsatz von adaptivem Fast-Forward [HHWH10] können Videos, je nach Aktivität innerhalb der Szene, während des Abspielens unterschiedlich stark beschleunigt werden. Die Implementierung von Methoden zur besseren Visualisierung dieser wechselnden Geschwindigkeiten soll in dieser Arbeit ebenfalls durchgeführt werden.
3. **Durchführung einer Benutzerstudie:** Nachdem eine Auswahl geeigneter Fast-Forward Visualisierungen getroffen wurde, sollen diese im Rahmen einer Benutzerstudie hinsichtlich ihrer Effektivität bei der Erkennung von Objekten und ihres Eindrucks auf die Teilnehmer analysiert werden. Die Eindrücke der Teilnehmer hinsichtlich der Geschwindigkeitsvisualisierungen sollen ebenfalls in einer eigenen Aufgabe der Benutzerstudie erfasst und untersucht werden.

Mithilfe der implementierten Funktionen werden Videos der einzelnen Visualisierungen erstellt, welche dann für die Aufgaben der Benutzerstudie verwendet werden können. Die Implementierung wurde mit C++, Qt und OpenCV durchgeführt.

1.3. Gliederung

Diese Arbeit ist in folgender Weise gegliedert:

Kapitel 1 - Einleitung: Dieses Kapitel bietet eine Einführung in das Thema der Arbeit. Durch eine Erläuterung der allgemeinen Motivation und eine Beschreibung der Aufgabenstellung soll für den Leser der Zusammenhang der nachfolgenden Kapitel nachvollziehbar werden.

Kapitel 2 - Related Work: Ein Überblick, welche Arbeiten mit ähnlicher Thematik bereits veröffentlicht wurden, wird durch dieses Kapitel geboten. Um den Inhalt eines Videos schneller erfassen zu können, ist es möglich, statische oder dynamische Methoden zu verwenden. Die Unterschiede dieser Methoden werden geklärt und ihre Verwendbarkeit zur Lösung der Aufgabenstellung erläutert.

Kapitel 3 - Fast-Forward Visualisierungen: In diesem Kapitel geht es um die Implementierung unterschiedlicher Methoden zum beschleunigten Abspielen von Videos. Alle verwendeten Methoden unterliegen bei der Implementierung einem Grundprinzip, welches als erstes genauer erläutert wird. Mit dem Frameskip und dem biologisch motivierten Blending werden dann zunächst zwei bereits existierende Methoden vorgestellt. Bei der Differenzen- und der Tracking-Methode handelt es sich um eigene Implementierungen, die im Rahmen dieser Arbeit entwickelt wurden. Für die Verwendung der Methoden in einer Benutzerstudie wurde zusätzlich ein Videoplayer mit speziellen Eigenschaften implementiert, die auf diesen individuellen Fall angepasst wurden. Die Vorstellung des Videoplayers bildet den Abschluss dieses Kapitels.

Kapitel 4 - Geschwindigkeitsvisualisierungen: Durch die Verwendung von adaptivem Fast-Forward können Videos mit variierender Geschwindigkeit wiedergegeben werden. In diesem Kapitel werden Methoden beschrieben, die implementiert wurden, um dem Benutzer den Eindruck dieser Geschwindigkeiten besser zu vermitteln. Bei den drei vorgestellten Methoden handelt es sich um einen Tachometer, einen Rahmen und Störstreifen.

Kapitel 5 - Studie: Nach der Implementierung der Methoden werden diese in einer Benutzerstudie miteinander verglichen. Dieses Kapitel handelt von der Planung, Vorbereitung und Durchführung dieser Studie. Zunächst werden die Hypothesen und Fragen erläutert, deren Überprüfung das Ziel der Studie sein soll. Darauf basierend werden Aufgaben formuliert, die der Erkennung von Objekten, Beurteilung der Bewegungsverfolgung und dem Vergleich der Geschwindigkeitsvisualisierungen dienen. Die Ergebnisse der Studie werden am Ende dieses Kapitels ausgewertet.

Kapitel 6 - Probleme und Verbesserungen: Nach der Auswertung der Studie und dem Sammeln neuer Erkenntnisse beschäftigt sich dieses Kapitel mit den Problemen, die im Verlauf der Arbeit zutage getreten sind. Dies bezieht sich insbesondere auf die Probleme der Studienteilnehmer mit den einzelnen Visualisierungen und mögliche Verbesserungen.

Kapitel 7 - Zusammenfassung und Ausblick: Abschließend werden die Ergebnisse der Arbeit in diesem Kapitel zusammengefasst.

Related Work

Im Bereich des „Video-Browsing“ gibt es zahlreiche Publikationen, die Methoden beschreiben, wie man den Inhalt eines Videos schneller als in der Originalzeit durchsehen kann.

2.1. Methoden des Video-Browsers

Um den Inhalt eines Videos schneller erfassen zu können, wird der Inhalt in den meisten Methoden durch ausgewählte Stellvertreter, den „Surrogates“, repräsentiert. Li et al.[LZT01] und Tse et al.[TMD⁺98] geben in ihren Arbeiten einen guten Überblick über verschiedene Surrogates und deren Verwendung. Dabei kann man bei der Repräsentation zwischen statischen und dynamischen Methoden unterscheiden[TMD⁺98]:

- **Statisch:**
 - Filmstreifen / Storyboards
 - Gelabelte Thumbnails
 - Hierarchisch angeordnete Surrogates
- **Dynamisch:**
 - Zeitlich beschleunigte Surrogates
 - Mehrfach parallele Surrogates
 - Slide Show

2.1.1. Statische Repräsentationen

Im Bereich der statischen Repräsentationen gibt es besonders im Bereich der Filmstreifen und Storyboards zahlreiche Veröffentlichungen, die sich mit der Auswahl und Anordnung möglichst informationsträchtiger Keyframes beschäftigen. Die Anordnung der Keyframes kann beispielsweise wie bei Uchihashi et al.[UFGB99], Yeung und Yeo[YY97], Hwang et al.[HLC⁺06] oder Calic und Campbell[CC07] ähnlich wie in einem Comic geschehen. Dabei lassen sich die Bilder auch mit zusätzlichen Informationen wie Richtungspfeilen versehen, wie es bei den Schematic Storyboards von Goldman et al.[GCSS06] der Fall ist.

Gelabelte Thumbnails finden sich beispielsweise auf der Seite von „The Open Video Project“[ope12]. Die zur Verfügung stehenden Videos in dieser Datenbank sind in verschiedene Kategorien unterteilt und lassen sich nach Genre, Dauer, Farbe / Schwarz-Weiss, mit /ohne Ton unterteilen. Jedes Video innerhalb dieser Unterteilung wird durch ein kleines Bild aus dem Video repräsentiert. Zusätzlich lassen sich auf dieser Seite auch Zusammenfassungen der Videos durch Storyboards und Fast-Forward Beschleunigung ansehen.

Für hierarchisch angeordnete Surrogates müssen Videoszenen oder Keyframes in Klassen unterteilt werden, die den gleichen oder ähnlichen Inhalt besitzen. Jede Klasse besitzt einen gemeinsamen Repräsentanten, der diesen Inhalt widerspiegeln soll. [ZZC96]

2.1.2. Dynamische Repräsentationen

Die Fast-Forward Beschleunigung von Videos findet sich in zahlreichen Mediaplayern wie dem „Media Player Classic“[med11] oder dem „VLC media player“[vlc11]. Hierbei handelt es sich auch um den Bereich, der im Verlauf dieser Arbeit genauer untersucht werden soll.

Slaughters et al.[SSM97] zeigten in ihrer Arbeit mehrere Slide Shows parallel. Slide Shows werden auch von Ding et al.[DMT97] verwendet. Dabei werden Keyframes wie bei einem Diavortrag hintereinander angezeigt. Die Abspielgeschwindigkeit kann hierbei variiert werden. Das in dieser Arbeit verwendete Frameskip-Verfahren (siehe Abschnitt 3.2) arbeitet sehr ähnlich. Wobei die Abspielgeschwindigkeit jedoch nicht vom Benutzer variiert werden kann, sondern der Framerate des Eingangsvideos entspricht. Die dabei verwendeten Keyframes können in uniformen Abständen, oder mittels Informationen des Adaptive Fast Forward von Höferlin et al.[HHWH10] ausgewählt werden (siehe Kapitel 4).

2.2. Studien zum Thema Video-Browsing

In den Benutzerstudien von Drucker et al.[DGDMW02], Schoeffmann und Boeszoermyi [SB09] und Wildemuth et al.[WMW⁺02][WMY⁺03] wird Fast-Forward mit verschiedenen Methoden der Autoren verglichen. In ähnlicher Weise sollen die implementierten Methoden dieser Arbeit mit der gängigen Methode für Fast-Forward in Videos und dem biologisch motivierten Blending von Höferlin et al.[HHWH10] verglichen werden.

In der Studie von Wildemuth et al.[WMY⁺03] wurde Fast-Forward dadurch realisiert, dass jedes n-te Bild aus dem Originalvideo verwendet wurde. Dann wurde das Verfahren und weitere Surrogates in Hinblick auf Objekterkennung, Aktionserkennung und Zusammenfassung des Inhalts der gezeigten Szenen untersucht. Um die Objekterkennung zu prüfen, wurden die Benutzer gebeten, gesehene Objekte aus Wörter- und Bildlisten herauszusuchen. Bei der Benutzerstudie, die in dieser Arbeit durchgeführt wurde, wird ebenfalls die Erkennung von Objekten geprüft werden. Aufgrund des monotonen Videomaterials wurde allerdings auf den Einsatz solcher Listen verzichtet. Es wurde stattdessen eine interaktive Methode verwendet, die dem Benutzer die Möglichkeit bietet, ein erkanntes Fremdobjekt direkt zu melden.

Als Ergebnis der Studie wurde empfohlen, mindestens $\frac{1}{64}$ der Originalbilder für Fast-Forward zu verwenden, um die getesteten Aufgaben noch mit einer Leistung zu erfüllen. Im Verlauf dieser Arbeit wird die größte Beschleunigung $\frac{1}{20}$ der Bilder bei 25 Bildern pro Sekunde verwendet.

Die Studie von Tse et al.[TVSM99] wurden Slideshows bei unterschiedlichen Abspielraten untersucht. Die Benutzer mussten für die Objekterkennung ebenfalls Listen durchsehen, auf denen Objekte gezeigt wurden, von denen die rausgesucht werden mussten, welche im Video vorkamen. Als Resultat der Studie wurde die Möglichkeit eingeräumt, Slideshows bei 8 Bildern pro Sekunde und mit jedem 150-ten Bild als Keyframe zu verwenden. Bei den getesteten Videos handelte es sich jedoch nicht um Überwachungsvideos. Aus diesem Grund dienen diese Ergebnisse ebenfalls als Orientierung, um nachgewiesene Schranken der Beschleunigung nicht zu übertreten.

Viele dieser Methoden verwenden Algorithmen, die bestimmen, welche Ausschnitte eines Videos von Bedeutung sind. Überwachungsvideos lassen sich auf diese Weise durchaus auch bearbeiten. Das Videomaterial einer Bahnhofshalle ist im Vergleich zu den meisten Hollywoodfilmen jedoch relativ eintönig. Deshalb sollte es meistens schwer fallen eine Unterteilung des Videos in Dialoge, Explosionen oder ähnliches zu treffen, wie es in der Arbeit von Lienhart et al.[LPE97] bei Filmen gemacht wurde. Wenn man nicht genau weiss, wonach man sucht, dann bleibt trotz aller Technik manchmal nur noch das Durchsehen des Videos ohne eine automatische Voranalyse des Inhalts.

2. Related Work

Diese Arbeit beschäftigt sich nun mit der Frage, ob es möglich ist, diese unangenehme Aufgabe für den Benutzer zu erleichtern und eventuell auch eine Steigerung der Effektivität bei einer Suchaufgabe zu erreichen.

Fast-Forward Visualisierungen

Um digitale Videos schneller betrachten zu können, wurden verschiedene Methoden eingesetzt. Dabei handelt es sich beim Frameskip um das üblichste Verfahren, welches sich auch technisch am einfachsten umsetzen lässt. Das Blending erzeugt Bilder mit Bewegungsunschärfe. Die Differenzen lassen Objekte mehrfach erscheinen. Beim Tracking werden mittels Pfeilen und hellen Schweifen an den Objekten, vergangene und zukünftige Bewegungsrichtungen angezeigt. Dieses Kapitel handelt von der Erklärung dieser Methoden und der technischen Umsetzung.

Die Einsatzmöglichkeiten der einzelnen Methoden lassen sich wie in Abbildung 3.1 einteilen.

Frameskip lässt sich hierbei auf beliebige Videos anwenden. Dabei spielt es keine Rolle, ob es sich um Filme, Musikvideos oder ähnliches handelt. Blending lässt sich ebenfalls ohne größere Einschränkungen verwenden. Viele Kameranews würden das gesamte Bild jedoch unscharf machen und somit zu unübersichtlichen Situationen führen. Da für die Differenzen-Methode ein statisches Hintergrundbild berechnet werden muss, sind diese nur für Videos einsetzbar, in denen die Kamera fixiert wurde, und dauerhaft denselben Bereich filmt. Jeder Kameranews oder Szenenwechsel bedarf einer Neuberechnung des Hintergrundbildes, was bei häufigem Wechsel keinen Sinn mehr ergeben würde. Die stärkste Einschränkung besitzt das Tracking, da hierfür noch eine genaue Erkennung der Objekte innerhalb der Szene nötig wird.

Um die Visualisierungen untereinander vergleichbar zu machen, muss deshalb mit Videos gearbeitet werden, die von allen Visualisierungen sinnvoll angezeigt werden können. Deshalb wurden in dieser Arbeit Videos mit fixierten Kamreas und vorhandenen Trackinginformationen verwendet [avs12][beh12].

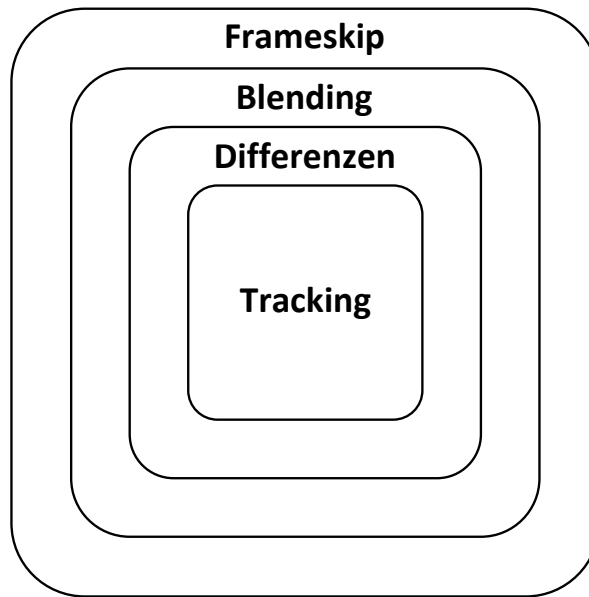


Abbildung 3.1.: Einsatzgebiete der Methoden

3.1. Prinzip der Implementierungen

Das Grundprinzip bei der Implementierung der einzelnen Fast-Forward Visualisierungen wurde von Höferlin et al.[HHWH10] übernommen. Dabei sind folgende Werte von Interesse:

- D_M Anzeigedauer des momentanen Bildes
- D_S Summierte Dauer der eingegangenen Bilder
- Sw Schwellenwert zur Anzeige des Bildes
- Fps Framerate des Videos
- Bf Beschleunigungsfaktor des Videos

Bei gegebener Framerate des Videos und einem Beschleunigungsfaktor, lässt sich die genaue Anzeigedauer eines Bildes in Millisekunden bestimmen(3.1):

$$D_M = \frac{1000}{Fps * Bf} \quad (3.1)$$

Die Anzeigedauer wird an die Renderfunktion der entsprechenden Visualisierung übergeben und dort aufsummiert(3.2).

$$D_S = \sum_i D_{M_i} \quad (3.2)$$

Entsprechend dem verwendeten Videomaterial wird eine Framerate von 25 Bildern pro Sekunde verwendet. Dies bedeutet, bei normaler Abspielgeschwindigkeit, wird alle 40 ms ein Bild angezeigt. Diese 40 ms bilden den Schwellenwert Sw , bei dessen Überschreitung ein Ausgabebild erzeugt wird. Anschließend wird Sw so oft von D_S abgezogen und ein Bild ausgegeben, bis $D_S < Sw$ ist. Auf diese Weise lassen sich nicht nur Beschleunigungen, sondern auch Verlangsamungen des Videos in einer Funktion umsetzen.

Die gerenderten Videos der einzelnen Methoden werden eine Framerate von 25 Bildern pro Sekunde erhalten. Dadurch wird es nötig, alle 40 ms ein Bild anzuzeigen. Diese 40 ms bilden also den Schwellenwert. Wird dieser beim Aufsummieren erreicht, so wird der Schwellenwert von der Summe abgezogen und das momentane Bild angezeigt. Dies geschieht solange, bis der Schwellenwert wieder unterschritten wird. Auf diese Weise lassen sich nicht nur Beschleunigungen, sondern auch Verlangsamungen in einer Funktion umsetzen. Diese Vorgehensweise bildet die Grundlage für alle Visualisierungen des Schnellvorlaufs, die in dieser Arbeit umgesetzt wurden.

Wenn nach dem Überschreiten des Schwellenwertes ein Bild ausgegeben wird, so soll dies Informationen über alle Bilder des vergangenen Zeitintervalls beinhalten. Es werden Möglichkeiten gesucht, die den Eindruck der Bewegung erhalten und das Objekt, die Bewegungsrichtung und die Bewegungsgeschwindigkeit erkennen lassen[Cuto2]. Abgesehen von der Frameskip-Methode erfüllen die drei vorgestellten Methoden teilweise oder sogar ganz diese Forderungen:

Methode	Erkennbarkeit	Bewegungsrichtung	Bewegungsgeschwindigkeit
Blending	-	+	-/+
Differenzen	+	+	-/+
Tracking	+	+	+

Tabelle 3.1.: Vergleich der vorgestellten Methoden

Die folgenden Abschnitte stellen nun die einzelnen Methoden vor und erläutern genauer, inwiefern sie diese Forderungen erfüllen.

3.2. Frameskip

Die Abspielgeschwindigkeit von digitalem Videomaterial befindet sich in den häufigsten Fällen zwischen 24-30 Bildern pro Sekunde. Eine Beschleunigung beim Abspielen kann erreicht werden, indem man das Video mit 50-60 Bildern pro Sekunde abspielt. Bei vielen Monitoren und Fernsehern liegt bei dieser Framerate die technische Grenze dieser Möglichkeit. Um höhere Geschwindigkeiten im Schnellvorlauf erreichen zu können, müssen Bilder übersprungen werden. In bekannten Videoplayern wie dem „VLC media player“ [vlc11] und dem „Media Player Classic“ [med11] lässt sich diese Methode beobachten. Bei Wildemuth et al.[WMY⁺03] wird der Begriff „Nth frame fast forward“ dafür verwendet. Da diese Methode als das allgemein übliche Vorgehen für den digitalen Schnellvorlauf in Videos anzusehen ist, darf es bei dem Vergleich mit neuen Methoden nicht übergangen werden. Aufgrund des Überspringens von Bildern, wird diese Methode hier als „Frameskip“ bezeichnet.

3.2.1. Implementierung

Die Implementierung des Frameskip nach dem vorab beschriebenen Grundprinzip ist denkbar einfach. Die Zeiten der einzelnen Bilder werden aufsummiert und wenn der Grenzwert erreicht ist, wird das momentane Bild ausgegeben. Es bleibt anzumerken, dass der Frameskip zum Beispiel durch die direkte Auswahl der zu zeigenden Bilder wesentlich schneller umgesetzt werden kann. Die hier verwendete Implementierung dient als das einfachste Beispiel, um das Verarbeitungsprinzip der einzelnen Bilder zu verdeutlichen. Durch diese Implementierung lassen sich auch die Gewichtungen der einzelnen Frames für das Adaptive Fast Forward in Aufgabe 3 (siehe Abschnitt 5.2.4) direkt verwenden.

3.2.2. Vorteile der Frameskip-Methode

Folgende Punkte sprechen für eine Verwendung der Frameskip-Methode:

- Die Frameskip-Methode reduziert lediglich die Anzahl der gezeigten Bilder. Bei geringeren Geschwindigkeiten können so Bilder übersprungen werden, ohne dass der Benutzer diese Sprünge im Video stark wahrnimmt.
- Es werden nur Bilder aus dem Originalvideo gezeigt. Da die Bilder nicht verarbeitet werden, muss der Benutzer sich nicht an fremde oder zusätzliche Informationen gewöhnen und wird nicht davon abgelenkt.
- Der entscheidende Vorteil des Frameskip ist, dass er sich auf jedes beliebige Video anwenden lässt. Ob es sich bei dem Video nun um einen Spielfilm, Werbespots oder Überwachungsaufnahmen handelt, spielt keine große Rolle.

3.3. Blending

Das biologisch motivierte Blending erstellt beim Schnellvorlauf Bilder, die im Grunde einer Langzeitbelichtung in der Fotografie entsprechen. Dabei lässt man die Kamera über einen längeren Zeitraum Licht aufnehmen, was vor allem bei schlechten Lichtverhältnissen zu einem helleren Bild führt. Wird während der Aufnahme eine Bewegung vor der Kamera ausgeführt, so wird diese in Form von Bewegungsschlieren auf dem Bild sichtbar. Das Blending erzeugt ebenfalls solche Bilder, in denen diese Bewegungsunschärfe deutlich zu sehen ist.

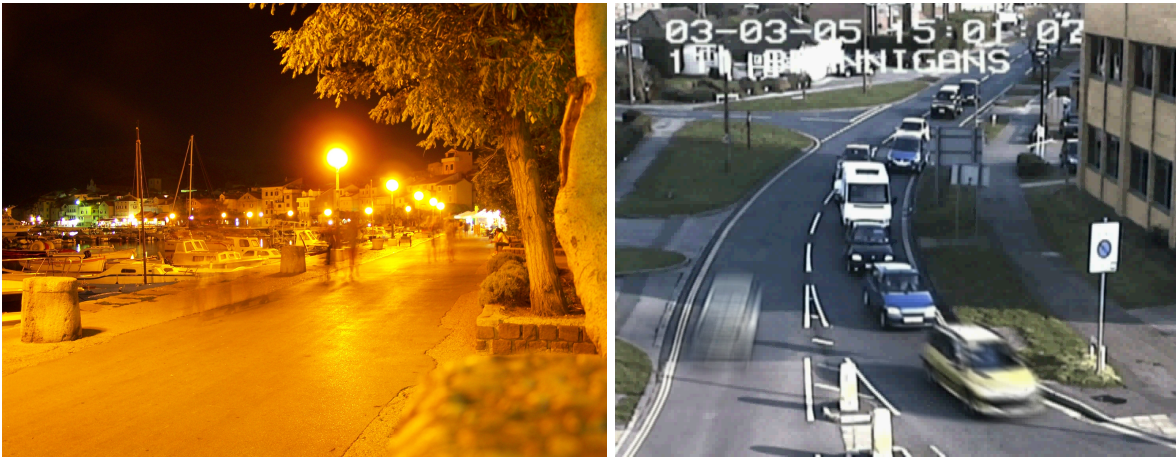


Abbildung 3.2.: links: Langzeitbelichtung mit einer Digitalkamera bei Nacht, rechts: Blending. Die Bewegungsunschärfe nimmt je nach Geschwindigkeit zu.

Das hier verwendete, biologisch motivierte Blending entspricht dem von Höferlin et al. [HHWH10] beschriebenen Verfahren. Hierbei besteht die Möglichkeit, die einzelnen Bilder mit einer unterschiedlichen Gewichtung zu verrechnen. Für eine konstante Beschleunigung handelt es sich dabei auch um einen einheitlichen Wert. Im Falle einer adaptiven Fast-Forward Visualisierung können dadurch die Prioritäten der einzelnen Bilder berücksichtigt werden.

Weshalb es sich hier um ein biologisch motiviertes Blending handelt, soll nun als nächstes geklärt werden.

3.3.1. Prinzip des biologisch motivierten Blendings

Es handelt sich bei dem verwendeten Verfahren um ein biologisch motiviertes Blending, da, wie bei dem biologischen Vorgang des Sehens, eine Integration der Reize vorgenommen wird. Da Aufnahmen einer Kamera aus technischen Gründen eine nichtlineare Gammakorrektur enthalten, aber die menschliche Wahrnehmung lineare RGB Signale benötigt, werden die

3. Fast-Forward Visualisierungen

verwendeten Bilder zunächst in dieses Format überführt. Im linearen RGB Format werden die Bilder dann zusammengerechnet und anschließend wieder eine Gammakorrektur vorgenommen, so dass die Bilder von dem entsprechenden Anzeigegerät wie bisher behandelt werden können (siehe Abbildung 3.3).

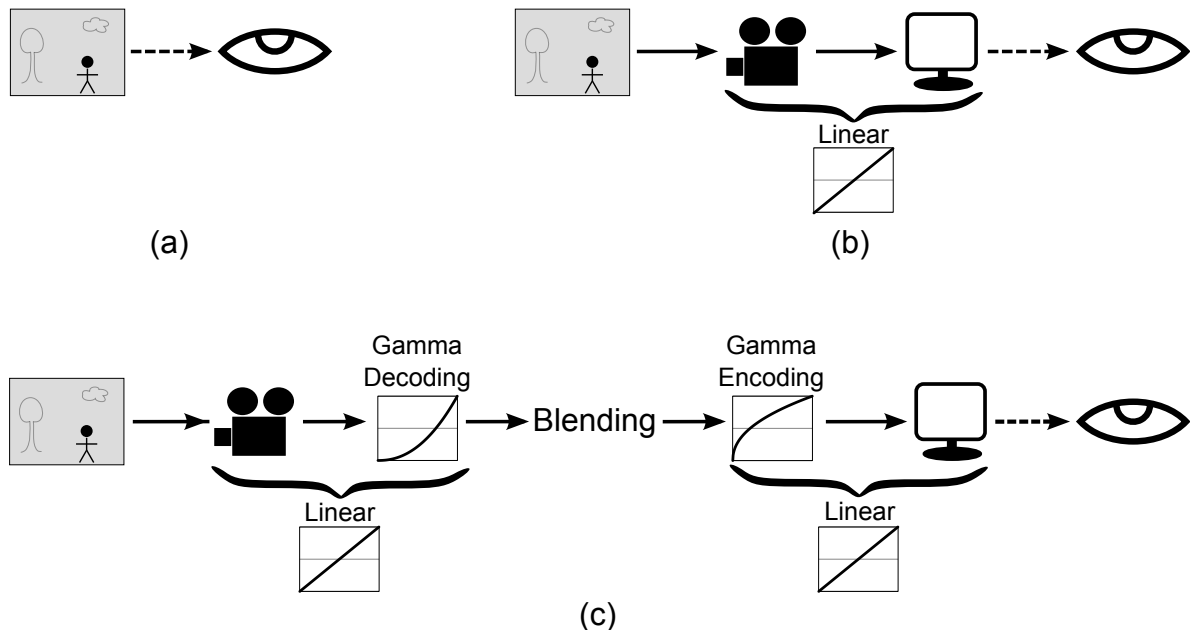


Abbildung 3.3.: Verarbeitung der Bilder [HHWH10]: a) Visuelle Stimuli erreichen das Auge unverändert. b) Die Bilder einer Kamera sind gammakodiert. Der Monitor macht die nicht-lineare Transformation rückgängig. c) Künstliche Integration von visuellen Stimuli, dem menschlichen Sehen nachempfunden. Vor dem Blending wird das Bild in das lineare RGB-Format umgewandelt und anschließend wieder gammakodiert.

3.3.2. Implementierung

Im Vergleich zum Frameskip werden beim Blending keine Bilder übersprungen. Das Prinzip des Aufsummierens der Anzeigedauer einzelner Bilder innerhalb eines bestimmten Intervalls findet auch in dieser Visualisierung wieder Verwendung. Hierbei wird jedoch nicht nur die Zeitinformation der einzelnen Bilder berücksichtigt. Gleichzeitig wird jedes einzelne Eingabebild in die Berechnung des nächsten Ausgabebildes miteinbezogen. In dieser Arbeit wurde die bereits vorhandene Implementierung des biologisch motivierten Blendings aus der Arbeit von Höferlin et al. [HHWH10] verwendet. Darin werden die einzelnen Bilder mit Berücksichtigung ihrer Anzeigedauer nach der bereits beschriebenen Methode verrechnet.

3.3.3. Vorteile der Blending-Methode

Für eine Anwendung der Blending-Methode sprechen folgende Vorteile:

- Bei dieser Methode werden die Informationen aus allen Bildern des Originalvideos verwendet und komprimiert in einer geringeren Anzahl von Bildern dargestellt.
- Schnelle Objekte erzeugen eine Bewegungsunschärfe, durch die der Benutzer das Objekt leichter verfolgen verfolgen kann.
- Durch seinen biologisch motivierten Ansatz entspricht Blending unter den verwendeten Methoden am ehesten der natürlichen Wahrnehmung von schnellen Bewegungen.

3.4. Differenzen

Die Visualisierung mittels Differenzbildern soll dem Unschärfeproblem des Blendings entgegenwirken. Die Informationen aus übersprungenen Bildern sollen auch hier durch einen Bewegungsverlauf angezeigt werden, dabei aber das bewegte Objekt gut erkennbar lassen. Das Ergebnis dieser Visualisierung ähnelt dem stroboskopischen Effekt, der beispielsweise von Agarwala et al.[ADA⁺04] für Bilder und von Rav-Acha et al.[RAPP06][PRAP08] für Videos präsentiert wurde. Es werden dabei mehrere Instanzen eines Objekts in Bewegung an verschiedenen Positionen angezeigt. Im Schnellvorlauf hat dies den Vorteil, dass ein Objekt nicht wie beim Frameskip nur einmal an einer Stelle zu sehen ist, sondern einzelne Positionen in mehreren Bildern angezeigt werden. In der hier vorgestellten Visualisierung werden diese alten Positionen durch transparente Instanzen der Objekte dargestellt. Auf diese Weise lässt sich klarer erkennen, wo sich die aktuelle Position des Objektes befindet. Je transparenter eine Instanz dabei ist, desto länger ist es her, dass sich das Objekt an dieser Stelle befunden hat. Eine Effekt, wie er auch in der Arbeit von Teodosio und Bender [TB93] zu sehen ist.

3.4.1. Implementierung

Die Implementierung setzt sich aus zwei Schritten zusammen:

1. Vordergrund und Hintergrund voneinander trennen.
2. Vordergrundinformationen für das Rendering verrechnen.

Die errechneten Bilder erzeugen dann ein neues, beschleunigtes Video, in dem die Objekte geisterhafte Spuren hinterlassen und somit mehrfach im Bild zu sehen sind. Einige alternative Methoden zur Verrechnung der Einzelbilder wurden während der Implementierung ebenfalls getestet, diese finden in diesem Abschnitt ebenfalls Erwähnung.

Hintergrundentfernung

Grundlage dieser Visualisierung ist eine Trennung der bewegten Objekte vom statischen Hintergrund. Eine einfache Methode dies zu erreichen, ist die Verwendung eines Hintergrundmodells. Dabei handelt es sich um ein Bild, das nur den statischen Hintergrund der Szene enthält. Zur Initialisierung wird ein entsprechendes, von Hand erstelltes Bild geladen. Um das Hintergrundmodell anpassen zu können, wird es mit der Formel des Running Gaussian Average (Gleichung (3.3)) [Pico4] aktualisiert.

$$\mu_t = \alpha \cdot I_t + (1 - \alpha) \cdot \mu_{t-1} \quad (3.3)$$

Dabei wird der Wert eines Pixels I_t zum Zeitpunkt t mit dem Gewicht α zu dem berechneten Wert μ_{t-1} aus dem vorherigen Schritt gerechnet. Dieser Schritt lässt sich auf alle Pixel des Bildes anwenden. Auf diese Weise können beispielsweise Änderungen der Lichtverhältnisse oder neue unbewegte Objekte in den Hintergrund aufgenommen werden. Bei der Differenzen-Methode wurde $\alpha = 0.01$ gewählt. Dadurch findet eine langsame Anpassung des Hintergrundes statt, was bei den verwendeten Videos geeignet erschien.

Das Ziel der Trennung von Vorder- und Hintergrund soll eine Maske der Silhouetten aller Objekte in Bewegung sein. Für jedes Videobild wird deshalb der Differenzbetrag mit dem Hintergrundmodell erstellt. Dies ergibt pixelgenaue Unterschiede der Bilder. Da es bereits aufgrund der Komprimierung von Videos zu kaum sichtbaren Unterschieden in den einzelnen Bildern kommen kann, sind die resultierenden Bilder von einem Rauschen durchzogen und noch nicht für die Visualisierung verwendbar. Störpixel, die aufgrund sehr kleiner Unterschiede entstanden sind, können durch einen geeigneten Grenzwert entfernt werden. In einem ersten Verarbeitungsschritt werden deshalb alle Pixel mittels einer Binarisierung entfernt, deren Grauwert unter 31% liegt. Dieser Wert hat sich für die verwendeten Videos als geeignet erwiesen.

An dieser Stelle lässt sich bereits erkennen, dass nur noch die gewünschten Objekte von der Maske erfasst werden. In den vorangegangenen Schritten wurden allerdings auch Teile innerhalb der Objekte entfernt, die eigentlich auch zur Maske gehören. Mithilfe der morphologischen Operatoren Erosion und Dillatation kann dieses Problem verringert werden.

Objektmasken verrechnen

Um nun mehrere transparente Instanzen der Objekte anzeigen zu können, genügt es, immer die letzten beiden Bilder, die für ein Überschreiten des minimalen Zeitintegrals gesorgt haben, übereinander zu blenden. Dadurch wird allerdings auch das Objekt an der akuten Position transparenter und es kann zu Überlappungen kommen. Dies lässt

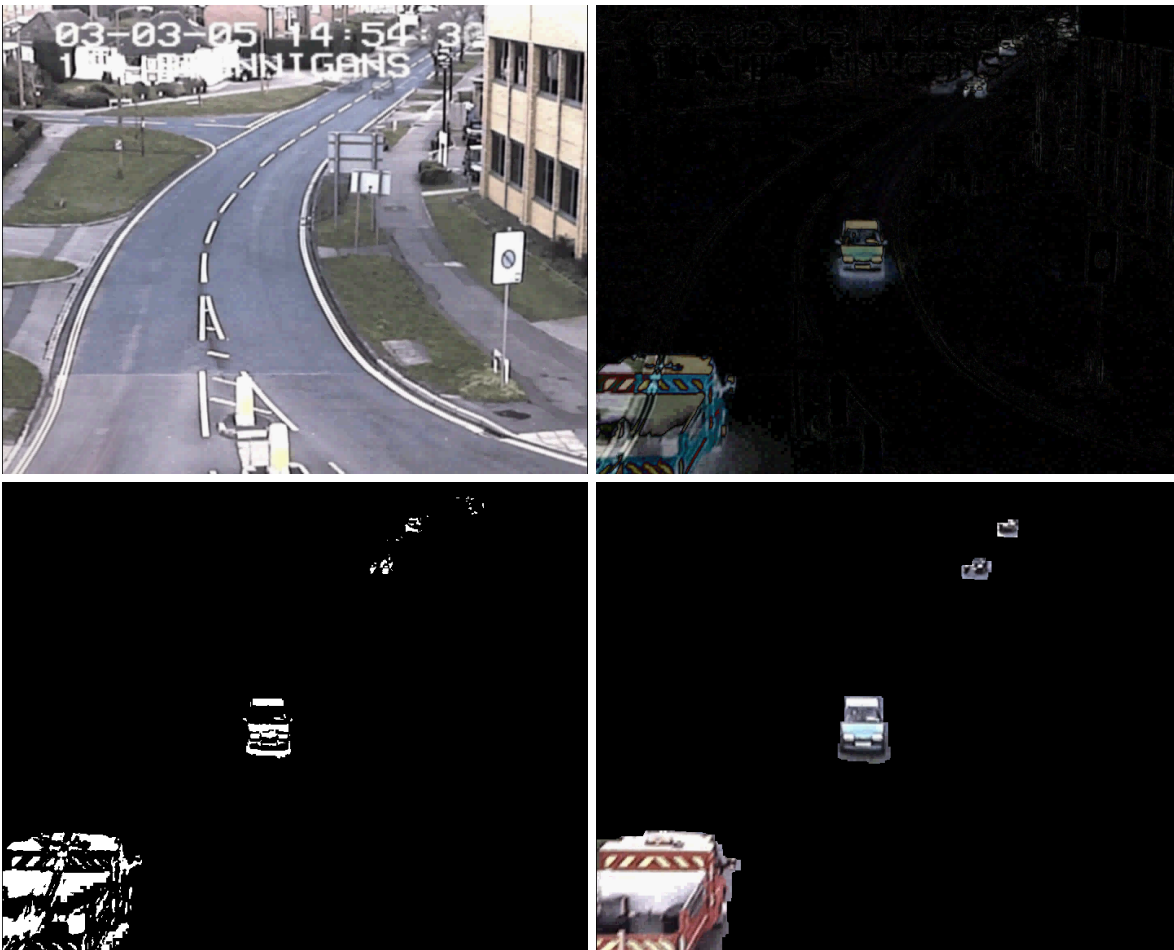


Abbildung 3.4.: Verarbeitungsschritte zum Separieren der Objekte vom Hintergrund: (links-oben: Das momentane Hintergrundmodell, rechts-oben: Differenzbild zwischen Hintergrundmodell und aktuellem Bild, links-unten: Binarisierte Maske, rechts-unten: Objektmaske nach der Anwendung von Erosion/Dilatation

die Objekte bereits schon besser erkennen, als beim Blending. Das Ziel ist allerdings, dass die aktuellen Objekte vollständig opak werden, und über allen anderen Instanzen liegen.

Zur Verrechnung der Objektmasken werden diese zunächst gesammelt, bis die benötigte Anzahl zusammengekommen ist, um ein neues Bild zu generieren. Jedes Differenzbild wird dann auf das Ergebnis der letzten Integralverarbeitung addiert. Der Quotient aus Bilddauer und minimalem Zeitintegral bestimmt dabei den Faktor, mit dem das Differenzbild davor multipliziert wird. Insgesamt führt dieses Vorgehen zu einer ähnlichen Unschärfe wie beim Blending. Das Ergebnis wird hierbei jedoch aufgehellt, was eine Hervorhebung der Objekte

3. Fast-Forward Visualisierungen

bewirken soll.

Der wichtigste Schritt besteht nun darin, das Differenzbild der aktuellen Positionen in das bisher berechnete Ergebnis zu kopieren. Dadurch werden die Objekte wieder annähernd in den Zustand vor der Bearbeitung versetzt und sind so gut erkennbar wie zu Beginn. Durch Artefakte in den Differenzbildern kann es zu unschönen Stellen innerhalb der Objekte kommen. Zur Verringerung des Effektes wurde das Bild noch ein Mal mit dem aktuellen Ursprungsbild übereinandergeblendet.

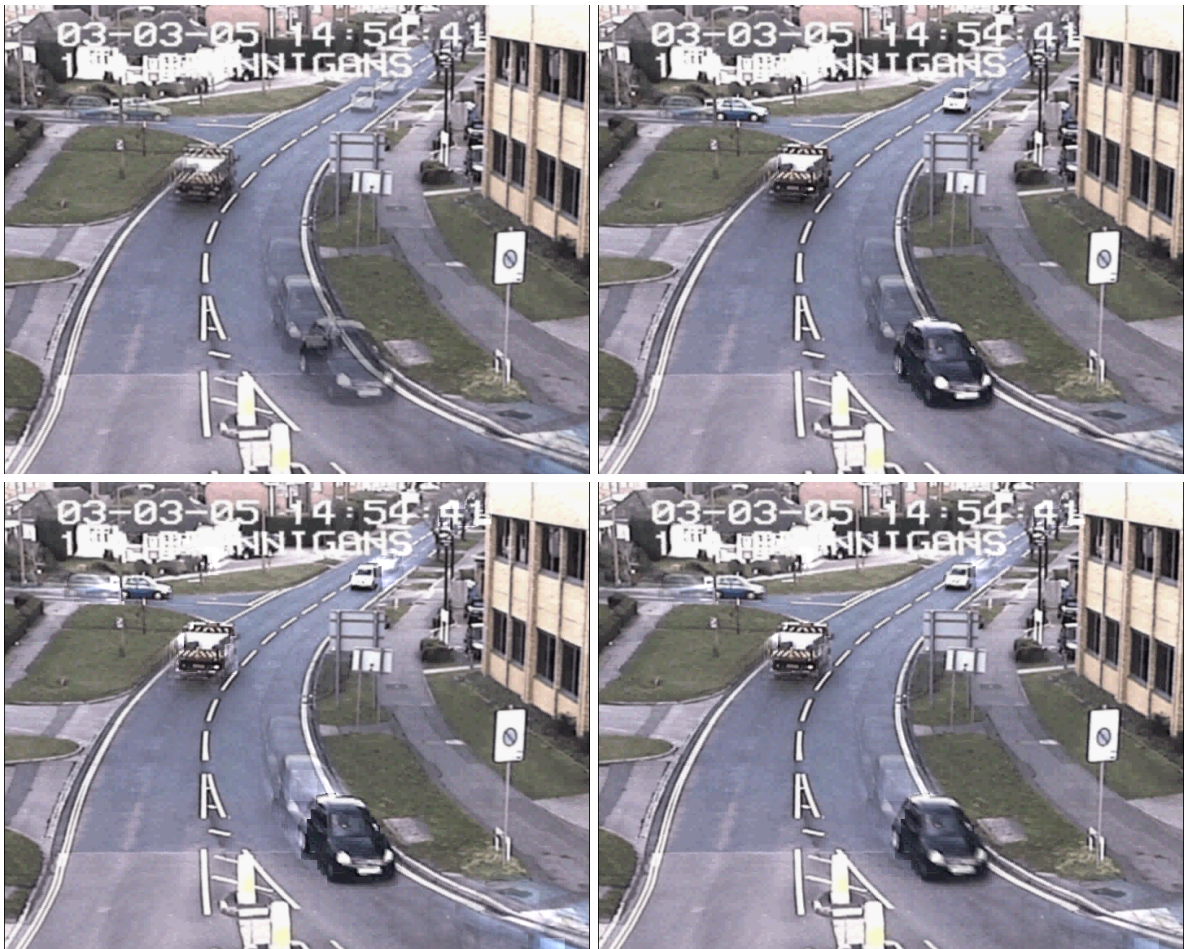


Abbildung 3.5.: Ergebnisse: links-oben: Geisterbilder allein, rechts-oben: Geisterbilder mit Objektmaske, links-unten: Aufhellung durch Verrechnung aller Objektmasken des Zeitintegrals, rechts-unten: Zusammenblenden mit dem originalen Videobild um Artefakte zu vermindern.

Während der Implementierungsphase dieser Methode wurden als Zwischenschritt die Geisterbilder erstellt. Diese hätten bereits in der Studie verwendet werden können. Die

Erweiterung durch die Bewegungsunschärfe und Aufhellung wurde gewählt, um den Bewegungseindruck zu verstärken und die Objekte etwas hervorzuheben.

3.4.2. Vorteile der Differenzen-Methode

Für eine Verwendung der Differenzen-Methode sprechen folgende Punkte:

- Durch die Geisterbilder ist ein Objekt länger an einer Stelle im Bild zu sehen, als in den anderen Visualisierungen. Dadurch hat der Benutzer mehr Zeit das Objekt zu erkennen, bevor es wieder verschwindet.
- Im Vergleich zum Blending werden die Objekte weiterhin scharf dargestellt.
- Auch bei dieser Methode ziehen die Objekte einen Schweif hinter sich her, ähnlich wie die Bewegungsunschärfe beim Blending. Durch dessen Aufhellung fallen die bewegten Objekte besser auf.

3.5. Tracking

Die Tracking-Methode versieht die Bilder im Video nicht nur mit Informationen der übersprungenen Bilder, es werden zusätzlich noch Informationen über zukünftige Positionen der Objekte eingeblendet. Bewegte Objekte ziehen einen hellen Schweif hinter sich her, dieser soll die vergangene Bewegungsrichtung anzeigen. Die zukünftige Bewegungsrichtung wird in Form von Pfeilen angezeigt, die dem Objekt vorausseilen. Diese werden in das vorhandene Bild integriert. In ähnlicher Art und Weise werden Pfeile zur Anzeige der Bewegungsrichtung in der Arbeit von Chao et al.[CT]10] verwendet.

Funktionieren die anderen Visualisierungen noch direkt auf beliebigen Überwachungsaufnahmen mit fixierter Kamera, so werden für das Tracking zusätzliche Daten benötigt. Ein Video, das mit dieser Visualisierung angezeigt werden soll, braucht eine Datei mit sämtlichen Bounding-Box Informationen der vorkommenden Objekte in der Szene. Diese Dateien wurden vorab erstellt. Bestehende Informationen des BEHAVE-Datensatzes[beh12] und des AVSS-Datensatzes[avs12] wurden dabei noch von Hand mit dem „ViPER-Toolkit“[vip12] ergänzt.

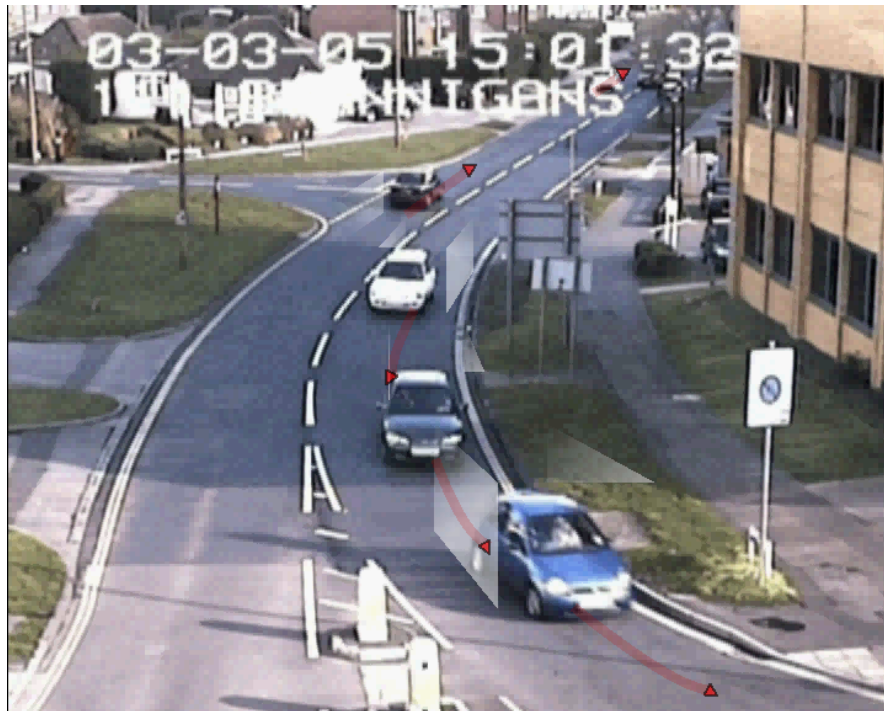


Abbildung 3.6.: Tracking: Durch die Schweife und Pfeile wird dem Benutzer vermittelt, woher die Objekte kommen und in welche Richtung sie sich bewegen werden.

3.5.1. Implementierung

Tracking Daten

Wie bereits erwähnt werden bei dieser Visualisierung zusätzliche Daten mit Informationen über die Objekte benötigt. Über jedes getrackte Objekt werden folgende Informationen aus einer dafür vorgesehenen Datei ausgelesen:

- **Startbild:** Die Nummer des ersten Videobildes, indem das Objekt zu sehen ist.
- **Dauer:** Die Anzahl der Videobilder, in denen das Objekt vorkommt.
- **Bounding-Boxes:** Liste mit Rechtecken, die das Objekt in jedem einzelnen Videobild umgeben.

Diese Datei und die darin enthaltenen Informationen wurden bereitgestellt. Dieses Dateiformat war bekannt und schon einmal verwendet worden.

Während beim Blending und den Differenzen noch Bilder verrechnet wurden, entsprechen die verwendeten Bilder beim Tracking, denen des Frameskip.

Pfeile und Schweife

Mit den gegebenen Informationen über die Position und Größe der getrackten Objekte lassen sich nun Kontrollpunkte erstellen, die für das Zeichnen der Pfeile und Schweife nötig sind. Prinzipiell geht man für beide Teile auf dieselbe Weise vor:

1. Objekt und Index der aktuellen Bounding-Box werden benötigt.
2. Bestimmen der Kontrollpunkte
3. Zeichnen der Kurven mittels Kontrollpunkte

Man iteriert durch alle vorhandenen Objekte und bestimmt anhand der aktuellen Framenummer, welche Bounding-Box des Objektes in diesem Frame gültig ist.

Die Bestimmung von Kontrollpunkten ist nötig, da zum Zeichnen der Pfeillinien und der Schweife quadratische Bézierkurven verwendet werden. Jede dieser Bézierkurven benötigt 3 Kontrollpunkte. Die Kurve, die zur Pfeilspitze führt, hat ihren ersten Kontrollpunkt mittig am Boden der aktuellen Bounding-Box. Die anderen beiden Kontrollpunkte werden bestimmt, indem man intervallweise im Index der Bounding-Boxen voranschreitet. Aufgrund verschiedenartiger Situationen in Überwachungsvideos, lässt sich nicht genau festlegen, wie groß das Intervall sein soll. Grundsätzlich bedeutet eine Vergrößerung des Intervalls auch eine Verlängerung der Kurven bei Objekten in Bewegung. In den verwendeten Videos hat sich ein Intervall von 20 für die Pfeile und 10 für die Schweife als geeignet erwiesen. Bei den Pfeillinien reichen 3 Kontrollpunkte, bei den Schweifen werden für jeden Eckpunkt der aktuellen Bounding-Box die Kontrollpunkte auf dieselbe Weise bestimmt.

Anschließend werden quadratische Bézierkurven mittels der Kontrollpunkte gezeichnet. Diese Methode wurde gewählt, da die resultierenden Kurven sehr viel eleganter aussehen als das bloße lineare Verbinden der Kontrollpunkte.

Zum Zeichnen der Pfeilspitzen wird der letzte Kontrollpunkt als Basismittelpunkt des zu zeichnenden Dreiecks verwendet. Da es durch die Intervallsprünge aber zu sehr unruhigen Pfeilen kommen kann, wird der letzte Kontrollpunkt für die Pfeillinien aus den vorherigen Positionen gemittelt. Für die verwendeten Videos hat ein gemittelter Wert aus 5 alten Kontrollpunkten zu ausgeglicheneren Bewegungen der Pfeillinien geführt. Für die Pfeilspitze wird dann noch ein zusätzlicher Punkt benötigt, der die Richtung des Dreiecks bestimmt. Hierfür wird ein gemittelter Wert der ersten Kontrollpunkte verwendet.

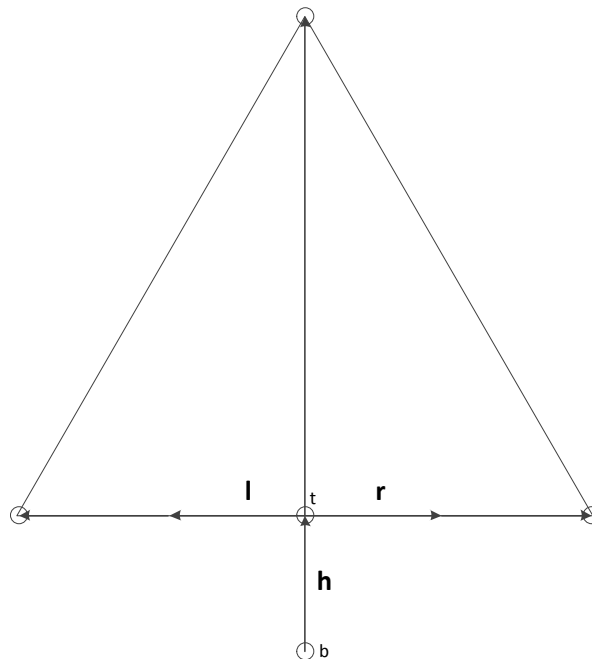


Abbildung 3.7.: Skizze der Pfeilspitze

$$\vec{h} = t - b$$

$$\vec{l} = ((-1) * \vec{h}_y, \vec{h}_x)^T$$

$$\vec{r} = ((\vec{h}_y, (-1) * \vec{h}_x)^T$$

Wie in Abbildung 3.7 zu sehen, wird durch \vec{h} der Vektor bestimmt, der die Pfeilrichtung angibt. Die Vektoren \vec{l} und \vec{r} stehen senkrecht dazu und bestimmen die Breite des Pfeils. Durch diese Vektoren lassen sich die Eckpunkte bestimmen, indem man den jeweiligen Vektor zur Mitte der Dreiecksbasis addiert und eine angemessene Skalierung vornimmt. Anschließend werden die Eckpunkte miteinander verbunden und die Fläche mit einem halbtransparenten Rot gefüllt.

Bei den Schweifen werden die Bézierkurven der linken und rechten Seite jeweils miteinander zu Flächen verbunden. Diese wird mit einem Gradientenverlauf gefärbt, der beim Objekt mit einem halbtransparenten Weiß beginnt und dann linear in die vollständige Transparenz übergeht.

Durch den Intervallsprung wird dem Benutzer zusätzlich zur Bewegungsrichtung noch Information über die Geschwindigkeit des jeweiligen Objektes gegeben. Bewegt sich ein

Objekt schnell durch das Bild, so werden die Positionen in dem festen Intervall weiter auseinander liegen als bei einem Objekt, das sich auf demselben Weg langsamer bewegt. Dadurch haben schnelle Objekte längere Pfeile und Schweife als langsame Objekte. Die Pfeile und Schweife werden solange gezeichnet, wie sich ein Intervallsprung zur Bestimmung der Kontrollpunkte durchführen lässt.

Alternativ zu den verwendeten Schweifen, wurden auch Funktionen getestet, die den Pfeil nach hinten weg verlängerten. Auf diese Weise bewegten sich die Objekte quasi immer auf diesem verlängerten Pfeil, ähnlich wie auf einer Schiene. Diese Methode wurde nicht verwendet, da die Schweife für einen stärkeren Bewegungseindruck sorgten. Auch Versuche mit geraden Speedlines, wie sie beispielsweise Hwang et al.[HLC⁺06] verwenden, wurden in der Implementierungsphase getestet. Dabei wurden in bestimmten Abständen Linien auf ähnliche Weise wie die Schweife, an den Bounding-Boxen angebracht. Da jedoch nur die Bounding-Box-Information vorhanden war, konnte diese Methode nicht überzeugen. Die Verwendung von exakten Silhouetten der Objekte könnte hier die Verwendung dieser Speedlines möglich machen.

3.5.2. Vorteile der Tracking-Methode

Folgende Punkte sprechen für eine Verwendung der Tracking-Methode:

- Die Schweife geben wie bei den Differenzen Auskunft über die Richtung, aus der das Objekt gekommen ist.
- Im Gegensatz zu den anderen Visualisierungen besitzt Tracking den Vorteil, dass man bereits von Beginn an weiß, zu welchem Zeitpunkt sich ein Objekt an einem bestimmten Ort befinden wird. Diese Informationen können für den Benutzer sichtbar gemacht werden, indem man ihm durch die Pfeile den zukünftigen Bewegungsverlauf zeigt. Es wird also mehr Information in einem Bild untergebracht, als in den anderen Visualisierungen.
- Die Pfeile und Schweife unterscheiden sich zudem in ihrer Farbwahl stark vom Inhalt der Videos und fallen deshalb besonders auf. Dies kann von Vorteil sein, da es jedes getrackte Objekt markiert, und somit die Aufmerksamkeit des Benutzers auf sich zieht. Störende Ablenkungen durch den Hintergrund können auf diese Weise verringert werden.

3.6. Videoplayer mit Eingabefunktion

Die Teilnehmer der Benutzerstudie müssen in ihrer ersten Aufgabe (siehe Abschnitt 5.2.2) Fremdobjekte im Video erkennen. Ein handschriftliches Festhalten der Ergebnisse lässt sich nicht sinnvoll umsetzen. Da die Objekte sehr schnell durch das Bild wandern, müsste der Studienleiter sämtliche Fremdobjekte genau kennen, um in wenigen Augenblicken zu entscheiden, ob ein Teilnehmer diese nun richtig erkannt hat, oder nicht. Zusätzlich würde das Aufschreiben der Ergebnisse den Teilnehmer ablenken.

Aus diesem Grund wurde für die Studie noch ein Programm implementiert, das in der Lage ist, die Videos der einzelnen Visualisierungen abzuspielen. Gleichzeitig können Eingaben des Benutzers angenommen und beurteilt werden. Für jedes Video, das bei dieser Aufgabe Verwendung findet, wurde eine Datei erstellt, in der der genaue Framebereich der einzelnen Fremdobjekte vermerkt ist. Diese wird beim Laden des Videos eingelesen und eine Liste mit allen Fremdobjekten erstellt. Bei einer Eingabe prüft das Programm, ob sich eines der gelisteten Objekte im aktuellen Frame befindet. Ist dies der Fall, so wird das entsprechende Objekt aus der Liste entfernt. Bei jeder Eingabe, richtig oder falsch, wird ein Vermerk in eine Datei geschrieben. Dabei wird festgehalten, bei welchem Frame gedrückt wurde und ob dies richtig war oder nicht. Abschließend wird an das Ende der Datei die Gesamtzahl der richtigen, falschen und verpassten Eingaben geschrieben. Für die Erstellung dieser Datei wird der Name des jeweiligen Videos und die Teilnehmer ID verwendet. Auf diese Weise lässt sich bei der Auswertung genau nachvollziehen, welche Ergebnisse von welchem Teilnehmer erbracht wurden.

Eingaben können mit der Leertaste über die Tastatur getätigt werden. Um die Teilnehmer möglichst lange motiviert zu halten, wurde ein weiteres Eingabegerät verwendet, das der Aufgabe einen spielerischeren Charakter verleiht. Den Benutzern wird ein Buzzer zur Verfügung gestellt, der aus dem Quizspiel Buzz[buz12] für die Playstation 2 stammt. Um die Eingabe der Buzzer in Qt verarbeiten zu können, wurde der „Qt SDL joystick wrapper“ von René Reucher[qt-08] verwendet.

Jedesmal, wenn von dem Benutzer der rote Knopf des Buzzers betätigt wird, erfolgt eine Prüfung, ob im aktuellen Frame ein Fremdobjekt existiert. Je nachdem ob dies der Fall war oder nicht, wird dem Benutzer ein direktes Feedback in Form von einem Symbol für Richtig oder Falsch angezeigt.

Geschwindigkeitsvisualisierungen

Viele Videos, die von Überwachungskameras aufgezeichnet werden, enthalten Szenen, in denen über einen längeren Zeitraum nichts passiert. Das kann zum Beispiel die Aufnahme eines Firmengeländes bei Nacht sein, oder eine S-Bahnstation, bei der nur zu bestimmten Uhrzeiten viele Leute zu sehen sind. Hier macht es Sinn, keine konstante Beschleunigung zu verwenden, sondern in unwichtigen Situationen zu beschleunigen, und wenn etwas Wichtiges passiert, wieder zu verlangsamen. Dies wird mittels der Adaptive Fast-Forward-Methode von Höferlin et al.[HHWH10] umgesetzt.

Die wechselnden Geschwindigkeiten können aber nicht immer direkt wahrgenommen werden. Aus diesem Grund wurden hierfür drei Visualisierungen implementiert, die dem Benutzer den Geschwindigkeitswechsel verdeutlichen sollen. Außerdem soll er den Eindruck vermittelt bekommen, dass ein Video schnell oder langsam läuft.

Ein Tachometer oder ein Rahmen um das Bild versuchen mittels Farben diese Eindrücke zu vermitteln. Als dritter Ansatz wurden Störstreifen gewählt, die, ähnlich wie bei analogen Videorekordern, je nach Geschwindigkeit vermehrt im Bild auftauchen.

4.1. Tachometer

Der Tachometer ist in Abbildung 4.1 dargestellt. Er wird im Video in der rechten unteren Ecke platziert. Ähnlich wie beim SmartPlayer von Cheng et al.[CLCC09] dient er zum Anzeigen der momentanen Abspielgeschwindigkeit in Form des Beschleunigungsfaktors.

Er besitzt eine Farbskala die von Weiss für normale, über Gelb und Rot bis Schwarz für die hohen Beschleunigungsfaktoren reicht. Mit einem Faktor $Skala = 40$ wird durch die Logarithmierung in Gleichung (4.1) der Ausschlag für zunehmende Geschwindigkeiten immer geringer, wodurch eine 10.000-fache Beschleunigung nötig wäre, um das Ende der Skala zu erreichen. Der Ausschlag der Tachonadel berechnet sich folgendermaßen:

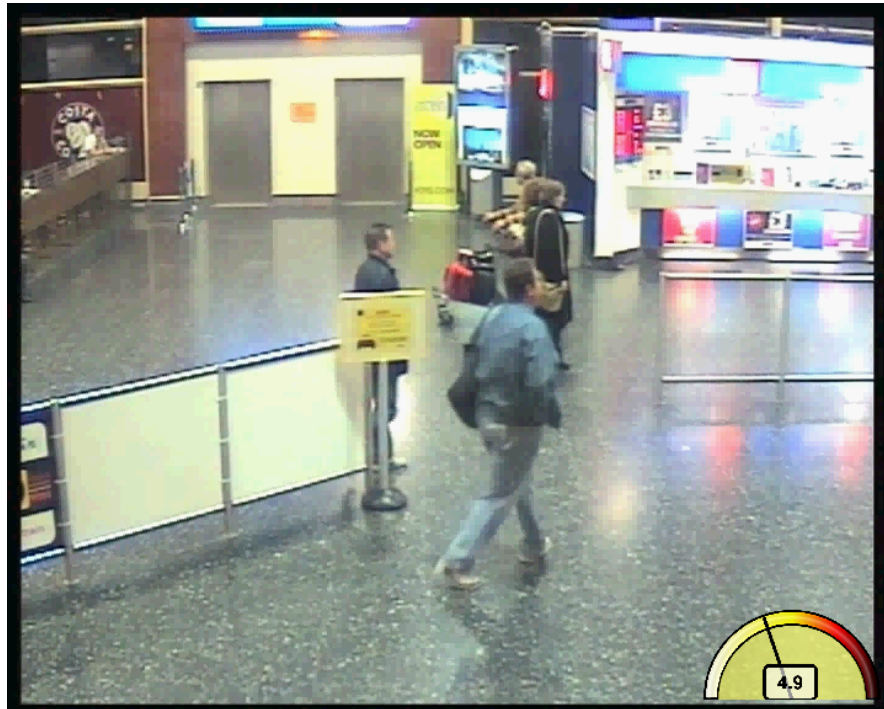


Abbildung 4.1.: Der Tachometer befindet sich rechts unten im Bild. Die momentane Geschwindigkeit wird zusätzlich als Dezimalzahl angezeigt.

$$\text{Geschwindigkeit}_T = \log(\text{Geschwindigkeit}_C) \cdot 10 \quad (4.1)$$

$$\beta = 180 - \text{Geschwindigkeit}_T \cdot (180 / \text{Skala})$$

$$\vec{p} = \left(\frac{w}{2} + r \cdot \cos(\beta), h - r \cdot \sin(\beta) \right)^T \quad (4.2)$$

Die Formel (4.1) wurde gewählt, um den Anstieg der Tachonadel bei zunehmender Geschwindigkeit zu verringern. Gleichung (4.2) berechnet dann aus den Polarkoordinaten mit β die kartesischen Koordinaten für den Punkt \vec{p} , an dem sich die Spitze der Tachonadel befindet. Geschwindigkeit_C ist die gedämpfte Geschwindigkeit, die als Wert angezeigt wird. Geschwindigkeit_T wird nur für den Ausschlag der Tachonadel verwendet. Auf eine genaue Beschriftung der Tachometerskala wurde verzichtet, da diese nur einer qualitativen Einschätzung der Geschwindigkeit dienen soll. Die Tachonadel erhält eine Dämpfung, um unruhige Sprünge der Tachonadel zu vermindern:

$$\text{Geschwindigkeit}_C(t) = 0.95 \cdot \text{Geschwindigkeit}_C(t-1) + 0.05 \cdot \text{Geschwindigkeit}_I$$

Dabei ist $Geschwindigkeit_t$ die ungedämpfte Geschwindigkeit zum Zeitpunkt t . Die von der Nadel angepeilte Farbe auf der Skala wird zusätzlich im Hintergrund des Tachometers angezeigt. Auf diese Weise soll es dem Benutzer ermöglicht werden, durch einen kurzen Blick auf Farbe und Tachonadel, eine Geschwindigkeitsinformation zu erhalten. Zusätzlich bietet der Tachometer als einzige der drei Visualisierungen eine genaue Anzeige der momentanen Geschwindigkeit mittels Dezimalwert.

4.2. Rahmen



Abbildung 4.2.: Der Rahmen umfasst das Bild und ändert je nach Geschwindigkeit seine Farbe.

Der Rahmen wird transparent um das Video herum gezeichnet. Er erhält seine Farbinformation auf die selbe Weise wie der Tachometer, die Dämpfung der Geschwindigkeit ist hier allerdings geringer. Es fehlt hier jedoch ein genauer Überblick, welche Position auf der Skala momentan angezeigt wird. Der Benutzer soll so wenig wie möglich vom Inhalt des Videos abgelenkt werden, indem er den Rahmen nur mittels peripheren Sehens wahrnimmt. Beim peripheren Sehen werden die Änderungen des Rahmens vorwiegend als Blinken wahrgenommen[Daho6]. Die Unterscheidung von Hell und Dunkel soll eine Zuordnung von langsamen und schnellen Geschwindigkeiten erlauben. Da die Farbskala des Tachometers

übernommen wurde, kann bei Bedarf der Rahmen fokussiert werden, um durch die Farbe genauere Informationen über die momentane Geschwindigkeit zu erhalten.

4.3. Störstreifen

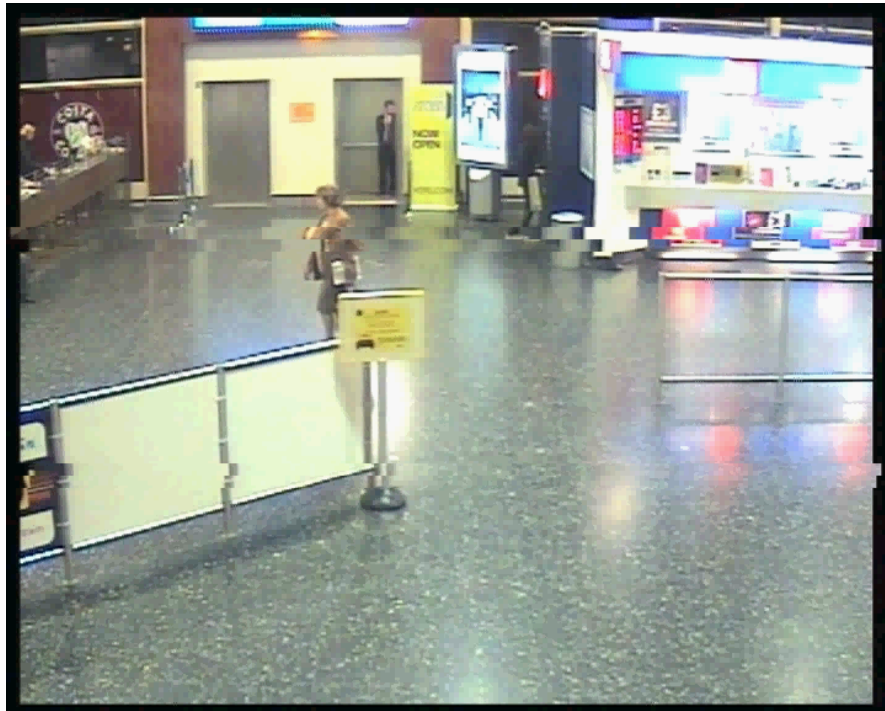


Abbildung 4.3.: Störstreifen in Stufe 6: 2 Streifen der Dicke 2.

Die Störstreifen sind inspiriert durch den Schnellvorlauf analoger Videorekorder. Dabei handelt es sich um horizontal verlaufende, schmale Ausschnitte des Videobildes. Die Höhe dieser Ausschnitte ist vorgegeben. Die im folgenden erwähnte Dicke der Störstreifen bezieht sich auf einen Wert zwischen 1 und 3. Dabei handelt es sich um die Anzahl der Ausschnitte, die einen Störstreifen bilden. Für jeden Ausschnitt des Bildes wird ein horizontaler Offset mit der Gleichung (4.3) berechnet. Dieser Wert gibt eine Verschiebung des Ausschnittes nach links oder nach rechts an. Durch die Zufallsvariable ändert sich diese Verschiebung von Bild zu Bild und es kommt zu den bewegten Störstreifen.

$$Offset_{horizontal} = (rand() \bmod 10) \cdot -1^{rand() \bmod 2}; \quad (4.3)$$

Die Abspielgeschwindigkeit wird insgesamt in 10 Stufen unterteilt. Dabei beschreiben folgende Formeln die Anzahl und Breite der Störstreifen:

$$(Stufe)x = \lfloor Geschwindigkeit \cdot \frac{10}{max.Geschwindigkeit} \rfloor \quad (4.4)$$

$$\text{Streifenanzahl}(x) = \lfloor (x + 1) / 3 \rfloor \quad (4.5)$$

$$\text{Streifendicke}(x) = ((x - 2) \bmod 3) + 1 \quad (4.6)$$

Aus der momentanen Geschwindigkeit wird in Gleichung (4.4) die benötigte Stufe der Beschleunigung bestimmt. Durch die maximale Geschwindigkeit kann der Bereich angegeben werden, auf den die 10 Stufen skaliert werden. Die Anzahl und Dicke der Streifen ergibt sich wie in den Gleichungen (4.5) und (4.6) beschrieben. Die erste Stufe bleibt dabei noch ohne Streifen. In den nächsten 3 Stufen entsteht ein Störstreifen mit ein- bis dreifacher Dicke. Dann entsteht auf den nächsten Stufen ein zweiter und die Streifen werden dicker. Bis in der 10-ten Stufe 3 Störstreifen mit dreifacher Dicke zu sehen sind.

Die höhere Priorität wurde hierbei der qualitativen Beurteilung der Geschwindigkeitsunterschiede zugewiesen. Durch eine Reduktion auf 10 diskrete Stufen wurde die vermittelte Geschwindigkeitsinformation drastisch gekürzt. Der Benutzer soll dadurch möglichst wenig Zeit mit der Interpretation der Streifen verbringen, um so nicht unnötig vom Inhalt des Videos abgelenkt zu werden. Ein weiterer Vorteil hierbei ist, dass der Benutzer nicht auf die Randbereiche des Bildes angewiesen ist, um Informationen zur Geschwindigkeit zu erhalten. Die Streifen befinden sich hier direkt im Fokus des Benutzers und können zusammen mit dem Video betrachtet werden.

Studie

Nach der Implementierung der Visualisierungen gilt es nun, diese auf ihre effektive Nutzbarkeit zu prüfen. Hierfür wird eine Benutzerstudie benötigt, deren Ziel die Überprüfung einiger Thesen und die Beantwortung einiger Forschungsfragen sein soll. Dieses Kapitel handelt Schritt für Schritt die Planung, Umsetzung und Auswertung einer solchen Studie ab.

5.1. Hypothesen und Fragen

Zunächst gilt es, durch die Formulierung von Hypothesen und Fragen eine genaue Planung der Studie zu ermöglichen. Ist einmal geklärt, was man durch die Studie zeigen, beziehungsweise erfahren möchte, so lassen sich entsprechende Aufgaben dafür erstellen. Die Fragestellung teilt sich hierbei in drei Bereiche auf:

1. Erkennung
2. Bewegungsverfolgung
3. Geschwindigkeitseindruck

Die Erkennung von Objekten spielt dabei die wichtigste Rolle, da der Sinn beim Durchsehen von Videos im Schnelldurchlauf meistens in der Erkennung einer bestimmten Situation oder eines Objektes liegt. Bei der Verfolgung von Objekten soll überprüft werden, welche der Visualisierungen im Vergleich zu den anderen bevorzugt wird. Der Geschwindigkeitseindruck bezieht sich auf die wechselnden Geschwindigkeiten des Adaptive Fast-Forward Verfahrens. Hierbei sollen subjektive Eindrücke über die dafür implementierten Visualisierungen gesammelt werden.

5.1.1. Erkennung

Es besteht Grund zu der Annahme, dass die Frameskip-Methode bei der Erkennung von Objekten gute Ergebnisse liefern wird. Schließlich findet dieses Verfahren sehr häufig Anwendung und die meisten Teilnehmer sollten es gewohnt sein, auf diese Weise etwas in einem Video zu suchen. Dagegen sollte besonders die Blending-Methode die Erkennung von Objekten erschweren. Durch die Differenzen-Methode werden die Objekte mehrfach an einer Stelle zu sehen sein, was die Erkennung verbessern sollte. Die Tracking-Methode sollte mithilfe der Pfeile und Schweife die Aufmerksamkeit auf die wichtigen Objekte lenken, um dem Benutzer so mehr Zeit für die Erkennung einzuräumen. Für die Erkennung von Objekten sollen folgende Hypothesen geprüft werden:

- **Hypothese 1:** Die Blending-Methode wird im Vergleich die schlechtesten Ergebnisse bei der Erkennung von Objekten liefern. Dies liegt vor allem an der zunehmenden Unschärfe der Objekte, je höher die Geschwindigkeit wird.
- **Hypothese 2:** Frameskip und Differenzen werden die besten Ergebnisse liefern. Durch die mehrfach sichtbaren Objekte bei der Differenzen-Methode könnte die Objekterkennung etwas besser ausfallen als beim Frameskip.
- **Hypothese 3:** Abhängig von der Szene wird Tracking unterschiedliche Ergebnisse liefern. Wenn es wie in Video 2 sehr viel Bewegung gibt, so könnte es durch die Methode zu zusätzlicher Verwirrung kommen.

Zusätzlich sollen durch die subjektiven Meinungen der Teilnehmer folgende Fragen beantwortet werden:

- In welcher Visualisierung hatten die Benutzer den Eindruck, dass sie die Objekte am besten erkennen konnten?
- Welche Visualisierung benötigt sehr viel Aufmerksamkeit?
- Welche Visualisierung wird am angenehmsten beim Zusehen empfunden?
- Welche Visualisierung wird am hilfreichsten bei der Erkennung empfunden?
- Welche Visualisierung wird bei der Erkennung bevorzugt?

5.1.2. Bewegungsverfolgung

Wie bereits erwähnt spielt die Change Blindness (siehe Abschnitt 1.1) durch die Unterbrechung von Bewegungsabläufen eine wichtige Rolle beim Betrachten von Überwachungsvideos. Da die Fragen zur Erkennung bereits formuliert wurden, soll hier die Bewegungsverfolgung separat betrachtet werden. Folgende Hypothese und Fragen sollen geprüft werden:

- **Hypothese 4:** Frameskip wird aufgrund der sprunghaften Positionswechsel von Objekten bei hohen Geschwindigkeiten den anderen Methoden unterlegen sein. Tracking bietet durch die Pfeile und Schweife genauere Informationen darüber, wo das Objekt war, wo es ist und wo es in den kommenden Bildern sein wird. Dadurch sollte eine einfache Verfolgung der Bewegungen möglich sein.

Folgende Fragen sind hierbei interessant:

- Welche Visualisierung wird im Vergleich zu den anderen bevorzugt gewählt?
- Gibt es Visualisierungen, die im direkten Vergleich häufig zu einer Unentschlossenheit beim Teilnehmer geführt haben?

5.1.3. Geschwindigkeitseindruck

Als letztes sollen noch der Tachometer, der Rahmen und die Störstreifen in Bezug auf die Vermittlung der wechselnden Geschwindigkeit geprüft werden. Hierfür gilt es folgende Hypothesen und Fragen zu überprüfen:

- **Hypothese 5:** Der Tachometer vermittelt die Geschwindigkeit und den Geschwindigkeitswechsel am besten.
- **Hypothese 6:** Die Ablenkung durch den Rahmen und die Störstreifen wird geringer sein, als die des Tachometers. Diese Methoden bieten weniger Information und müssen nicht extra fokussiert werden.

Welche Visualisierung letztendlich am geeignetsten zur Vermittlung der wechselnden Geschwindigkeiten ist, gilt es herauszufinden.

5.2. Aufgabenstellungen

Besonders zur Prüfung der effektiven Erkennung bedarf es einiger Vorbereitungen. Nachdem die Fragestellung erörtert wurde, gilt es die Variablen zu bestimmen, die bei der Aufgabenstellung beeinflusst werden können. Mit dieser Grundlage lassen sich dann entsprechende Aufgaben formulieren, die den erhofften Informationsgehalt liefern.

5.2.1. Variablen

Bei der Erkennungsaufgabe lassen sich folgende Variablen beeinflussen:

- **Fast-Forward Visualisierungen:** Frameskip, Blending, Differenzen und Tracking sind die vier verwendeten Methoden, um ein Video zu beschleunigen.
- **Beschleunigungsfaktor des Videos:** Die Abspielgeschwindigkeit der Videos kann erhöht werden, indem der Beschleunigungsfaktor für die Visualisierungen geändert wird. Starke Beschleunigungen führen zu einem erhöhten Schwierigkeitsgrad, da zu erkennende Objekte in immer weniger Frames vorhanden sein werden.
- **Videoszene:** Mit der Wahl der Videoszene lässt sich der Schwierigkeitsgrad der Aufgabe beeinflussen. Ein geringer Blickwinkel der Kamera und wenig Bewegung lässt sich leichter überblicken als ein weiter Blickwinkel mit sehr viel Bewegung.

Für die Verfolgung der Bewegungen gelten dieselben Variablen. Zur Prüfung der Geschwindigkeitsvisualisierungen müssen jedoch neue Variablen bedacht werden:

- **Geschwindigkeitsvisualisierungen:** Tachometer, Rahmen und Störstreifen sind hier die drei verwendeten Methoden, die zum Betrachten eines Videos mit variierender Geschwindigkeit verwendet werden.
- **Beschleunigungsfaktor des Videos:** Die Geschwindigkeit des Videos variiert zwar, dies geschieht jedoch immer relativ zu einem konstanten Beschleunigungsfaktor. Durch die Wahl dieses Faktors lässt sich die Abspielgeschwindigkeit beeinflussen. Eine Variation des Faktors verspricht keine großen Änderungen am Gesamteindruck der Studienteilnehmer von den einzelnen Visualisierungen. Deshalb wird dieser Faktor ebenfalls konstant gehalten werden.
- **Videoszene:** Das Geschehen innerhalb einer Szene kann die Wahrnehmung wechselnder Geschwindigkeiten beeinflussen. Eine Szene in der sich dauernd Menschen durch das Bild bewegen, lässt den Beobachter leichter erahnen, ob diese nun gerade beschleunigt oder verlangsamt werden. Diese Variable wird konstant durch ein Video (siehe Abschnitt 5.2.4) vertreten, das sich durch seinen Wechsel zwischen viel und wenig Personenverkehr besonders gut eignet.

Da nun die beeinflussbaren Faktoren geklärt wurden, lassen sich nun geeignete Aufgaben definieren. Insgesamt setzt sich die Benutzerstudie aus 3 Aufgaben zusammen:

1. Erkennung von Fremdobjekten innerhalb zweier verschiedener Szenen durch die verschiedenen Visualisierungen und anschließende Bewertung seitens der Teilnehmer.
2. Direkter Vergleich der Visualisierungen bezüglich der Verfolgung von Bewegungen.
3. Direkter Vergleich der Geschwindigkeitsvisualisierungen seitens der Teilnehmer.

5.2.2. Erste Aufgabe - Fremdobjekte erkennen

Für die erste Aufgabe werden als Ausgangsmaterial zwei Ausschnitte aus den Datensätzen AVSS[avs12] und BEHAVE[beh12] verwendet. Das BEHAVE Video hat eine Länge von 15:16 Minuten. Im Verlauf dieser Arbeit wird es als „Video 1“ bezeichnet werden. Das AVSS Video hat eine Länge von 17:49 Minuten, es wird fortan als „Video 2“ bezeichnet.



Abbildung 5.1.: links: BEHAVE - Video 1, rechts: AVSS - Video 2

Wie in Abbildung 5.1 zu sehen ist, umfasst die Szene, die in Video 1 zu sehen ist, einen Parkplatz mit Straße und ein Stück Grünanlage mit einem Gehweg. Die Szene in Video 2 bietet dagegen einen recht weiten Überblick über eine Straße mit relativ viel Verkehr, inklusive der umliegenden Gehwege, Gebäude und Seitenstraßen. Da die verschiedenen Visualisierungen auf dieses Material angewendet werden sollen, muss zunächst die Beschleunigung der Videos an ihren Schwierigkeitsgrad angepasst werden. Im Rahmen der durchgeführten Pilotstudie (siehe Abschnitt 5.3.1) haben sich Beschleunigungsfaktoren von 20 für Video 1 und 10 für Video 2 als sinnvoll erwiesen. Diese Beschleunigungsfaktoren machen Video 1 0:57 Minuten und Video 2 1:46 Minuten lang. Der Benutzer bekommt diese beschleunigten Videos in allen Visualisierungen zu sehen und zwar für Video 1 und Video 2. Da dies trotz der Beschleunigung noch immer eine recht lange Zeit zur Beobachtung des Videomaterials ist, wurde die erste Aufgabe mit einigen spielerischen Elementen versehen, um die Motivation des Benutzers möglichst lange aufrecht zu erhalten.

Um nun zu prüfen, wie sich die Erkennung von Objekten innerhalb der einzelnen Visualisierungen verhält, wurden Fremdobjekte als Stimuli in die Szenen eingefügt, die der Benutzer entdecken muss. Als Fremdobjekt wurde eine Comicfigur in die Videos eingefügt, die der von Martin Handford erfundenen Figur „Wally“ (Waldo, Walter)[wal12] nachempfunden ist.

Dem ein oder anderen Studienteilnehmer könnten die hierzulande unter dem Namen „Wo ist Walter?“ bekannten Wimmelbilder ein Begriff sein. Die Figur wurde gewählt, um



Abbildung 5.2.: links: Waldo im Original[wal12], rechts: Die angepasste Kopie für die Studie

der Aufgabe einen kleinen Suchspielcharakter zu verleihen. Wie in Abbildung 5.2 zu sehen ist, wurden Anpassungen vorgenommen, die diese Figur weniger auffällig in die Szene integrieren sollten. Dafür wurde eine weniger auffällige Farbwahl der Kleidung getroffen, und die konturgebenden Outlines entfernt. Anschließende Schritt-Animationen von links, rechts, vorne und hinten machen es dann möglich, die Figur so in das Video zu integrieren, dass sie die gleichen Wege beschreiten kann wie die realen Personen in den Videos.

Beide Testvideos wurden nun präpariert, indem die Figur mehrfach in die Videos eingefügt wurde. In Video 1 taucht die Figur insgesamt 13 mal, in Video 2 insgesamt 8 mal auf. Auch wenn sich die eingefügten Figuren unterschiedlich bewegen, so bleiben die Personen und Fahrzeuge der ursprünglichen Videos an denselben Stellen, was zu einem Lerneffekt beim Benutzer führen kann. Durch die erhöhten Geschwindigkeiten und die Konzentration des Benutzers auf die Suchaufgabe, sollte sich dieser Effekt verringern. Um dies jedoch noch besser zu verhindern, wurden weitere Vorbereitungen getroffen: Von beiden Videos wurden jeweils vier präparierte Versionen erstellt. Diese unterscheiden sich in der Reihenfolge und zu den Zeitpunkten, an denen sich die Figur durch das Bild bewegt. Auf diese Weise kann man in der Studie dem Benutzer das Video in allen Visualisierungen zeigen, ohne dass er jedesmal das Gleiche zu sehen bekommt. Jede der vier Versionen wird mit allen Visualisierungen gerendert. Insgesamt werden also sowohl für Video 1 als auch für Video 2 jeweils 16 Videos erstellt. Um alle Visualisierungen sinnvoll vergleichen zu können, muss auch die Reihenfolge variiert werden, in der sie der Benutzer zu sehen bekommt. Es macht

keinen Sinn, eine bestimmte Visualisierung stets am Anfang und eine andere am Ende zu zeigen. Dies hätte mit Sicherheit den Effekt, dass die Ergebnisse bei beiden Visualisierungen verfälscht wären. Am Anfang durch die Unbekanntheit der Aufgabe und am Ende durch den unvermeidlichen Lerneffekt. Deshalb wird die Reihenfolge, in der die Benutzer die Visualisierungen zu sehen bekommen, systematisch permutiert. Daraus ergibt sich letztendlich auch die Zahl der Studienteilnehmer. Vier Visualisierungen in allen möglichen Permutationen ergibt $4! = 24$ Versuchsteilnehmer. Auf diese Weise wird jede Visualisierung gleich oft an jeder Stelle der Reihenfolge gezeigt. Abhängig von der Reihenfolge wird dann die jeweilige Version des Videos in der entsprechenden Visualisierung verwendet.

Die Studienteilnehmer sollen jedesmal, wenn sie die Figur im Bild erkennen, den Buzzer drücken. Da aufgrund der hohen Beschleunigung die Reaktionszeit der meisten Teilnehmer nicht ausreichen würde, um in dem Moment zu drücken, wenn die Figur zu sehen ist, wurde eine verzögerte Reaktion innerhalb von 2 Sekunden nach dem Verschwinden der Figur erlaubt. Dabei wird bei jedem Tastendruck ein direktes Feedback gegeben, ob man richtig oder falsch lag (siehe Abschnitt 3.6).

Im Anschluss wurde ein Fragebogen ausgeteilt um die subjektive Eindrücke der Teilnehmer zu erfassen (siehe Abschnitt 5.4.1).

5.2.3. Zweite Aufgabe - Bewegungsverfolgung vergleichen

In der zweiten Aufgabe soll der Teilnehmer im Paarvergleich bestimmen, in welcher Visualisierung er die Bewegungen einzelner Objekte besser verfolgen kann. Hierfür werden ihm immer zwei Videos mit demselben Ausschnitt in zwei verschiedenen Visualisierungen vorgeführt und anschließend notiert, welche er bezüglich der Bewegungsverfolgung besser findet. Unentschlossenheit im positiven wie auch im negativen Sinn wird dabei als Enthaltung festgehalten. Für den Vergleich jeder Visualisierung mit den anderen ergeben sich nach Abzug doppelter Vergleiche noch 6 Paare, die von jedem Benutzer verglichen werden müssen. Um auch hier eine abwechselnde Reihenfolge zu erhalten, werden die Paare zunächst in eine gemischte Reihenfolge gebracht. Die Elemente werden zunächst immer zyklisch nach links verschoben, dann gespiegelt und wieder zyklisch verschoben. Nach der Hälfte der Teilnehmeranzahl werden dann die Vergleichspaare vertauscht und das Tauschverfahren ein weiteres Mal angewendet.

5.2.4. Dritte Aufgabe - Geschwindigkeitsvisualisierungen vergleichen

In der dritten Aufgabe wird dem Benutzer ein Video gezeigt, das durch Adaptive Fast-Forward beschleunigt wurde. Dabei handelt es sich um ein Video aus der „Imagery Library

for Intelligent Detection Systems(i-LIDS)“[i-L12], es wird von hier an als Video 3 bezeichnet. Für dieses Video waren bereits Daten für das Adaptive Fast-Forward von Höferlin et al.[HHWH10] vorhanden, und konnten direkt verwendet werden. Dieses Video wurde den Studienteilnehmern in jeder der drei Geschwindigkeitsvisualisierungen gezeigt. Hierbei wurde auch die Reihenfolge der Visualisierungen permutiert, um möglichen Einflüssen durch Ermüdung oder anderen Konzentrationsproblemen entgegen zu wirken. Anschließend wurden die Eindrücke des Benutzers in einem Fragebogen festgehalten. Dabei waren folgende Punkte zu beantworten:

- **Effektivität:** Wie gut hat die jeweilige Visualisierung geholfen, den Geschwindigkeitswechsel besser einzuschätzen?
- **Beanspruchung:** Wie anstrengend war es die Visualisierung anzusehen?
- **Ablenkung:** Wurde der Benutzer durch die Visualisierung vom Inhalt des Videos abgelenkt?

Abschließend wurde noch vermerkt, welche der Visualisierungen die Geschwindigkeit am besten vermittelt hat.

5.3. Durchführung

Nach dem Entwurf und der Planung der Benutzerstudie musste diese nun durchgeführt werden. Um eventuelle Probleme in der Aufgabenstellung und den Fragebögen frühzeitig erkennen zu können, wurde zunächst eine Pilotstudie mit 4 Kandidaten durchgeführt. Nach Korrekturen einiger Probleme wurde dann die eigentliche Benutzerstudie mit der vollen Teilnehmerzahl begonnen.

5.3.1. Pilotstudie

Vor der eigentlichen Benutzerstudie wurde das Studiendesign anhand von 4 Teilnehmern getestet. Aufgrund dieser Ergebnisse wurden folgende Punkte für die Benutzerstudie angepasst:

- **Anpassung der Likert-Skala:**
Die ursprüngliche Likert-Skala bestand aus 21 Werten. Um klarere Differenzierungen zu erhalten und eine Häufung neutraler Aussagen in der Mitte zu vermeiden, wurde die Skala auf eine gerade Anzahl von Werten reduziert. Zur besseren Übersicht für den Benutzer besitzt die neue Skala nun die Werte von 1 (für Widerspruch) bis 10 (für Zustimmung).
- **Anpassung des Schwierigkeitsgrades im BEHAVE-Video:**
Bei Betrachtung der F1-Scores (siehe Abschnitt 5.4.1) von Video 1 stellte sich heraus, dass zu hohe Ergebnisse in allen Visualisierungen erreicht wurden. Vorwiegend Ergebnisse im Bereich um die 100% machen einen Vergleich der Visualisierungen jedoch nicht möglich. Während der Pilotstudie lief das BEHAVE-Video noch mit 16facher Beschleunigung. Die Geschwindigkeit wurde zur Erhöhung des Schwierigkeitsgrades auf 20fache Beschleunigung angepasst. Zusätzlich wurden weitere Figuren eingefügt, die sich an schwerer zu erkennenden Stellen bewegen. Beispielsweise in der Gruppe von Menschen oder kurz am Rand entlang.
- **Orientierungsblatt für die Visualisierungen:** Während des Tutorial und beim Ausfüllen der Fragebögen kam es häufiger zu dem Problem, die Bezeichnungen der Visualisierungen ihren Ergebnissen zuzuordnen. Aus diesem Grund wurde ein Orientierungsblatt erstellt, das zu den neu eingeführten Fast-Forward Visualisierungen entsprechende Bilder und kurze Beschreibungen enthält. Die darauf enthaltenen Bilder stammen aus Video 2 und wurden so gewählt, dass die gleiche Stelle im Video durch die einzelnen Visualisierungen repräsentiert wurde.

5.3.2. Benutzerstudie

Nach der Anpassung der genannten Punkte im Studiendesign wurde die richtige Benutzerstudie durchgeführt.

Umgebungsbedingungen

Die Studie wurde in einem Raum mit künstlicher Beleuchtung durchgeführt. Die Studienteilnehmer wurden an einen Tisch gesetzt, auf dem sich lediglich ein Laptop und für die Studie relevante Materialien befanden. Diese wurden den Teilnehmern jedoch erst sichtbar gemacht, wenn es die Aufgabenstellung erforderte. So wurden beispielsweise die bunten Buzzer für die erste Aufgabe erst gezeigt, als die entsprechende Folie im Tutorial präsentiert wurde. Auch ein Handyverbot und ein Schild an der Tür mit der Aufschrift „Visualisierungsstudie - Bitte nicht stören!“ fanden Verwendung. Durch diese Maßnahmen sollten ablenkende Störfaktoren möglichst vermieden werden.

Für die Durchführung wurde ein Laptop mit einer Auflösung von 1366 X 768 Pixeln verwendet. Je nach Sitzhaltung betrug der Abstand der Teilnehmer zum Monitor zwischen ca. 50cm - 80cm. Durch einen Sehtest[mco12], der aus 2,50m Entfernung zum Monitor durchgeführt wurde, konnte sichergestellt werden, dass eine eventuelle Sehbehinderung die Ergebnisse nicht verfälscht.

Teilnehmer

Insgesamt wurde die Studie mit 24 Teilnehmern durchgeführt. Davon waren 7 weibliche und 17 männliche Teilnehmer. Die erste Hälfte der Teilnehmer trat im Zeitraum vom 19. bis zum 22. Dezember 2011 zu der Studie an. Die zweite Hälfte in dem Zeitraum nach Weihnachten bis zum 12.1.2012. Das Alter der Teilnehmer reichte von 20 bis 31 Jahren, mit einem Durchschnittsalter von 25,5 Jahren.

Die Fernsehzeit, wobei auch das Ansehen von Videos am Computer beachtet wurde, reichte von 2 bis 35 Stunden in der Woche. Im Durchschnitt 9,5 Stunden pro Woche. Die Zeit vor dem Computer insgesamt reichte von 8 bis 60 Stunden in der Woche, mit einem Durchschnitt von 31,2 Stunden in der Woche. Computerspiele nahmen hierbei 0 bis 20 Stunden mit einem Durchschnitt von 4,5 Stunden ein.

12 Teilnehmer gaben an ein Hobby zu haben, das eine erhöhte Aufmerksamkeit erfordert.

11 Teilnehmer trugen eine Brille oder Kontaktlinsen.

Alle Teilnehmer besaßen eine allgemeine Hochschulreife, beziehungsweise eine Fachhochschulreife. Bei der Mehrzahl der Teilnehmer handelte es sich um Studenten unterschiedlicher Studiengänge. Mit 7 Softwaretechnikern war dieser Studiengang am häufigsten vertreten. 2

Teilnehmer hatten kein Studium abgeschlossen.

Als Aufwandsentschädigung erhielt jeder Studienteilnehmer 10 Euro.

Studienablauf

Insgesamt dauerte es zwischen 45 und 55 Minuten, bis ein Teilnehmer die komplette Studie beendet hatte. Der Verlauf der Studie lässt sich dabei in folgende Punkte gliedern:

- **Ablaufplan und Erklärungen zur Studie:** Zu Beginn der Studie wurde jeder Teilnehmer über den Ablauf und die Rahmenbedingungen der Studie aufgeklärt. Dabei wurde den Teilnehmern eingeräumt, dass sie die Studie jederzeit abbrechen können. Ausserdem wurde auf die Anonymität der Ergebnisse durch die zufällige Vergabe von Teilnehmer IDs hingewiesen.
- **Allgemeine Informationen:** Es wurden allgemeine Informationen zur Person abgefragt. Diese betrafen Alter, Geschlecht, Fernseh- und Computerzeiten, Hobbies und Ausbildungsinformationen.
- **Sehtest:** Der bereits erwähnte Sehtest[mco12] wurde nun durchgeführt, da ab diesem Zeitpunkt die Nutzung des Laptops nötig wurde.
- **Tutorial 1:** Der Teilnehmer konnte sich dann ein Tutorial ansehen, das ihm zunächst die Fast-Forward Visualisierungen mit Videoausschnitten aus Video 1 vorstellte. In diesen Ausschnitten tauchten bereits die, als Stimuli verwendeten, Waldos vereinzelt auf. Dafür wurden die einzelnen Visualisierungen mit einer Kombination von Waldos gerendert, die nur zu Testzwecken diente und bei der eigentlichen Aufgabe keine Verwendung fand. Während des Tutorials war es dem Teilnehmer gestattet, zu jeder Zeit Fragen zu stellen. Häufig waren die Unterschiede zwischen einigen Visualisierungen nicht sofort ersichtlich. Besonders die Unterschiede zwischen Frameskip, Blending und Differenzen mussten erklärt werden. An dieser Stelle wurde den Teilnehmern das Orientierungsblatt, das nach der Pilotstudie erstellt wurde, ausgehändigt. Anhand der darauf enthaltenen Bilder und einem erneuten Blick auf die entsprechenden Folien im Tutorial konnten so die Unterschiede erklärt werden.
- **Aufgabe 1:** Die erste Aufgabe wurde den Teilnehmern dann auf entsprechenden Folien des Tutorials erklärt. Zunächst wurde die Comicfigur vorgestellt und in einem Video, das mit Normalgeschwindigkeit lief, gezeigt. Als Ziel der Aufgabe sollten die Teilnehmer jedesmal, wenn sie die Figur sehen, auf den Buzzer drücken. Dieser wurde ihnen in diesem Augenblick zum ersten Mal präsentiert, da er davor nur als mögliche Ablenkungsquelle gedient hätte. Den Teilnehmern wurde hier ebenfalls erklärt, dass ihnen 2 Sekunden zusätzliche Zeit zum Drücken des roten Buzzerknopfes zur Verfügung steht und sie ein direktes visuelles Feedback erhalten, ob ihr Knopfdruck

nun richtig war oder nicht. Im Anschluss an diese Einführung wurde das Tutorial verlassen und zu dem für die Studie entwickelten Videoplayer gewechselt. Entsprechend der Reihenfolge, die für die momentane ID festgelegt war, wurden dann die Visualisierungen für Video 1 geladen. Vor jedem Video erfuhr der Teilnehmer, welche Visualisierung er als nächstes zu sehen bekommt. Durch einen Knopfdruck auf den Buzzer konnte er dann das Video starten, sobald er sich dazu bereit fühlte. Während der einzelnen Visualisierungen war es nicht möglich, das Video zu pausieren. Kurze Pausen zwischen den Visualisierungen waren jedoch möglich, da es den Benutzern selbst überlassen war, wann sie die nächste Visualisierung starten. Nachdem alle vier Visualisierungen beendet waren, bekam der Teilnehmer den Fragebogen für Video 1 zum Ausfüllen ausgehändigt.

Dem Teilnehmer wurde die neue Szene in Video 2 anhand des Orientierungsblattes erklärt: „Da es in dieser Szene mehr Bewegung gibt und insgesamt eine größere Fläche überblickt werden muss, läuft dieses Video nicht mehr so schnell ab wie Video 1.“ Ein Hinweis, dass Video 1 mit 20-facher Beschleunigung und Video 2 mit 10-facher Beschleunigung läuft, sollte dem Teilnehmer eine Vorstellung davon geben, was ihn erwarten wird. Anschließend wurde wie bei Video 1 der Ablauf wiederholt. Entsprechend der ID wurden die Visualisierungen von Video 2 geladen, betrachtet und dann der Fragebogen ausgefüllt.

- **Aufgabe 2:** Eine Folie mit der Beschreibung der zweiten Aufgabe wurde dem Teilnehmer als nächstes präsentiert. Mittels einer zuvor erstellten Abspielliste konnten dann die Vergleichspaare der Visualisierungen nacheinander im Media Player Classic[med11] abgespielt werden. Dabei bekam der Teilnehmer keine Namen der Visualisierungen zu sehen. Diese waren lediglich mit 1 und 2 betitelt. Nach jedem Paar sollte der Teilnehmer sich dann darüber äußern, bei welchem Ausschnitt er die Bewegungen der Personen und Objekte besser verfolgen konnte. Jedes Visualisierungspaar wurde den Teilnehmern nacheinander mit einem Ausschnitt aus Video 1 und Video 2 gezeigt. Die Äußerungen des Teilnehmers wurden schriftlich auf dem dafür vorgesehenen Bogen festgehalten. Bei längerem Überlegen des Teilnehmers wurde das Video pausiert.
- **Kommentare zu den Fast-Forward Visualisierungen:** Nachdem der Teilnehmer sich durch die ersten beiden Aufgaben eine Meinung zu den Visualisierungen gebildet hatte, konnte er diese nun von sich geben. Mit speziellem Fokus auf Stärken, Schwächen und Verbesserungsideen für die einzelnen Visualisierungen wurden die Kommentare vom Studienleiter schriftlich festgehalten.
- **Tutorial 2:** Bevor die dritte Aufgabe durchgeführt werden konnte, musste dem Teilnehmer zunächst einmal erklärt werden, dass es bei der kommenden Aufgabe neue Visualisierungen geben wird, die sich mit einem anderen Aspekt des Fast-Forward in Videos auseinandersetzen. Im Tutorial wurde darauf hingewiesen, dass bei dieser Aufgabe die variierende Geschwindigkeit des Videos im Vordergrund steht, und auf welche Aspekte der Teilnehmer bei dieser Aufgabe achten muss. In den folgenden

Folien wurden dann die Geschwindigkeitsvisualisierungen in der Reihenfolge Rahmen, Störstreifen und Tachometer vorgestellt und erklärt. Hierfür wurden kleine Ausschnitte der Visualisierungen in Endlosschleifen präsentiert. Sämtliche Ausschnitte stammen aus Video 3.

- **Aufgabe 3:** Die Reihenfolge der drei Visualisierungen konnte auch in dieser Aufgabe mittels einer Abspielliste des Media Player Classic[med11] geladen werden. Während diese geladen wurde, konnte der Teilnehmer einen Blick auf den anschließenden Fragebogen (siehe Abschnitt 5.4.3) werfen, um sich zu erinnern, worauf er bei den Videos zu achten hat. Dann wurde das Video 3 in den einzelnen Visualisierungen gezeigt. Alle drei Videos liefen hintereinander ohne Pause. Im Anschluss füllte der Teilnehmer den der Aufgabe entsprechenden Fragebogen aus.
- **Kommentare zu den Geschwindigkeitsvisualisierungen:** Wie schon für die Fast-Forward Visualisierungen konnten an dieser Stelle auch Kommentare zu den soeben gesehenen Geschwindigkeitsvisualisierungen abgegeben werden. Auch hier lag der Fokus auf den Stärken und Schwächen der einzelnen Visualisierungen.
- **Geldausgabe und Beantwortung offener Frage:** Nachdem alle Aufgaben abgeschlossen waren, konnten sich die Teilnehmer in eine Liste eintragen, um den Erhalt des Geldes zu bestätigen. Dann erhielten sie ihre 10 Euro Aufwandsentschädigung. Je nach Interesse des Teilnehmers und der freien Zeit am Schluss konnten noch offene Fragen gestellt werden.

5.4. Ergebnisse

Als auszuwertende Ergebnisse wurden die Fragebögen und Ergebnisdateien von 24 Studienteilnehmern verwendet. Die Studie wurde insgesamt mit 25 Teilnehmern durchgeführt. Bei einem Teilnehmer kam es jedoch während des zweiten Videos in der ersten Aufgabe zu einer Überforderung seiner Aufmerksamkeit. Dies hatte zur Folge, dass er auf dem Fragebogen völlig undifferenziert bei allen Fragen immer dieselbe niedrige Punktzahl vergab. Aus diesem Grund wurde die Studie mit einem weiteren Teilnehmer und derselben Visualisierungsreihenfolge durchgeführt.

5.4.1. Auswertung der ersten Aufgabe

Mit Hilfe der gewonnenen Ergebnisse bei der ersten Aufgabe soll geprüft werden, wie sich die einzelnen Visualisierungen auf die Erkennungsrate der Fremdobjekte ausgewirkt hat. Zusätzlich sollen die im Fragebogen gestellten Fragen ausgewertet werden.

F-Scores

Um die Erkennungsgenauigkeit der Studienteilnehmer mit den verschiedenen Visualisierungen zu prüfen, muss ein Maß für die gespeicherten Ergebnisse verwendet werden. Bei der Erkennung der Figuren wurden richtige Treffer r , falsche Treffer f und verpasste Treffer m jedes Teilnehmers gespeichert. Anhand dieser Werte lassen sich F-Scores ermitteln, die die Genauigkeit jedes Teilnehmers durch einen Wert zwischen 0 und 1 repräsentieren.

$$Precision = \frac{r}{r + f}$$

$$Recall = \frac{r}{r + m}$$

$$F_{\beta} = (1 + \beta^2) \cdot \frac{Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \quad (5.1)$$

Wie in Gleichung (5.1)[MKSW99] zu sehen, setzt sich der Wert F aus $Precision$, $Recall$ und β zusammen. $Precision$ ergibt sich aus der Anzahl der richtigen Treffer, dividiert durch die Summe von richtigen und falschen Treffern. $Recall$ ergibt sich aus der Anzahl der richtigen Treffer, dividiert durch die Summe der richtigen und verpassten Treffer[BYRN⁺99]. Die korrekte Erkennung von den gesuchten Objekten wurde sowohl durch das Übersehen als auch

durch Verwechslungen von Objekten oder Personen beeinflusst. Aus diesem Grund sollen *Precision* und *Recall* gleich stark gewichtet werden [BYRN⁺99]. Somit wird $\beta = 1$ gewählt und es ergibt sich die Gleichung (5.2) [MKS^W99], die für die Messwerte verwendet wird. Es bleibt zu vermerken, dass in der Praxis dem Übersehen mehr Bedeutung zugesprochen werden sollte, als der Verwechslung von Personen.

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (5.2)$$

Aufgrund der unterschiedlichen Geschwindigkeit und der Bewegungsdichte werden Video 1 und Video 2 getrennt voneinander betrachtet. Abbildung 5.3 zeigt die entsprechenden Box-Plots der F1-Scores.

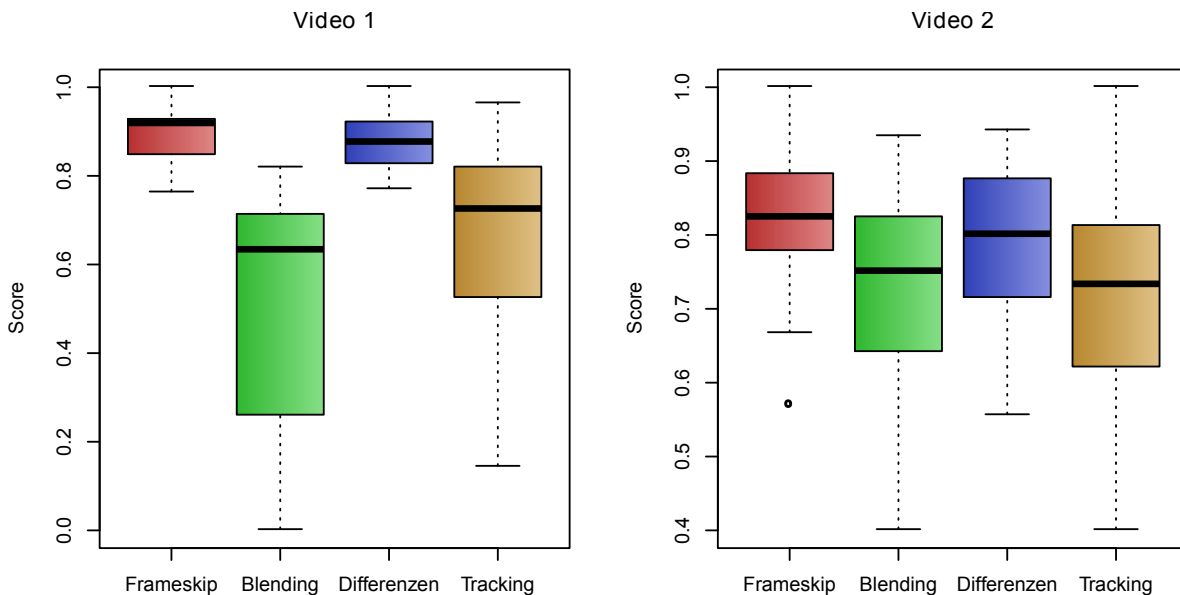


Abbildung 5.3.: F1-Scores von Video 1(links),Video 2(rechts)

Signifikante Unterschiede der Visualisierungen bezüglich der Scores konnten nachgewiesen werden (siehe Abschnitt 5.4.1). Es lässt sich beobachten, dass in beiden Videos die effektive Erkennung der Fremdobjekte durch Frameskip und Differenzen am Besten gewährleistet wird. Sowohl beim Blending als auch beim Tracking sinkt die Erkennungsrate im Vergleich dazu. In Video 1, dessen Beschleunigung sehr hoch war, schneidet besonders Blending im Vergleich zu den anderen Methoden um einiges schlechter ab. Der hohe Quartilabstand ist beim Blending und beim Tracking durch die permutierte Reihenfolge der Videos zu erklären. So waren besonders in Video 1 die Personen, welche Blending oder Tracking als erste Methode präsentiert bekamen, völlig von der Aufgabe überfordert. Dieser Effekt schlägt sich auf die Verteilung der Werte nieder (siehe Abbildung A.1). Bei einer geringeren Geschwindigkeit, wie sich in Video 2 beobachten lässt, nähern sich die Ergebnisse langsam an.

5. Studie

Was die effektive Erkennung von Objekten betrifft, so ist die Frameskip-Methode am besten. Die Differenzen-Methode konnte leider keine Steigerung erreichen, wirkt sich aber auch nicht negativ darauf aus. Die Ergebnisse der Tracking-Methode deuten darauf hin, dass die zusätzlichen Informationen zu Ablenkungen geführt haben. Je geringer die Beschleunigung des Videos wird, desto weniger Bewegungsunschärfe ist beim Blending vorhanden. Dadurch lässt sich eine Verbesserung der Ergebnisse zwischen Video 1 und Video 2 erklären.

Signifikanztests

Um signifikante Unterschiede der Ergebnisse nachweisen zu können, muss zunächst geprüft werden, welche statistischen Testverfahren bei den gemessenen Werten verwendet werden können. Zur Durchführung einer einfaktoriellen Varianzanalyse (ANOVA) müssten folgende Bedingungen erfüllt sein[Zöfo3]:

- Normalverteilung in den Stichproben
- Varianzhomogenität über die Stichproben hinweg

Aus diesem Grund müssen die vorhandenen Ergebnisse zunächst einmal auf ihre Verteilungsform innerhalb der einzelnen Visualisierungen getestet werden. Hierfür wurde der Shapiro-Wilk Test verwendet. Dieser prüft, ob sich die gegebenen Werte mit dem Modell einer Normalverteilung beschreiben lassen und zeichnet sich durch eine hohe Zuverlässigkeit aus[SHo6].

Die Testergebnisse (in R mittels Funktion *shapiro.test()* berechnet) sind in den Tabellen 5.1 und 5.2 für die einzelnen Videos zu sehen. Bei einem Signifikanzniveau von 5% bedeutet dies für die Ergebnisse, wenn der Test einen p-Wert < 0.05 ergibt, dass es sich um keine Normalverteilung handelt. Die betroffenen Werte wurden rot markiert.

	Frameskip	Blending	Differenzen	Tracking
W	0.9526	0.8398	0.96	0.9056
p	0.3086	0.001422	0.4382	0.02838

Tabelle 5.1.: Aufgabe 1: Shapiro-Wilk Test Video 1

	Frameskip	Blending	Differenzen	Tracking
W	0.9642	0.8914	0.9472	0.9815
p	0.5294	0.0142	0.2358	0.9217

Tabelle 5.2.: Aufgabe 1: Shapiro-Wilk Test Video 2

In Video 1 liegen die p-Werte für Blending und Tracking weit unter dem Signifikanzniveau ($p = 0.05$). In Video 2 ist dies beim Blending auch der Fall. In diesen Fällen kann davon ausgegangen werden, dass es sich nicht um eine Normalverteilung handelt. Ein Blick auf die entsprechenden Histogramme (Abbildung A.1) zeigt, wie bereits erwähnt, dass sich die Werte bei beiden Visualisierungen in zwei Bereichen häufen.

Einmal die Teilnehmer, die Blending oder Tracking als eine der ersten Visualisierungen gesehen haben und entsprechend schlechte Ergebnisse erzielt haben. Die zweite Häufung bei den besseren Ergebnissen kommt durch die Teilnehmer zustande, die Blending oder Tracking zuletzt gesehen haben. Diese kommen besonders beim Blending durch die gesammelte Erfahrung in den anderen Visualisierungen zustande. Leider hat sich dieser Effekt nicht durch die Permutation der Reihenfolge ausgleichen lassen. Aus diesem Grund konnte keine Normalverteilung erreicht werden. Bezüglich der F1-Scores lassen sich nun folgende Voraussetzungen für einen Test zusammenfassen:

- Unabhängige Stichproben: 4 Visualisierungen wurden nacheinander an einem Teilnehmer hinsichtlich ihrer Effektivität getestet. Durch die Permutation der Reihenfolge kann insgesamt von einer Unabhängigkeit der Stichproben ausgegangen werden.
- Die vorliegenden Werte sind ordinalskaliert.
- Es liegt keine Normalverteilung der Werte vor.
- Jede Stichprobe ist in sich unabhängig, da jeder Teilnehmer nur einmal pro Visualisierung gemessen wurde.

Da nicht alle Werte einer Normalverteilung folgen, ist die Durchführung einer ANOVA nicht möglich. Stattdessen bedarf es in diesem Fall eines parameterfreien Test, der unter den gegebenen Voraussetzungen anwendbar ist. Der H-Test nach Kruskal und Wallis [Zöfo3] erweist sich dabei als der Test, der am häufigsten in diesem Fall verwendet wird. Dieser benötigt keine Normalverteilung der Werte und dient zum Vergleich der zentralen Tendenzen von mehr als zwei Stichproben. Der Test prüft die Hypothese, die Stichproben entstammen der selben Grundgesamtheit. P-Werte unterhalb des Signifikanzniveaus von 5% verwerfen diese Hypothese, was bedeutet, dass mindestens 2 der Stichproben sich voneinander unterscheiden [SHo6]. Die Auswertung mittels H-Test (in R mittels `kruskal.test()`) ergab signifikante Unterschiede (Video 1: $\chi^2 = 55.5964, df = 3, p - value = 5.122e - 12$; Video 2: $\chi^2 = 10.2893, df = 3, p - value = 0.01626$) zwischen den Visualisierungen.

Da durch den H-Test davon ausgegangen wird, dass sich mindestens zwei der Visualisierungen signifikant unterschiedliche Werte geliefert haben, lassen sich nun Post-Hoc-Tests durchführen. Diese vergleichen die einzelnen Visualisierungen paarweise miteinander auf signifikante Unterschiede. Der U-Test nach Mann und Whitney wurde für die gegebenen Voraussetzungen als geeignet befunden (in R mittels `pairwise.wilcox.test()`). Dabei wurden die approximierten Ergebnisse mit Bonferroni-Korrektur für multiple Vergleiche [Abdo7], verwendet.

	Frameskip	Blending	Differenzen
Blending	2.7e-08	-	-
Differenzen	1.000	6.9e-08	-
Tracking	3.2e-05	0.13866	0.00036

Tabelle 5.3.: p-Werte des U-Test Video 1

	Frameskip	Blending	Differenzen
Blending	0.038	-	-
Differenzen	1.000	0.402	-
Tracking	0.088	1.000	0.741

Tabelle 5.4.: p-Werte des U-Test Video 2

Wie aus den Tabellen 5.3 und 5.4 ersichtlich wird, konnten im ersten Video signifikante Unterschiede zwischen fast allen Visualisierungen nachgewiesen werden. Lediglich zwischen Frameskip mit Differenzen und Blending mit Tracking konnte mittels U-Test kein Unterschied nachgewiesen werden. Im zweiten Video lässt sich dagegen nur noch zwischen Frameskip und Blending ein signifikanter Unterschied nachweisen. Diese Ergebnisse sprechen dafür, dass bei einer niedrigeren Geschwindigkeit wie in Video 2 noch keine starke Beeinflussung des Benutzers durch die Visualisierungen stattfindet. Wird die Geschwindigkeit und damit der Schwierigkeitsgrad erhöht, so wird ersichtlich, dass Frameskip und Differenzen sich von Blending und Tracking, was die effektive Erkennung von Fremdobjekten betrifft, deutlich abhebt.

Auswertung der Fragebögen

Für Untersuchung subjektiven Eindrücke wurden nach Video 1 und Video 2 Fragebögen an die Studienteilnehmer ausgeteilt. Folgende Punkte galt es zu beantworten:

1. **Erkennung:** Mit dieser Visualisierung konnte ich die gesuchten Objekte gut erkennen.
2. **Aufmerksamkeit:** Diese Visualisierung benötigte viel Aufmerksamkeit/Anstrengung zum Lösen der Aufgabe.
3. **Angenehm:** Ich empfand es angenehm, diese Visualisierung zu verwenden.
4. **Hilfreich:** Diese Visualisierung hat mir geholfen, die Aufgabe gut zu lösen.

Diese Punkte konnten auf einer Likert-Skala bewertet werden. Dabei entsprach der Wert 1 der Aussage „Stimmt nicht“ und der Wert 10 der Aussage „Stimmt“.

Bei der Analyse der Antworten, konnte mittels Shapiro-Wilk Test nicht in allen Fällen eine Normalverteilung nachgewiesen werden. Aus diesem Grund wurde, wie bereits bei den F1-Scores, für die einzelnen Fragen der H-Test durchgeführt. Bei allen Fragen zu Video 1 und Video 2 konnten signifikante Unterschiede in der Beurteilung der Methoden nachgewiesen werden (siehe Tabellen A.1 und A.3). Deshalb wurden die Methoden wieder post-hoc mittels U-Test untersucht (siehe Tabellen A.2 und A.4). Die Ergebnisse lassen sich folgendermaßen zusammenfassen:

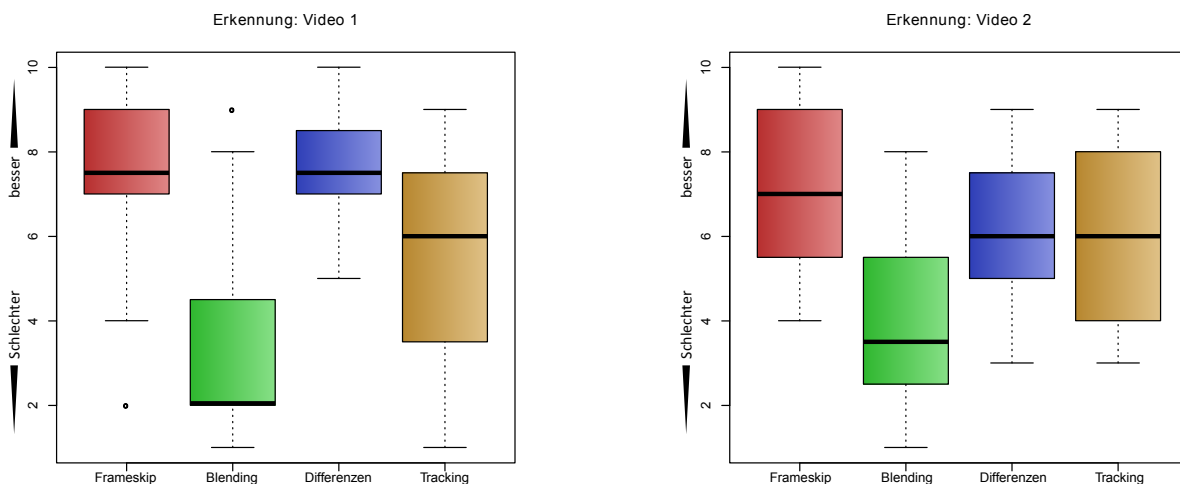


Abbildung 5.4.: Aufgabe 1: Erkennung

- **Erkennung:**

- Video 1: Einzig bei der Bewertung von Frameskip und Differenzen konnte kein signifikanter Unterschied nachgewiesen werden. Die Verteilung der Bewertung für diese beiden Methoden ist auch sehr ähnlich (siehe Abbildung A.2). Nach der Meinung der Studienteilnehmer konnten sie mit diesen beiden Methoden auch die Objekte besser erkennen als beim Tracking und Blending. Besonders schlecht wurde die Erkennung beim Blending befunden.
- Video 2: Hier wurden signifikante Unterschiede nur im Vergleich von Blending mit allen anderen Methoden festgestellt. Es lässt sich eine Angleichung der Bewertungen erkennen, wie sie sich auch schon bei der Betrachtung der F1-Scores zu diesem Video zeigte. Dies lässt darauf schließen, dass bei der geringeren Geschwindigkeit die Erkennung nicht mehr so stark von den einzelnen Methoden abhängt. Einzig Blending wurde im Vergleich zu den anderen Methoden deutlich schlechter bewertet.

- **Aufmerksamkeit:**

- Video 1: Insgesamt konnte nur zwischen Blending und Differenzen ein signifikanter Unterschied bei der Bewertung nachgewiesen werden. Blending wurde

5. Studie

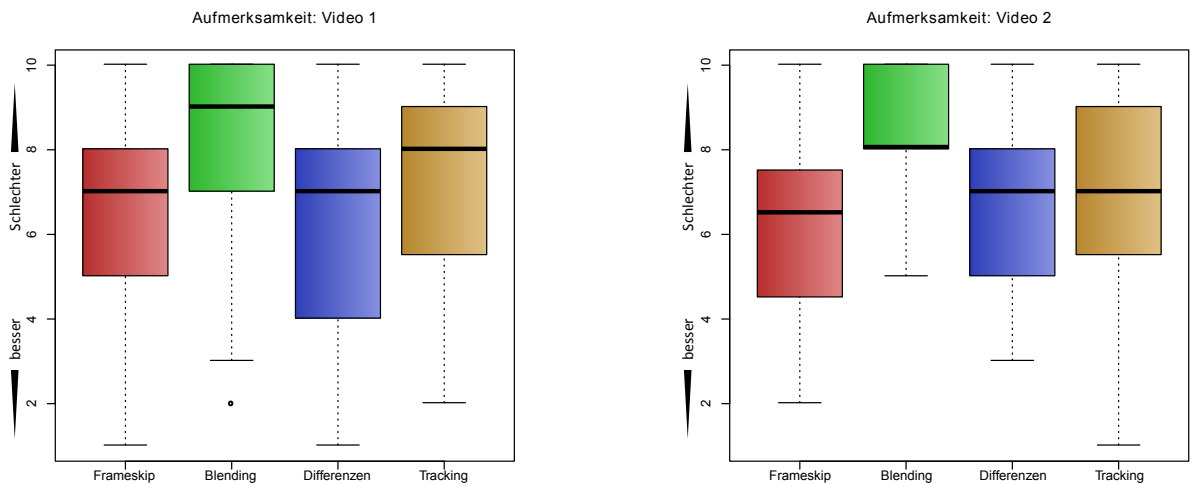


Abbildung 5.5.: Aufgabe 1: Aufmerksamkeit

hierbei wieder insgesamt am schlechtesten bewertet. Tracking benötigte auch leicht mehr Aufmerksamkeit als Frameskip und Differenzen, welche wieder sehr ähnlich bewertet wurden (siehe auch Abbildung A.3).

- Video 2: Beim Blending wurde im Vergleich mit allen anderen Methoden ein signifikanter Unterschied nachgewiesen. Bei den restlichen Methoden konnte kein signifikanter Unterschied bei der Bewertung nachgewiesen werden. Der Boxplot ist dem von Video 1 sehr ähnlich. Blending benötigt hierbei wieder die meiste Aufmerksamkeit.

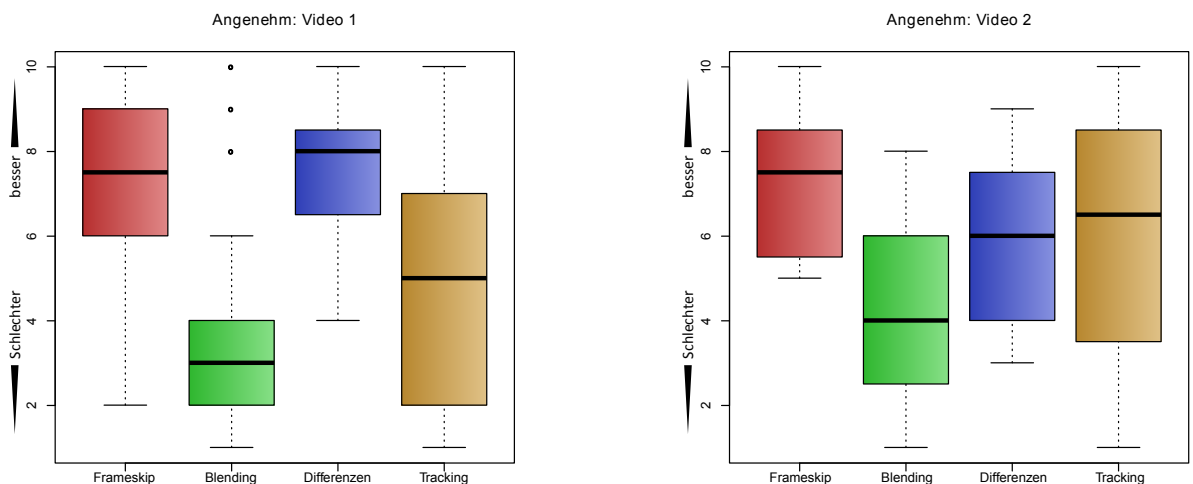


Abbildung 5.6.: Aufgabe 1: Angenehm

- **Angenehm:**

- Video 1: Keine signifikanten Unterschiede gab es hier bei Frameskip mit Differenzen und Blending mit Tracking. Dabei wurden Frameskip und Tracking am angenehmsten empfunden, während Blending am unangenehmsten war, abgesehen von 3 Ausreißern. Die Streuung der Werte beim Tracking ist stärker als bei den anderen Methoden. Ein Blick auf das entsprechende Histogramm (siehe Abbildung A.4) zeigt, dass es beim Tracking eine Häufung bei den schlechteren Bewertungen gab, während Werte größer als 3 eher gleichverteilt vergeben wurden.
- Video 2: Blending unterschied sich zu Frameskip und Differenzen. Beim Vergleich von Blending mit Tracking konnte dagegen kein signifikanter Unterschied festgestellt werden. Auch hier lässt sich wieder eine Angleichung der Werte zwischen den Visualisierungen beobachten. Frameskip wurde insgesamt leicht besser bewertet als die anderen Methoden. Blending und Tracking wurden im Vergleich zu Video 1 nicht mehr so unangenehm empfunden. Es lässt sich beobachten, dass Tracking in Video 2 mindestens genauso angenehm empfunden wurde, wie die Differenzen. Eine deutliche Verbesserung gegenüber Video 1.

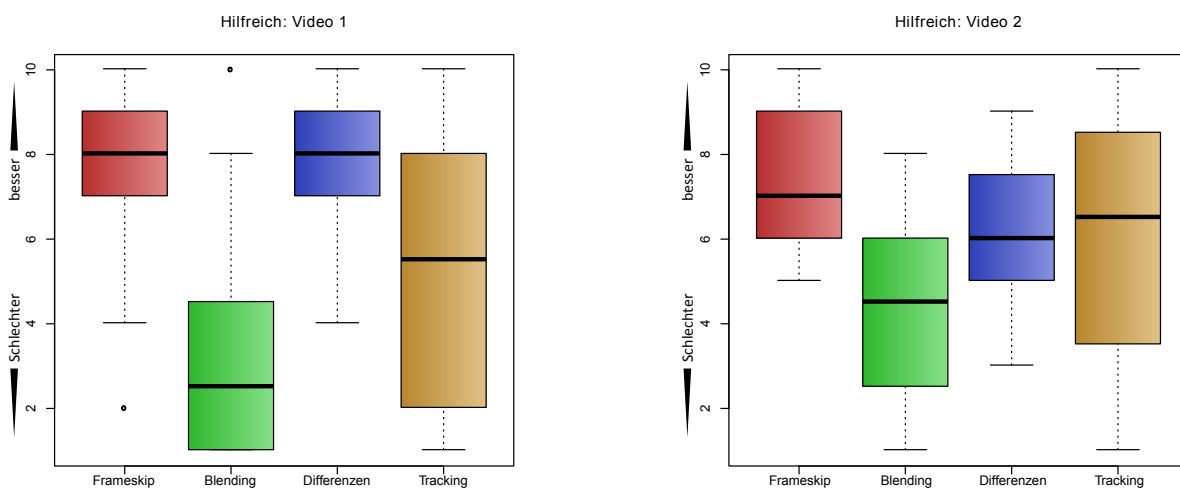


Abbildung 5.7.: Aufgabe 1: Hilfreich

- **Hilfreich:**

- Video 1: Auch hierbei gab es nur zwischen Frameskip mit Differenzen und Blending mit Tracking keine signifikanten Unterschiede. Die Bewertung entspricht in etwa der Erkennung. Blending bleibt die am wenigsten hilfreiche Methode und Tracking sorgt für eine geteilte Meinung bei den Teilnehmern (siehe Abbildung A.5).
- Video 2: Blending unterschied sich zu Frameskip und Differenzen. Beim Vergleich von Blending mit Tracking konnte dagegen kein Unterschied festgestellt werden.

Auch hier wurde Blending als am wenigsten hilfreich bewertet. In Video 1 waren Frameskip und Differenzen noch gleich bewertet worden, in Video 2 wurde Frameskip dagegen als etwas hilfreicher eingeschätzt.

Bei der Frage, welche der Visualisierungen die Teilnehmer für die Durchführung der Aufgabe bevorzugten, entschieden sich 46% für Frameskip, 25% für Differenzen, 22% für Tracking und 7% für Blending bei Video 1. Bei Video 2 entschieden sich 38% für Frameskip, 35% für Tracking, 19% für Differenzen und 8% für Blending.

Nach der Auswertung der Ergebnisse lässt sich rückblickend über die anfangs aufgestellten Hypothesen (siehe Abschnitt 5.1) folgendes sagen:

- **Hypothese 1:** Wie erwartet konnten die Objekte in beiden Videos beim Blending am schlechtesten erkannt werden. In Video 2 war Tracking zwar noch etwas schlechter als Blending, jedoch ohne signifikanten Unterschied.
- **Hypothese 2:** Die Ergebnisse bei Frameskip und Differenzen unterscheiden sich in beiden Videos nicht signifikant voneinander. Eine Verbesserung durch die mehrfache Darstellung der Objekte mittels Geisterbilder brachte keine Verbesserung.
- **Hypothese 3:** Die Ergebnisse beim Tracking sind bei Video 1 besser ausgefallen als in Video 2. Da die Geschwindigkeit in Video 2 geringer war, liegt der Grund für die Verschlechterung der Ergebnisse wahrscheinlich in der hohen Anzahl von Objekten.

Freie Kommentare

Folgende Kommentare wurden von den Studienteilnehmern zu den Methoden am häufigsten geäußert:

- „Beim Frameskip sind schnelle Bewegungen schwer zu erkennen.“
- „Frameskip ist am wenigsten überladen mit Informationen und deshalb am besten.“
- „Blending hatte zu unscharfe Bilder um etwas erkennen zu können.“
- „Bei den Differenzen bleiben die Objekte gut erkennbar, die Spur aus Geisterbildern hilft bei der Erkennung.“
- „Die Artefakte bei den Differenzen stören.“
- „Die Pfeile beim Tracking helfen dabei die Bewegungen zu verfolgen.“
- „Die Pfeile heben die Objekte hervor für eine schnellere Erkennung.“
- „Die Schweife beim Tracking erschweren die Erkennung.“
- „Tracking enthält manchmal zuviel Information auf einmal.“

- „Tracking ist gut wenn nicht zuviel los ist.“

5.4.2. Auswertung der zweiten Aufgabe

Die Ergebnisse der zweiten Aufgabe liegen in Form von Vergleichen der einzelnen Visualisierungen untereinander vor. Dabei wurde von den Studienteilnehmern für jeden Vergleich angegeben, bei welcher der Visualisierungen sie Objekte besser verfolgen konnten, oder ob sie unentschieden waren. Bei 4 Visualisierungen mussten $\binom{4}{2} = 6$ Vergleiche durchgeführt werden. Die Ergebnisse werden in den folgenden Tabellen zusammengefasst. Dabei wurde beim Vergleich von Zeile zu Spalte immer ein Punkt vergeben, wenn die Zeile bevorzugt wurde, ein halber Punkt bei Unentschieden und kein Punkt, wenn die Spalte bevorzugt wurde [BLBo8]. Die Summe der Zeilen ergibt dann eine Gesamtpunktzahl der einzelnen Visualisierungen, womit sich ein Ranking formulieren lässt. Zusätzlich werden die Rankings von Video 1 und Video 2 in Abbildung A.6 veranschaulicht.

	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	17	14.5	11	42.5
Blending	7	0	7	5	19
Differenzen	9.5	17	0	7.5	34
Tracking	13	19	16.5	0	48.5

Tabelle 5.5.: Vergleichspaare Video 1

Für Video 1 wurde also Tracking am besten für die Verfolgung von Bewegungen empfunden. Frameskip folgt mit einem Abstand von 6 Punkten. An dritter Stelle befindet sich die Differenzen-Methode mit einem Abstand zu Frameskip von 8.5 Punkten. Den letzten Platz erhält Blending mit einem deutlicheren Abstand von 15 Punkten.

	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	14	16	14	44
Blending	10	0	13	9.5	32.5
Differenzen	8	11	0	6	25
Tracking	10	14.5	18	0	42.5

Tabelle 5.6.: Vergleichspaare Video 2

In Video 2 erhält Frameskip mit 44 Punkten den ersten Platz. Mit nur 1.5 Punkten Abstand folgt Tracking. Die Differenzen-Methode folgt mit einem Abstand von 10 Punkten zum Tracking. Auf dem letzten Platz befindet sich wieder Tracking mit 7.5 Punkten Abstand zu Differenzen.

5. Studie

	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	31	30,5	25	86,5
Blending	17	0	20	14,5	51,5
Differenzen	17,5	28	0	13,5	59
Tracking	23	33,5	34,5	0	91

Tabelle 5.7.: Vergleichspaare Gesamt

Für einen Gesamteindruck der Bewegungsverfolgung wurden die Bewertungen der einzelnen Videos zusammengerechnet. Dabei kam das endgültige Ranking von

$$\textit{Tracking} > \textit{Frameskip} > \textit{Differenzen} > \textit{Blending}$$

zustande. Abbildung 5.8 zeigt dieses Ranking mit den relativen Häufigkeiten der Siege, Unentschieden und Niederlagen.

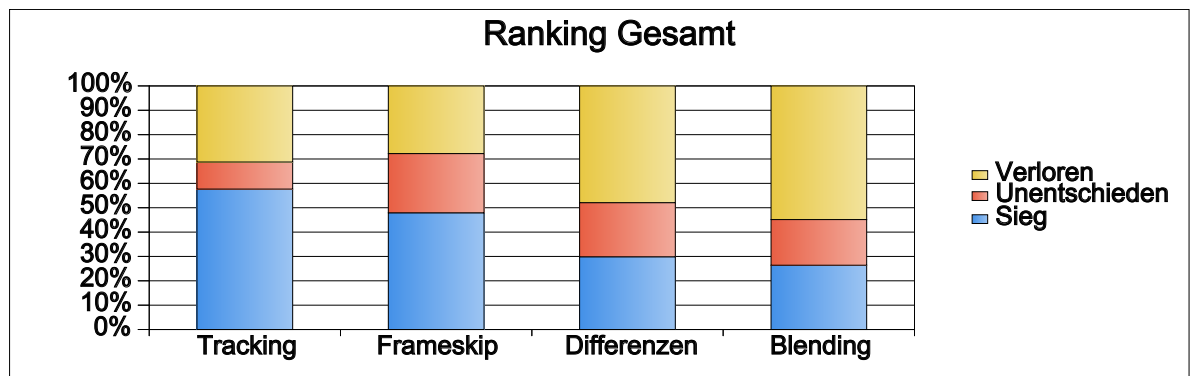


Abbildung 5.8.: Ranking der Visualisierungen

Das schlechte Abschneiden von Blending ist hierbei etwas überraschend, da die leichtere Verfolgung von Bewegungen ein Vorteil dieser Methode sein sollte. Die Analyse der Vergleichspaare nach Kendall[BLBo8] (siehe Abbildung A.7) ergab insgesamt, dass die Paarvergleiche überzufällig übereinstimmend waren ($p < 0.05$), jedoch in einem sehr geringen Maß ($A = 0.07$).

Wie in **Hypothese 4** (siehe Abschnitt 5.1.2) prognostiziert, liegt Tracking insgesamt vor der Frameskip-Methode. Differenzen und Blending wurden bei dieser Aufgabe nicht wie erwartet der Frameskip-Methode vorgezogen.

Freie Kommentare

Folgende Kommentare wurden von den Studienteilnehmern zur Bewegungsverfolgung am häufigsten geäußert:

- „Beim Frameskip konnten die Objekte am besten verfolgt werden.“
- „Beim Frameskip war die Verfolgung zu ruckelig.“
- „Beim Blending hilft der Verlauf der Bewegungsunschärfe bei der Verfolgung.“
- „Bei der Differenzen-Methode helfen die Geisterbilder bei der Verfolgung.“
- „Beim Tracking haben die Pfeile und Schweife geholfen, die Bewegungen zu verfolgen.“

5.4.3. Auswertung der dritten Aufgabe

Für die dritte Aufgabe wurde ein Fragebogen verwendet, der folgende Punkte enthielt:

- **Effektivität:** Diese Visualisierung hat mir geholfen, den Geschwindigkeitswechsel besser einzuschätzen.
- **Beanspruchung:** Ich empfand es anstrengend, die Visualisierung anzusehen.
- **Ablenkung:** Meine Aufmerksamkeit wurde durch die Visualisierung vom Inhalt des Videos abgelenkt.

Diese Punkte waren wieder mit einer Likert-Skala von 1 bis 10, wie in der ersten Aufgabe, für die drei Visualisierungen Tachometer, Rahmen und Störstreifen zu beurteilen. Dabei unterschied sich jeder Fragebogen durch die individuelle Reihenfolge der Visualisierungen, in der sie die Teilnehmer gesehen hatten. Statistisch signifikante Unterschiede der Werte bei den einzelnen Fragen konnten mittels H-Test nicht nachgewiesen werden. Folglich konnten keine Post-Hoc-Tests durchgeführt werden.

- **Effektivität:** $\chi^2 = 1.8808, df = 2, p - value = 0.3905$
- **Beanspruchung:** $\chi^2 = 2.1723, df = 2, p - value = 0.3375$
- **Ablenkung:** $\chi^2 = 2.4677, df = 2, p - value = 0.2912$

Die Ergebnisse werden in Abbildung 5.9 zusammengefasst. Zu den einzelnen Fragen lässt sich folgendes anmerken:

- **Effektivität:** Was die Einschätzung des Geschwindigkeitswechsels angeht, so wurden die Störstreifen leicht besser empfunden als der Tachometer und der Rahmen. Wobei die Streuung der Werte für den Tachometer geringer ist, also entsprechend mehr Leute ihre Bewertung in der Nähe des Medians abgegeben haben. Die Störstreifen erhalten eine höhere Streuung. Es gab Teilnehmer, die diese Streifen als sehr störend empfanden und dadurch entsprechend niedrigere Bewertungen abgegeben haben.
- **Beanspruchung:** Bei der Beanspruchung konnten die Störstreifen ebenfalls leicht besser abschneiden als die beiden anderen Methoden. Auffällig ist hierbei wieder, dass Rahmen und Störstreifen eine höhere Streuung der Werte aufweisen, als der Tachometer.
- **Ablenkung:** Der Tachometer wird hier insgesamt schlechter bewertet als die beiden anderen Methoden. Rahmen und Störstreifen sind hier recht gleich bewertet worden. Die Streuung der Werte ist bei allen drei Methoden ziemlich gleich.

Die hohen Interquartilabstände sind ein Maß für die hohe Streuung der Werte. Ein zusätzlicher Blick auf die Histogramme (siehe A.8) lässt besonders bei der Ablenkung eine Gleichverteilung der Werte vermuten. Insgesamt lässt sich zumindest sagen, dass die Rahmen im Vergleich zu dem Tachometer und den Störstreifen schlechter abgeschnitten haben. Die statistische Signifikanz dieser Ergebnisse konnte leider nicht nachgewiesen werden.

Die Teilnehmer sollten am Schluss noch unter der Fragestellung, durch welche Visualisierung die Geschwindigkeit am besten vermittelt wird, eine Wahl treffen. 40% der Teilnehmer entschieden sich für den Tachometer, 36% für die Störstreifen und 24% für den Rahmen.

Hypothese 5 (siehe Abschnitt 5.1.3) hat den Tachometer als den Sieger bei der Vermittlung der Geschwindigkeit prognostiziert. Dies lässt sich ohne statistische Signifikanz der Ergebnisse jedoch nicht sicher bestätigen. Überraschenderweise liegen die Störstreifen ziemlich gleich mit dem Tachometer.

Hypothese 6 prognostizierte eine stärkere Ablenkung durch den Tachometer als durch die Rahmen und die Störstreifen. Die entstandenen Ergebnisse lassen die Vermutung zu, dass diese Aussage zutrifft.

Freie Kommentare

Folgende Kommentare wurden von den Studienteilnehmern zu den Geschwindigkeitsvisualisierungen am häufigsten geäußert:

- „Die Einschätzung der Geschwindigkeit ist beim Tachometer sehr gut.“
- „Der Tachometer lenkt zu sehr vom Inhalt ab.“

- „Der Tachometer ist zu klein und zieht den Blick immer in die untere Ecke.“
- „Der Rahmen ist immer sichtbar und benötigt keine so genaue Beobachtung wie der Tacho.“
- „Die Farbwechsel des Rahmens stören zu sehr, da es zu sehr blinkt.“
- „Die Störstreifen im Bild lenken nicht den Blick ab und stören nicht.“
- „Die Störstreifen lassen die Geschwindigkeit nicht gut erkennen.“
- „Der Geschwindigkeitswechsel bei den Störstreifen wird zu langsam bemerkt.“

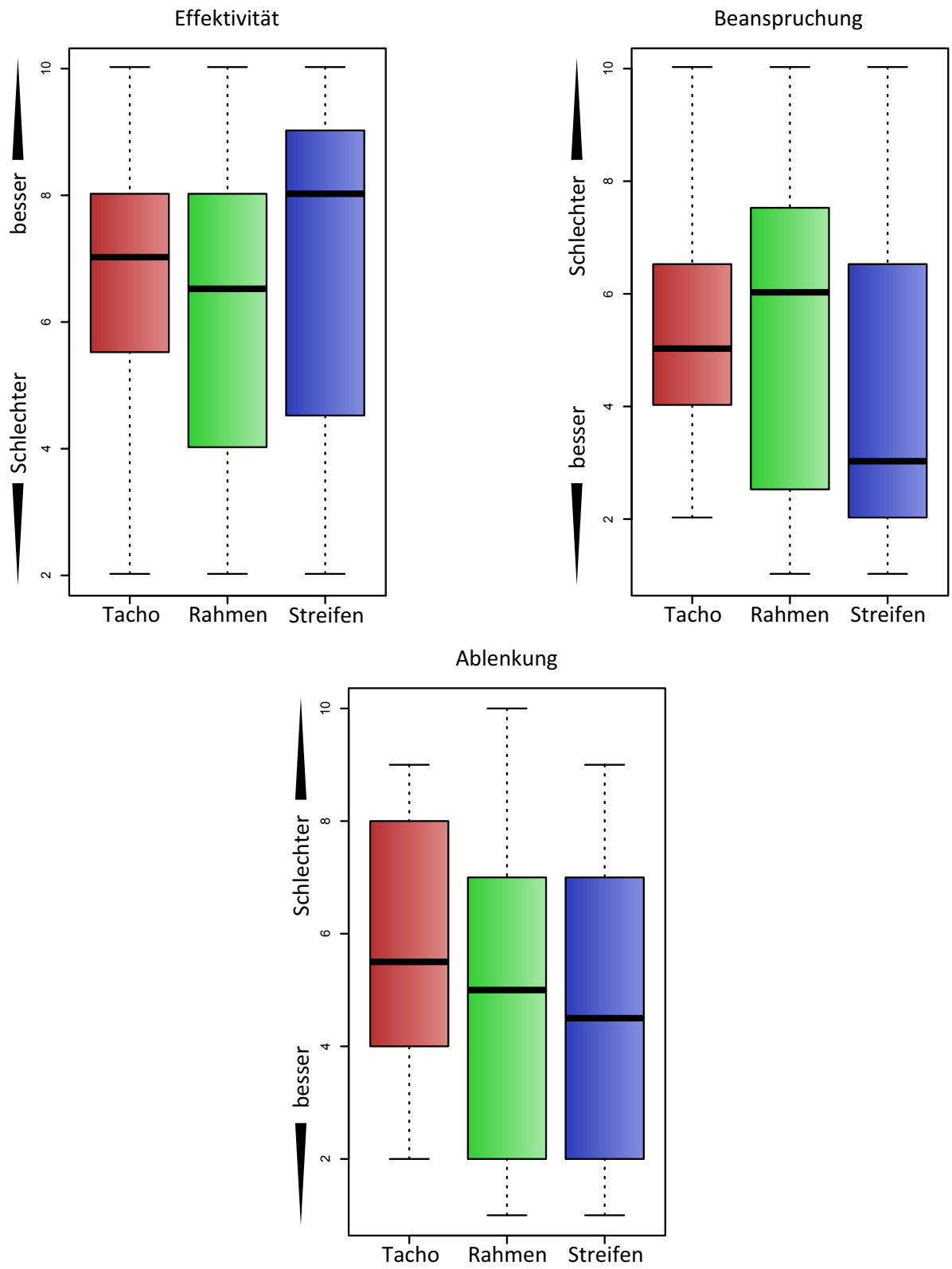


Abbildung 5.9.: Boxplots der Fragen zu Aufgabe 3

Probleme und Verbesserungen

6.1. Fast-Forward Visualisierungen

Einige der Probleme von Fast-Forward Visualisierungen waren bereits im Vorfeld bekannt. So waren die Probleme der Frameskip-Methode überhaupt erst der Grund für die Entwicklung der alternativen Methoden, die in dieser Arbeit vorgestellt wurden. Auch die Probleme der Blending-Methode waren bereits im voraus abzuschätzen. Die Probleme der Differenzen- und Tracking-Methode ergaben sich vor allem aus den Ergebnissen der Benutzerstudie und den Kommentaren der einzelnen Teilnehmer. In folgenden Abschnitten soll noch einmal zusammengefasst dargelegt werden, welche Probleme während der Arbeit erkennbar wurden und wie sich diese verringern oder gar beseitigen lassen.

Frameskip

Wie bereits in Kapitel 1.1 erwähnt, liegen die Probleme der Frameskip-Methode im Überspringen der Bilder. Mit zunehmender Geschwindigkeit werden immer mehr Bilder ausgelassen. Dies hat zur Folge, dass die Sprünge auch im Bild sichtbar werden. Beobachtet man beispielsweise eine Person, die sich durch das Bild bewegt, so ändert sich deren Position mit zunehmender Geschwindigkeit immer sprunghafter, was eine genauere Beobachtung und Verfolgung der Bewegung erschweren kann und „Change Blindness“ hervorruft.

Blending

Die Blending-Methode führt insbesondere bei höheren Geschwindigkeiten zu einer viel zu starken Unschärfe der Objekte. Wie es sich in der Studie gezeigt hat, ist eine 20-fache Beschleunigung bereits ausreichend, um eine durchgehende Erkennung der Objekte unmöglich zu machen. Lediglich Objekte, die zeitweise an einer Stelle stehen bleiben,



Abbildung 6.1.: Je höher die Geschwindigkeit, desto größer sind die Positionssprünge der Personen von Bild zu Bild. Hier sind 3 Folgebilder bei 16-facher Beschleunigung zu sehen.

können auf diese Weise noch erkannt werden. Die verwendeten Fremdobjekte befanden sich dagegen dauerhaft in Bewegung durch das Bild und konnten nur sehr schlecht erkannt werden. Besonders die Studienteilnehmer, die das Blendingvideo als erstes gesehen hatten, konnten nur einen Bruchteil der vorkommenden Fremdobjekte erkennen.

Die Ergebnisse des zweiten Videos deuten darauf hin, dass eine 10-fache Beschleunigung langsam die Grenze für eine sinnvolle Verwendung der Blending-Methode erreicht. Hier beginnt die effektive Erkennung bereits abzunehmen, die Ergebnisse fallen aber nicht so schlecht aus, wie bei der doppelten Beschleunigung. Eine Verbesserung der Blending-Methode wurde mit den Differenzen versucht. Ein sinnvoller Einsatz der Methode bleibt auf niedrigere Geschwindigkeiten beschränkt.

Differenzen

Das Hauptproblem der Differenzen-Methode liegt in der Bildung von Artefakten. Diese kommen zum einen durch ungenaue Differenzbilder zustande, welche durch die Kernelgröße der morphologischen Operatoren eine Blockform erhalten. Durch die Aufhellung des Bewegungsschweifs werden diese Artefakte noch hervorgehoben und fallen besonders auf, wenn Objekte ihre Geschwindigkeit reduzieren, oder ganz zum Stehen kommen. Solange das Objekt in diesem Moment noch nicht in den Hintergrund aufgenommen wurde, sind die genannten Artefakte hinter dem Objekt zu sehen.

Besonders macht sich dies in Video 2 bemerkbar, wenn sich auf der Straße der Verkehr staut. Hier bewegen sich die Objekte nur noch in Schrittgeschwindigkeit und halten immer wieder an. Diese Zeit reicht nicht aus, um die Objekte in den Hintergrund miteinzubeziehen und sie somit nicht mehr als Differenz zu erkennen.



Abbildung 6.2.: Die Differenzen-Methode erzeugt bei langsameren Abspielgeschwindigkeiten häufiger unschöne Artefakte.

Durch präzisere Verfahren zur Trennung von Vorder- und Hintergrund könnte sich dieses Problem verringern lassen. Eine möglichst genaue Erkennung der Silhouette von Objekten würde verhindern, dass Pixel aufgehellt werden, die gar nicht zu dem Objekt gehören. Angesichts der Ergebnisse der Studie, scheint die Aufhellung für eine bessere Verfolgung der Bewegungen keinen großen Einfluss auf die Teilnehmer gehabt zu haben. Daher könnte man für die Verbesserung der Differenzen-Methode auf das Verrechnen der Bewegungsschweife verzichten und das Teilergebnis verwenden, in dem nur die Geisterbilder vorkommen. Dabei fallen Fehler in den Differenzbildern nicht sehr stark auf und die hellen Artefakte kommen nicht mehr zustande.

Tracking

Probleme treten bei der Tracking-Methode auf, wenn in der verwendeten Szene sehr viele Objekte zu sehen sind. Dadurch kommt es häufiger vor, dass einige Objekte sich räumlich nahe beieinander befinden. Werden nun beide Objekte mit Pfeilen und Schweifen versehen, so kann es zu Überdeckungen kommen, die Erkennungsaufgaben erschweren.

Bei der Studie ergab es sich ebenfalls, dass einige Teilnehmer von der Visualisierung so abgelenkt waren, dass sie gar nicht auf das Objekt achteten, welches sich zwischen Pfeil und



Abbildung 6.3.: Überdeckungen von Personen durch die Schweife können diese schwer erkennbar machen.

Schweif befand. Insgesamt lässt sich sagen, dass besonders die verwendeten weißen Schweife viel zu auffällig sind. Eine sehr starke Erhöhung der Transparenz wäre eine Möglichkeit dem entgegen zu wirken. Es wäre ebenfalls möglich, auf den Schweif zu verzichten und stattdessen den Pfeil auf die hinteren Punkte zu verlängern, wie es in der Entwicklung bereits implementiert wurde. Von vielen Studienteilnehmern wurde die Markierung der Objekte durch den Pfeil als positiv bewertet. Hier wäre noch eine Untersuchung interessant, wie unterschiedliche Markierungen der Objekte sich im Vergleich zu den Pfeilen auf die Erkennung auswirken.

6.2. Geschwindigkeitsvisualisierungen

Die Probleme der einzelnen Geschwindigkeitsvisualisierungen ergeben sich vor allem durch die Kommentare der Studienteilnehmer und ihre Bewertungen.

Tachometer

Trotz seiner präzisen Geschwindigkeitsangabe und der Möglichkeit, die Geschwindigkeit in Relation zu minimalem und maximalem Wert zu sehen, hat der Tachometer ein großes

Problem: Die Ablenkung. Um genauere Informationen über die momentane Geschwindigkeit zu erhalten, musste der Tachometer von den Studienteilnehmern fokussiert werden. Viele Teilnehmer beklagten dabei, dass sie die Szene nicht mehr richtig beobachten konnten und ihre Aufmerksamkeit immer zwischen Szene und Tachometer wechselte.

Da dieses Problem bereits im Vorfeld erahnt wurde, findet im Hintergrund des Tachometers eine Einfärbung der Fläche in der Farbe der momentanen Geschwindigkeit auf der Skala statt. Auf diese Weise sollte der Teilnehmer schneller durch die Farbe über die Geschwindigkeit informiert werden, ohne dabei die Szene lange aus dem Fokus zu verlieren. Für die genaueren Informationen, die diese Visualisierung als einzige bietet, mussten die Teilnehmer dennoch länger hinsehen. Der Verwendungszweck des Tachometers, wie auch der anderen beiden Visualisierungen hängt stark davon ab, wie genau die Information sein soll, die der betrachtenden Person vermittelt wird. Der Tachometer bietet in dieser Form sehr präzise quantitative wie auch qualitative Informationen über die Geschwindigkeit. Wenn dies erwünscht ist, so muss entschieden werden, ob die entstehende Ablenkung in dem konkreten Fall vernachlässigt werden kann, oder nicht.

Rahmen

Der Rahmen als Objekt an sich führte im Allgemeinen zu keiner Ablenkung von der Szene. Wie erwartet konnten die Teilnehmer den Wechsel der Rahmenfarbe in Form von blinkenden Helligkeitsunterschieden sehen. Dabei wurden jedoch auch geringe Schwankungen der Geschwindigkeit bereits als starkes Blinken wahrgenommen, was das Beobachten der Visualisierung für viele Teilnehmer sehr anstrengend machte. Da man anhand der Rahmenfarbe keine Relation zwischen der momentanen und der maximalen Geschwindigkeit herstellen konnte, ließ sich hauptsächlich nur der qualitative Unterschied zwischen langsam und schnell beurteilen.

Wenn es bei der Verwendung genügt, nur zwischen langsam und schnell zu unterscheiden, so lässt sich die Rahmenvisualisierung verwenden. Hierbei könnte man allerdings auf eine aufwändige Farbskala verzichten, und einige wenige Graustufen mit hohem Kontrast zueinander verwenden.

Störstreifen

Bei der Betrachtung der Störstreifen fiel besonders auf, dass die zehn verwendeten Geschwindigkeitsstufen nicht klar voneinander unterschieden werden konnten. Besonders die Dicke der Streifen spielte keine große Rolle. Die Anzahl der Streifen und der damit verbundene Positionswechsel waren für die Studienteilnehmer besser wahrzunehmen. Aus diesem

Grund konnte nur noch ein Teil der ohnehin schon reduzierten Geschwindigkeitsstufen eingeschätzt werden.

Um die Dicke der Störstreifen auffälliger zu gestalten, könnten beispielsweise Aufhellungen oder Verdunklungen der entsprechenden Bereiche vorgenommen werden. Dabei besteht jedoch das Risiko, dass die Störstreifen stärker vom Inhalt der Szene ablenken.

6.3. Probleme der Studie

- **Erkennung beim Blending:** Die Erkennung der Fremdobjekte beim Blending wurde zum Teil geraten. Studienteilnehmer, die Blending als eines der letzten der vier Methoden gesehen haben wussten bereits, an welchen möglichen Positionen die Figur auftauchen konnte. Auf diese Weise mussten sie nicht unbedingt jedesmal die Figur erkennen, um richtig zu liegen. Obwohl die Reihenfolge der Visualisierungen variiert wurde, konnte dieser Effekt nicht ausgeglichen werden. Tatsächlich ist Erkennung der Objekte beim Blending also um einiges schlechter zu bewerten, als es die Ergebnisse vermuten lassen.
- **Blending mit Adaptive Fast-Forward:** Blending mit adaptive Fast-Forward ermöglicht beim Beobachten eine bessere Erkennung der Objekte, da diese nicht konstant beschleunigt werden, wie es bei der Studie der Fall war. Um der Blending-Methode eine fairere Chance gegenüber den anderen einzuräumen, wäre eine Erkennungsaufgabe mit adaptivem Fast-Forward nötig gewesen.
- **Aufgabe für Geschwindigkeitsvisualisierungen:** Der Vergleich der Geschwindigkeitsvisualisierungen basiert auf den subjektiven Eindrücken der Studienteilnehmer. Eine Aufgabe, die beispielsweise Messwerte zur Ablenkung durch die einzelnen Visualisierungen liefert, hätte mehr Aufschluss gegeben. Leider hätte die Durchführung einer solchen Aufgabe den Zeitrahmen der Studie und wahrscheinlich auch die Aufmerksamkeit der Teilnehmer stark überschritten.

Die gewählten Aufgaben in der Studie könnten mit Sicherheit noch erweitert werden. Dabei bleibt zu bedenken, innerhalb einer Stunde wurden drei Aufgaben von den Teilnehmern erfüllt. Die erste Aufgabe erforderte dabei höchste Konzentration und viele der Teilnehmer waren sichtbar erleichtert, als sie diese Aufgabe hinter sich hatten. Wenn die anderen beiden Aufgaben nun auch eine so hohe Konzentration erfordern würden, dann müssten längere Pausen zwischen den Aufgaben eingeführt werden. Dies würde den gesetzten Zeitrahmen von einer Stunde jedoch weit überschreiten. Aus diesem Grund wurde die höchste Priorität auf die Erkennung von Objekten gelegt und bei der Bewegungsverfolgung sowie den Geschwindigkeitsaufgaben auf Aufgaben verzichtet, die eine sehr hohe Konzentration erfordern.

Zusammenfassung und Ausblick

Im Verlauf dieser Arbeit wurden verschiedene Methoden zum beschleunigten Abspielen von Überwachungsvideos vorgestellt. Dabei wurde Frameskip als der Standard vorgestellt, der in den meisten Fällen verwendet wird. Biologisch motiviertes Blending wie es in der Arbeit von Höferlin et al.⁷ vorgestellt wurde, erzeugt Bilder mit einem natürlichen Eindruck von Bewegungsunschärfe und vermeidet auf diese Weise die sprunghafte Positionsänderung von bewegten Objekten, wie sie bei Frameskip zu beobachten ist. Differenzen dienen als eine Art Verbesserung des Blendings. Die Objekte in den erzeugten Videos erhalten auch einen Bewegungsschweif, dieser wird allerdings mit einem scharfen Differenzbild überlagert. Durch Überlagern der erzeugten Bilder werden zusätzlich Geisterbilder der Objekte an vergangenen Positionen erzeugt. Beim Tracking wurden die Daten einer vorangehenden Objekterkennung verwendet, um zusätzliche Informationen in den erzeugten Videos einfügen zu können. Dabei wurden erkannte Objekte mit künstlich erzeugten Bewegungsschweifern versehen, um den Bewegungsverlauf wie in den vorhergehenden Visualisierungen sichtbar zu machen. Zusätzlich wurden rote Pfeile verwendet, um Informationen über die zukünftige Bewegung der Objekte zu visualisieren. Sämtliche Methoden erzeugten Videos, die im Vergleich zum Original in der Abspielzeit reduziert waren.

Um die wechselnden Geschwindigkeiten von Videos, die mittels adaptivem Fast-Forward beschleunigt wurden, besser sichtbar zu machen, wurden drei Visualisierungen vorgestellt. Der Tachometer enthält sowohl quantitative als auch qualitative Informationen über die Geschwindigkeit. Seine Skala gibt Informationen über die momentane Geschwindigkeit in Relation zum maximalen Wert. Der transparente Rahmen umgibt das Bild, zeigt aber nur noch die Farben der Tachoskala an. Die Anzahl der nostalgisch wirkenden Störstreifen ändert sich mit der Geschwindigkeit des Videos. Durch die insgesamt geringe Anzahl der Störstreifen bieten diese am wenigsten qualitative Information über die aktuelle Geschwindigkeit des Videos.

In einer Benutzerstudie wurden dann die unterschiedlichen Visualisierungen miteinander verglichen. Anhand von drei Aufgaben konnten die Studienteilnehmer zur Beantwortung der Forschungsfragen beitragen. Die Erkennung von Fremdobjekten wurde anhand eines Suchspiels gemessen. Dabei tauchte eine Comicfigur in den gezeigten Videos der einzelnen Visualisierung immer wieder auf und musste mittels Drücken eines Buzzers erkannt werden. Zusätzliche Fragebögen nach dem Anschauen der Videos erfassten Informationen über die persönlichen Eindrücke der Teilnehmer hinsichtlich der Visualisierungen.

Um bestimmen zu können, in welche der Visualisierungen die Objekte am besten verfolgt werden konnten, wurden Paarvergleiche zwischen den Methoden durchgeführt. Dabei konnten sich die Teilnehmer entscheiden, bei welcher der Visualisierungen sie die Objekte besser verfolgen konnten, oder unentschieden bleiben.

Für eine Auswertung der Geschwindigkeitsvisualisierungen wurden ebenfalls Fragebögen von den Studienteilnehmern ausgefüllt, nachdem sie sich ein Video in allen drei Visualisierungen angesehen hatten.

Eine Auswertung der in der Studie gesammelten Informationen ergab, dass sich die verwendeten Methoden bei der Erkennung von Objekten in zwei Gruppen einteilen lassen, die sich signifikant voneinander unterscheiden. Die eine Gruppe bilden Frameskip und Differenzen. Diese liefern sehr ähnliche Ergebnisse, wodurch sich zeigt, dass die Differenzen-Methode nicht wie erhofft zu einer verbesserten Erkennung geführt hat. Tracking und Blending bildet die zweite Gruppe, wobei Blending eindeutig die schlechtesten Ergebnisse liefert.

Bei der Bewegungsverfolgung hat sich Tracking als eine vielversprechende Methode erwiesen. Blending konnte an dieser Stelle nicht überzeugen.

Überraschenderweise wurden der Tachometer und die Störstreifen bei den Geschwindigkeitsvisualisierungen sehr ähnlich bewertet. Ob eine Vertrautheit der Störstreifen aus Zeiten der analogen Videorekorder dafür verantwortlich ist, lässt sich an dieser Stelle nur vermuten. Ein Einsatz der verschiedenen Visualisierungen ist mit Sicherheit von der geplanten Aufgabe abhängig. Eine Verwendung der Rahmen wird aufgrund der vorliegenden Ergebnisse allerdings nicht empfohlen.

Ausblick

Wenn man sich nun die Frage stellt, ob die Frameskip-Methode nun durch eine der vorgeschlagenen Alternativen ersetzt werden könnte, so lässt sich dies nicht mit Sicherheit sagen. Eine genauere Vergleich zwischen Differenzen und Frameskip wäre nötig, um zu sehen, ob das Studiendesign für die ähnlichen Ergebnisse der beiden Methoden verantwortlich war, oder ob die Unterschiede zu gering sind, um irgendwelche Vor- oder Nachteile zu erzeugen. Um den Bewegungseindruck auch bei hohen Geschwindigkeiten besser zu erhalten, als beim reinen Überspringen von Bildern, kann die Tracking-Methode weiterentwickelt werden.

Prinzipiell waren während des Tutorials viele der Teilnehmer von der Vorschau der Tracking-Methode begeistert, als sie damit aber die Aufgabe lösen sollten, änderten einige jedoch ihre Meinung. Wenn die erwähnten Probleme mit der Überdeckung verringert werden und die Informationen etwas dezenter im Video plaziert werden, dann könnten sich mit der Tracking-Methode sicherlich bessere Ergebnisse bei der Objekterkennung erzielen lassen, als es bei der durchgeführten Benutzerstudie der Fall war. Abschließend lässt sich sagen, dass Blending für den Fast-Forward Einsatz nicht geeignet ist, wenn es darum geht, eine möglichst gute Erkennung der Objekte zu garantieren. Die Differenzen-Methode und Tracking lassen sich noch verbessern und könnten dann bei höheren Geschwindigkeiten eine Alternative zur Frameskip-Methode sein.

Anhang A

Tabellen und Histogramme

	H-Test nach Kruskal und Wallis
Erkennung	$\chi^2 = 34.4722, df = 3, p - value = 1.575e - 07$
Aufmerksamkeit	$\chi^2 = 10.3095, df = 3, p - value = 0.01611$
Angenehm	$\chi^2 = 30.345, df = 3, p - value = 1.168e - 06$
Hilfreich	$\chi^2 = 35.4582, df = 3, p - value = 9.748e - 08$

Tabelle A.1.: Aufgabe 1 Fragen: H-Test Video 1

Erkennung	Frameskip	Blending	Differenzen
Blending	3.6e-05	-	-
Differenzen	1.000	7.10e-06	-
Tracking	0.025	0.039	0.014
Aufmerksamkeit	Frameskip	Blending	Differenzen
Blending	0.107	-	-
Differenzen	1.000	0.028	-
Tracking	1.000	0.565	0.666
Angenehm	Frameskip	Blending	Differenzen
Blending	0.00012	-	-
Differenzen	1.000	2.20e-05	-
Tracking	0.03256	0.53922	0.01138
Hilfreich	Frameskip	Blending	Differenzen
Blending	1.20e-05	-	-
Differenzen	1.000	5.0e-06	-
Tracking	0.021	0.15	0.012

Tabelle A.2.: Aufgabe 1 Fragen: p-Werte des U-Test Video 1

	H-Test nach Kruskal und Wallis
Erkennung	$\chi^2 = 21.1475, df = 3, p - value = 9.811e - 05$
Aufmerksamkeit	$\chi^2 = 15.7057, df = 3, p - value = 0.001303$
Angenehm	$\chi^2 = 20.6593, df = 3, p - value = 0.0001239$
Hilfreich	$\chi^2 = 18.3175, df = 3, p - value = 0.0003783$

Tabelle A.3.: Aufgabe 1 Fragen: H-Test Video 2

Erkennung	Frameskip	Blending	Differenzen
Blending	0.00023	-	-
Differenzen	1.000	0.00479	-
Tracking	0.36261	0.02893	1.000
Aufmerksamkeit	Frameskip	Blending	Differenzen
Blending	0.0018	-	-
Differenzen	1.000	0.0154	-
Tracking	1.000	0.0482	1.000
Angenehm	Frameskip	Blending	Differenzen
Blending	4.40e-05	-	-
Differenzen	0.123	0.029	-
Tracking	0.972	0.099	1.000
Hilfreich	Frameskip	Blending	Differenzen
Blending	0.0002	-	-
Differenzen	0.2453	0.0258	-
Tracking	1.000	0.0929	1.000

Tabelle A.4.: Aufgabe 1 Fragen: p-Werte des U-Test Video 2

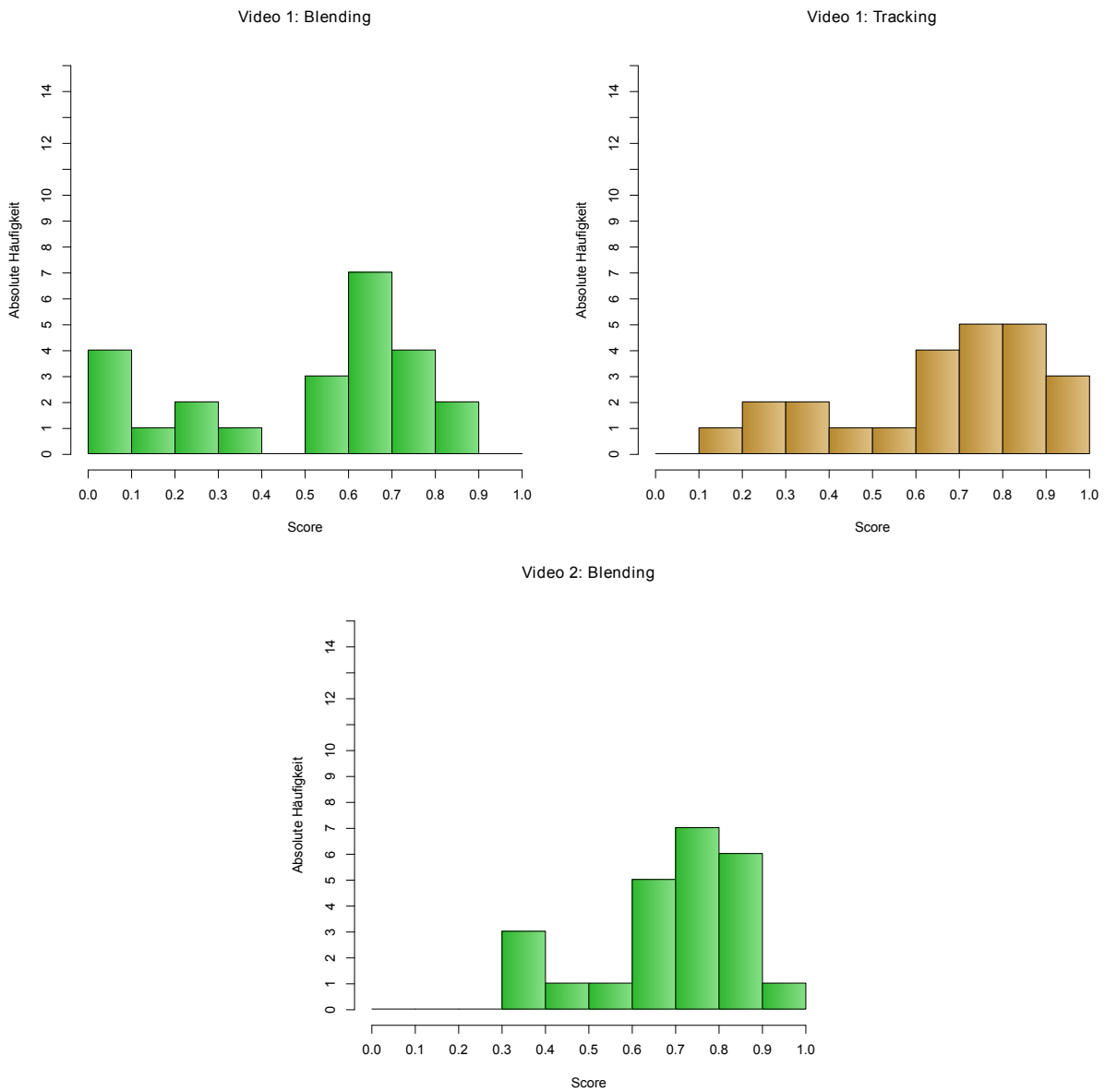


Abbildung A.1.: Histogramme der Werte ohne Normalverteilung

A. Tabellen und Histogramme

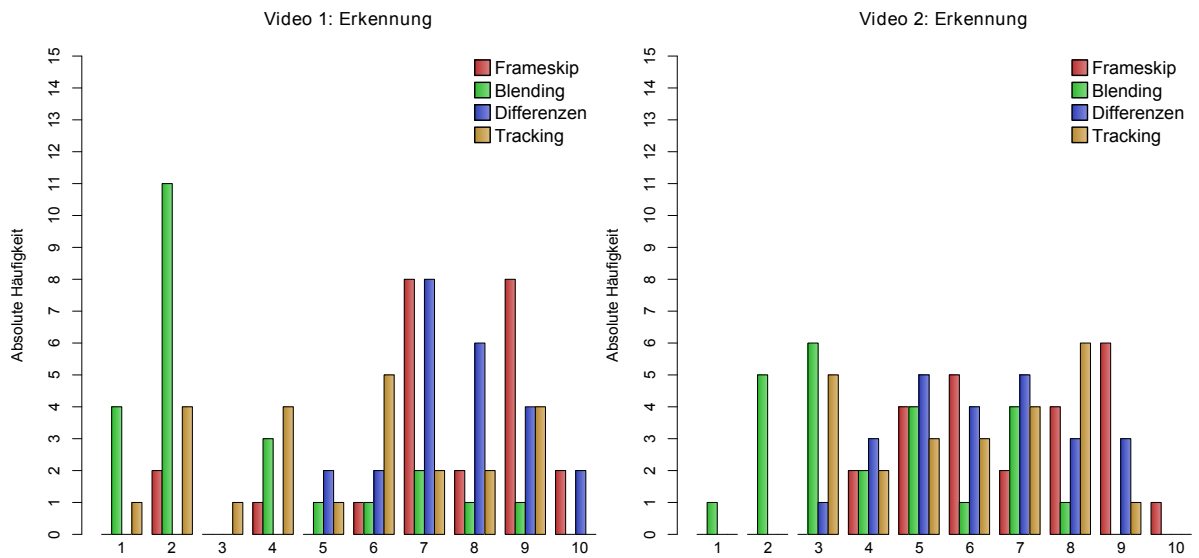


Abbildung A.2.: Aufgabe 1 Fragen: Erkennung Histogramm

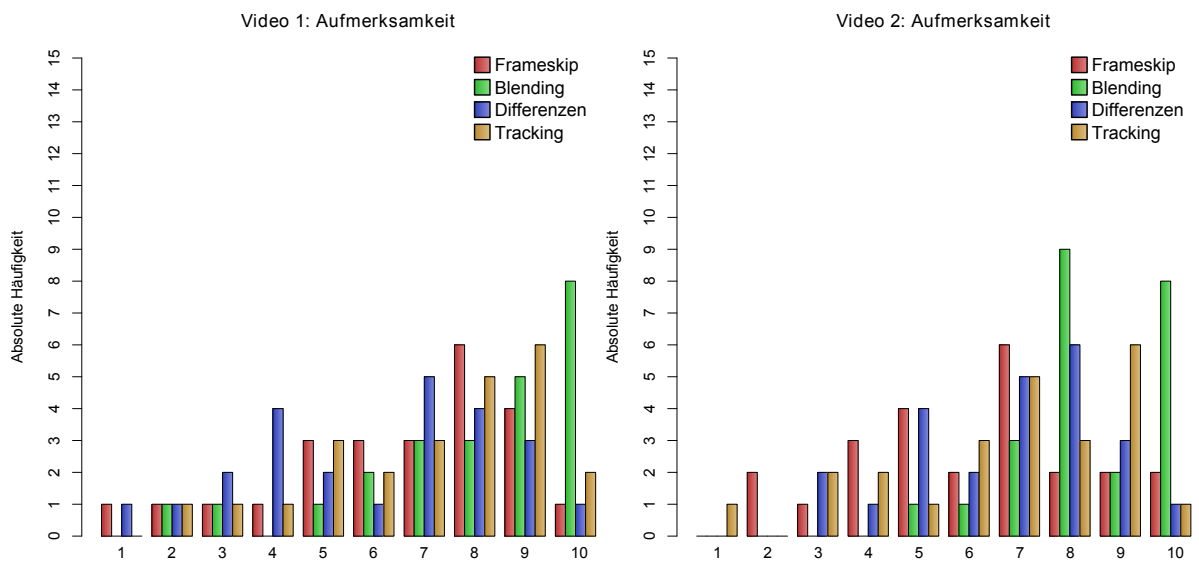


Abbildung A.3.: Aufgabe 1 Fragen: Aufmerksamkeit Histogramm

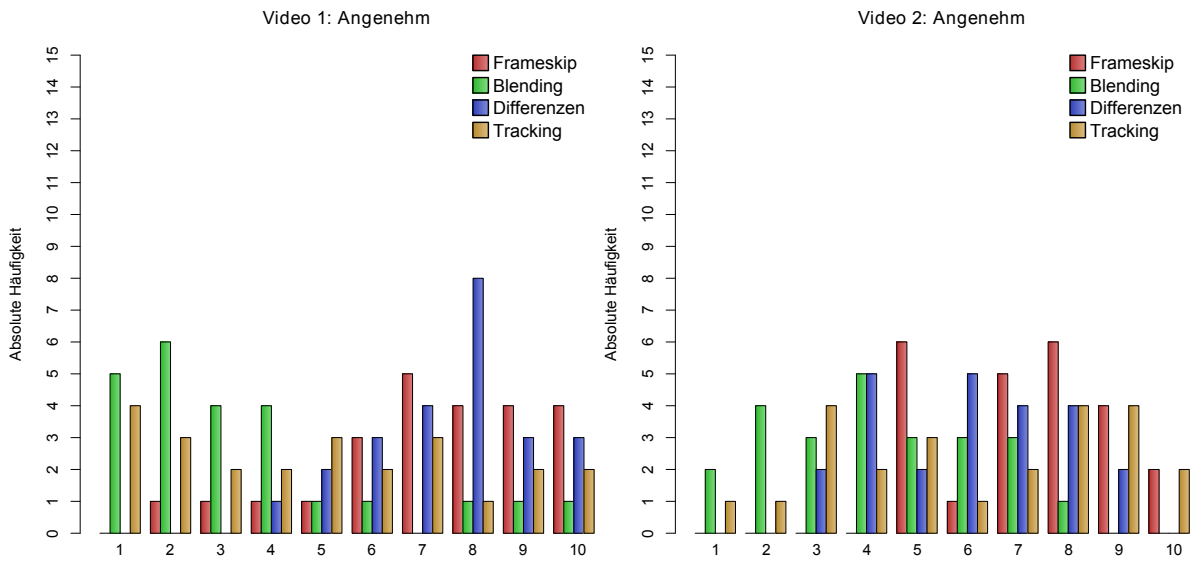


Abbildung A.4.: Aufgabe 1 Fragen: Angenehm Histogramm

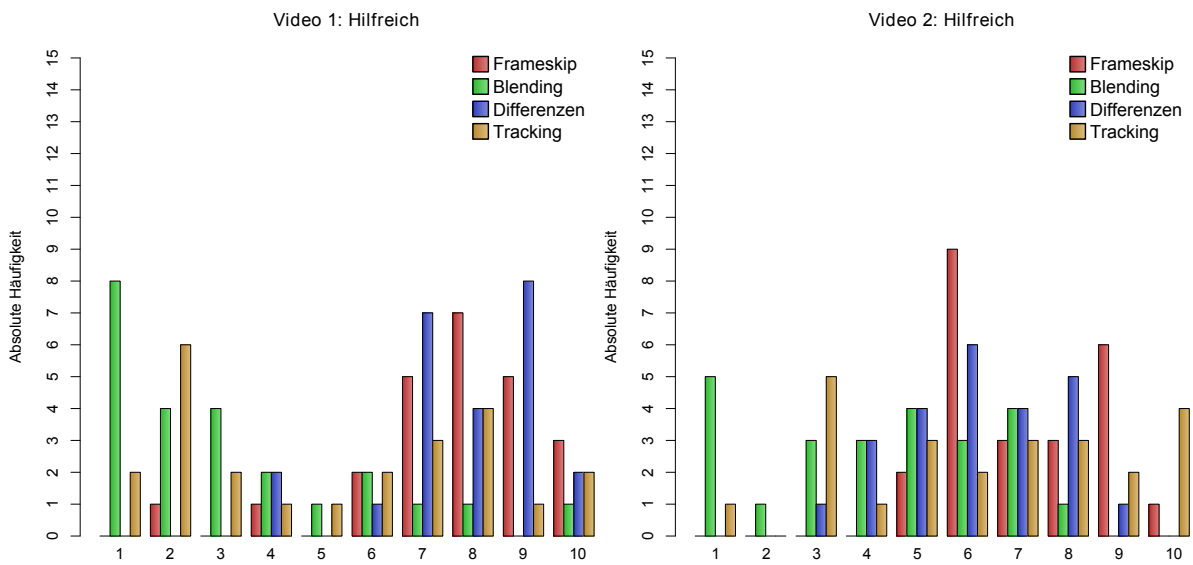


Abbildung A.5.: Aufgabe 1 Fragen: Hilfreich Histogramm

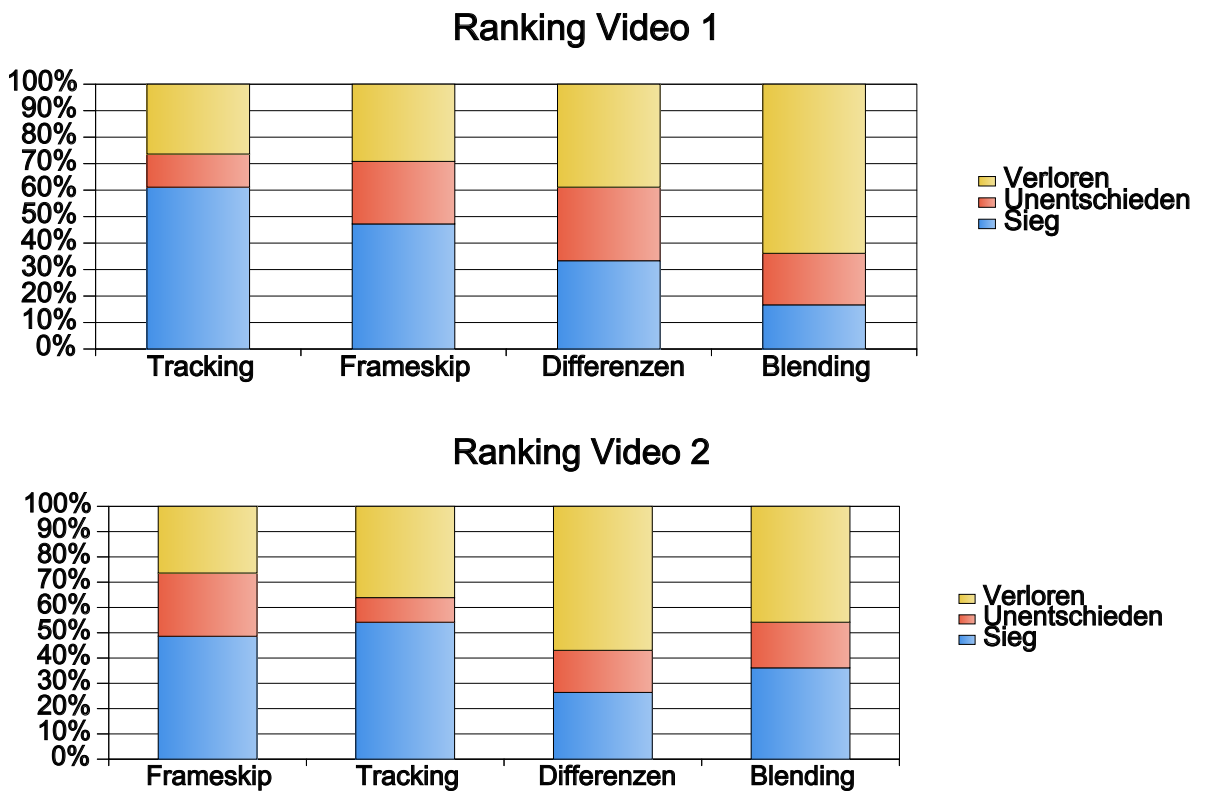


Abbildung A.6.: Aufgabe 2: Rankings

Aufgabe 2

Video 1	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	17	14,5	11	42,5
Blending	7	0	7	5	19
Differenzen	9,5	17	0	7,5	34
Tracking	13	19	16,5	0	48,5
Summe	29,5	53	38	23,5	144
Quadratsumme	308,25	939	531,5	202,25	$\Sigma=1981$
J	918,5				
A	0,1092995169				
FG	6,8429752066				
Chi ²	23,1157024793				
alpha	0,005				

Video 2	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	14	16	14	44
Blending	10	0	13	9,5	32,5
Differenzen	8	11	0	6	25
Tracking	10	14,5	18	0	42,5
Summe	28	39,5	47	29,5	144
Quadratsumme	264	527,25	749	322,25	$\Sigma=1862,5$
J	859,25				
A	0,0377415459				
FG	6,8429752066				
Chi ²	12,3429752066				
alpha	0,1				

Video 1+2	Frameskip	Blending	Differenzen	Tracking	Summe
Frameskip	0	31	30,5	25	86,5
Blending	17	0	20	14,5	51,5
Differenzen	17,5	28	0	13,5	59
Tracking	23	33,5	34,5	0	91
Summe	57,5	92,5	85	53	288
Quadratsumme	1124,25	2867,25	2520,5	1017,5	$\Sigma=7529,5$
J	3620,75				
A	0,0699615839				
FG	6,3969754253				
Chi ²	26,8969754253				
alpha	0,0005				

Abbildung A.7.: Aufgabe 2: Berechnung der Urteilskonkordanz nach Kendall[BLBo8]

A. Tabellen und Histogramme

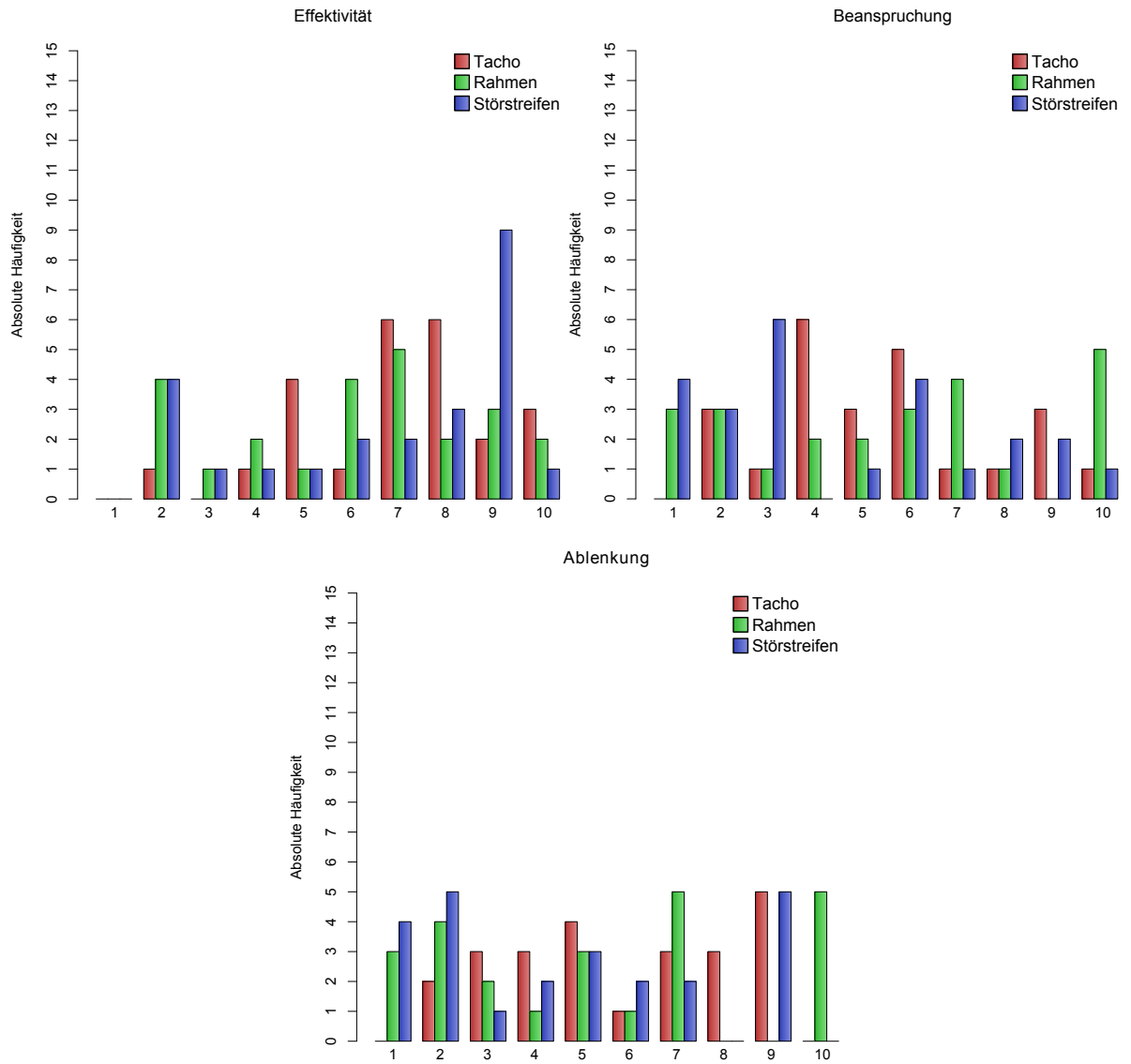


Abbildung A.8.: Aufgabe 3: Histogramme zu den Fragen

Literaturverzeichnis

- [Abdo7] H. Abdi. *NJ Salkind (ed.): Encyclopedia of Measurement and Statistics*. Sage, 2007. (Zitiert auf Seite 51)
- [ADA⁺04] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, M. Cohen. Interactive digital photomontage. In *ACM Transactions on Graphics (TOG)*, volume 23, pp. 294–302. ACM, 2004. (Zitiert auf Seite 19)
- [avs12] 2007 IEEE International Conference on Advanced Video and Signal based Surveillance. http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.html, 2012. (Zitiert auf den Seiten 13, 23 und 39)
- [beh12] BEHAVE Interactions Test Case Scenarios. <http://groups.inf.ed.ac.uk/vision/BEHAVEDATA/INTERACTIONS/>, 2012. (Zitiert auf den Seiten 13, 23 und 39)
- [BLBo8] J. Bortz, G. Lienert, K. Boehnke. *Verteilungsfreie Methoden in der Biostatistik*. Springer Verlag, 2008. (Zitiert auf den Seiten 57, 58 und 79)
- [Bro95] B. Brown. CCTV in town centres: Three case studies. *Crime Detection and Prevention Series Paper 68*, 1995. (Zitiert auf Seite 5)
- [buz12] BUZZ! <http://www.buzzthegame.com/de-de/>, 2012. (Zitiert auf Seite 28)
- [BYRN⁺99] R. Baeza-Yates, B. Ribeiro-Neto, et al. *Modern information retrieval*. Addison-Wesley New York, 1999. (Zitiert auf den Seiten 48 und 49)
- [CC07] J. Calic, N. W. Campbell. Compact visualisation of video summaries. *EURASIP Journal on Advances in Signal Processing*, 2007(24):1–14, 2007. (Zitiert auf Seite 10)
- [CLCC09] K. Cheng, S. Luo, B. Chen, H. Chu. Smartplayer: user-centric video fast-forwarding. In *Proceedings of the 27th international conference on Human factors in computing systems*, pp. 789–798. ACM, 2009. (Zitiert auf Seite 29)

- [CTJ10] G. Chao, Y. Tsai, S. Jeng. Augmented keyframe. *Journal of Visual Communication and Image Representation*, 21(7):682–692, 2010. (Zitiert auf Seite 23)
- [Cuto2] J. Cutting. Representing motion in a static image: constraints and parallels in art, science, and popular culture. *Perception London*, 31(10):1165–1194, 2002. (Zitiert auf Seite 15)
- [Daho6] M. Dahm. *Grundlagen der Mensch-Computer-Interaktion*. Pearson Studium, 2006. (Zitiert auf Seite 31)
- [DGDMWo2] S. Drucker, A. Glatzer, S. De Mar, C. Wong. SmartSkip: consumer level browsing and skipping of digital video content. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pp. 219–226. ACM, 2002. (Zitiert auf Seite 11)
- [DMT97] W. Ding, G. Marchionini, T. Tse. Previewing video data: Browsing key frames at high rates using a video slide show interface. In *Proceedings of the International Symposium on Research, Development & Practice in Digital Libraries*, pp. 151–158. 1997. (Zitiert auf Seite 10)
- [GCSSo6] D. Goldman, B. Curless, D. Salesin, S. Seitz. Schematic storyboarding for video visualization and editing. *ACM Transactions on Graphics (TOG)*, 25(3):862–871, 2006. (Zitiert auf Seite 10)
- [GSo5] M. Gill, A. Spriggs. Assessing the impact of CCTV. *Home Office*, (February):175, 2005. (Zitiert auf Seite 5)
- [HFo4] F. Helten, B. Fischer. Reactive attention: video surveillance in Berlin shopping malls. *Surveillance & Society*, 2(2/3):323–346, 2004. (Zitiert auf Seite 5)
- [HHWH10] B. Höferlin, M. Höferlin, D. Weiskopf, G. Heidemann. Information-based adaptive fast-forward for visual surveillance. *Multimedia Tools and Applications*, pp. 1–24, 2010. (Zitiert auf den Seiten 7, 10, 11, 14, 17, 18, 29 und 42)
- [HLC⁺o6] W. Hwang, P. Lee, B. Chun, D. Ryu, H. Cho. Cinema comics: Cartoon generation from video stream. *Proc. of GRAPP*, pp. 299–304, 2006. (Zitiert auf den Seiten 10 und 27)
- [i-L12] Imagery Library for Intelligent Detection Systems. <http://www.homeoffice.gov.uk/science-research/hosdb/i-lids/>, 2012. (Zitiert auf Seite 42)
- [LPE97] R. Lienhart, S. Pfeiffer, W. Effelsberg. Video abstracting. *Communications of the ACM*, 40(12):54–62, 1997. (Zitiert auf Seite 11)
- [LZTo1] Y. Li, T. Zhang, D. Tretter. An overview of video abstraction techniques. Technical report, HP Laboratories Palo Alto, 2001. (Zitiert auf Seite 9)

- [mco12] m.c optique Online Sehtest. http://www.mc-optique.de/Auge___Sehen/Sehtest/sehtest.html, 2012. (Zitiert auf den Seiten 44 und 45)
- [med11] Media Player Classic - Home Cinema. <http://mpc-hc.sourceforge.net/>, 2011. (Zitiert auf den Seiten 10, 16, 46 und 47)
- [MKSW99] J. Makhoul, F. Kubala, R. Schwartz, R. Weischedel. Performance measures for information extraction. In *Proceedings of DARPA Broadcast News Workshop*, pp. 249–252. 1999. (Zitiert auf den Seiten 48 und 49)
- [ope12] The Open Video Project. <http://www.open-video.org>, 2012. (Zitiert auf Seite 10)
- [Pico04] M. Piccardi. Background subtraction techniques: a review. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 4, pp. 3099–3104. Ieee, 2004. (Zitiert auf Seite 20)
- [PRAP08] Y. Pritch, A. Rav-Acha, S. Peleg. Nonchronological video synopsis and indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:1971–1984, 2008. (Zitiert auf Seite 19)
- [qt-08] UMRK Weblog -Blog Archive- Qt SDL joystick wrapper. <http://www.batcom-it.net/?p=59>, 2008. (Zitiert auf Seite 28)
- [RAPP06] A. Rav-Acha, Y. Pritch, S. Peleg. Making a long video short: Dynamic video synopsis. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pp. 435–441. IEEE, 2006. (Zitiert auf Seite 19)
- [ROC97] R. Rensink, J. O'Regan, J. Clark. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5):368, 1997. (Zitiert auf Seite 6)
- [SA05] D. Simons, M. Ambinder. Change blindness. *Current Directions in Psychological Science*, 14(1):44, 2005. (Zitiert auf Seite 6)
- [SB09] K. Schoeffmann, L. Boeszoermyeni. Video browsing using interactive navigation summaries. In *Content-Based Multimedia Indexing, 2009 Seventh International Workshop on*, pp. 243–248. IEEE, 2009. (Zitiert auf Seite 11)
- [SBC07] K. Scott-Brown, P. Cronin. An Instinct for Detection: Psychological Perspectives on CCTV Surveillance. *The Police Journal*, 80(4):287–305, 2007. (Zitiert auf Seite 6)
- [SH06] L. Sachs, J. Hedderich. *Angewandte Statistik*. Springer, 2006. (Zitiert auf den Seiten 50 und 51)
- [SL97] D. Simons, D. Levin. Change blindness. *Trends in cognitive sciences*, 1(7):261–267, 1997. (Zitiert auf Seite 6)

- [SL98] D. Simons, D. Levin. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5(4):644–649, 1998. (Zitiert auf Seite 6)
- [SSM97] L. Slaughters, B. Shneiderman, G. Marchionini. Comprehension and object recognition capabilities for presentations of simultaneous video key frame surrogates. *Research and Advanced Technology for Digital Libraries*, pp. 41–54, 1997. (Zitiert auf Seite 10)
- [TB93] L. Teodosio, W. Bender. Salient video stills: Content and context preserved. In *Proceedings of the first ACM international conference on Multimedia, MULTIMEDIA '93*, pp. 39–46. ACM, 1993. (Zitiert auf Seite 19)
- [TMD⁺98] T. Tse, G. Marchionini, W. Ding, L. Slaughter, A. Komlodi. Dynamic key frame presentation techniques for augmenting video browsing. In *Proceedings of the working conference on Advanced visual interfaces, AVI '98*, pp. 185–194. ACM, 1998. (Zitiert auf Seite 9)
- [TVSM99] T. Tse, S. Vegh, B. Shneiderman, G. Marchionini. An Exploratory Study of Video Browsing, User Interface Designs and Research Methodologies: Effectiveness in Information Seeking Tasks. In *Proceedings of the ASIS Annual Meeting*, volume 36, pp. 681–692. Information Today, 1999. (Zitiert auf Seite 11)
- [UFGB99] S. Uchihashi, J. Foote, A. Girgensohn, J. Boreczky. Video manga: generating semantically meaningful video summaries. In *Proceedings of the seventh ACM international conference on Multimedia (Part 1)*, pp. 383–392. ACM, 1999. (Zitiert auf Seite 10)
- [vip12] ViPER: The Video Performance Evaluation Resource. <http://viper-toolkit.sourceforge.net/>, 2012. (Zitiert auf Seite 23)
- [vlc11] Official page for VLC media player, the Open Source video framework! <http://www.videolan.org/vlc/>, 2011. (Zitiert auf den Seiten 10 und 16)
- [VV05] M. Valera, S. Velastin. Intelligent distributed surveillance systems: a review. *Image Rochester NY*, 152(2):192–204, 2005. (Zitiert auf Seite 5)
- [wal12] The Official Home Of Where's Waldo? <http://findwaldo.com/>, 2012. (Zitiert auf den Seiten 39 und 40)
- [WMW⁺02] B. Wildemuth, G. Marchionini, T. Wilkens, M. Yang, G. Geisler, B. Fowler, A. Hughes, X. Mu. Alternative surrogates for video objects in a digital library: users' perspectives on their relative usability. *Research and Advanced Technology for Digital Libraries*, pp. 143–160, 2002. (Zitiert auf Seite 11)

- [WMY⁺03] B. Wildemuth, G. Marchionini, M. Yang, G. Geisler, T. Wilkens, A. Hughes, R. Gruss. How fast is too fast?: evaluating fast forward surrogates for digital video. In *Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries*, pp. 221–230. IEEE Computer Society, 2003. (Zitiert auf den Seiten 11 und 16)
- [YY97] M. Yeung, B. Yeo. Video visualization for compact presentation and fast browsing of pictorial content. *Circuits and Systems for Video Technology, IEEE Transactions on*, 7(5):771–785, 1997. (Zitiert auf Seite 10)
- [Zöfo3] P. Zöfel. *Statistik für Psychologen im Klartext*. Im Klartext. Pearson Studium, 2003. (Zitiert auf den Seiten 50 und 51)
- [ZZC96] D. Zhong, H. Zhang, S. Chang. Clustering methods for video browsing and annotation. In *Proc. SPIE*, volume 2670, pp. 239–246. Citeseer, 1996. (Zitiert auf Seite 10)

Alle URLs wurden zuletzt am 08.02.2012 geprüft.

Erklärung

Hiermit versichere ich, diese Arbeit selbständig verfasst und nur die angegebenen Quellen benutzt zu haben.

(Kuno Kurzhals)