

**Auswahl von Ankerpunkten  
für die topologische Navigation  
mit Hilfe des Laptev-Operators**

**Diplomarbeit  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen**

**vorgelegt von**

**Gerrit Ecke**

**Tübingen, Januar 2014**



Erklärung:

Hiermit erkläre ich,

- dass ich diese Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe
  
- dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Tübingen, den



## Zusammenfassung

Mit topologischen Verfahren kann anhand eindeutiger Merkmale und Handlungsanweisungen navigiert werden. Solche Methoden ermöglichen eine Wegfindung, die weniger aufwändig und fehleranfällig ist, als metrische, physikalisch korrekte Karten zu etablieren. Eine Vielzahl von Studien spricht dafür, dass sie ein gutes Modell für verschiedene biologische Systeme sind, die aus den verschiedensten Gründen auf eine sichere Navigationsleistung angewiesen sind.

Für diese Arbeit wurde die Frage untersucht, ob eine große zeitliche Änderung visueller Information während einer Fortbewegung eindeutige und nützliche Hinweise auf die räumliche Konfiguration enthalten. Ausgangspunkt war die Vorstellung, dass sich vor und nach dem Durchschreiten einer Tür wenig verändert, genau im Bereich der Tür dagegen, in relativ kurzem Zeitraum, der Raum aus dem man kommt verschwindet und der nächste einsehbar wird. Man kann sich vorstellen, dass ein solcher Ort eine Bedeutung als Knotenpunkt für die Orientierung haben kann.

Konkret wurde der Laptev-Operator genutzt, mit dem große räumlich-zeitliche Änderungen aus Bildfolgen extrahiert werden können, um ein Maß für die Änderungssalienz zu schaffen. Er ist eine Erweiterung des etablierten und verbreiteten Harris-Kanten- und Ecken-Detektors: eine Bildfolge wird als dritte, zeitliche Dimension aufgefasst. Vereinfacht ausgedrückt kann mit ihm eine „räumlich-zeitliche Ecke“ detektiert werden, also zum Beispiel eine Ecke, die ihre Richtung ändert oder die in einer zeitlichen Folge aus dem Bild verschwindet oder erscheint.

Es konnte gezeigt werden, dass, wenn eine räumliche Geometrie im Vordergrund etwas dahinterliegendes verdeckt und sich beides beim Fortbewegen optisch gegeneinander verschiebt, eine hohe Änderungssalienz auftritt. Das Ansprechverhalten ist in einer statischen Umgebung eindeutig für die beschriebene Situation. Es ist denkbar, solche Orte als Ankerpunkte für die topologische Navigation zu verwenden.



## Danksagung

Wissenschaftliches Arbeiten, so auch das Erstellen dieser Diplomarbeit, geht mit anspruchsvoller und beharrlicher Betätigung einher. Ohne die Geduld und Unterstützung von Familie, Freunden und Mitarbeitern wäre mir der Abschluss dieser Ausarbeitung nicht möglich gewesen. Dafür möchte ich mich herzlich bei Euch bedanken.

Großem Dank verpflichtet bin ich meinem Betreuer, Herrn Prof. Dr. Mallot, dafür, dass er mir die Bearbeitung dieser interessanten wissenschaftlichen Fragestellung ermöglicht hat. Die Gespräche und Hilfestellungen zu diesem Projekt waren eine große Bereicherung und haben mir den richtigen Weg gewiesen. Danke, dass Sie mir die Möglichkeit geben, auch weiterhin an Ihrem Lehrstuhl an einem Themengebiet zu forschen und zu es zu studieren, das meinen persönlichen Neigungen nicht besser entgegenkommen könnte.

Ein besonderes Wort des Dankes möchte ich an Georgiana, meiner Frau richten. Ohne Dich wäre es mir nicht möglich gewesen mein Studium mit dieser Abschlussarbeit zu beenden. Danke für Deine große Geduld und Unterstützung bei meiner lang währenden beruflichen Orientierungssuche.

An dieser Stelle danke ich außerdem Boris Benkner, für den ich in den letzten Jahren als wissenschaftlicher Assistent tätig war. Nicht nur hast Du mir durch Deine große Flexibilität erst die abschließende Phase meines Studiums ermöglicht. Auch war es mir nur durch unsere gemeinsame Tätigkeit möglich, eine Vorstellung für meine berufliche Zukunft zu entwickeln.

Weiterhin bedanke ich mich bei Dr. Hansjürgen Dahmen, Dr. Gregor Hardieß, Marc Halfmann, Stephan Lancier und allen anderen Mitwirkenden des Lehrstuhls für Kognitive Neurowissenschaften für die fruchtbaren, anregenden Diskussionen und das außerordentlich angenehme Arbeitsklima.

## Inhaltsverzeichnis

|  |           |
|--|-----------|
| <b>1. Einleitung</b> .....   | <b>1</b>  |
| 1.1. Motivation.....   | 1         |
| 1.2. Navigation mit dem Schnappschussverfahren.....                    | 2         |
| 1.2.1. Konzepte der Schnappschussnavigation.....                       | 2         |
| 1.2.2. Topologie mit Schnappschüssen.....                              | 4         |
| 1.2.3. Optimierung der Auswahl von Orten für Schnappschüsse.....       | 5         |
| 1.3. Grundlagen zur räumlich-zeitlichen Bildanalyse.....               | 7         |
| 1.3.1. Der Moravec-Eckpunktdetektor.....                               | 7         |
| 1.3.2. Der Harris-Kanten-und-Ecken-Detektor.....                       | 8         |
| 1.3.3. Der Laptev-Detektor für räumlich-zeitliche Ereignispunkte.....  | 10        |
| 1.4. Fragestellung.....  | 12        |
| <b>2. Methoden</b> .....   | <b>13</b> |
| 2.1. Auswertung.....   | 13        |
| 2.1.1. Detektion von Änderungssalienz.....                             | 13        |
| 2.1.2. Implementation der Auswertung.....                              | 13        |
| 2.1.3. Auswahl von Ankerpunkten für die topologische Navigation.....   | 14        |
| 2.2. Simulation virtueller Umgebungen.....                             | 15        |
| 2.2.1. Design der 3D-Welten.....                                       | 15        |
| 2.2.2. Rendern und Ablaufsteuerung.....                                | 16        |
| 2.3. Experimente.....  | 17        |
| 2.3.1. Systematische Untersuchung anhand kurzer Routen.....            | 17        |
| 2.3.2. Untersuchung anhand ausgewählter längerer Routen.....           | 18        |
| 2.3.3. Evaluation anhand von Routen durch enge Passagen.....           | 19        |
| <b>3. Ergebnisse</b> .....   | <b>21</b> |
| 3.1. Allgemeines Ansprechverhaltens des Detektors.....                 | 21        |
| 3.2. Änderungssalienz-Profile ausgewählter längerer Routen.....        | 23        |
| 3.3. Auswahl von Ankerpunkten in engen Passagen.....                   | 27        |
| <b>4. Diskussion</b> .....   | <b>31</b> |
| 4.1. Änderungssalienz lässt Rückschlüsse auf die Raumgeometrie zu..... | 31        |
| 4.2. Offene Fragestellungen .....                                      | 32        |
| 4.3. Anwendung in der topologischen Navigation.....                    | 33        |

|   |           |
|---|-----------|
| <b>A. Anhang</b> .....  | <b>37</b> |
| A.1. Datensatz der Routen aus Experiment 2.....                   | 37        |
| A.1.1. Änderungssalienzprofile bei Texturzusammenstellung 1.....  | 37        |
| Routen in der Nähe der westlichen, in den Raum ragenden Wand..... | 37        |
| Routen durch den mittigen Durchgang.....                          | 38        |
| Routen in der Nähe der südöstlichen Säule.....                    | 38        |
| A.1.2. Änderungssalienzprofile bei Texturzusammenstellung 2.....  | 39        |
| Routen in der Nähe der westlichen, in den Raum ragenden Wand..... | 39        |
| Routen durch den mittigen Durchgang.....                          | 40        |
| Routen in der Nähe der südöstlichen Säule.....                    | 40        |
| A.2. Datensatz der Routen aus Experiment 3.....                   | 41        |
| A.2.1. Änderungssalienzprofile bei Texturzusammenstellung 1.....  | 41        |
| Routen in Korridorpassage (Raum 1).....                           | 41        |
| Routen in der Türpassage (Raum 2).....                            | 41        |
| Routen in der spitz zulaufenden Passage (Raum 3).....             | 42        |
| Routen in der Passage mit gekrümmter Wand (Raum 4).....           | 43        |
| A.2.2. Änderungssalienzprofile bei Texturzusammenstellung 2.....  | 44        |
| Routen in Korridorpassage (Raum 1).....                           | 44        |
| Routen in der Türpassage (Raum 2).....                            | 44        |
| Routen in der spitz zulaufenden Passage (Raum 3).....             | 45        |
| Routen in der Passage mit gekrümmter Wand (Raum 4).....           | 46        |
| A.3. Weitere Ankerpunkte aus Experiment 3.....                    | 47        |
| <b>B. Literaturverzeichnis</b> .....                              | <b>49</b> |

## Abbildungsverzeichnis

|   |    |
|---|----|
| Abbildung 1.1: Dressur des Bienenwols auf den Reizkomplex „Zapfenkreis“. Aus Tinbergen (1932).....  | 2  |
| Abbildung 1.2: Vergleich von einem Schnappschuss einer 360°-Aufnahme mit der aktuellen Bildinformation. Aus Franz et al. (1998), nach Cartwright and Collett (1983).... | 3  |
| Abbildung 1.3: Grundriss einer Umgebung mit topologischem Netzwerkgraphen. Aus Franz et al. (1998).....   | 5  |
| Abbildung 1.4: Anwendung des Harris-Operators auf ein Beispielbild.....   | 10 |
| Abbildung 1.5: Resultate für die Anwendung des Laptev-Operators auf synthetische Bildfolgen. Aus Laptev (2005).....   | 12 |
| Abbildung 2.1: Typisches gerendertes Einzelbild in der virtuellen Umgebung mit Texturzusammenstellung 1.....  | 16 |
| Abbildung 2.2: Derselbe Raum wie in Abbildung 2.1 mit Texturzusammenstellung 2...   | 16 |
| Abbildung 2.3: Blick in den zweiten Raum mit Texturzusammenstellung 1.....  | 18 |
| Abbildung 2.4: Grundriss der virtuellen Innenraum-Umgebung mit den in Experiment 1 getesteten, kurzen Routen.....   | 18 |
| Abbildung 2.5: Grundriss der Innenraum-Umgebung mit für Experiment 2 ausgewählten Routen.....   | 19 |
| Abbildung 2.6: Grundriss der vier Räume mit engen Passagen unterschiedlicher Geometrie.....   | 21 |
| Abbildung 3.1: Detektorantwort für Texturzusammenstellung 1.....  | 21 |
| Abbildung 3.2: Detektorantwort für Texturzusammenstellung 2.....  | 22 |
| Abbildung 3.3: Nur Detektorwerte, auf die in gleicher Richtung ein kleinerer Wert folgte, Texturzusammenstellung 1.....   | 22 |
| Abbildung 3.4: Nur Detektorwerte, auf die in gleicher Richtung ein kleinerer Wert folgte, Texturzusammenstellung 2.....   | 23 |
| Abbildung 3.5: Änderungssalienz-Profil der 180°-Route durch den mittigen Durchgang (O → W, Texturzusammenstellung 2).....   | 24 |
| Abbildung 3.6: Änderungssalienz-Profil für die schräg verlaufende 135°-Route durch den Durchgang (SO → NW, Texturzusammenstellung 2).....                               | 24 |
| Abbildung 3.7: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 1) für die Routen durch den Durchgang und in der Nähe der Säulen.....                    | 25 |

|  |       |
|--|-------|
| Abbildung 3.8: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 1) für die Routen im Bereich der in den Raum ragenden Wand.....   | 26    |
| Abbildung 3.9: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 2) für die Routen durch den Durchgang und in der Nähe der Säulen.....   | 26    |
| Abbildung 3.10: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 2) für die Routen im Bereich der in den Raum ragenden Wand.. ..  | 27    |
| Abbildung 3.11: Ankerpunkte in der Korridorpassage (Raum 1, Texturzusammenstellung 1).....   | 28    |
| Abbildung 3.12: Ankerpunkte in der Korridorpassage (Raum 1, Texturzusammenstellung 2).....   | 28    |
| Abbildung 3.13: Ankerpunkte in der Türpassage (Texturzusammenstellung 2).....  | 29    |
| Abbildung 3.14: Ankerpunkte in der spitz zulaufenden Passage (Texturzusammenstellung 1).....   | 29    |
| Abbildung 3.15: Ankerpunkte in der spitz zulaufenden Passage (Texturzusammenstellung 2).....   | 30    |
| Abbildung 3.16: Mit Texturzusammenstellungen 2 wurde auf den schräg verlaufenden Routen das Verschwinden beider Wände aus dem Sichtfeld einzeln erkannt: jeweils etwa mittig das der am Anfang der Route näheren Wand und zum Ende der Route das der am Ende der Route näheren Wand..... | 30    |
| Abbildung A.1 – A108: Änderungssalienzprofile aller Routen.....  | 37-47 |
| Abbildung A.109: Ankerpunkte in der Türpassage (Texturzusammenstellung 1).....   | 47    |
| Abbildung A.110: Ankerpunkte in der spitz zulaufenden Passage (Texturzusammenstellung 2).....  | 48    |



# 1. Einleitung

## 1.1. Motivation

Seit geraumer Zeit wird, insbesondere im Zusammenhang mit dem Navigationsverhalten von Insekten, die Idee der Wegfindung anhand von topologischer Repräsentation diskutiert. Solche Verfahren folgen dem Prinzip, eindeutige Anhaltspunkte und Handlungsanweisungen, zu etablieren, anhand derer ein Ort wiedergefunden werden kann. Dafür werden keine genauen räumlichen Informationen benötigt. Visuelle Merkmale sind für die topologische Navigation ein häufig genutzter Anhaltspunkt. Der Ansatz steht im Gegensatz zur Etablierung metrischer Karten, die eine Umgebung physikalisch korrekt wiedergeben. Für solche exakte Karten sind sehr genaue, zeitaufwändige und datenintensive Messungen erforderlich und ihr Aufbau ist mit bestehenden Methoden häufig labil und fehleranfällig.

Einen Ort kann man wiedererkennen, indem man sich die Ansichten seiner lokalen Objekte und / oder Konfiguration dieser merkt. Dafür ist entscheidend, dass die Objekte unveränderlich und eindeutig identifizierbar sind. Es ist auch möglich, einen Ort zu erkennen ohne den Bildinhalt zu klassifizieren: ein Schnappschuss der Szene hat einfache, aber häufig eindeutige Merkmale wie die Farbverteilung oder Kontrastmuster, die ebenso einzigartig sein und einen Ort eindeutig beschreiben können. In weiträumigem Terrain dagegen können große, weit entfernte Objekte als Landmarken zur räumlichen Orientierung dienen. Für eine Abschätzung, an welchem Ort man sich befindet, lässt sich in einem weiten Feld ihre Richtung, die Größe in der sie erscheinen und das Verhältnis in dem sie zueinander stehen als wertvolle Informationsquellen nutzen. Größter Vorteil topologischer Verfahren ist, dass die Navigation nicht auf eine metrische Repräsentation angewiesen ist und mit nicht sehr rechenintensiven und dennoch robusten Verarbeitungsschritten arbeitet, wodurch sich der Verfahrensaufwand deutlich reduziert.

Die biologische Relevanz visueller Information für das Navigationsverhalten konnte am Beispiel von Insekten schon sehr früh gezeigt werden. Bereits 1932 zeigte Tinbergen [15] in Experimenten, in denen das Anflugverhalten untersucht wurde, dass Bienenwölfe (*Philanthus triangulum*, eine Grabwespenart) anhand der Umgebungskonfiguration ihre Nestöffnung wiederfinden. Tinbergen positionierte in der Nähe der

Nester markante Landmarken, nämlich auch in natürlicher Umgebung vorkommender Zapfen. Versob er die Landmarken nach nur kurzer Dressurzeit zu einem Scheinnest in der Nähe, zeigte der Bienenwolf sein Suchverhalten an entsprechender Stelle (Abbildung 1.1).

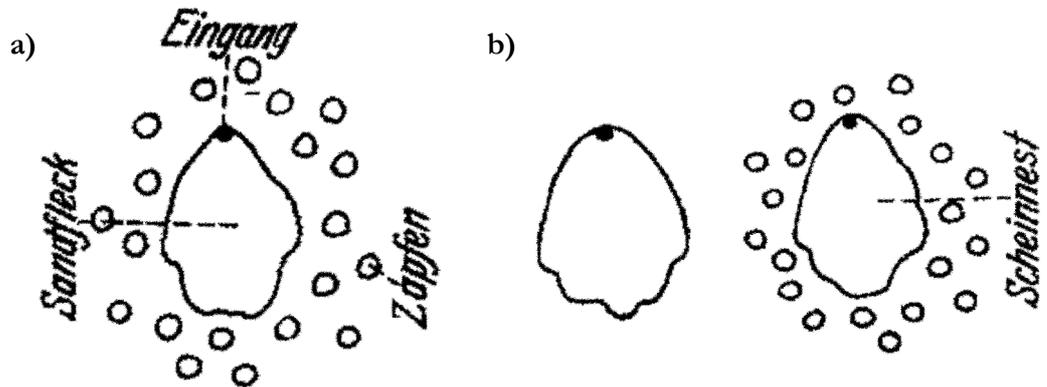


Abbildung 1.1: Der Bienenwolf kann auf einen Zapfenkreis als entscheidendes Merkmal für die Nestsuche dressiert werden. a) Dressuranordnung, b) Versuchsanordnung. Aus Tinbergen (1932).

## 1.2. Navigation mit dem Schnappschussverfahren

### 1.2.1. Konzepte der Schnappschussnavigation

Ein großer Fortschritt beim Verständnis der Navigationsleistung von Insekten waren die Arbeiten von Cartwright und Collett [2]. Bienen wurden darauf konditioniert, eine Futterstelle in der Nähe einer Landmarkenkonfiguration, bestehend aus drei Säulen, aufzusuchen. Durch geschickte Verschiebung der Säulen zueinander und Analyse der Aufenthaltsverteilung während der Suche konnte gezeigt werden, dass aus Perspektive des Tieres die Winkel der Orientierungspunkte zueinander das entscheidende Merkmal für die Ortserkennung sind und nicht zum Beispiel der Abstand des Tieres zu den einzelnen Landmarken. Erklärt wurde der Befund mit einem einfachen Modell: die drei Säulen des erinnerten Schnappschusses werden den korrespondierenden Säulen im aktuellen Sichtfeld zugeordnet. Die Bewegungsrichtung wird so gewählt, dass eine bessere Übereinstimmung der Winkel zu erwarten ist. Vereinfacht lautet die Anweisung, sich auf zu eng stehende Säulen zuzubewegen und sich von zu weit voneinander separierten zu entfernen. Stimmen die Abstände aller Landmarken zueinander mit dem Schnappschuss überein, wurde der gesuchte Ort gefunden.

Nachteil der Methode ist die Komplexität des zu lösenden Korrespondenzproblems und der daraus resultierende Rechenaufwand, sowie eine große Anfälligkeit für fehlerhafte Zuordnungen. Franz et al. [4] schlugen ein Modell vor, das ohne Auffinden korrespondierender Landmarken auskommt. An der aktuellen Position erstellte ein Agent pro Iterationsschritt eine 360°-Aufnahme der Umgebung und errechnete, wie sich diese Abbildung bei Bewegungen in verschiedene Richtungen verändern würde. Dazu musste für den Abstand der Objekte eine Annahme getroffen werden, weil mögliche Methoden der Entfernungsschätzung, wie zum Beispiel das Stereosehen, nicht angewendet wurden. Unter der Annahme, dass alles Sichtbare gleich weit entfernt ist, ließ sich die Berechnung vereinfachen: eine Verzerrungsoperation führte zu Verschiebung, Expansion und Kontraktion von Bildbereichen (Abbildung 1.2). Im Anschluss wurde der Schnappschuss mit den resultierenden Bildern verglichen, indem die Korrelation der Pixelwerte zueinander berechnet wurde. Die Richtung, in der die höchste Korrelation gemessen wurde, wurde für eine kurze Strecke eingeschlagen. Das Vorgehen wurde bis zum Erreichen des Ziels, definiert durch sehr große Ähnlichkeit der Aufnahme mit dem Schnappschuss, wiederholt durchlaufen. In Experimenten hatten nach diesem Verfahren arbeitende Roboter große Fangbereiche: das Umfeld, in dem das Anwenden der Handlungsanweisung dazu führt, das Ziel zu erreichen.

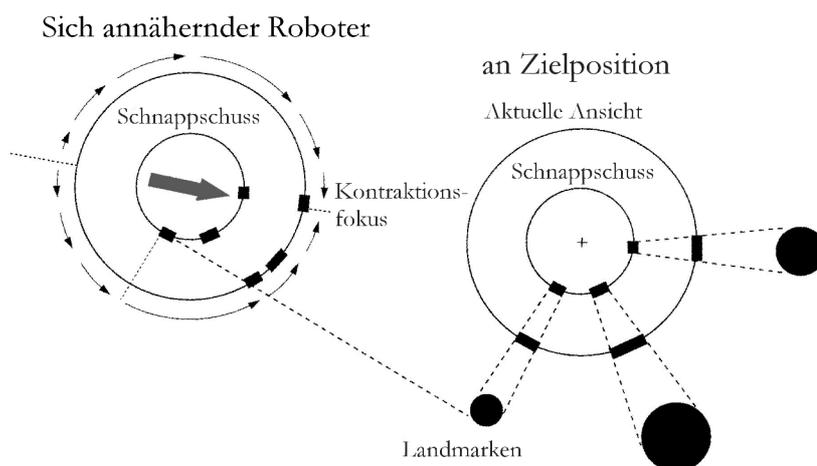


Abbildung 1.2: Vergleich von einem Schnappschuss einer 360°-Aufnahme mit der aktuellen Bildinformation. An anderen Orten ist die Abbildung der Landmarken verschoben, veranschaulicht durch Pfeile. Einfache Verzerrungsoperationen ermöglichen eine Vorhersage dafür, wie sich die Aufnahme in der gewählten Richtung verändert. Aus Franz et al. (1998), nach Cartwright and Collett (1983).

### 1.2.2. Topologie mit Schnappschüssen

Um in komplexem, weiträumigen Terrain etwas wiederzufinden reicht ein einziger Schnappschuss vom Zielort nicht aus. Schon in geringem Abstand ist die Bildähnlichkeit zum Zielschnappschuss nicht mehr hinreichend um zurückzufinden. Eine Möglichkeit zu navigieren ist, Aufnahmen in regelmäßigen Abständen zu erstellen, um so ein Gitternetz von Schnappschüssen zu bekommen, an denen entlang navigiert werden kann. Zeil et al. [17] haben Schnappschüsse in natürlichem Umfeld systematisch an Rasterpunkten aufgenommen und gezeigt, dass Korrelationsminima aneinander gereiht werden können. Diese Strategie ist sinnvoll für Aufgaben, die die vollständige Kartierung eines überschaubaren Gebiets erfordern. In letzter Zeit wurden beispielsweise Ideen entwickelt, Staubsaugerroboter mit diesem Ansatz zu betreiben [5].

Für viele Navigationsaufgaben ist eine flexiblere, effizientere und weniger aufwändige Strategie wünschenswert. Eine Variante, die mit deutlich weniger gespeicherten Schnappschüssen auskommt um ein Gebiet zu kartieren, wurde von Franz et al. [3] vorgestellt. Während der Exploration eines Agenten mit 360°-Kamera wurde die Entscheidung, eine neue Aufnahme zu erstellen, immer dann getroffen, wenn der Bildunterschied zum letzten Schnappschuss (die Korrelation der Pixelwerte) eine bestimmte Schwelle überschritt. Diese wurde experimentell so gewählt, dass der Unterschied groß genug war, um die zwei aufeinanderfolgenden Bilder eindeutig voneinander zu unterscheiden, aber klein genug, um sicher von einem Aufnahmeort zum nächsten navigieren zu können. Die Verbindungen zwischen zwei nebeneinander liegenden Orten wurden in einem Netzwerkgraph gespeichert. Eine Verzweigung entstand, wenn ein bereits aufgesuchter Ort wiedererkannt wurde. Dafür wurde das aktuelle Sichtfeld ständig mit allen gespeicherten Bildern verglichen und der zugehörige Ort angesteuert, sobald eine hohe Ähnlichkeit festgestellt wurde. So entstand eine vollwertige topologische Karte mit Schnappschüssen als Knotenpunkten. Vorteil gegenüber einer metrischen Rasterung ist, dass viel weniger Aufnahmen nötig werden um eine Umgebung zu kartieren.

### 1.2.3. Optimierung der Auswahl von Orten für Schnappschüsse

Es ist naheliegend, die Entscheidung, einen neuen Schnappschuss aufzunehmen und zu speichern, mit der Bedeutung des Ortes für die Navigation zu verknüpfen. Zum einen kann es sinnvoll sein sich eine Region einzuprägen, weil dort etwas wichtiges, wie zum Beispiel eine Futterstelle, zu finden ist. Zum anderen kann es sich um einen Ort handeln, der die Funktion einer Weggabelung hat, an dem also eine Handlungsentscheidung getroffen werden muss um zum Ziel zu gelangen. Mit der zuvor vorgestellten Methode wurde der Aufnahmeort auf Basis des Vergleichs mit vorherigen Aufnahmen bestimmt. Das hat zur Folge, dass in gleicher Umgebung mit unterschiedlichen Startpositionen, selbst bei ansonsten absolut gleichen Routen, andere Karten entstehen (Abbildung 1.3). Diese sind zwar in sich optimiert, können aber wegen ihrer Verschiedenheit nicht ideal sein. Gibt es ideale Orte, um Schnappschüsse aufzunehmen und abzuspeichern und wenn ja, wie lassen sie sich bestimmen bzw. nach welchen Kriterien lässt sich die Entscheidung dazu treffen?

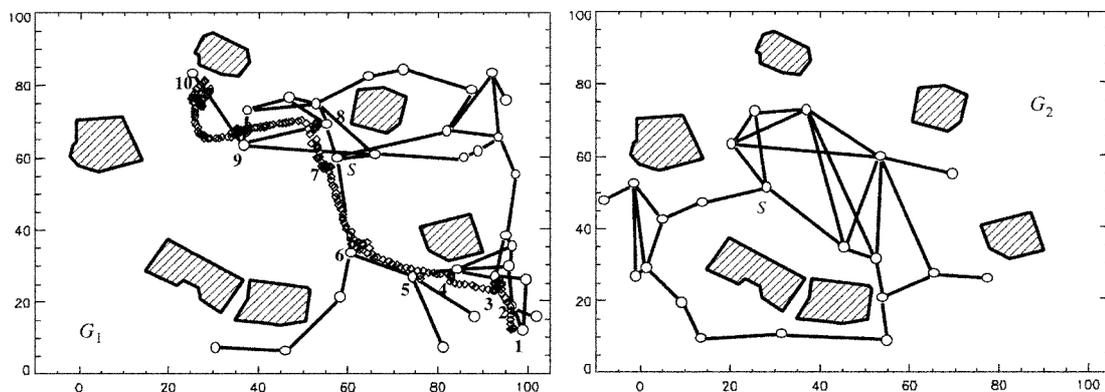


Abbildung 1.3: Grundriss einer Umgebung, in der ein Roboter topologische Netzwerkgraphen erstellt hat. Kreise stehen für Orte, an denen Schnappschüsse aufgenommen wurden, Linien für gespeicherte Verbindungen. Diamanten beschreiben eine Beispieltrajektorie von Position 1 zu Position 10. Unterschiedliche Startpositionen  $S$  führen zu verschiedenen topologischen Karten  $G_1$  und  $G_2$ . Aus Franz et al. (1998).

Orte, an denen sich die visuelle Information stark ändert, wenn sie auf einer Route passiert werden, haben das Potenzial bedeutend für die Schnappschussnavigation zu sein. Beispielsweise könnte das Durchschreiten einer Tür als wichtiger Orientierungspunkt und Ankerpunkt für den Eintritt in eine neue Umgebung dienen. Bewegt man sich auf eine geöffnete Tür zu, ändert sich zunächst nur sehr wenig. Kommt man dem Durchgang näher, wird kontinuierlich mehr vom dahinter liegenden Raum

sichtbar. Auf Höhe der Tür kann schließlich der vorher verdeckte Bereich eingesehen werden, eine weitere Bewegung in dieselbe Richtung ändert jetzt wieder sehr wenig an der Bildinformation. Wäre man in der Lage, so eine Situation eindeutig zu detektieren, besäße man ein wertvolles Werkzeug für die Schnappschussnavigation.

Die Unterschiede, die sich im Blickfeld ergeben, werden von der räumliche Konfiguration im direkten Umfeld und maßgeblich von der zurückgelegten Strecke bestimmt. Indirekt ist die Änderung zeitlich, wenn man die Eigengeschwindigkeit berücksichtigt. In dieser Arbeit wird der Laptev-Operator, mit dem die zeitliche Änderung räumlicher Strukturen in einer Bildfolge ermittelt werden kann, auf den Nutzen für die Schnappschussnavigation untersucht [9].

### 1.3. Grundlagen zur räumlich-zeitlichen Bildanalyse

#### 1.3.1. Der Moravec-Eckpunktdetektor

Die Entwicklung des Laptev-Detektors für räumlich-zeitliche Ereignispunkte reicht zurück zum Moravec-Eckpunktdetektor [10]. Dieser einfache und effiziente Algorithmus ist Teil früherer Arbeiten zur Hindernisvermeidung für Roboter. Bei der Auswertung zweidimensionaler Aufnahmen, der Projektion dreidimensionalen Raums mit reduziertem Informationsgehalt, entsprechen große lokale Kontraste häufig tatsächlichen Objektgrenzen. Solche Kanten haben einen kontinuierlichen Charakter. Um Punkte in Bildfolgen eindeutig zuzuordnen zu können, sind jedoch diskrete Ortseigenschaften gefragt. Geeignet sind Elemente, die bei verschiedenen Blickwinkeln eindeutig identifiziert werden können. Die Endpunkte von Kanten, im Folgenden als Ecken bezeichnet, zeigen dieses bedeutende Attribut. Entscheidendes Merkmal ist eine große Kontrastvarianz in zwei orthogonale Richtungen.

Um solche Punkte zu extrahieren, wird über Rasterpunkte im Bild ein quadratisches Fenster  $w(u,v)$  gelegt (typische Fenstergröße:  $4^2$  Pixel, Abstand der Rasterpunkte: halbe Fensterbreite) und der Durchschnitt der Intensitätsänderung  $E$  benachbarter Pixel in jeweils vier Richtungen  $(x,y) = \{(1,0), (1,1), (0,1), (-1,1)\}$  (horizontal, vertikal und in zwei Richtungen diagonal) betrachtet.

$$E_{x,y} = \sum_{u,v} w_{u,v} (I_{x+u, y+v} - I_{u,v})^2 \quad (1)$$

Drei Fälle können unterschieden werden: a) das Areal zeichnet sich nicht durch eine Änderung der Lichtintensität aus, was in einen geringen Wert in alle Richtungen resultiert; b) im Bereich der Fensterfunktion liegt eine Kante, die Verschiebung ergibt je nach Orientierung in eine oder zwei Richtungen eine große, in andere Richtungen jedoch nur eine geringe Änderung; c) im Fensterbereich liegt eine Ecke oder ein isolierter Punkt, alle Verschiebungen zeigen einen großen Änderungswert.

Fall c) beschreibt die gesuchten Attribute. Entsprechende Bildkoordinaten lassen sich extrahieren, indem man jeweils die kleinste der vier Änderungen ( $\min(E)$ ) betrachtet. An lokalen Maxima dieser Werte befindet sich ein Eckpunkt.

### 1.3.2. Der Harris-Kanten-und-Ecken-Detektor

Aus der Anwendung des Moravec-Eckpunktdetektors resultiert kein isotropes Ergebnis, weil nur diskrete Richtungen in 45°-Schritten betrachtet werden. Um alle Richtungen gleich zu gewichten, entwickelten Harris und Stephens [6] eine Methode, die auf eine analytische Erweiterung für kleine Verschiebungen beruht

$$\begin{aligned} E_{x,y} &= \sum_{u,v} w_{u,v} (I_{x+u, y+v} - I_{u,v})^2 \\ &= \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial x} x + \frac{\partial I}{\partial y} y + O(x^2, y^2) \right)^2 \quad (2) \end{aligned}$$

mit Annäherung der ersten Ableitungen der Bildintensität (Abbildung 1.4 a, b)

$$\begin{aligned} \frac{\partial I}{\partial x} &\approx I * (-1 \ 0 \ 1) \\ \frac{\partial I}{\partial y} &\approx I * \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad (3) \end{aligned}$$

Der Ausdruck kann in eine quadratische Form überführt werden. Sie beschreibt einen parabolischen Potentialtopf mit elliptischem Querschnitt

$$\begin{aligned} E_{x,y} &= x^2 \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial x} \right)^2 + 2xy \sum_{u,v} w_{u,v} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} + y^2 \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial y} \right)^2 \\ &= (x \ y) M \begin{pmatrix} x \\ y \end{pmatrix} \quad (4) \end{aligned}$$

mit der 2 x 2 Matrix

$$M = \begin{pmatrix} \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial x} \right)^2 & \sum_{u,v} w_{u,v} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \\ \sum_{u,v} w_{u,v} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} & \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial y} \right)^2 \end{pmatrix} \quad (5)$$

Analog zum Moravec-Operator lassen sich anhand der Eigenwerte  $\lambda_1$  und  $\lambda_2$  der Matrix M drei Fälle unterscheiden: a) handelt es sich um einen Bildabschnitt mit gerin-

gen Helligkeitsunterschieden, sind beide Eigenwerte klein. b) Befindet sich im Fenster eine beliebig orientierte, gerade Kante, ergeben Verschiebungen parallel zur Kante für E nur kleine Intensitätsänderungen., orthogonal dazu sind die Änderungen dagegen groß, was sich in einem kleinen und einem zweiten großen Eigenwert der Matrix M ausdrückt. c) Ist im untersuchten Bereich eine Ecke, ergeben Verschiebungen in alle Richtungen einen großen Wert für E, weswegen beide Eigenwerte groß sind.

Um zu bestimmen, ob es sich um eine Ecken- oder eine Kantenregion handelt und um ein Maß dafür zu schaffen, führten die Autoren die Kenngröße R ein

$$R = \lambda_1 \lambda_2 - k (\lambda_1 + \lambda_2)^2 \quad (6)$$

Der Term wird positiv für eine Eckenregion, negativ für eine Kantenregion und bleibt klein, wenn es in dem Bereich keine nennenswerten Intensitätsänderungen gibt. Die Eigenwerte müssen nicht explizit berechnet werden, weil

$$\begin{aligned} \lambda_1 \lambda_2 &= \det(M) = \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial x} \right)^2 \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial y} \right)^2 - 2 \sum_{u,v} w_{u,v} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \\ \lambda_1 + \lambda_2 &= \text{trace}(M) = \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial x} \right)^2 + \sum_{u,v} w_{u,v} \left( \frac{\partial I}{\partial y} \right)^2 \end{aligned} \quad (7)$$

und R (Abbildung 1.4 c) somit ausgedrückt werden kann mit

$$R = \det(M) - k \text{trace}^2(M) \quad (8)$$

Lokale Maxima in R können als Eckpunkte extrahiert werden (Abbildung 1.4 d). R beinhaltet ein Maß dafür, wie stark der Kontrast der gefundenen Ecke ist. Er kann für einen Schwellenwert genutzt werden um kleine Extrema herauszufiltern und hat eine Aussagekraft bezüglich der Beständigkeit des Eckpunktes.

Um Signalrauschen besser zu unterdrücken wurde außerdem vorgeschlagen, anstelle der binären, rechteckigen Fensterfunktion eine runde, gleichmäßig abfallende zu verwenden, beispielsweise eine Gaußsche Glockenkurve

$$w_{u,v;\sigma} = \frac{1}{\sqrt{(2\pi)^2 \sigma^4}} \exp(-(u^2 + v^2)/2\sigma^2) \quad (9)$$

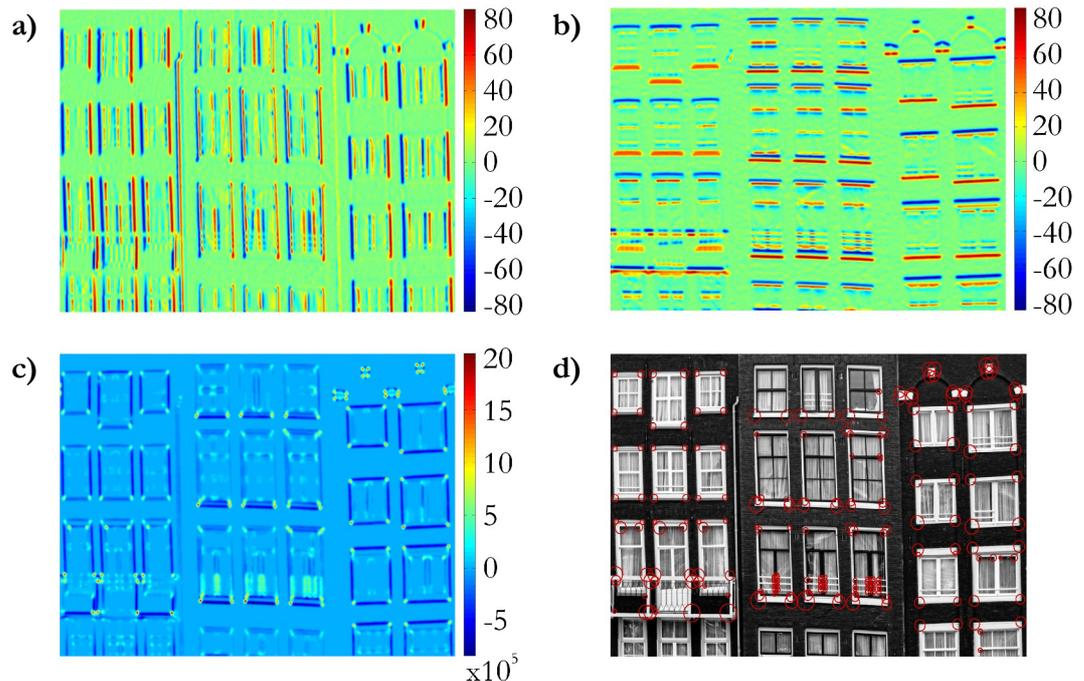


Abbildung 1.4: Anwendung des Harris-Operators auf ein Beispielbild. Änderung der Helligkeitswerte in  $x$ -Richtung (a) und in  $y$ -Richtung (b). c) Kenngröße  $R$  für  $k=0,04$ . Kanten erhalten negative Werte, Ecken positive Werte. d) extrahierte lokale Maxima in  $R$  sind Harris-Eckpunkte (Zusätzlicher Verarbeitungsschritt: Gauß-Filter mit Varianz 5 Pixel auf Ausgangsbild).

Ausgangsbild: Dominik Bartsch, 365/18 - Tilted Houses, CC-Lizenz (BY 2.0), <http://creativecommons.org/licenses/by/2.0/de/deed.de>, Quelle: [www.piqs.de](http://www.piqs.de)

### 1.3.3. Der Laptev-Detektor für räumlich-zeitliche Ereignispunkte

Laptev [9] beschrieb eine Erweiterung des Harris-Operators um Bildsegmente großer raum-zeitlicher Varianz in zusammenhängenden Bildfolgen zu detektieren. Zur Repräsentation der Zeit wird die Bildsequenz als dritte Dimension aufgefasst, Bildpunkte werden als räumlich-zeitliche Voxel betrachtet.

Die Fensterfunktion bleibt für die beiden Bildachsen mit  $\sigma$  gleich und wird in der Zeitachse um  $\tau$  erweitert

$$w_{u,v,t; \sigma, \tau} = w_{st} = \frac{1}{\sqrt{(2\pi)^3 \sigma^4 \tau^2}} \exp(-(u^2+v^2)/2\sigma^2 - t^2/2\tau^2) \quad (10)$$

Die Matrix  $M$  der quadratischen Form lässt sich schreiben als

$$M = \begin{pmatrix} \sum_{u,v} w_{st} \left( \frac{\partial I}{\partial x} \right)^2 & \sum_{u,v} w_{st} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} & \sum_{u,t} w_{st} \frac{\partial I}{\partial x} \frac{\partial I}{\partial t} \\ \sum_{u,v} w_{st} \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} & \sum_{u,v} w_{st} \left( \frac{\partial I}{\partial y} \right)^2 & \sum_{u,v} w_{st} \frac{\partial I}{\partial y} \frac{\partial I}{\partial t} \\ \sum_{u,t} w_{st} \frac{\partial I}{\partial x} \frac{\partial I}{\partial t} & \sum_{u,v} w_{st} \frac{\partial I}{\partial y} \frac{\partial I}{\partial t} & \sum_{u,v} w_{st} \left( \frac{\partial I}{\partial t} \right)^2 \end{pmatrix} \quad (11)$$

mit den Eigenwerten  $\lambda_1, \lambda_2, \lambda_3$ . Ein Punkt hat die Eigenschaften eines räumlich-zeitlichen Ereignisses, wenn alle drei Eigenwerte groß sind. Analog zum Harris-Detektor lassen sich solche anhand lokaler Maxima in der angepassten Kenngröße  $R$  ermitteln

$$\begin{aligned} R &= \lambda_1 \lambda_2 \lambda_3 - k (\lambda_1 + \lambda_2 + \lambda_3)^3 \\ &= \det(M) - k \text{trace}^3(M) \end{aligned} \quad (12)$$

Um Bildstrukturen auf verschiedenen Skalen gerecht zu werden, wurde der Operator außerdem um die Verwendung verschiedener Skalenräume erweitert [16]. Als erster Arbeitsschritt wird das Bild mit einem Gauß-Kernel gefaltet, dessen Halbwertsbreiten, abgesehen von einem Faktor  $s$ , den Halbwertsbreiten der Fensterfunktion  $w_{st}$  entsprechen ( $\sigma_{\text{Skalenraum}}^2 = s\sigma^2, \tau_{\text{Skalenraum}}^2 = s\tau^2$ ).

Mit der Methode werden beschleunigte Harris-Eckpunkte detektiert. Das lässt im Allgemeinen auf eine Änderung der Bewegungsrichtung von Objekten schließen. Darüber hinaus führt das Zusammentreffen von Strukturen zu großen räumlich-zeitlichen Änderungen und einem Ansprechen des Detektors (Abbildung 1.5).

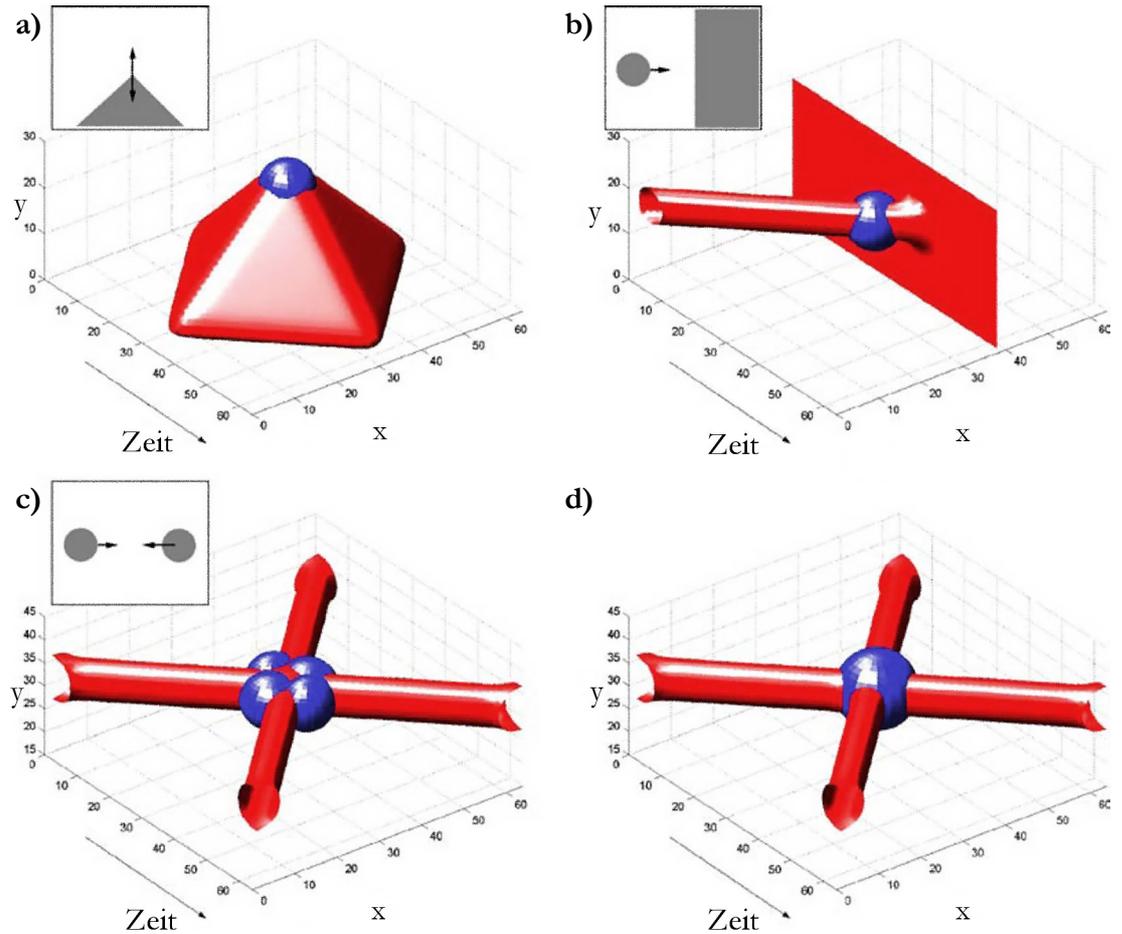


Abbildung 1.5: Resultate für die Anwendung des Laptev-Operators auf synthetische Bildfolgen. Rote Flächen symbolisieren Konturen. Lokale Maxima in  $R$  sind mit blauen Kugeln markiert. a) Richtungswechsel einer Ecke. b) Ein Kreis kollidiert mit einer Wand. c) Kollision zweier Kreise, Parameter  $\sigma^2 = \tau^2 = 8$ . d) Die gleiche Sequenz wie in c) mit  $\sigma^2 = \tau^2 = 16$ . Aus Laptev (2005).

## 1.4. Fragestellung

In dieser Arbeit wurde untersucht, ob mit Hilfe des Laptev-Operators ideale Orte für die topologische Navigation festgelegt werden können. Von Interesse war, ob das Ansprechen des Operators mit der räumlich Geometrie korreliert und insbesondere ob die Idee, Engstellen zu erkennen, damit verwirklicht werden kann. Die Spezifität der Methode für solche Situationen wurde genauer untersucht. Die Korrelation von Bild zu Bild würde für eine eindeutige Bestimmung beispielsweise nicht ausreichen, weil diese auch groß ist, wenn man sich nah an einer Wand befindet und sich auf sie zubewegt. Eine räumliche Engstelle würde sich im Ergebnis davon nicht abgrenzen. Das Vorgehen war explorativ, ausgehend von einer generellen Charakterisierung in einer allgemein gehaltenen Innenraum-Umgebung, hin zu genauer Untersuchung der Detektorantwort an Engstellen mit unterschiedlicher Geometrie.

## 2. Methoden

### 2.1. Auswertung

#### 2.1.1. Detektion von Änderungssalienz

Für die Hypothese, dass Orte, an denen sich die Bildinformation stark ändert, eine Bedeutung für die Schnappschussnavigation haben könnten, wurde für diese Arbeit ein Kriterium für die Änderungssalienz bestimmt. Maßgeblich war die größte Änderung, die für jedes Einzelbild einer Sequenz erreicht wurde, also jeweils der größte Wert vom Feld  $R = R_{\max}$ . Für eine sinnvolle Skalierung wurde außerdem ein logarithmischer Zusammenhang hergestellt:  $\text{Änderungssalienz} = \ln(1 + R_{\max})$ .

Ungerichtete Eigenbewegungen, wie sie zum Beispiel durch Instabilität der Kameraaufhängung entstehen, hätten Pseudobeschleunigung und damit ungewolltes Ansprechverhalten generiert. In der Praxis hätte die entstandene Bildsequenz einer Route daher vorverarbeitet werden müssen, indem die Einzelbilder zueinander ausgerichtet werden. Eine beschleunigte Kamerafahrt hätte ebenfalls zu ungewünschten Ergebnissen geführt, weswegen der Abstand zwischen zwei konsekutiven Aufnahmen mit der zurückgelegten Strecke hätte abgeglichen werden müssen. Für eine prinzipielle Analyse wurde die Ausgangslage der Experimente daher vereinfacht: in virtuellen, computergenerierten Innenräumen wurden geradlinige Strecken mit konstanter Geschwindigkeit simuliert. Weiteres ungewolltes Ansprechverhalten entstand in Randbereichen, wenn Eckpunkte ins Sichtfeld geraten oder verschwinden, weswegen ein Rahmen ignoriert werden muss.

#### 2.1.2. Implementation der Auswertung

Die Änderungssalienz wurde in C++ auf der Grundlage eines bestehenden, angepassten Quelltextes von Ross Messing, University of Rochester, berechnet (STIP-Velocity-Features, Quelle: <http://www.codeforge.com/s/0/stipdet>). Er basierte auf Programmcode von Schuldt, Laptev und Caputo. Hauptsächlich wurde die frei verfügbare Programmbibliothek OpenCV 2.4.5 (open source computer vision) verwendet, in welcher die gängigen Bildverarbeitungsalgorithmen integriert sind.

Die Implementation des Skalenraums war darin aus Effizienzgründen in einer Pyramidenrepräsentation mit vier Auflösungsstufen realisiert. Das Ausgangsbild wurde mit einem  $3 \times 3$  Gauß-Kernel mit Standardabweichung  $\sigma = 0,95$  gefaltet und die Anzahl der Pixel in Breite und Höhe im Anschluss mit der Nächster-Nachbar-Methode halbiert. Das Vorgehen wurde mit dem so erstellten Bild wiederholt, bis, inklusive Ausgangsbild, vier Bilder entstanden waren. Für jede Bildebene wurde das Feld mit der Änderungskenngröße  $R$  für alle Parameterkombinationen der Fensterfunktion  $w_{\sigma, \tau}$  mit  $\sigma^2 = 4$ ,  $\sigma^2 = 8$  und  $\tau^2 = 2$ ,  $\tau^2 = 4$  und mit  $k = 0,0005$  berechnet, und ein Rand von 10 Pixeln verworfen.

Der Quellcode wurde zunächst von OpenCV 1 in C auf die C++-Bibliothek OpenCV 2 angepasst. Zusätzlich wurde die Analyse der verschiedenen Routen automatisiert indem pro Einzelbild jeweils das Maximum von  $R$  extrahiert und protokolliert wurde. Die Größe der Fenstermaske wurde für eine Effizienzverbesserung beschränkt auf  $2,3 * \sigma$  bzw.  $2,3 * \tau$ .

### 2.1.3. Auswahl von Ankerpunkten für die topologische Navigation

Um zu überprüfen, ob sich die erzeugten Daten für eine Detektion des Verschwindens raumverdeckender Strukturen im Sichtfeld bei der Fortbewegung eignen, das heißt um die Position zu erkennen, an der das angestiegene Signal wieder verschwindet, wurde ein automatischer Entscheidungsweg zur Detektion von topologischen Ankerpunkten entwickelt. Dessen Parameter wurden auf die Gesamtheit der durchgeführten Versuche angepasst und überall gleich angewendet. Als Datengrundlage diente für jedes Einzelbild einer Route das Maximum der Änderungssalienz von allen protokollierten Parameterkombinationen. Wegen großer Schwankungen wurde der Verlauf zunächst mit einem Gauß-Kernel ( $\text{Sigma} = 20$ , Fensterbreite = 39) stark geglättet. Im Datensatz wurde dann nach „Ereignissen“ gesucht: Positionen, an denen ein bestimmter Schwellenwert überschritten wurde. Im weiteren Routenverlauf wurde dann die erste Position, auf der wieder ein Detektorwert unter einem zweiten, niedrigeren Schwellenwert erreicht wurde, als Ankerpunkt definiert (farbige Pfeile mit rotem Fadenkreuz visualisieren diese Positionen in den Abbildungen). Aufgrund der stark variierenden Detektorwerte je nach Texturierung hat sich eine Kombination aus absoluten und relativen Schwellen als nützlich erwiesen: als obere, erste Schwelle

wurde festgelegt, dass 30 % des Maximalwertes auf der gegebenen Route erreicht werden musste und dass mindestens ein Wert von  $R_{\max} = 0,15$  erreicht werden musste. Die untere, zweite Schwelle wurde so festgelegt, dass der Mittelwert aus zwei Kriterien gebildet wurde: aus 5 % des Maximalwertes auf der gegebenen Route und aus  $R = 0,02$ . Zusätzlich wurden nur Ereignisphasen mit einer Länge von höchstens 300 Einzelbildern der Bildsequenz zugelassen. Weil die Daten mit der Gauß-Glättungsmaske um jeden Punkt herum miteinander verrechnet wurden, wurde der zu detektierende starke Abfall des Signals nach hinten verschoben. Im letzten Schritt wurde daher, unter Berücksichtigung der Fensterbreite von 39 Einzelbildern, die Position für den Ankerpunkt um 20 Einzelbilder zurückversetzt.

## **2.2. Simulation virtueller Umgebungen**

### **2.2.1. Design der 3D-Welten**

Sämtliche Experimente wurden in simulierten, virtuellen Umgebungen durchgeführt, die mit dem freien 3D-Modellierungs- und Animationsprogramm Blender 2.68.0 (Blender Foundation, GPL) entworfen wurden. Texturen wurden unter dem Gesichtspunkt möglichst hoher Kontraste ausgewählt und waren im allgemeinen schwarzweiß. Jedes Umgebungsmodell wurde mit je zwei unabhängigen Texturzusammenstellungen ausgestattet, damit die Invarianz der Methode bei geometrisch gleichen Konfigurationen untersucht werden konnte (Abbildung 2.1, Abbildung 2.2). Alle Materialien hatten in den drei Farbkanälen (RGB) die gleichen Reflexionseigenschaften (ambient: 82 %, diffus: 18 %, spiegelnd: 0 %). Die Deckenhöhe betrug in allen Räumen 2,5 Meter.

Die Modelle wurden ins Wavefront-Dateiformat (Wavefront Technologies) in ASCII-Kodierung exportiert. In den Dateien lagen die zu rendernden Vertex-, Flächen-, Material-, sowie Textureigenschaften in OpenGL-kompatiblen (Khronos Group) Format vor.

### 2.2.2. Rendern und Ablaufsteuerung

Experimente wurden mittels gerendeter Bildsequenzen vorgegebener Routen durchgeführt. Anfangs- und Endpunkte wurden vorher in MATLAB 2013a (The MathWorks, Inc.) definiert und sämtliche Koordinaten für das Rendern der Einzelbilder mit den festgelegten Parametern Geschwindigkeit (0,25 m/s) und Bildabtastrate (24 Hz) berechnet und im Anschluss ins CSV-Dateiformat exportiert.

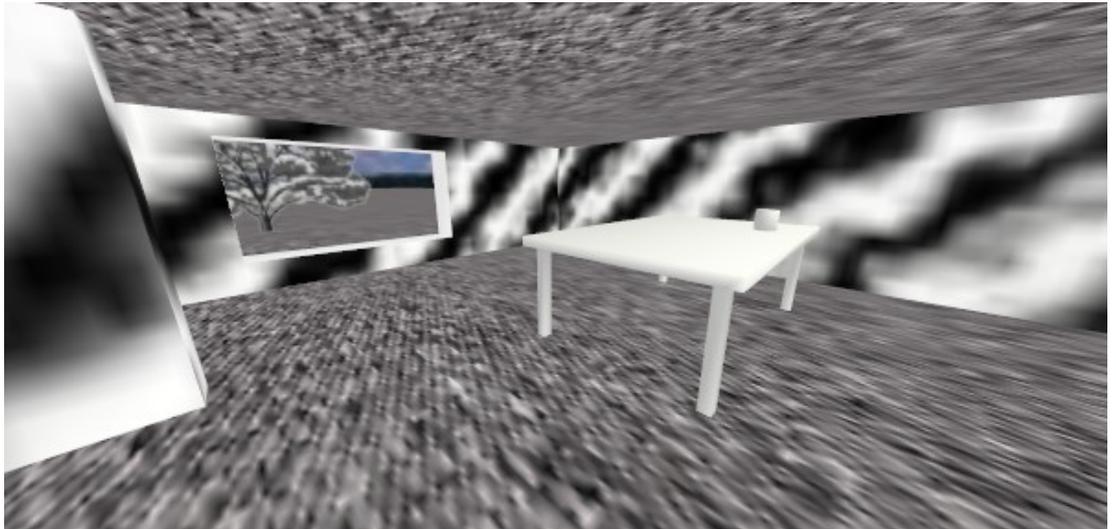


Abbildung 2.1: Typisches gerendertes Einzelbild in der virtuellen Umgebung mit Texturzusammenstellung 1.

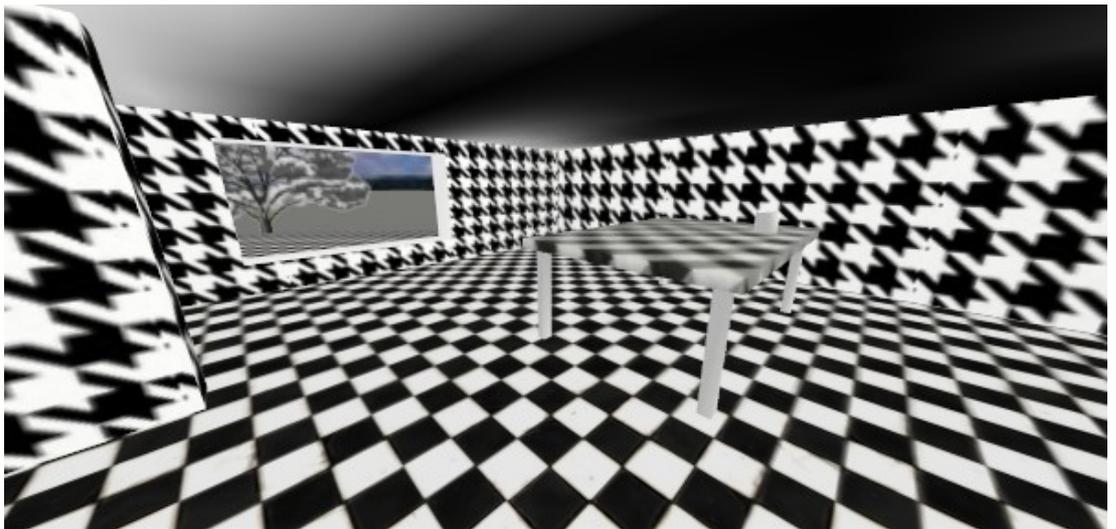


Abbildung 2.2: Derselbe Raum wie in Abbildung 2.1 mit noch kontrastreicherer Texturzusammenstellung 2. Sämtliche Experimente wurden mit den zwei verschiedenen Texturzusammenstellungen durchgeführt.

Das Rendern erfolgte mit einem OpenGL-Szenengraphsystem, der quelloffenen Programmibliothek OpenSceneGraph 3.0.1 (OpenSceneGraph Public License), in

der Programmiersprache C++ (Abbildung 2.1, Abbildung 2.2). Als Entwicklungsumgebung diente Visual Studio 2008 (Microsoft Corporation, Compiler: Visual C++). Das Gesichtsfeld wurde auf  $120^\circ$  in horizontaler Richtung und  $80^\circ$  in vertikaler Richtung bei einer Neigung des Blickwinkels um  $15^\circ$  nach unten festgelegt, mit einem Abstand zum Boden von 1,6 Metern. Die Auflösung betrug  $600 \times 290$  Pixel. Die Fortbewegungsrichtung war immer parallel zum Boden in Richtung des Zentrums des Gesichtsfeldes. Um Treppeneffekte und damit falsche Eckpunkte zu vermeiden, wurden die Szenen mit Antialiasing berechnet (Nvidia 16xQ CSAA).

Die Einzelbilder wurden im Portable-Network-Graphics-Format (PNG) gespeichert und anschließend mit MATLAB zu einem Video der jeweiligen Route zusammengefügt (komprimierte AVI-Datei, MJPEG).

## 2.3. Experimente

### 2.3.1. Systematische Untersuchung anhand kurzer Routen

Für eine generelle und systematische Analyse der Detektorantwort wurde eine virtuelle Innenraum-Umgebung, bestehend aus zwei mit einem Durchgang verbundenen Räumen, erstellt (Abbildung 2.4). Raum 1 (Westen, Abbildung 2.1, Abbildung 2.2) hatte eine rechteckige Grundfläche von  $9,76 \times 9,42$  Metern. An der dem Durchgang gegenüberliegenden, westlichen Seite befand sich ein Fenster von 3,25 Metern breite und 1,75 Metern Höhe. Dahinter war ein Baum (mit teilweise transparenten Bereichen) positioniert und in weiter Ferne der Himmel. An der südwestlichen Seite teilte eine in den Raum ragende Wand von 2,87 Metern Länge einen Bereich von 2,78 Metern Breite ab. Nordwestlich stand ein Tisch von 3 Metern Länge, 2 Metern Breite und 1 Meter Höhe. Die rechteckige Grundfläche von Raum 2 (Osten, Abbildung 2.3) hatte eine Abmessung von  $9,76 \times 9,57$  Metern. Dort befanden sich, in etwa mittig, zwei Säulen mit einem Durchmesser von 0,93 Metern. Eine der Säulen ging in einen 1,15 Meter hohen Kasten von 3,25 Metern Länge und 0,86 Metern Breite über. Die Beleuchtung bestand hauptsächlich aus einer ambienten Lichtquelle (OpenGL) ohne räumlicher Abschwächung. Für eine leichte Schattierung sorgte je eine, etwa mittig unter der Decke angebrachte, diffuse Lichtquelle (OpenGL) pro Raum mit quadratischer räumlicher Abschwächung.

Systematisch wurden an 188 Positionen, in regelmäßigen Abständen, kurze Routen ausgewertet (Abbildung 2.4). An jeder der Positionen wurden Bildsequenzen in acht Richtungen in 45°-Schritten generiert, indem je 11 Einzelbilder davor und dahinter aufgenommen und für die Berechnung der Änderungssalienz an der Referenzposition verwendet wurden.

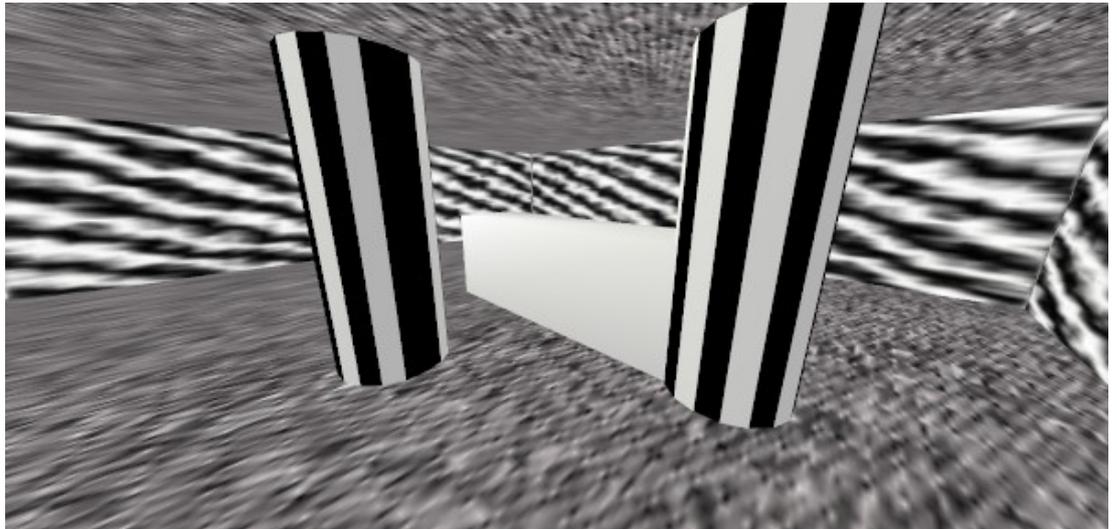


Abbildung 2.3: Blick in den zweiten Raum mit Texturzusammenstellung 1.

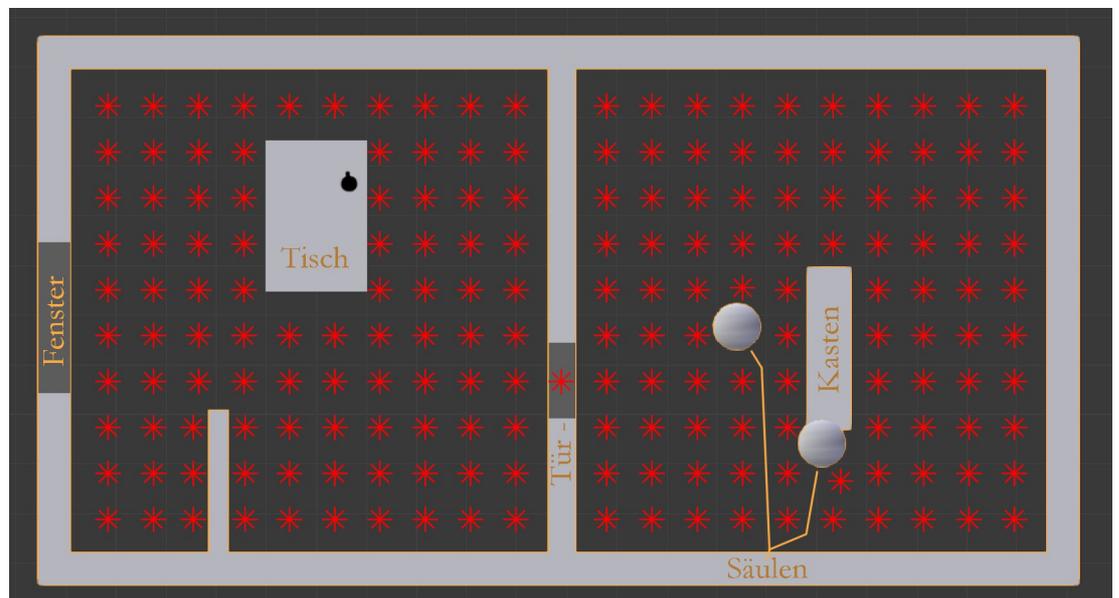


Abbildung 2.4: Grundriss der virtuellen Innenraum-Umgebung mit den in Experiment 1 getesteten, kurzen Routen (rote Linien).

### 2.3.2. Untersuchung anhand ausgewählter längerer Routen

Um den zeitlichen Verlauf des Detektorsignals zu evaluieren wurden ausgewählte längere, etwa vier Meter weite, Routen in derselben Innenraum-Umgebung von Ex-

periment 1 bestimmt, die Bildsequenz jeweils in beide Richtungen erzeugt und ausgewertet (Abbildung 2.5). Entscheidendes Merkmal für die ausgewählten Strecken war, dass sie nah an räumlichen Strukturen entlang verliefen.

Die Durchfahrt durch den mittige Durchgang (Tür) wurde orthogonal (mittig darauf zusteuernd), sowie schräg im  $45^\circ$ -Winkel (von rechts und von links darauf zusteuernd) simuliert. Kreuzpunkt all dieser Routen war genau in der Mitte des Durchgangs. Südöstlich, neben eine der Säulen, wurde die Vorbeifahrt parallel zur Wand getestet. Eine hohe Dichte verschiedener Routen wurde in der Nähe der südöstlichen, in den Raum ragenden Wand simuliert: parallel zu beiden Seiten der Wand, im  $90^\circ$ -Winkel an ihrem Endpunkt vorbei, sowie im  $45^\circ$ -Winkel daran vorbei.

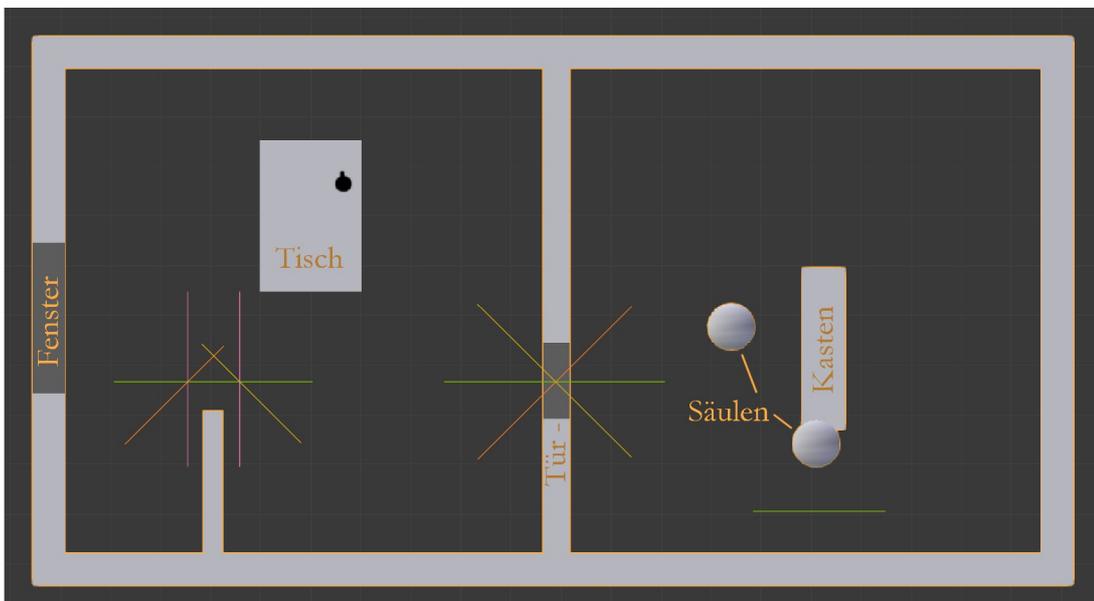


Abbildung 2.5: Grundriss der Innenraum-Umgebung mit für Experiment 2 ausgewählten Routen (farbig richtungskodierte Linien). Die Routen wurden jeweils in beide Richtungen simuliert.

### 2.3.3. Evaluation anhand von Routen durch enge Passagen

Zur Untersuchung des Einflusses verschiedener Raumgeometrien auf die Detektorantwort wurden vier simple virtuelle Räume (15 x 14 Meter Grundfläche) mit engen Passagen erstellt (Abbildung 2.6). Jede Engstelle war an der schmalsten Stelle 2 Meter breit. Raum 1 war der einzige nicht-symmetrische und bestand aus einem Korridor, der in ein größeres Zimmer mündete. In Raum 2 befand sich ein Durchgang, der einer Tür zwischen zwei Zimmern nachgeahmt war. Raum 3 hatte gerade, sich zur

Engstelle hin zuspitzende Seitenwände. Raum 4 war ähnlich wie Raum 3 aufgebaut, mit dem Unterschied, dass die Wände einen kreisförmigen Querschnitt hatten.

Um die Durchfahrten möglichst detailliert zu charakterisieren wurden Routen getestet, die orthogonal zur Engstelle genau mittig durch die Passage verliefen, sowie Routen, die um 0,66 Meter nach links oder rechts versetzt waren. Sie führten also jeweils näher an einer Seite der Engstellen entlang als an der anderen. Außer im ersten Raum, wo die Raumgeometrie es nicht zuließ, wurden außerdem um  $20^\circ$  schräg verlaufende Routen getestet (von links und rechts auf den Durchgang zusteuernd), die die Engstelle jeweils mittig kreuzten.

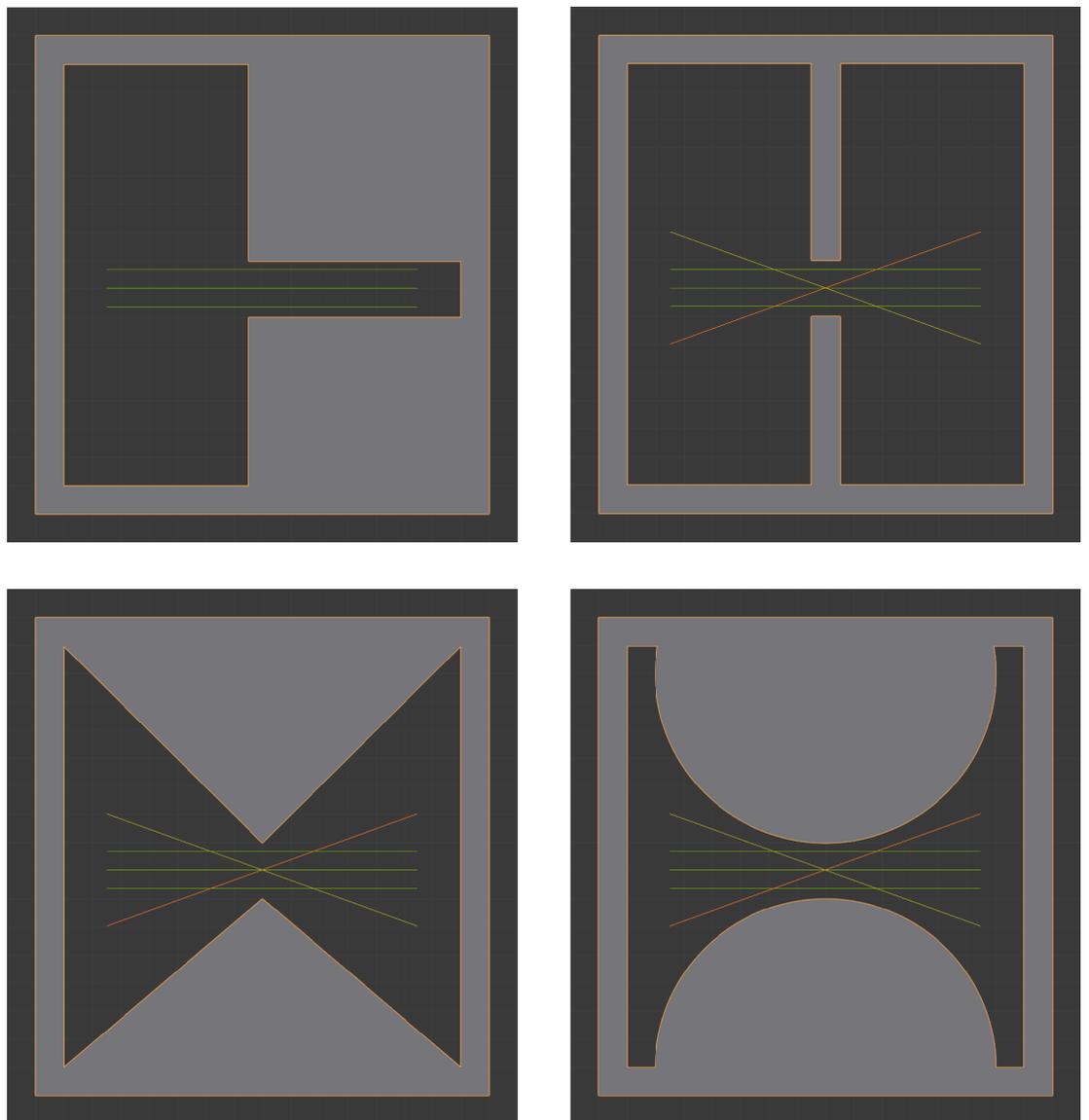


Abbildung 2.6: Grundriss der vier Räume mit engen Passagen unterschiedlicher Geometrie. Die Routen wurden jeweils in beide Richtungen simuliert (farbig richtungskodierte Linien).

### 3. Ergebnisse

#### 3.1. Allgemeines Ansprechverhaltens des Detektors

Um die Stärke der Detektorantwort darzustellen, wurden in den Grundriss der virtuellen Räume Pfeile in Bewegungsrichtung eingetragen. Maßgeblich war das Maximum der protokollierten Maxima von Kenngröße R aller Parameterkombinationen und Bildpyramidenniveaus. Die Pfeillänge stand in einem logarithmischen Verhältnis zur Signalstärke (Pfeillänge in Metern in virtueller Umgebung =  $10^4 * \ln(1+R_{\max})$ , Pfeillängen > 22 cm wurden gleich 22 cm gesetzt). Der Detektor generierte einen hohen Wert, wenn eine nahe Geometrie im Sichtfeld eine weiter entfernte verdeckte und sich optisch beides gegeneinander verschob, so zum Beispiel in der Nähe des Durchgangs, der Säulen und der in den Raum ragenden Wand (Abbildung 3.1). Dagegen ergab er einen niedrigen Wert, wenn direkt auf eine Wand zugesteuert wurde.

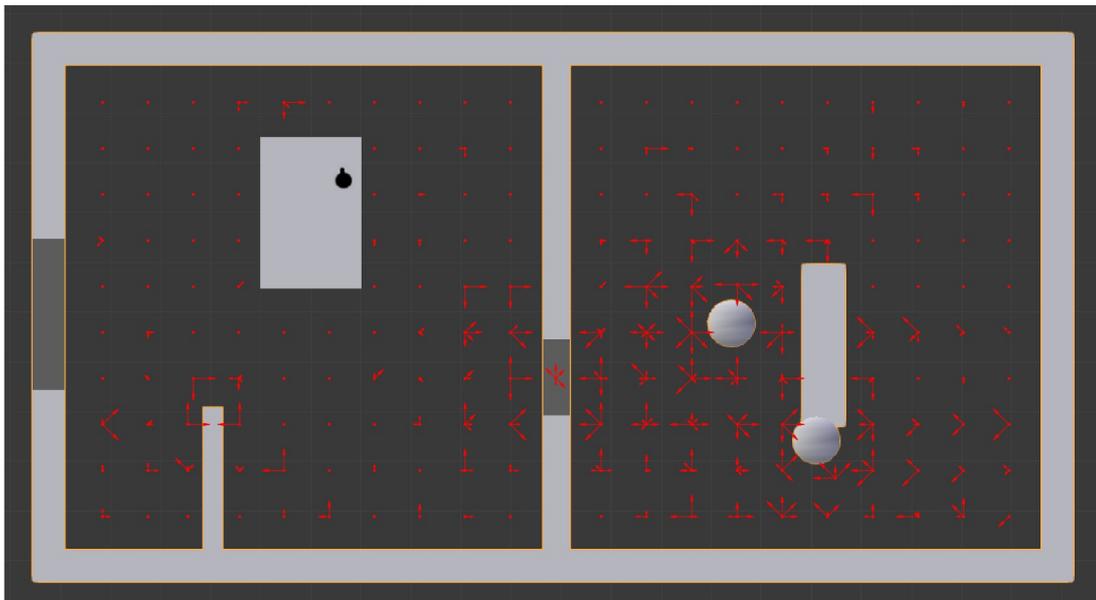


Abbildung 3.1: Detektorantwort für Texturzusammenstellung 1 (Pfeillänge proportional zu  $\ln(1+R_{\max})$ ). Nahe räumlicher Geometrien, die sich bei der Fortbewegung optisch gegenüber weiter entfernten verschoben, war die Änderungsalienz hoch, in der Nähe von Wänden dagegen niedrig.

Wichtig für ein zuverlässiges Signal waren starke Kontraste. In Texturzusammenstellung 1 war der Tisch weiß, entsprechend niedrig fielen in seinem Umfeld die Detektorwerte aus. Mit Texturzusammenstellung 2 (Abbildung 3.2) waren sie dagegen sehr hoch, weil der Tisch stark texturiert war. Der Befund konnte weiter eingegrenzt werden: in der Nähe des Fensters reichte ein starker Kontrast nicht aus. Die Wände waren im ersten Fall zwar kontrastreich aber mit geradlinigen Streifen texturiert. Eine

starke Detektorantwort gab es nur im zweiten Fall, in dem die Wandtexturen Eckenstrukturen hatten.

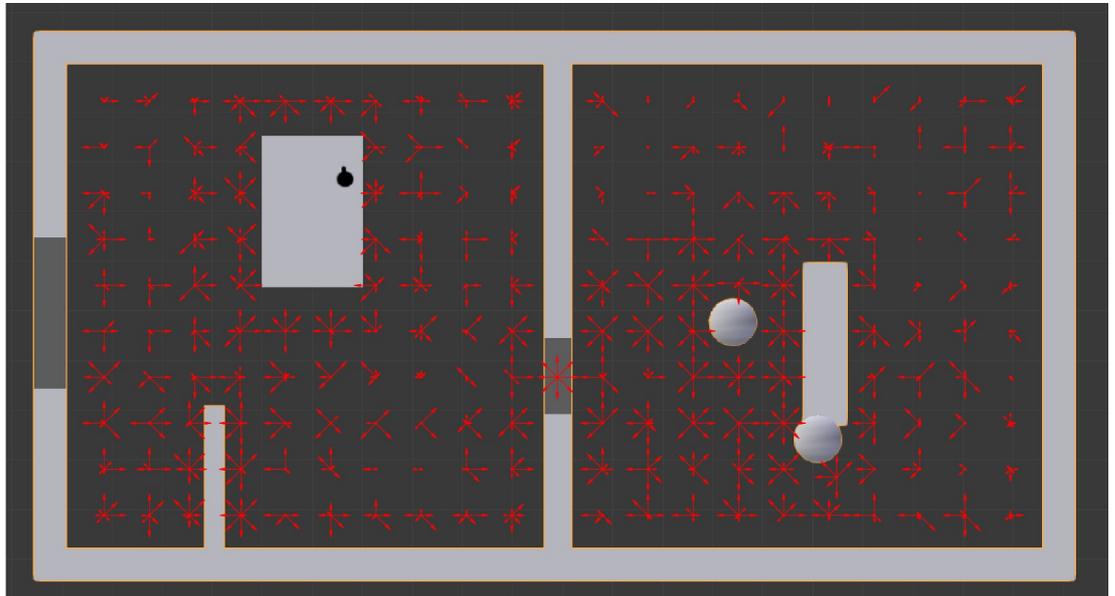


Abbildung 3.2: Detektorantwort für Texturzusammenstellung 2. Stärkere Kontraste bzw. Texturen mit Ecken erzeugten ein stärkeres Signal. Räumliche Strukturen wurden offenbar, die zuvor zu keiner Antwort geführt hatten.

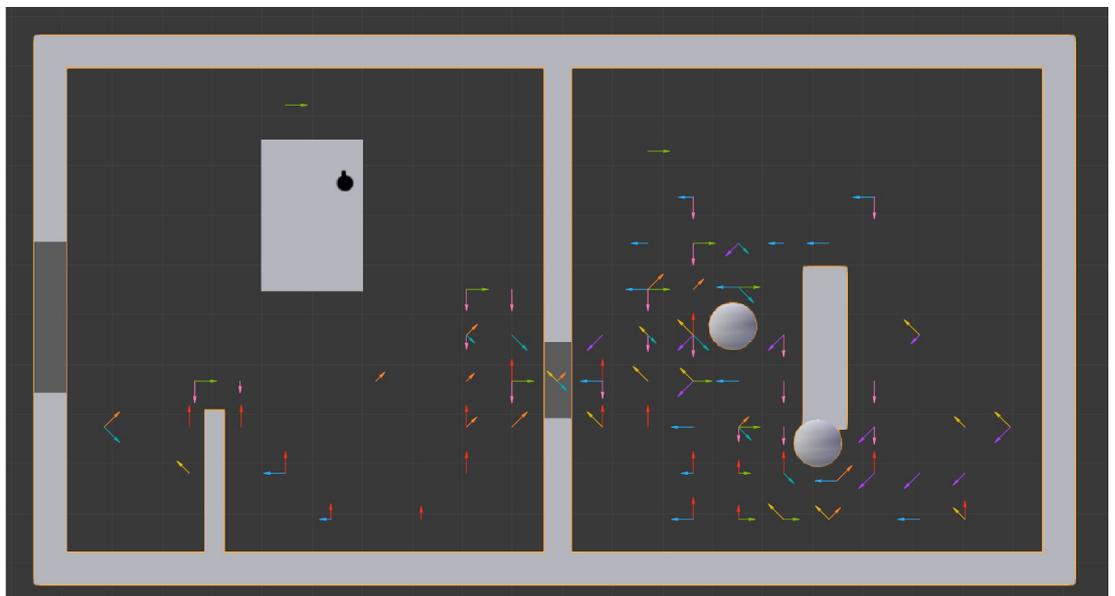


Abbildung 3.3: Texturzusammenstellung 1: Nur Detektorwerte, auf die in gleicher Richtung ein kleinerer Wert folgte und die einen Schwellenwert von  $2 \cdot 10^5$  überstiegen wurden eingetragen. Die Bewegungsrichtung ist farblich gekennzeichnet. Nach einem Durchgang bzw. nachdem eine räumliche Struktur passiert wurde war häufig ein Abfall eines vorher starken Signals zu beobachten.

Anhand derselben Daten wurde überprüft, ob das Durchschreiten eines Durchgangs bzw. das Vorbeilaufen an einer räumlichen Struktur durch ein Abfallen eines vorher starken Signals detektiert werden konnte (Abbildung 3.3, Abbildung 3.4). Dafür wur-

den nur Pfeile eingezeichnet, für die der nächste Detektorwert in dieselbe Richtung geringer ausfiel. Es zeigte sich, dass dies in den meisten Fällen kurz vorm Passieren besagter Raumstrukturen der Fall war. Solche Orte waren zum Beispiel in der Nähe des mittigen Durchgangs, neben der in den Raum ragenden, südwestlichen Wand und neben den südöstlichen Säulen. Umgekehrt ließen sich von den übriggebliebenen Positionen aus die korrespondierenden räumlichen Strukturen finden: sie verschwanden an dieser Stelle aus der Sicht. Das lässt sich nachvollziehen, wenn man das Gesichtsfeld von  $120^\circ$  berücksichtigt. Weiterhin konnte man sehen, dass Pfeile in eine Bewegungsrichtung (farblich kodiert), häufig etwa  $45^\circ$  zur Fortbewegungsrichtung, nebeneinander lagen und linear miteinander verbunden werden konnten.

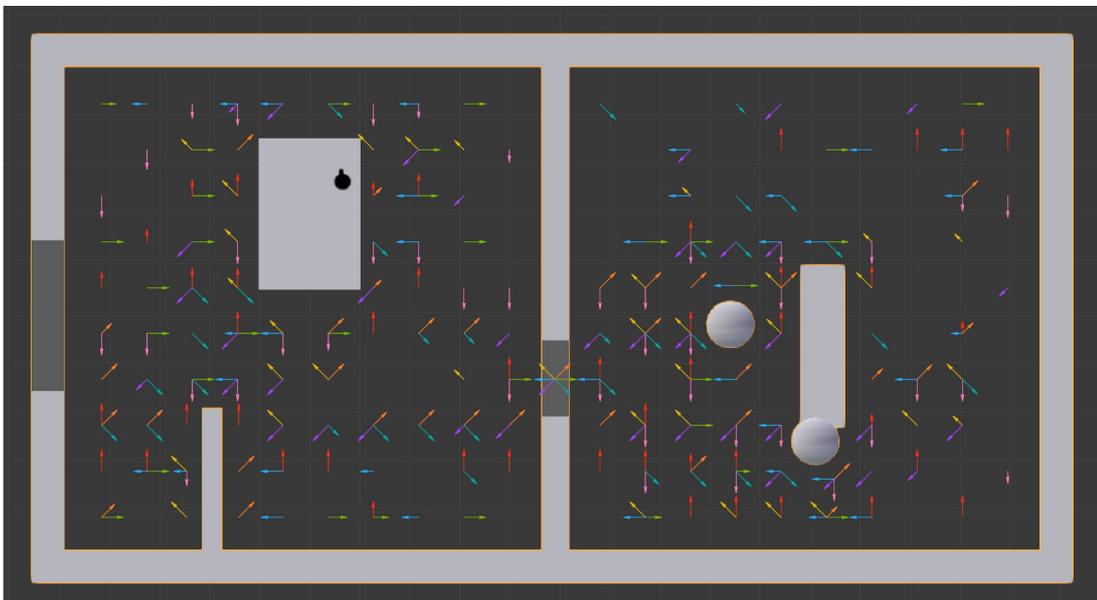


Abbildung 3.4: Texturzusammenstellung 2, Nur Detektorwerte, auf die in gleicher Richtung ein kleinerer Wert folgt, werden dargestellt. Pfeile in eine Richtung (farblich kodiert) liegen an vielen Stellen räumlich bezogen auf ihre korrespondierenden räumlichen Strukturen nebeneinander und lassen sich miteinander verbinden.

### 3.2. Änderungssalienz-Profile ausgewählter längerer Routen

Die Änderungssalienz stieg bei allen ausgewählten Routen langsam an und verringerte sich sehr schnell wieder, sobald die korrespondierende räumliche Struktur aus dem Blickfeld verschwunden war (Abbildung 3.5, alle Datensätze im Anhang). Das Profil des Verlaufs war in allen Fällen sehr ähnlich. Trotz logarithmischer Skalierung war der Maximalwert jedoch sehr unterschiedlich, abhängig zum Beispiel davon, wie ausgeprägt die Texturierung war (Texturzusammenstellung 2 führte prinzipiell zu größeren Werten) und vom Winkel, unter dem eine räumliche Geometrie angesteuert wur-

de. Es zeigte sich außerdem eine zeitliche periodische Schwankung des Signals, die auf die Texturmuster zurückzuführen ist.

Aufschlussreich war das Ergebnis von schrägen Routen durch den mittigen Durchgang. Hier gab es jeweils zwei korrespondierende Geometrien, nämlich die rechte und die linke Wand, die zu unterschiedlichen Zeiten aus dem Blickfeld verschwanden (Abbildung 3.6). Im Datenverlauf lassen sich beide Ereignisse nachvollziehen und voneinander unterscheiden.

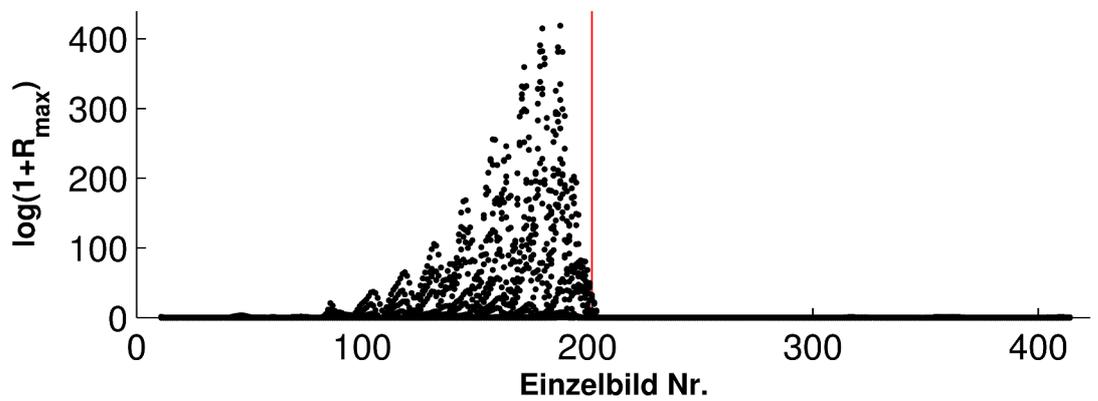


Abbildung 3.5: Änderungssaliensz-Profil der 180°-Route durch den mittigen Durchgang (O → W, Texturzusammenstellung 2). Die Datenpunkte stellen pro Einzelbild der Bildsequenz die Maximalwerte bei allen Parameterkombinationen dar. Der Detektorwert stieg bis etwa zur roten Linie, die das letzte Einzelbild mit Durchgang im Sichtfeld markiert, an und fällt dann stark ab.

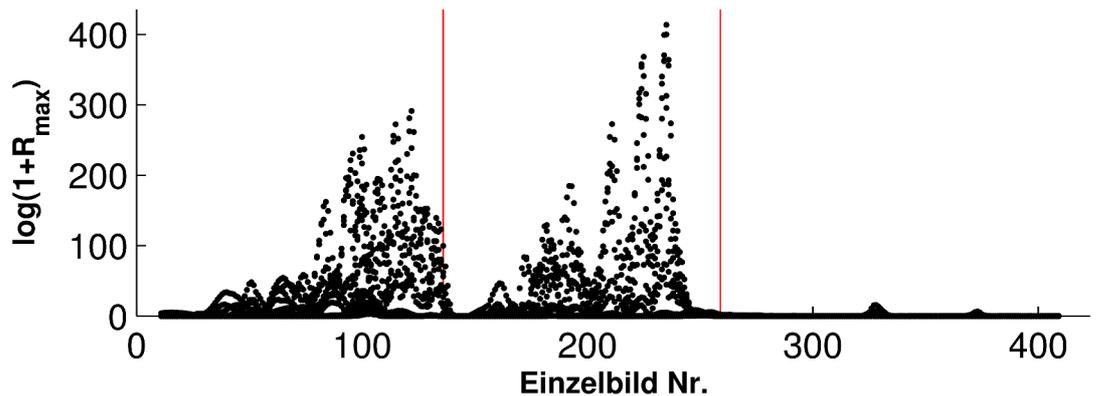


Abbildung 3.6: Änderungssaliensz-Profil für die schräg verlaufende 135°-Route durch den Durchgang (SO → NW, Texturzusammenstellung 2). Die erste rote Linie markiert das Verschwinden des linksseitigen Rahmens, die zweite das Verschwinden des rechtsseitigen Rahmens aus dem Sichtfeld.

Qualitativ unterschied sich das Ergebnis für automatisch erkannte Ankerpunkte bei unterschiedlichen Raumtexturen nicht (Abbildung 3.7 - Abbildung 3.10). Es gab kleinere Unterschiede: mit Texturzusammenstellung 1 wurde auf der 45°-Route (SW → NO) durch den mittigen Durchgang das zuvor beschriebene Verschwinden

beider korrespondierender Wände aus dem Sichtfeld zu unterschiedlichen Zeitpunkten erkannt (Abbildung 3.7, orange Pfeile). Aufgrund der erforderlichen starken Glättung der Ausgangsdaten ist dies bei den anderen schrägen Routen nicht der Fall. Mit stärkerer Texturierung wurde auf der gleichen Route (Abbildung 3.9) gar kein Ankerpunkt erkannt, weil mit Verschwinden des Durchgangs aus dem Sichtfeld schon die dahinterliegende Säule Änderungssalienz verursachte (siehe Abbildung A.31).

Dieselbe Säule löste mit Texturzusammenstellung 2 auch auf der 180°-Route südlich der zweiten Säule in relativ weiter Entfernung eine Detektorantwort aus (Abbildung 3.9, O → W, blauer Pfeil). Der detektierte Ankerpunkt korrespondierte mit dem Verschwinden dieser Säule aus dem Sichtfeld. Im Gegensatz dazu korrespondierte der Ankerpunkt bei der ersten Texturvariante mit der näheren Säule (Abbildung 3.7).

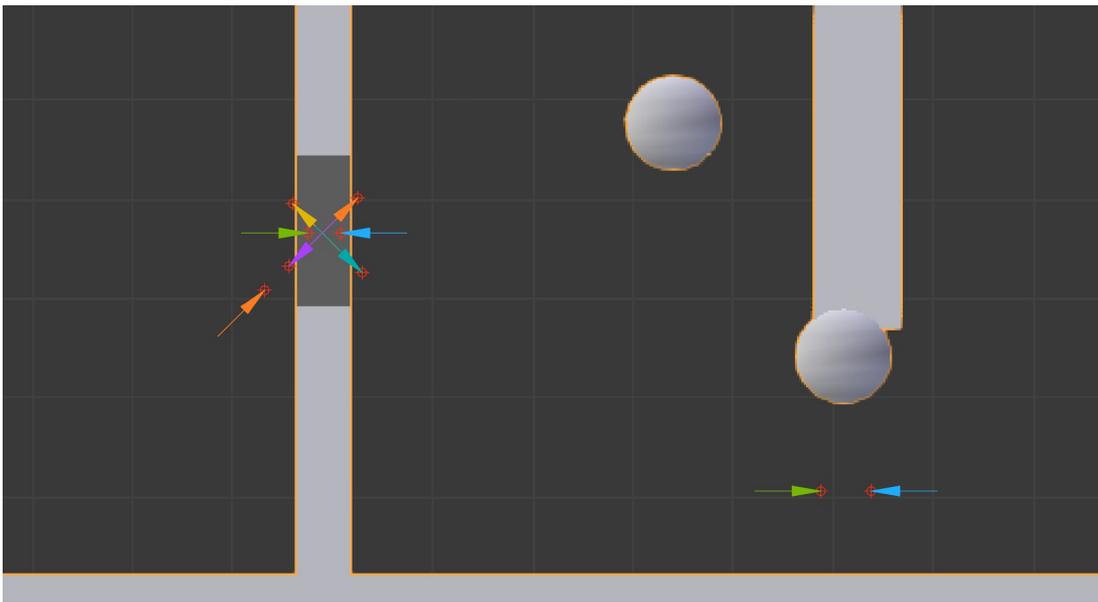


Abbildung 3.7: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 1) für die Routen durch den Durchgang und in der Nähe der Säulen. Rote Fadenkreuze markieren die Positionen, die Bewegungsrichtung ist mit farbigen Pfeilen illustriert.

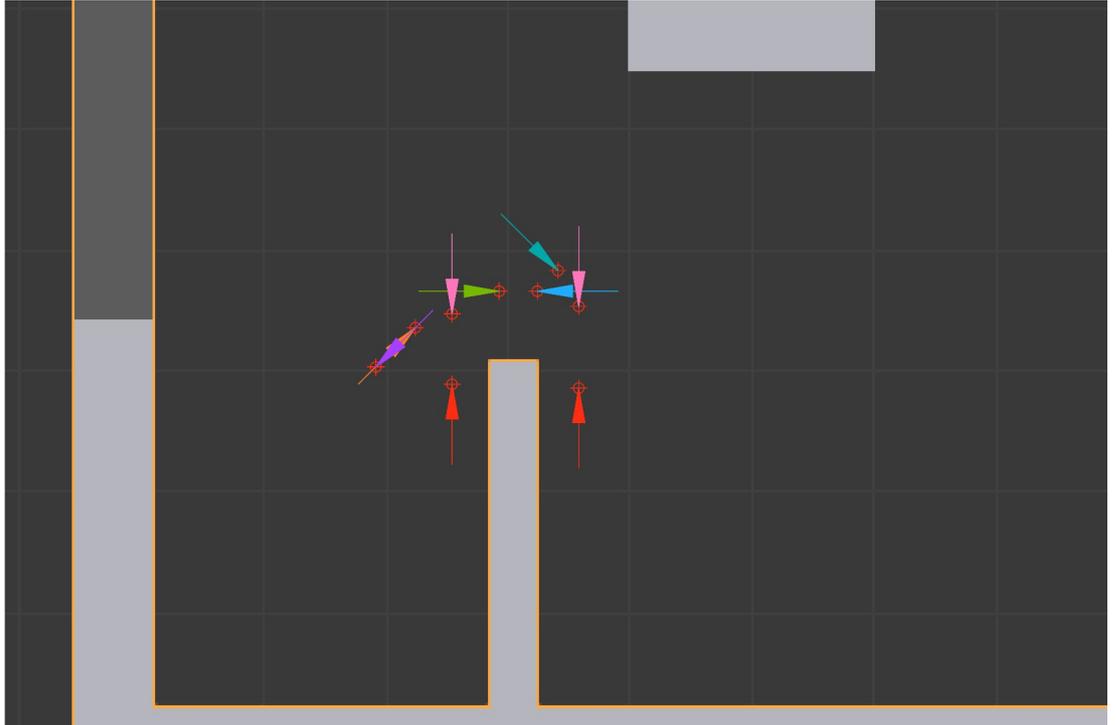


Abbildung 3.8: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 1) für die Routen im Bereich der in den Raum ragenden Wand. Das Verschwinden des verdeckten Raums aus dem Sichtfeld wird genauso erkannt ( $N \rightarrow S$ , rosa Pfeile) wie das Verschwinden der Wand aus dem Sichtfeld ( $S \rightarrow N$ , rote Pfeile)

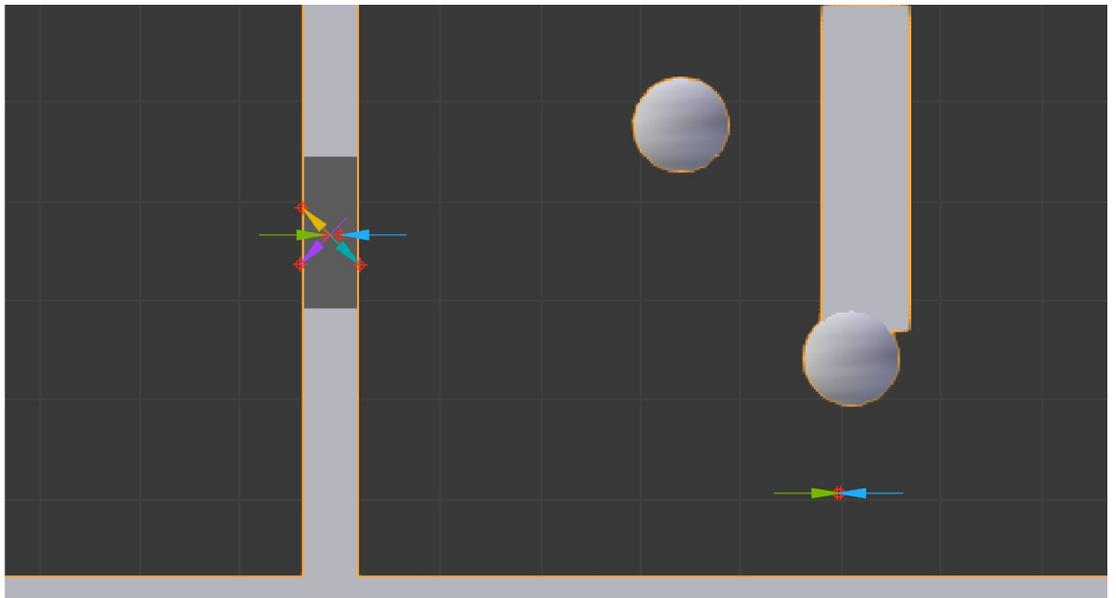


Abbildung 3.9: Automatische Detektion von Ankerpunkten für die Routen durch den Durchgang und in der Nähe der Säulen. Eine stärkere Texturierung (Texturzusammenstellung 2) führte dazu, dass die  $45^\circ$ -Route durch die Tür nicht erkannt wurde, die dahinterliegende Säule verursachte ein überlagertes Signal. Mit derselben, weiter entfernten Säule korrespondiert der Ankerpunkt der  $180^\circ$ -Route an der Säule vorbei ( $O \rightarrow W$ ), anders als bei der ersten Texturzusammenstellungen (Abbildung 3.7).

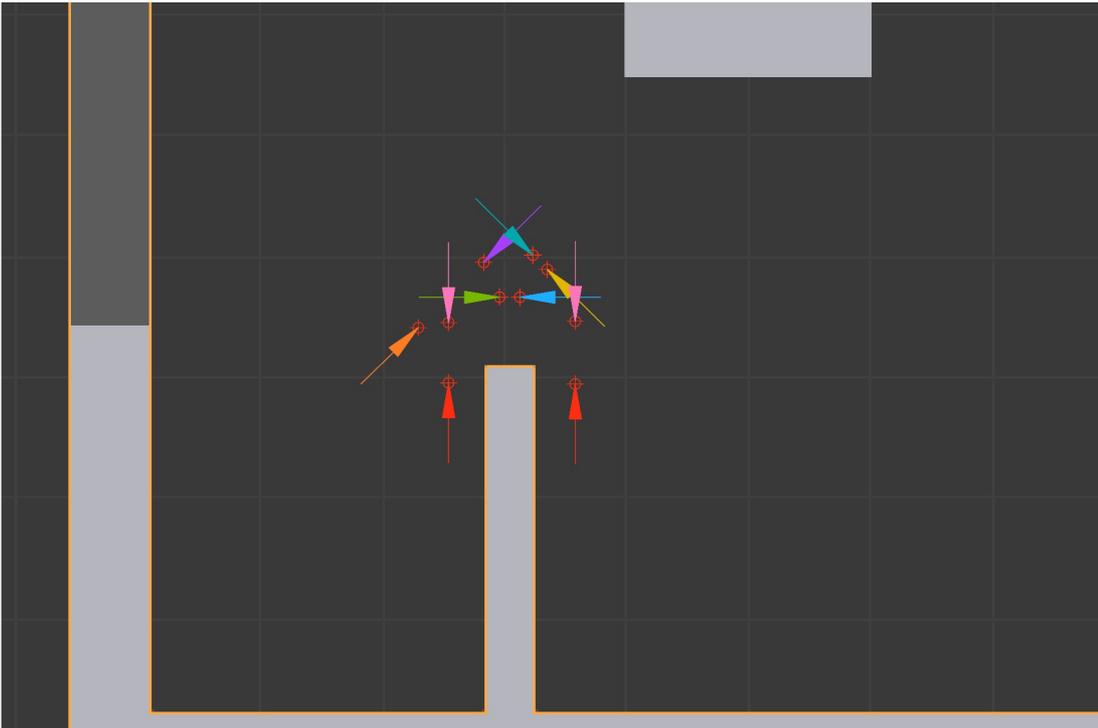


Abbildung 3.10: Automatische Detektion von Ankerpunkten (Texturzusammenstellung 2) für die Routen im Bereich der in den Raum ragenden Wand.

### 3.3. Auswahl von Ankerpunkten in engen Passagen

Die Befunde aus Experiment 2 wurden bei Routen durch enge Passagen mit verschiedenen räumlichen Geometrien grundsätzlich bestätigt (Abbildung 3.11 - Abbildung 3.16). Die Änderungssalienz stieg bei der Bewegung in Richtung der Engstelle langsam an und fiel schnell ab, wenn die zugehörige Wand aus dem Blickfeld verschwand und dahinter liegender Raum nicht mehr verdeckt wurde (alle Datensätze im Anhang).

Bei parallelen, nah nebeneinander verlaufenden Routen lagen in allen Fällen die detektierten Ankerpunkte annähernd auf einem Kreisbogen. Beide Endpunkte des Kreisbogens lagen in den Räumen mit kantigem Profil auf der engsten Stelle, nur im Raum mit gekrümmten Wänden lagen sie viel weiter in Bewegungsrichtung. Die weiteste Ausdehnung des Kreisbogens war immer auf der genau mittig verlaufenden Route, die Ankerpunkte wurden also früher auf dieser Route erkannt. Dass diese Anordnung robust bei allen Raumgeometrien gefunden wurde, ließ sich durch das Sichtfeld von  $120^\circ$  erklären. Beim mittigen durchfahren einer Engstelle verschwanden

den beide Wände am frühesten, je näher an einer Seite entlang gesteuert wurde, desto später verschwand die zugehörige Wand aus dem Sichtfeld.

Für die automatische Detektion der Ankerpunkte zeigte sich ein Abwägungskonflikt für die Parametereinstellungen. Bei starker Texturierung führte eine parallel zur Wand verlaufende Route zu einer schwachen Änderungssalienz, sodass eine höhere Schwelle hätte gewählt werden müssen (Abbildung 3.12). Mit einer niedrigeren Schwelle wären dagegen weitere sinnvolle Ankerpunkte in weniger kontrastreicher Umgebung gefunden worden (Abbildung 3.15). Allerdings war der Anfahrtswinkel in diesen Fällen sehr klein.

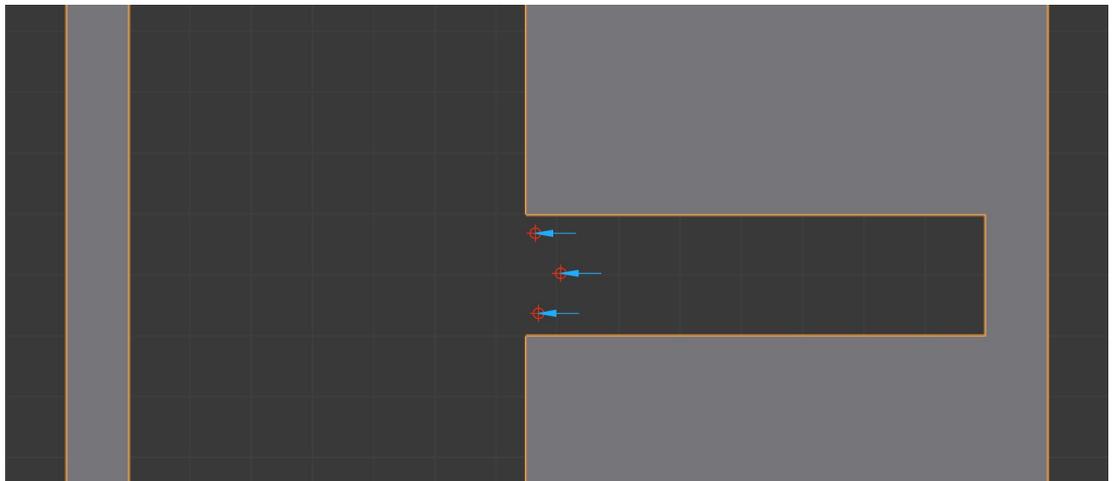


Abbildung 3.11: Bei einer Bewegung aus einem Korridor heraus (Texturzusammenstellung 1) wurden, bei parallel verlaufenden Routen, Ankerpunkte in kreisbogenförmiger Anordnung detektiert. In den Korridor hinein entstand keine Änderungssalienz, weil kein Raum im Sichtfeld verdeckt wurde.

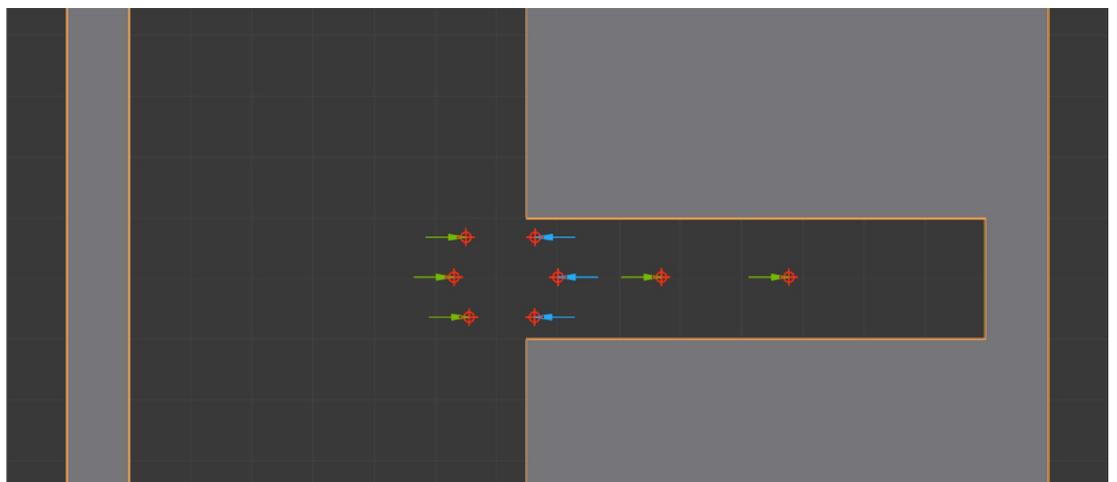


Abbildung 3.12: Entgegen der Variante mit Texturzusammenstellung 1 (Abbildung 3.11) wurden bei der stärker gemusterten Texturzusammenstellung 2 bei Bewegung in den Korridor hinein Ankerpunkte detektiert. Routen sehr nahe und parallel zu Wänden führten zu einer schwachen Änderungssalienz (siehe auch Abbildung 3.15).

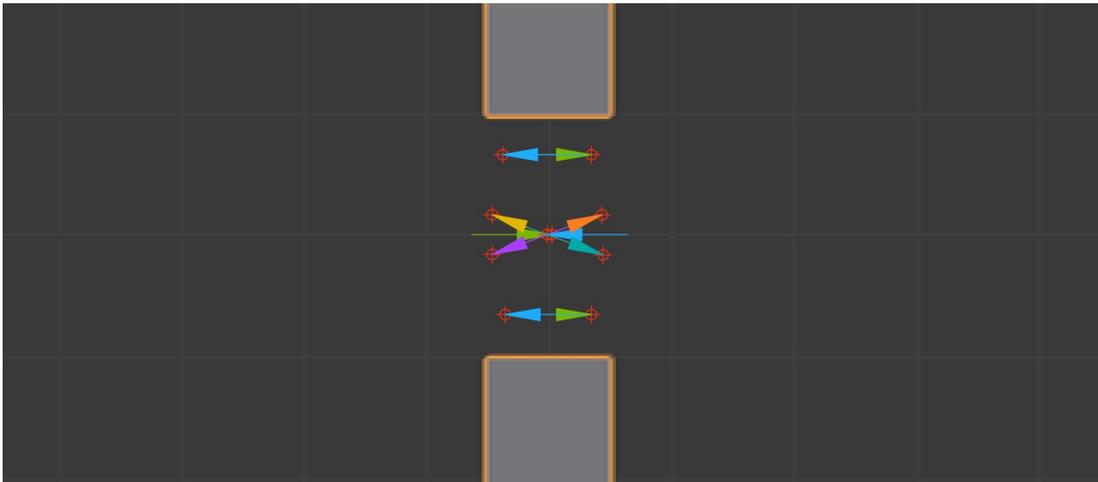


Abbildung 3.13: Bei paralleler Route durch den türähnlichen Durchgang (Texturzusammenstellung 2) ergab sich auch hier eine kreisbogenförmige Anordnung der Ankerpunkte. Die schräge Durchfahrt ergab, entsprechend dem späteren Verschwinden der rechtsseitigen Wand aus dem Sichtfeld, einen späteren Ankerpunkt. Die Methode war hier gegenüber der Texturzusammenstellung invariant (Abbildung A.109).

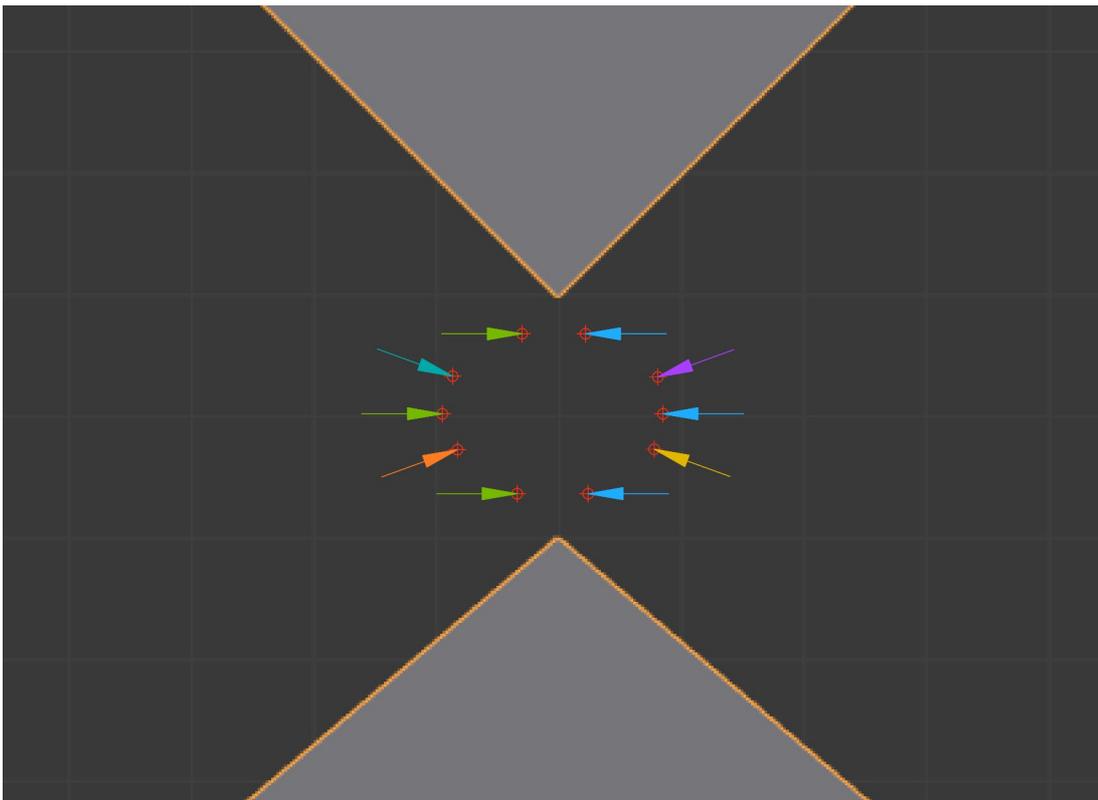


Abbildung 3.14: Der Durchgang mit sich zuspitzenden Wänden ergab für alle Bewegungsrichtungen die markante, kreisbogenförmige Anordnung der Ankerpunkte (Texturzusammenstellung 1). Die Methode war hier gegenüber der Texturzusammenstellungen invariant (Abbildung A.110).

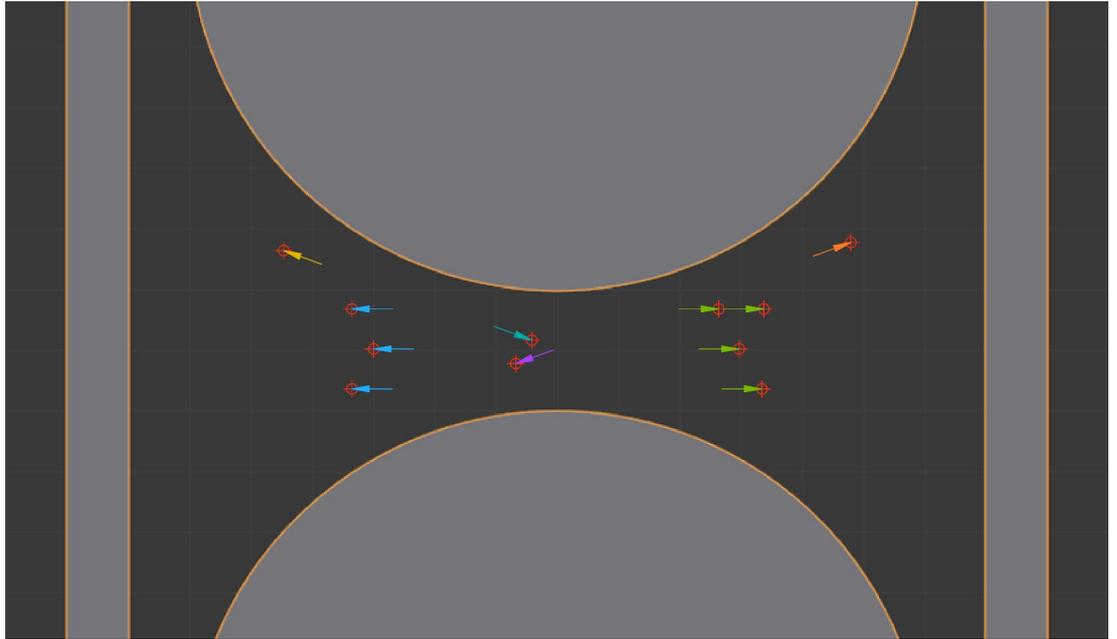


Abbildung 3.15: Bei geraden Routen durch die Passage mit gekrümmten Wänden ergab sich ebenfalls die charakteristische Kreisbogenanordnung. Es zeigte sich wieder der Abwägungskonflikt wie beim parallelen Laufen entlang der Wand im Fall des Korridors (Abbildung 3.12) Mit Texturzusammenstellungen 1 wurden wegen der Parameterkombinationen einige sinnvolle Ankerpunkte nicht detektiert (siehe Abbildung A.69 - Abbildung A.72).

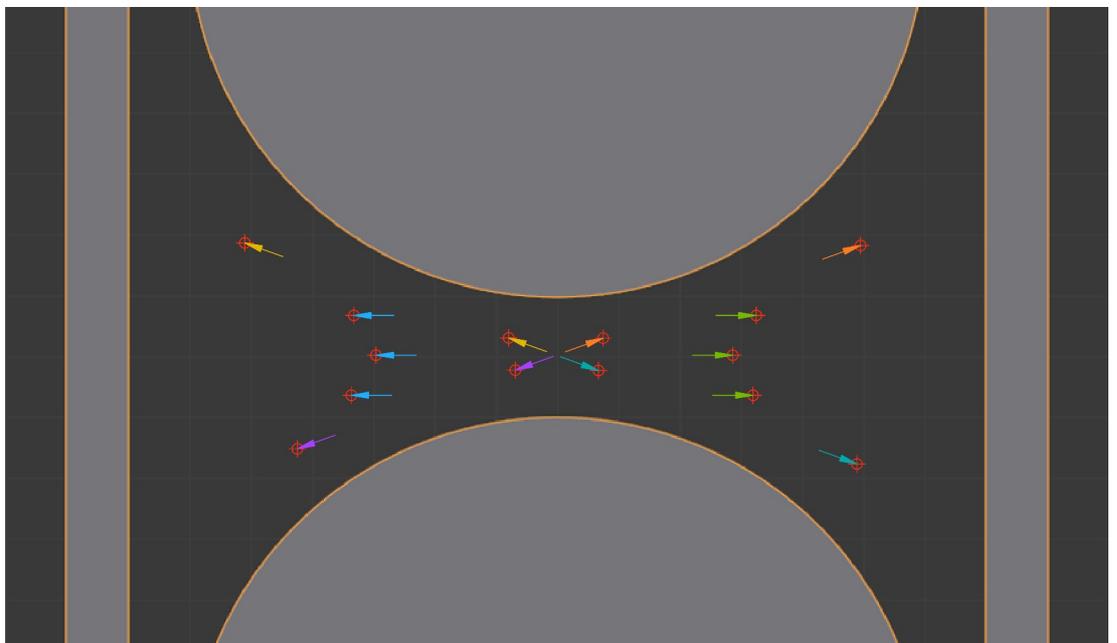


Abbildung 3.16: Mit Texturzusammenstellungen 2 wurde auf den schräg verlaufenden Routen das Verschwinden beider Wände aus dem Sichtfeld einzeln erkannt: jeweils etwa mittig das der am Anfang der Route näheren Wand und zum Ende der Route das der am Ende der Route näheren Wand.

## 4. Diskussion

### 4.1. Änderungssalienz lässt Rückschlüsse auf die Raumgeometrie zu

Mit dem Laptev-Operator kann eine sinnvolle Größe, nämlich die Änderungssalienz der Bildinformation bei der Fortbewegung, bestimmt werden. Sie ist hoch, wenn eine nahe räumliche Struktur etwas, das sich hinter ihr befindet, verdeckt und sich optisch beides gegeneinander verschiebt. In solchen Fällen treffen Bildstrukturen zusammen, ähnlich den einfachen geometrischer Formen aus Abbildung 1.5. Kann man die Voraussetzungen zu schaffen, von denen für diese Arbeit ausgegangen wurde – also zueinander ausgerichtete Einzelbilder sowie ein Abgleich mit der Geschwindigkeit – ist das der einzige Fall, in dem die Änderungssalienz in einer statischen Umwelt groß wird. Die Vorgehensweise, die vielleicht naheliegend wäre, den reinen Bildunterschied zu werten, würde dagegen auch einen hohen Wert ergeben, wenn man sich direkt auf eine Wand zubewegt.

Eine Ausnahme ist das parallele, nahe Entlangfahren an einer Kontrastreichen Wand, bei dem sehr geringe Änderungssalienz detektiert wird. Dem könnte man begegnen, indem man eine höhere Schwelle für ein Ereignis fordert als in dieser Arbeit. Man würde einzelne Ereignisse mit niedriger Änderungssalienz bei kleinem Anfahrtswinkel verwerfen, obwohl die Position sinnvoll ist. Ein geringer Wert kann bei der Exploration einer Umgebung aber immer noch ein Hinweis sein, den Ort genauer zu erkunden.

Wahrscheinlich ist der Effekt einer geringen Änderungssalienz in der Nähe einer Wand auf die Beschleunigung von Bildstrukturen nach den Gesetzen des optischen Flusses zurückzuführen. Falls notwendig bestünde eine Möglichkeit dem zu begegnen bestünde daher unter Umständen darin, beim Angleichen der Bilder zueinander zusätzlich Entzerrungsoperationen durchzuführen, sodass die Beschleunigung ausgeglichen würde. Die Verschiebung von Vordergrund zu Hintergrund bliebe davon unbeeinflusst.

Die aus dem Laptev-Operator abgeleitete Änderungssalienz kann für die Bestimmung von Ankerpunkten für die topologische Navigation genutzt werden. Diese werden an Engstellen einer Raumgeometrie gesetzt, wie in dieser Arbeit gezeigt werden konnte unabhängig von der genauen Geometrie des Ortes. Darüber hinaus

könnte man, wenn man neben dem gefundenen Ankerpunkt parallele Routen in geringem Abstand zueinander abfährt, ungefähr abschätzen, wo die Mitte der Engstelle ist. Das ist möglich, wenn man die robust in allen Konfigurationen entstandene Anordnung der Ankerpunkte im Kreisbogen annimmt. Kann solch ein Kreisbogen nicht gefunden werden, handelt es sich dagegen um eine einseitige in den Raum ragende Geometrie.

Möglicherweise könnten mit Flussfeldern ähnliche Informationen gewonnen werden. Der große Vorteil des Laptev-Operators ist, dass es im Gegensatz zu Algorithmen für den optischen Fluss nicht notwendig ist ein Korrespondenzproblem zu lösen. Möglicherweise ist die Methode daher robuster. Der zunächst größer erscheinende Rechenaufwand – für jeden Pixel in den Einzelbildern muss ein großer, dreidimensionaler Datenraum in die Rechnung einbezogen werden – relativiert sich, weil die Berechnung sehr gut parallelisiert werden kann. Nicht nur wird paralleles Rechnen in der Computertechnik immer bedeutender, auch sind biologische Systeme sehr stark parallel ausgelegt, weswegen dieser Detektor vielleicht ein gutes Modell darstellt.

In der Simulation dauerte die Berechnung der Änderungssalienz für jedes Einzelbild einer Bildsequenz etwa eine Sekunde. Es ist davon auszugehen, dass die Effizienz der Implementation deutlich optimiert werden kann. So würde eine Parallelisierung unter Verwendung einer GPU die Auswertung sicherlich beschleunigen. Des Weiteren ist es wahrscheinlich nicht notwendig, die Auswertung auf allen Auflösungsstufen der Bildpyramide und mit allen Parameterkombinationen der Fensterfunktionen durchzuführen. Darüber hinaus können die Fensterfunktionen möglicherweise beispielsweise auf die Halbwertsbreite beschnitten werden und der Detektor dennoch gute Ergebnisse erzielen. Eine Anwendung zu entwickeln, die die Bildinformation in Echtzeit analysiert, ist somit denkbar.

## 4.2. Offene Fragestellungen

Im Laufe der Arbeit sind weitere Fragen aufgekommen, die zu klären vielversprechend erscheint. Die Experimente wurden mit einem an die menschliche Sicht angelehnten Sichtfeld von  $120^\circ$  durchgeführt. Viele Insekten und auch Wirbeltiere haben

dagegen ein Sichtfeld von  $360^\circ$ . Es stellt sich die Frage, wie der Detektor in einem solchen Fall ansprechen würde. Es ist davon auszugehen, dass grundsätzlich die selbe Raumgeometrie detektiert werden würde. Allerdings müsste die Durchfahrt durch eine Engstelle zu einem symmetrischen Änderungssalienzprofil führen, weil die gesamte Umgebung gleichermaßen sichtbar ist und der Algorithmus richtungsinvariant ist. Dem entsprechend müsste eine vorwärts- wie rückwärtsfahrt zum selben Ergebnis führen. Das starke Abfallen der Änderungssalienz, mit dem in dieser Arbeit Ankerpunkte detektiert wurden, ist darauf zurückzuführen, dass die starke Änderung ab diesem Zeitpunkt hinter einem stattfindet und damit außerhalb des Sichtfeldes. Eine naheliegende Annahme wäre, dass genau in der Engstelle ein Maximum der Änderungssalienz beobachtet wird. Man könnte sich aber auch vorstellen, dass sie für ein kurzes Stück gering wird, weil kurzzeitig beide Räume fast vollständig eingesehen werden können und nicht mehr verdeckt wird.

Auf jeden Fall würde die Änderungssalienz zunächst in Bewegungsrichtung groß sein und, nachdem die Engstelle passiert wurde, in die entgegengesetzte Richtung. Diese Überlegung führt zu der Erkenntnis, dass in der Ergebnisgröße  $R$  zusätzlich zur Information, dass es einen verdeckten Raum gibt, auch die Information enthalten ist, in welcher Richtung in egozentrischen Koordinaten sich die korrespondierende Raumstruktur befindet. Diese Information weiter auszuwerten könnte zu Modellen führen, wie eine topologische Karte des Umfeldes zu erstellen ist und zu denkbaren Handlungsanweisungen für exploratives Verhalten.

### **4.3. Anwendung in der topologischen Navigation**

Ankerpunkte, die eindeutig mit der Raumgeometrie zusammenhängen, implizieren, dass man den Raum mit ihrer Hilfe zu segmentiert. Das erinnert an die ortssensitiven Neuronen in den Hippocampi von Ratten [12]. Ihre Feuerrate ist eine Funktion des Ortes, an dem sich das Tier aufhält. Diese Neurone zeigen immer dann Aktivität, wenn sich die Ratte an einer bestimmten Stelle in einer begrenzten Umgebung befindet. Vergleichbar mit einem Flickenteppich existieren viele sich überlappenden Ortsfelder, die jeweils durch ein Neuron repräsentiert werden. Ändert man die Konfiguration, die meist aus wenigen geometrischen Formen außerhalb der Versuchsanordnung besteht, ändern sich die Ortsfelder auf eine Weise, die Hinweise auf eine topo-

logische Repräsentation enthalten. So haben die Felder beispielsweise eine unterschiedliche räumliche Ausdehnung. Zwei Neuronen, die im Hippocampus nebeneinander liegen, korrespondieren nicht unbedingt mit nebeneinander liegenden Ortsfeldern. Die Felder ändern außerdem ihre Größe entsprechend so, wie die Begrenzung der Umgebung künstlich vergrößert oder verkleinert wird [11].

Penna und Wu [13] präsentierten eine mögliche Erklärung für Ortsfelder, die sie mit einem künstlichen neuronalen Netz realisierten. Ihre Variante basierte auf ein „Qualnav“ genanntes Modell von Kuipers und Levitt [8], das für offene Umgebungen wie eine Wiese mit einzelnen, weit verteilten Landmarken wie Bäumen, Felsen oder Häusern entwickelt wurde. Es wird davon ausgegangen, dass diese Objekte eindeutig identifiziert und unterschieden werden können. Je zwei von ihnen werden mit einer gedachten, unendlichen Linie verbunden. Wird das für alle Kombinationen von Landmarken durchgeführt entsteht ein Netz, das die Umgebung in Ortsfelder unterteilt. Man kann eindeutig bestimmen, in welchem Feld man sich befindet, indem man für jede der Verbindungslinien bestimmt, ob man sich rechts oder links von ihr befindet (die Seite wird eindeutig definiert, indem die Verbindungslinien gerichtet bestimmt werden). Eine Grenze zwischen zwei Landmarken, wird im einfachen Fall genau dann überschritten, wenn die zweite Landmarke einen Winkel von  $180^\circ$  relativ zur ersten überschreitet. Die Grenze wird außerhalb der Konfiguration, überschritten, wenn beide Landmarken sich übereinander schieben, also die zweite Landmarke einen Winkel von  $0^\circ$  zur ersten überschreitet. Bestimmen kann man das Überschreiten also einfach, indem man die Reihenfolge ermittelt, in der eine zweite Landmarke in dem Bereich von  $-180^\circ$  bis  $+180^\circ$  zur ersten im Sichtfeld erscheinen.

In der Praxis ist diese Methode beschränkt, weil gefordert wird, dass immer alle Landmarken sichtbar sind und wiedererkannt werden können. Darüber hinaus wird der Rechenaufwand schnell groß, wenn mehr Landmarken dazu kommen. Eine weitere Arbeit, die von ortssensitiven Neuronen in den Hippocampi von Ratten inspiriert ist, ist das Modell von Bachelder und Waxman, dessen Realisierung mit künstlichen neuronalen Netzen in den Roboter MAVIN integriert wurde [1]. Es basiert auf der Vorstellung, dass dreidimensionale Objekte durch eine Anzahl charakteristischer Ansichten repräsentiert werden könnten [7], [14]. Anstatt das Objekt zu drehen, um die verschiedenen Blickwinkel zu lernen, fährt der Roboter um das Objekt herum. In

den Experimenten waren das Modelle von Häusern und Fabriken in verkleinertem Maßstab. An den Wänden waren, zur vereinfachten Mustererkennung, kleine Leuchtquellen in unterschiedlicher Anordnung angebracht. Man kann sich vorstellen, dass der Algorithmus im Allgemeinen eine neue Ansicht erkennt, wenn eine vorher verdeckte Wand sichtbar wird. Diese Ansichten, deren Repräsentationen eindeutige Grenzen zueinander haben, können anders herum auch als Ortsfelder interpretiert werden, auf denen sich der Roboter befindet. Zu jeder Ansicht gehört also ein räumlicher Bereich, auf dem sich der Agent aufhält

Um das Modell auf mehrere Landmarken zu erweitern, wurde außerdem auch die Richtung, in der sich ein Objekt befand, für die Kodierung von Ortsfeldern verwendet, ähnlich der Methode von Penna und Wu [13]. In einer solchen Kombination mehrerer Systeme könnten auch die beschriebenen Laptev-Ankerpunkte aufgehen. In dem Schnappschussmodell von Franz et al. [4] fährt der Roboter beispielsweise zwischen kreisförmig angeordneten Modellhäusern. Der Fangbereich mit der Schnappschussmethode ist für einen etwa mittig aufgenommenen Schnappschuss sehr groß. Wird er aber an einer Engstelle aufgenommen, ergibt sich nur ein sehr kleiner Fangbereich. Man könnte die Ankerpunkte beispielsweise nutzen, um eine Regionengrenze zu definieren. Zwischen zwei solcher Regionengrenzen könnte man einen Schnappschuss erstellen, um einen möglichst großen Fangbereich zu erhalten und so die notwendige Anzahl an Schnappschüssen zu reduzieren.

Die in dieser Arbeit beschriebene Methode, Änderungssalienz mit Hilfe des Laptev-Operators zu bestimmen, ist sehr spezifisch für bestimmte räumliche Geometrien. Mit ihr kann detektiert werden, wenn ein Raum von einem Objekt oder einer anderen räumlichen Struktur verdeckt wird. Weitere Forschung dazu, die Änderungssalienz, bzw. die zugehörigen Ankerpunkte für die topologische Navigation, zu nutzen, hat ein großes Potential dafür, dass robuste und nicht sehr rechenintensiven Methoden für die topologische Navigation entwickelt werden können.



## A. Anhang

### A.1. Datensatz der Routen aus Experiment 2

#### A.1.1. Änderungssalienzprofile bei Texturzusammenstellung 1

##### Routen in der Nähe der westlichen, in den Raum ragenden Wand

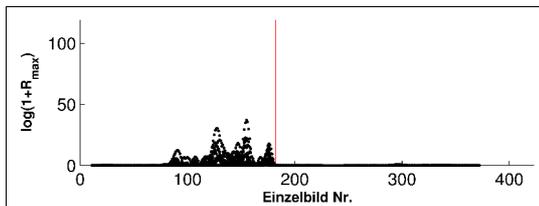


Abbildung A.1: Änderungssalienzprofil, 0°-Route nördlich der Wand ( $W \rightarrow O$ ).

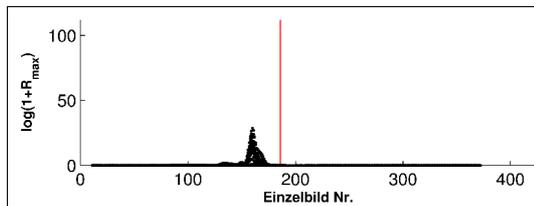


Abbildung A.2: Änderungssalienzprofil, 180°-Route nördlich der Wand ( $O \rightarrow W$ ).

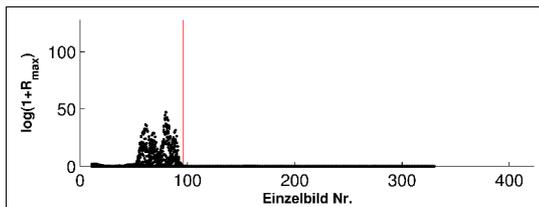


Abbildung A.3: Änderungssalienzprofil, 90°-Route östlich der Wand ( $S \rightarrow N$ ).

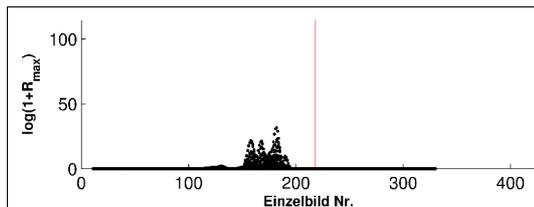


Abbildung A.4: Änderungssalienzprofil, 270°-Route östlich der Wand ( $N \rightarrow S$ ).

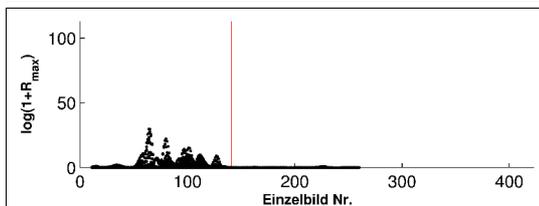


Abbildung A.5: Änderungssalienzprofil, 45°-Route östlich der Wand ( $SW \rightarrow NO$ ).

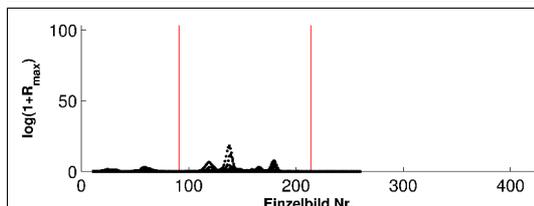


Abbildung A.6: Änderungssalienzprofil, 225°-Route östlich der Wand ( $NO \rightarrow SW$ ).

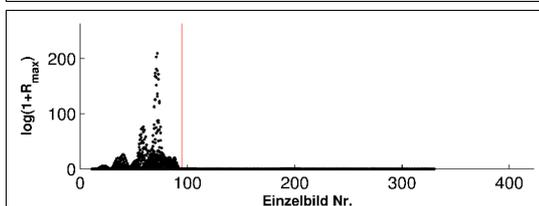


Abbildung A.7: Änderungssalienzprofil, 90°-Route westlich der Wand ( $S \rightarrow N$ ).

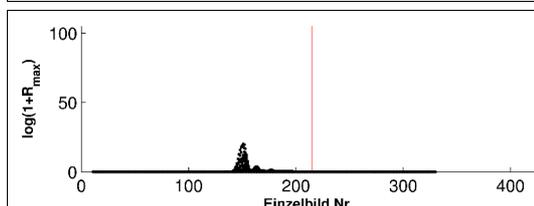
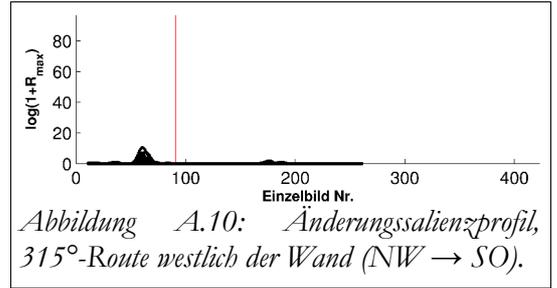
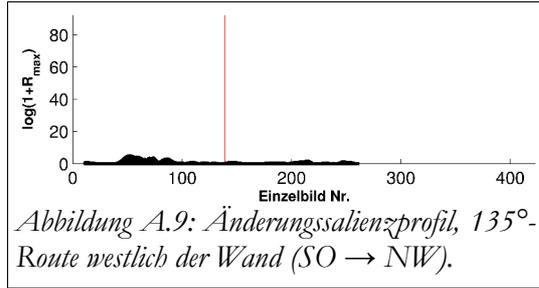
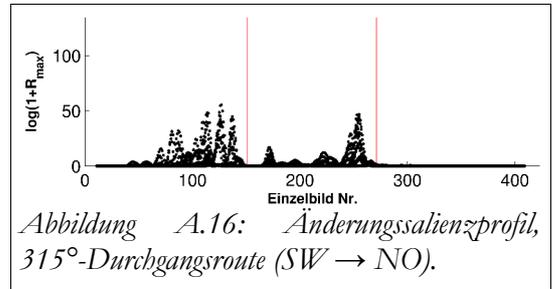
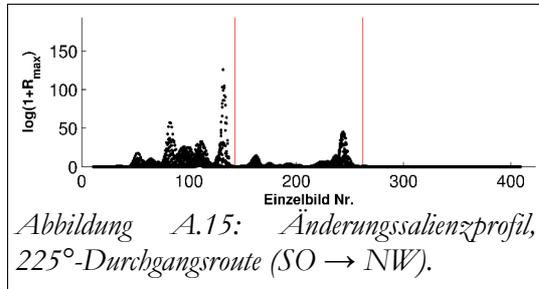
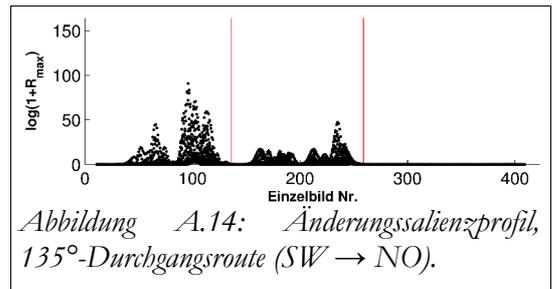
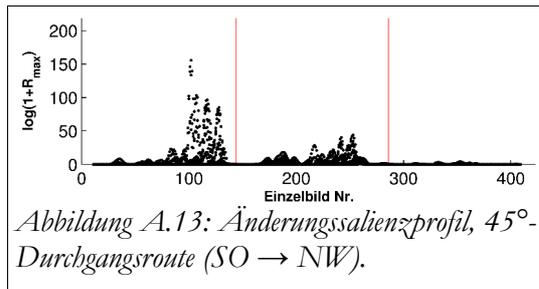
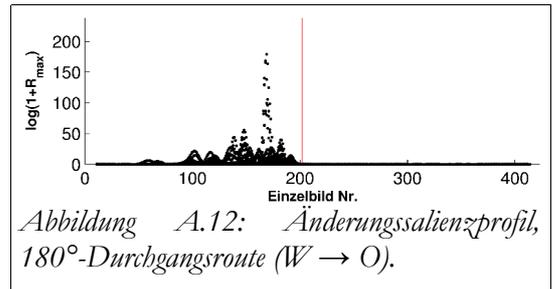
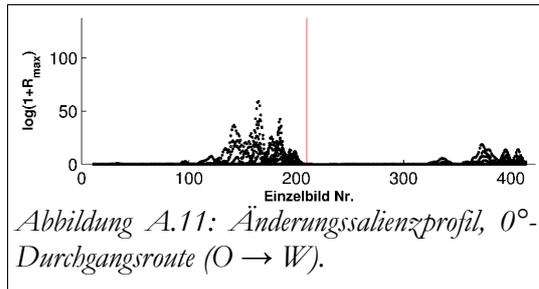


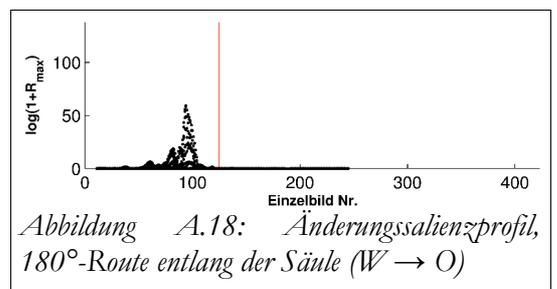
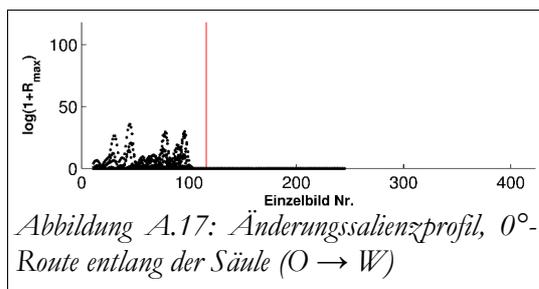
Abbildung A.8: Änderungssalienzprofil, 270°-Route westlich der Wand ( $N \rightarrow S$ ).



**Routen durch den mittleren Durchgang**

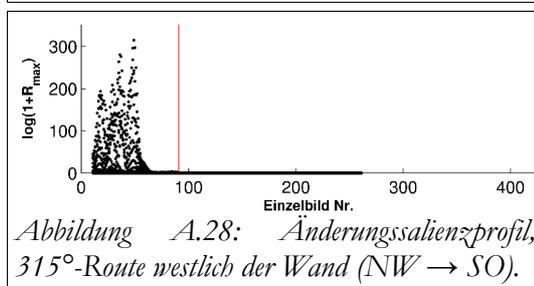
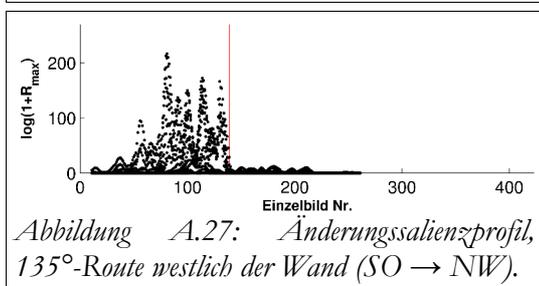
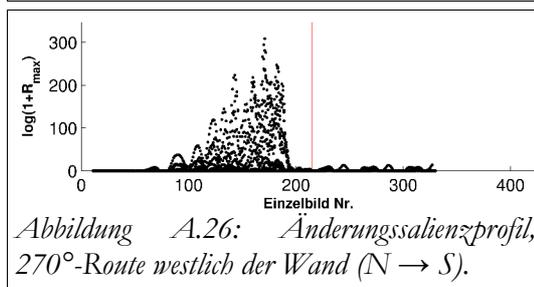
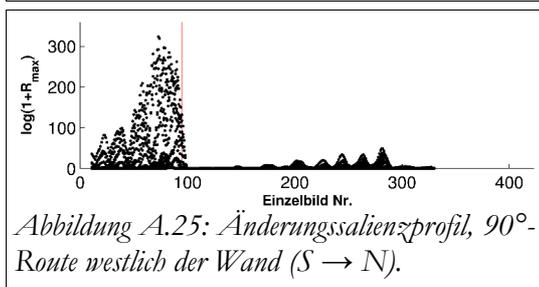
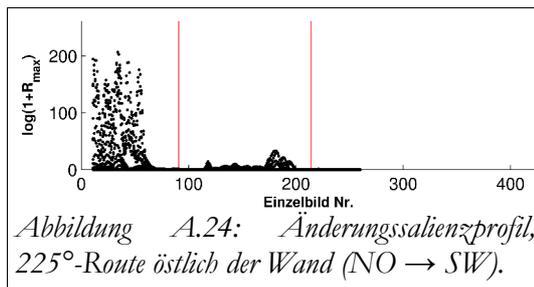
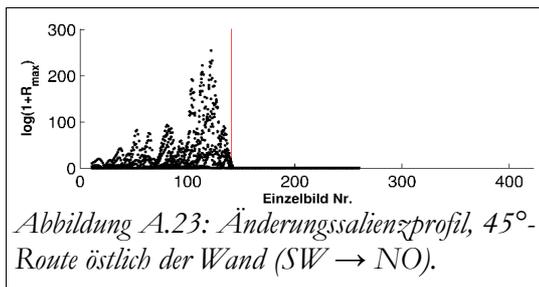
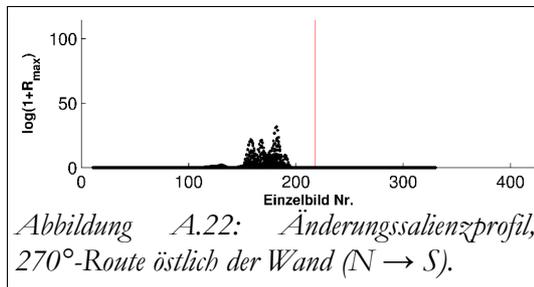
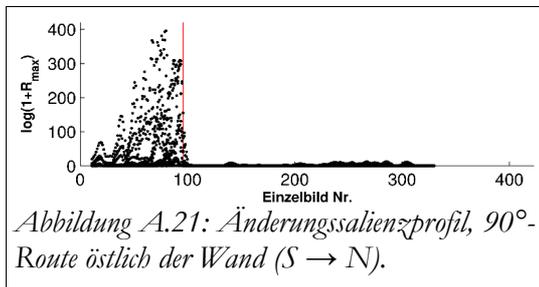
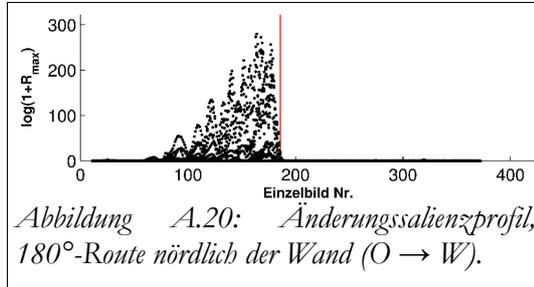
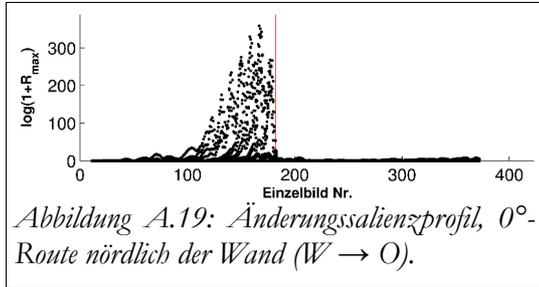


**Routen in der Nähe der südöstlichen Säule**



## A.1.2. Änderungssalienzprofile bei Texturzusammenstellung 2

### Routen in der Nähe der westlichen, in den Raum ragenden Wand



## Routen durch den mittigen Durchgang

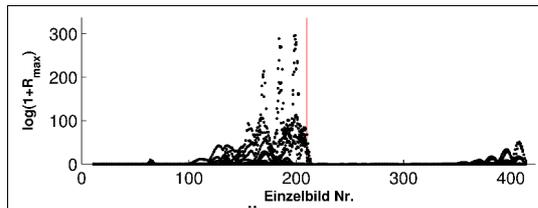


Abbildung A.29: Änderungssalienzprofil, 0°-Durchgangsrouten (O → W).

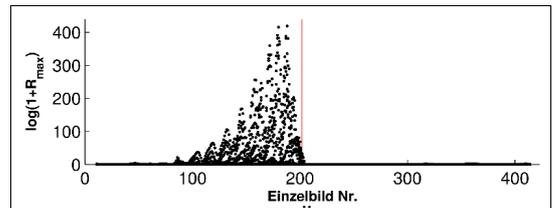


Abbildung A.30: Änderungssalienzprofil, 180°-Durchgangsrouten (W → O).

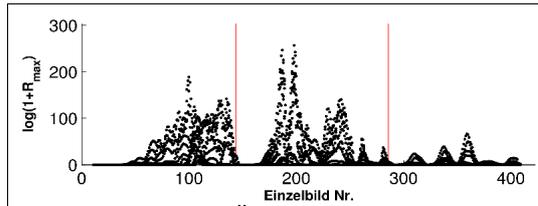


Abbildung A.31: Änderungssalienzprofil, 45°-Durchgangsrouten (SO → NW).

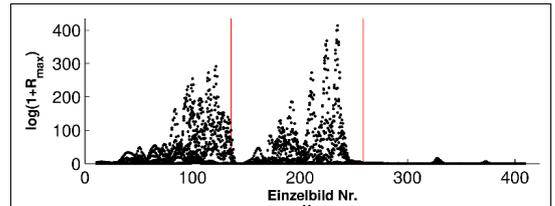


Abbildung A.32: Änderungssalienzprofil, 135°-Durchgangsrouten (SW → NO).

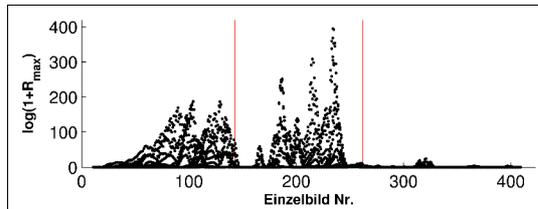


Abbildung A.33: Änderungssalienzprofil, 225°-Durchgangsrouten (SO → NW).

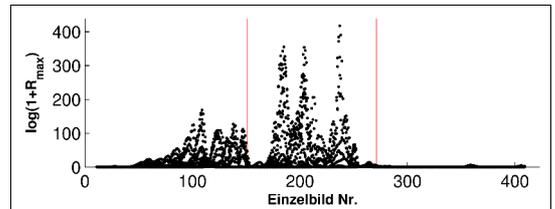


Abbildung A.34: Änderungssalienzprofil, 315°-Durchgangsrouten (SW → NO).

## Routen in der Nähe der südöstlichen Säule

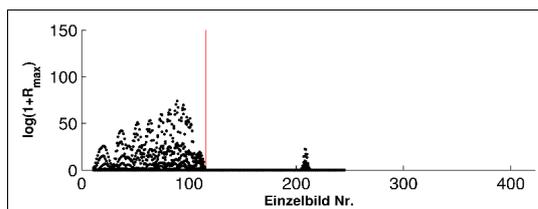


Abbildung A.35: Änderungssalienzprofil, 0°-Route entlang der Säule (O → W)

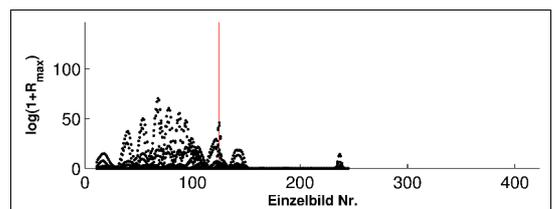


Abbildung A.36: Änderungssalienzprofil, 180°-Route entlang der Säule (W → O)

## A.2. Datensatz der Routen aus Experiment 3

### A.2.1. Änderungssalienzprofile bei Texturzusammenstellung 1

#### Routen in Korridorpassage (Raum 1)

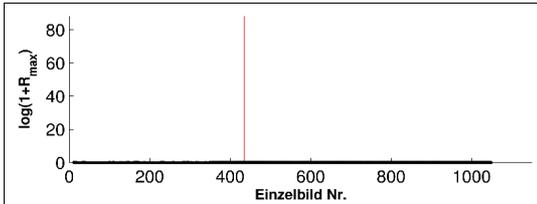


Abbildung A.37: Änderungssalienzprofil, 0°-Route, mittig (O → W)

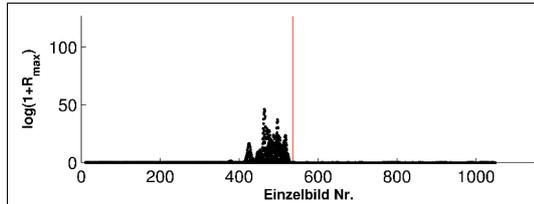


Abbildung A.38: Änderungssalienzprofil, 180°-Route, mittig (W → O)

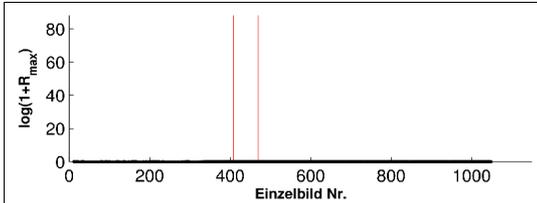


Abbildung A.39: Änderungssalienzprofil, 0°-Route, nördlich versetzt (O → W)

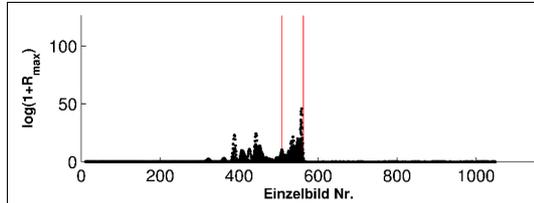


Abbildung A.40: Änderungssalienzprofil, 180°-Route, nördlich versetzt (W → O)

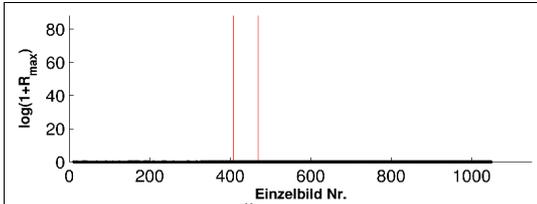


Abbildung A.41: Änderungssalienzprofil, 0°-Route, südlich versetzt (O → W)

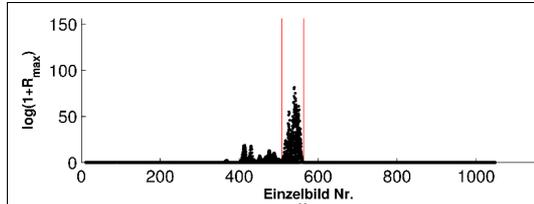


Abbildung A.42: Änderungssalienzprofil, 180°-Route, südlich versetzt (W → O)

#### Routen in der Türpassage (Raum 2)

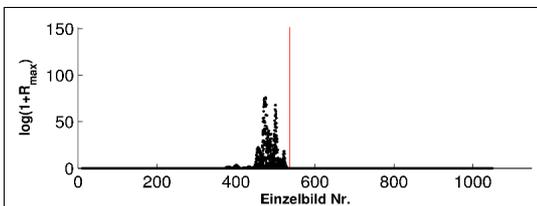


Abbildung A.43: Änderungssalienzprofil, 0°-Route, mittig (O → W)

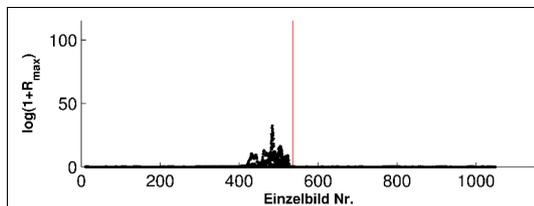


Abbildung A.44: Änderungssalienzprofil, 180°-Route, mittig (W → O)

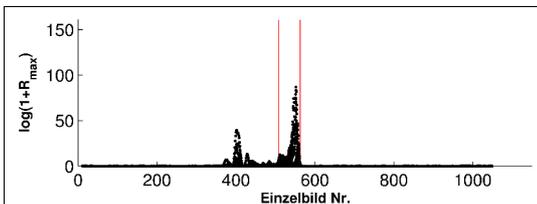


Abbildung A.45: Änderungssalienzprofil, 0°-Route, nördlich versetzt (O → W)

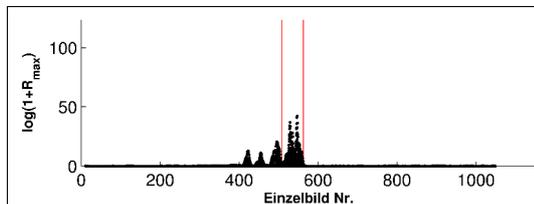


Abbildung A.46: Änderungssalienzprofil, 180°-Route, nördlich versetzt (W → O)

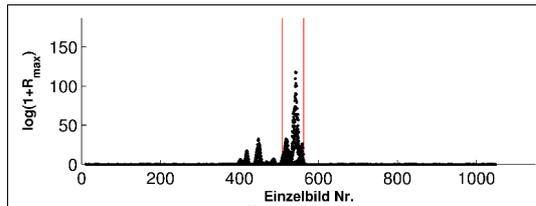


Abbildung A.47: Änderungssalienzprofil, 0°-Route, südlich versetzt (O → W)

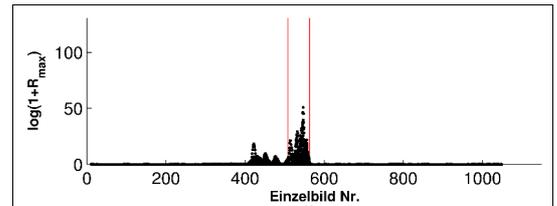


Abbildung A.48: Änderungssalienzprofil, 180°-Route, südlich versetzt (W → O)

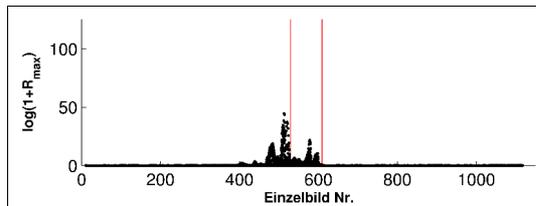


Abbildung A.49: Änderungssalienzprofil, 20°-Route (W SW → O NO)

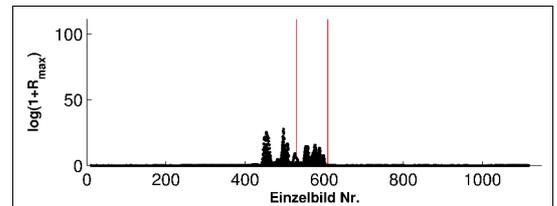


Abbildung A.50: Änderungssalienzprofil, 200°-Route (O NO → W SW)

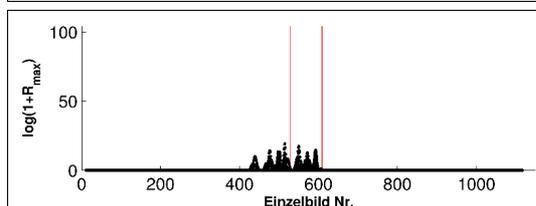


Abbildung A.51: Änderungssalienzprofil, 160°-Route (O SO → W NW)

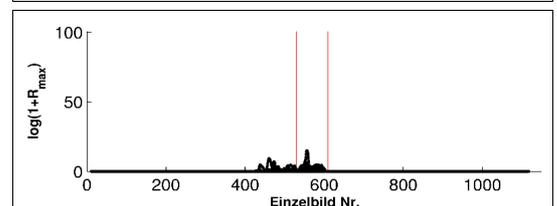


Abbildung A.52: Änderungssalienzprofil, 340°-Route (W NW → O SO)

### Routen in der spitz zulaufenden Passage (Raum 3)

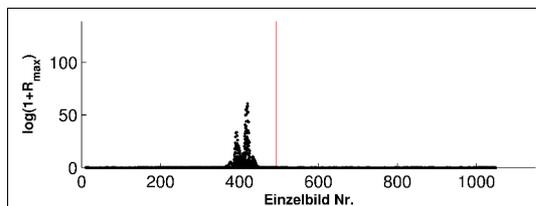


Abbildung A.53: Änderungssalienzprofil, 0°-Route, mittig (O → W)

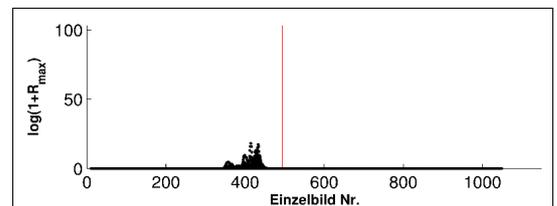


Abbildung A.54: Änderungssalienzprofil, 180°-Route, mittig (W → O)

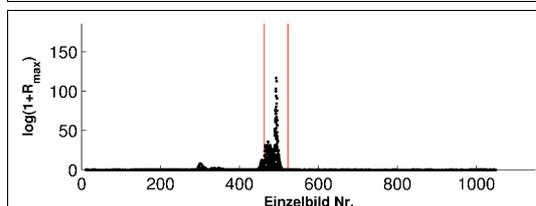


Abbildung A.55: Änderungssalienzprofil, 0°-Route, nördlich versetzt (O → W)

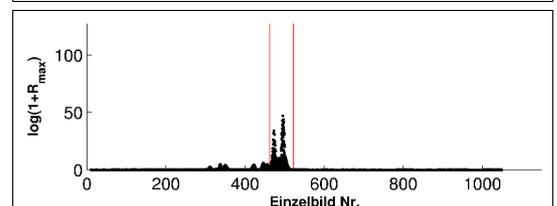


Abbildung A.56: Änderungssalienzprofil, 180°-Route, nördlich versetzt (W → O)

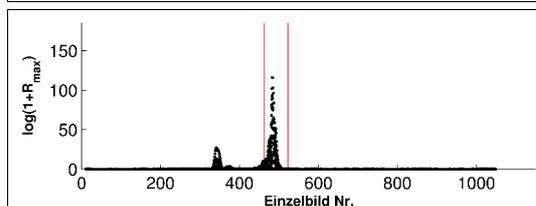


Abbildung A.57: Änderungssalienzprofil, 0°-Route, südlich versetzt (O → W)

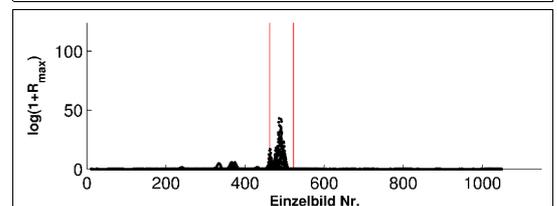


Abbildung A.58: Änderungssalienzprofil, 180°-Route, südlich versetzt (W → O)

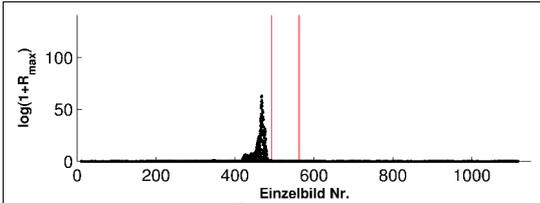


Abbildung A.59: Änderungssalienzprofil, 20°-Route (W SW → O NO)

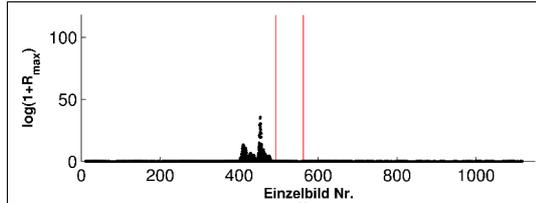


Abbildung A.60: Änderungssalienzprofil, 200°-Route (O NO → W SW)

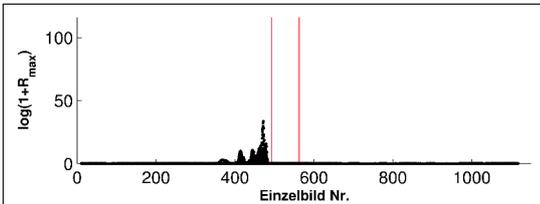


Abbildung A.61: Änderungssalienzprofil, 160°-Route (O SO → W NW)

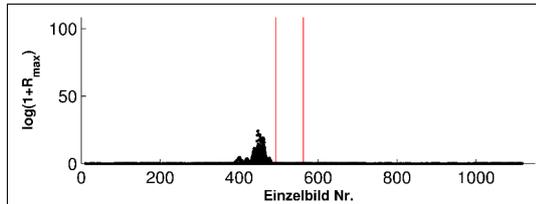


Abbildung A.62: Änderungssalienzprofil, 340°-Route (W NW → O SO)

### Routen in der Passage mit gekrümmter Wand (Raum 4)

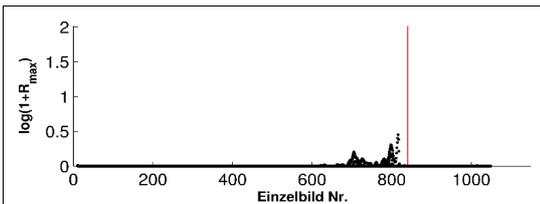


Abbildung A.63: Änderungssalienzprofil, 0°-Route, mittig (O → W)

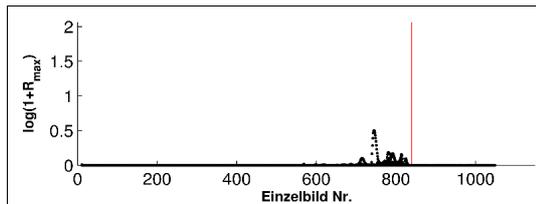


Abbildung A.64: Änderungssalienzprofil, 180°-Route, mittig (W → O)

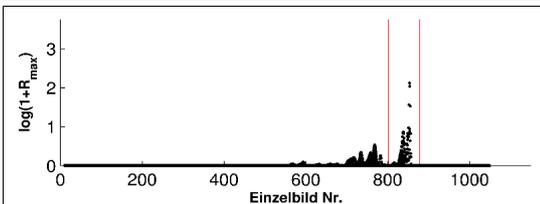


Abbildung A.65: Änderungssalienzprofil, 0°-Route, nördlich versetzt (O → W)

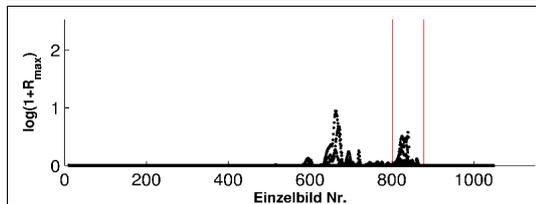


Abbildung A.66: Änderungssalienzprofil, 180°-Route, nördlich versetzt (W → O)

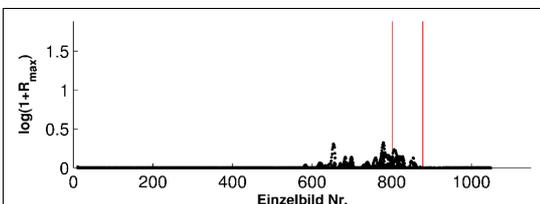


Abbildung A.67: Änderungssalienzprofil, 0°-Route, südlich versetzt (O → W)

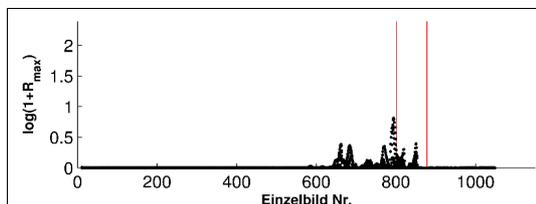


Abbildung A.68: Änderungssalienzprofil, 180°-Route, südlich versetzt (W → O)

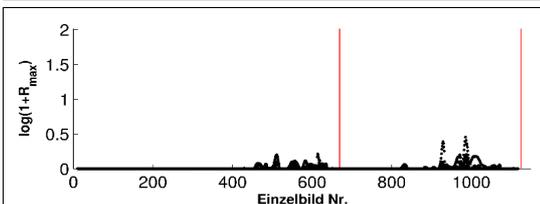


Abbildung A.69: Änderungssalienzprofil, 20°-Route (W SW → O NO)

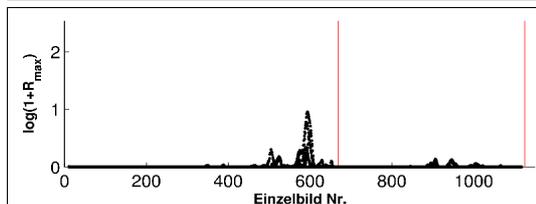
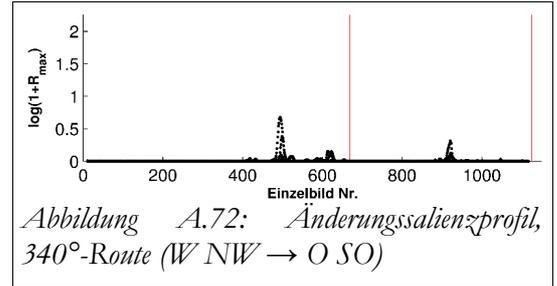
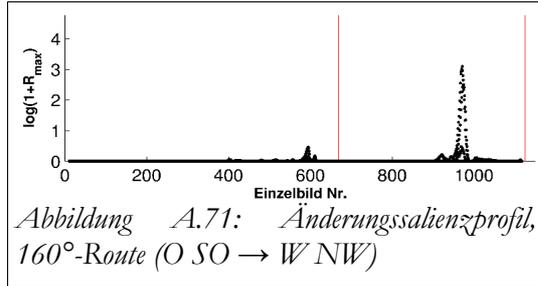
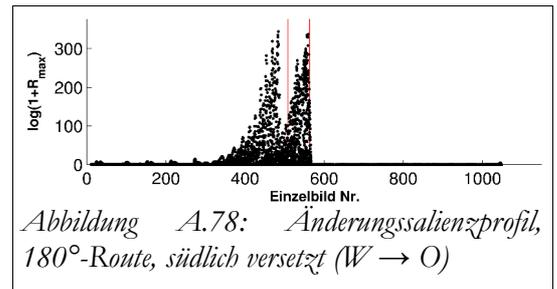
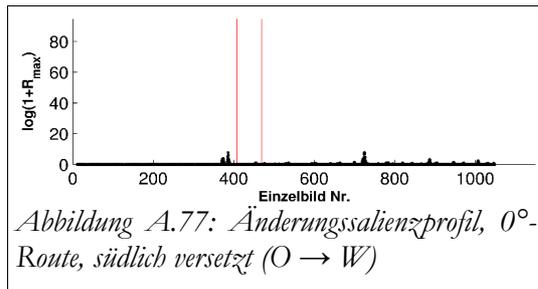
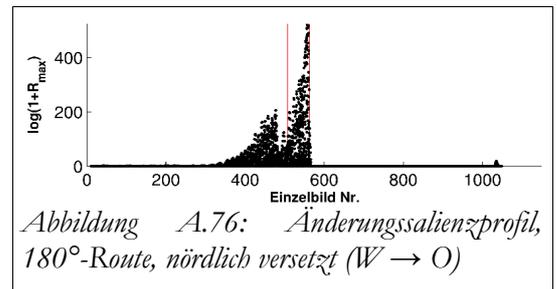
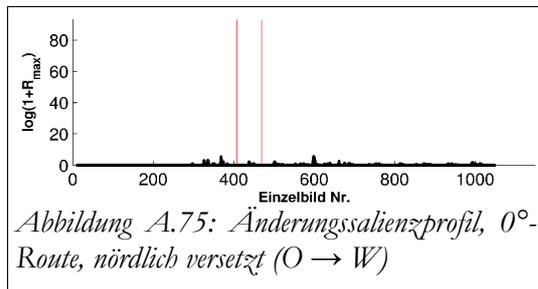
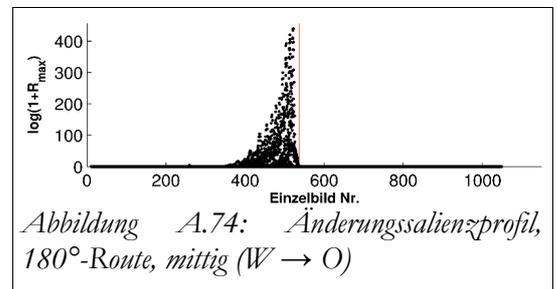
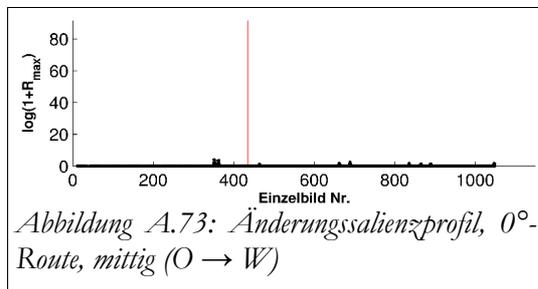


Abbildung A.70: Änderungssalienzprofil, 200°-Route (O NO → W SW)

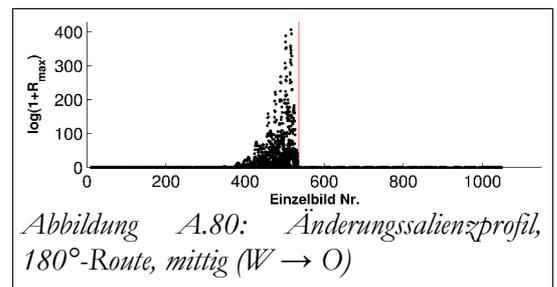
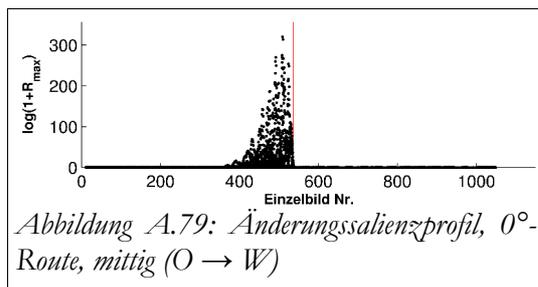


## A.2.2. Änderungssalienzprofile bei Texturzusammenstellung 2

### Routen in Korridorpassage (Raum 1)



### Routen in der Türpassage (Raum 2)



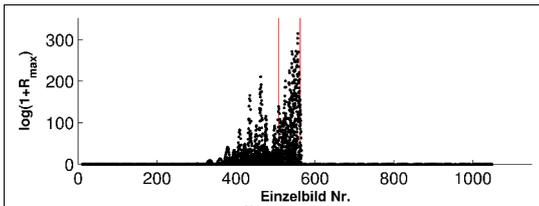


Abbildung A.81: Änderungssaliensprofil, 0°-Route, nördlich versetzt (O → W)

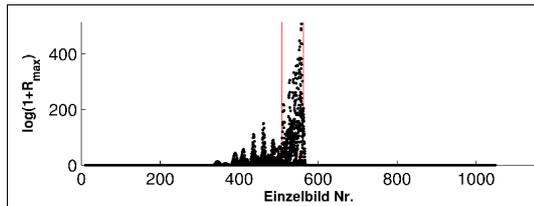


Abbildung A.82: Änderungssaliensprofil, 180°-Route, nördlich versetzt (W → O)

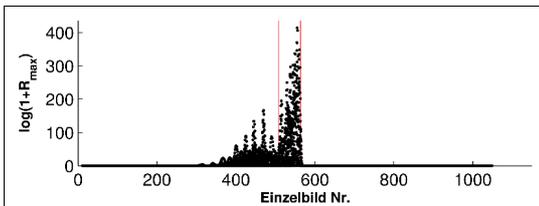


Abbildung A.83: Änderungssaliensprofil, 0°-Route, südlich versetzt (O → W)

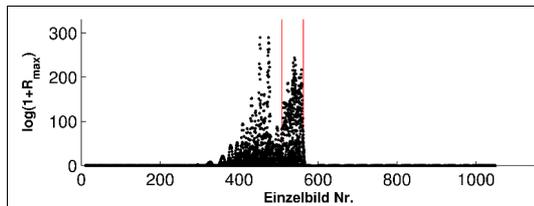


Abbildung A.84: Änderungssaliensprofil, 180°-Route, südlich versetzt (W → O)

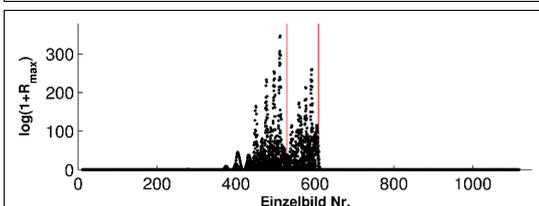


Abbildung A.85: Änderungssaliensprofil, 20°-Route (W SW → O NO)

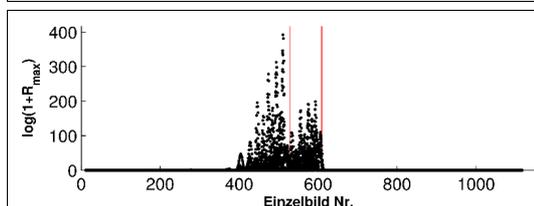


Abbildung A.86: Änderungssaliensprofil, 200°-Route (O NO → W SW)

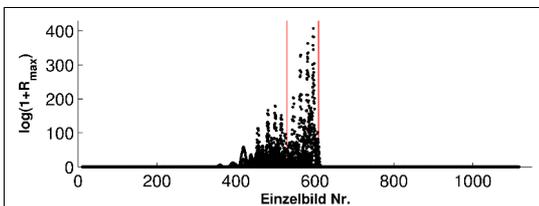


Abbildung A.87: Änderungssaliensprofil, 160°-Route (O SO → W NW)

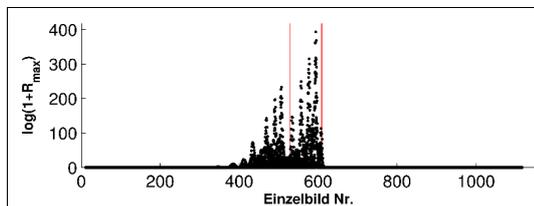


Abbildung A.88: Änderungssaliensprofil, 340°-Route (W NW → O SO)

### Routen in der spitz zulaufenden Passage (Raum 3)

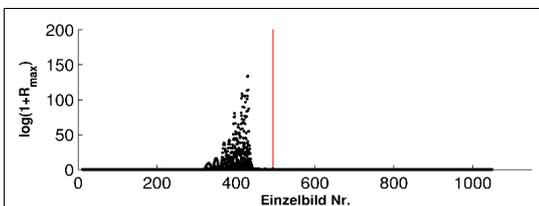


Abbildung A.89: Änderungssaliensprofil, 0°-Route, mittig (O → W)

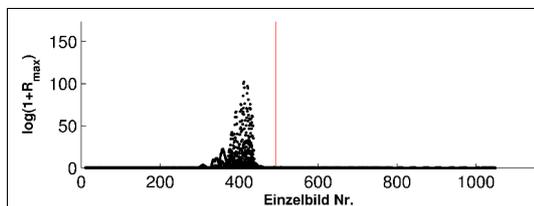


Abbildung A.90: Änderungssaliensprofil, 180°-Route, mittig (W → O)

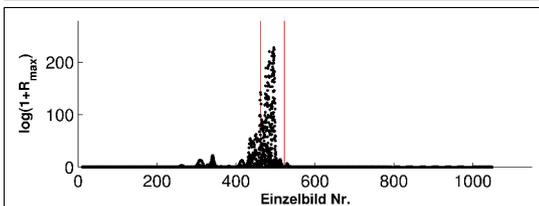


Abbildung A.91: Änderungssaliensprofil, 0°-Route, nördlich versetzt (O → W)

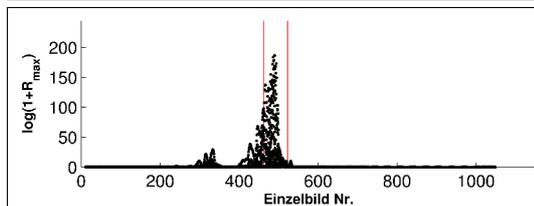
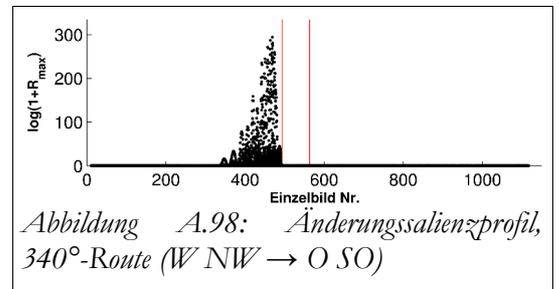
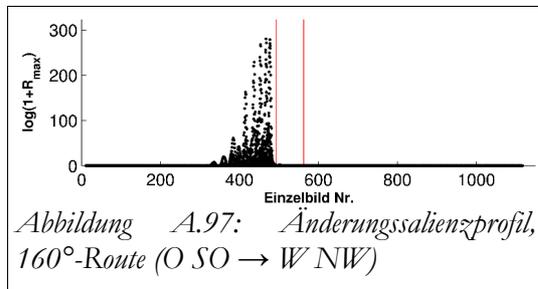
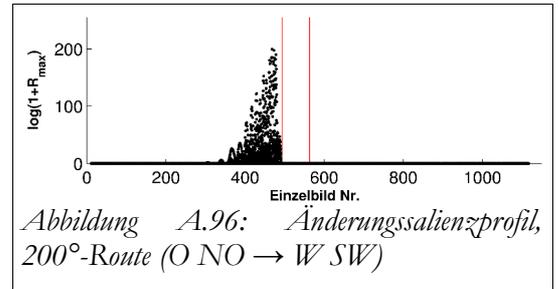
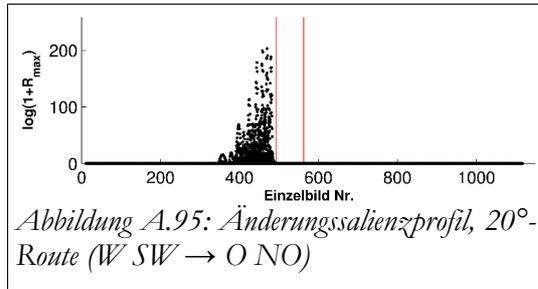
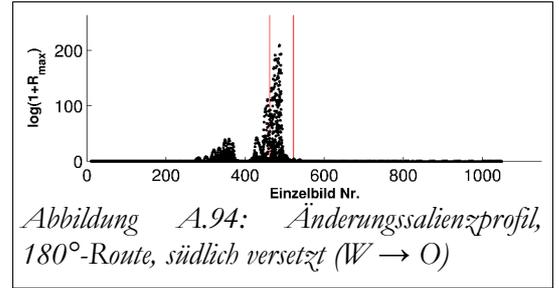
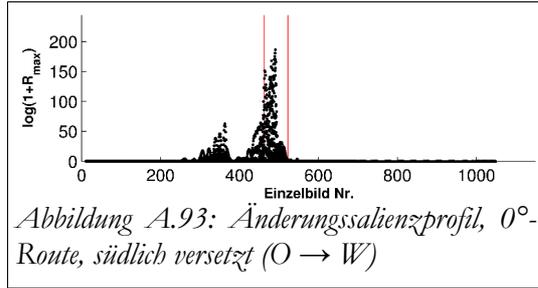
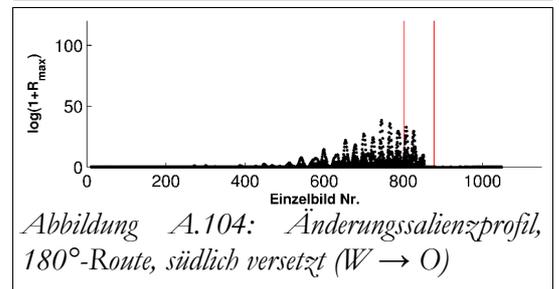
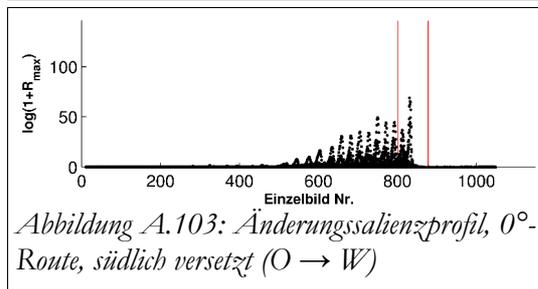
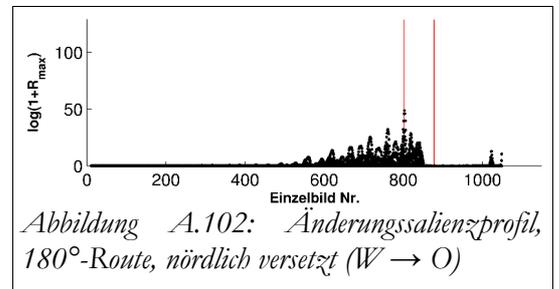
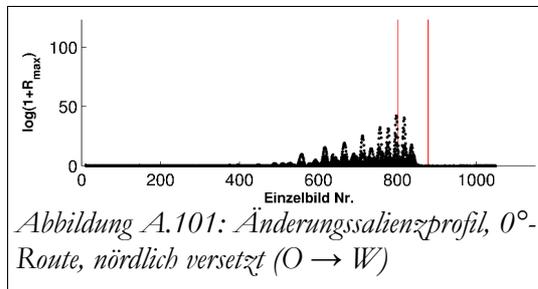
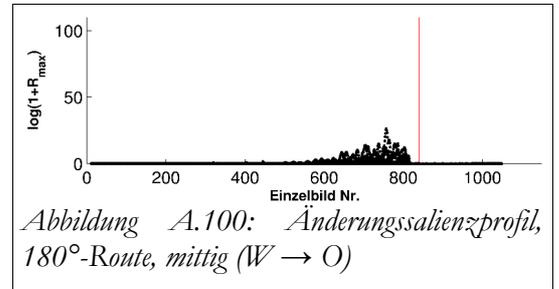
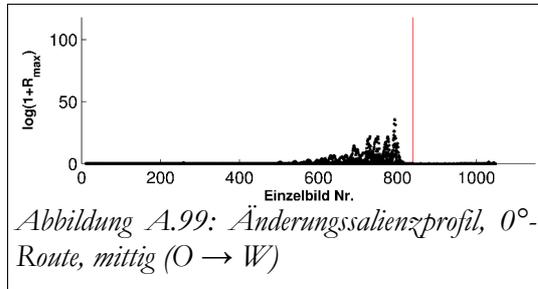


Abbildung A.92: Änderungssaliensprofil, 180°-Route, nördlich versetzt (W → O)



**Routen in der Passage mit gekrümmter Wand (Raum 4)**



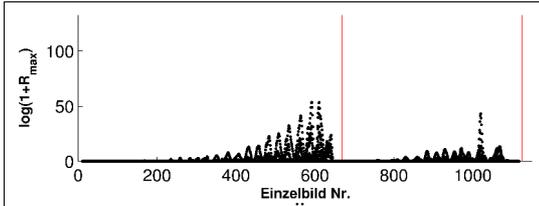


Abbildung A.105: Änderungssalienzprofil, 20°-Route (W SW → O NO)

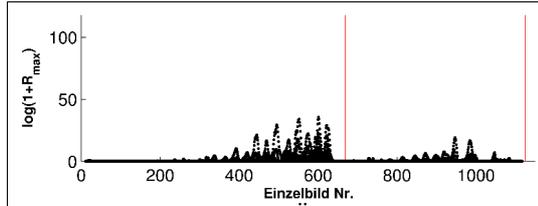


Abbildung A.106: Änderungssalienzprofil, 200°-Route (O NO → W SW)

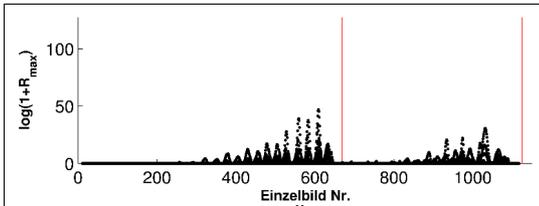


Abbildung A.107: Änderungssalienzprofil, 160°-Route (O SO → W NW)

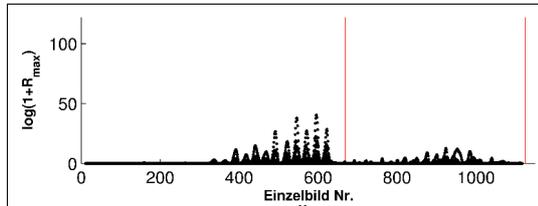


Abbildung A.108: Änderungssalienzprofil, 340°-Route (W NW → O SO)

### A.3. Weitere Ankerpunkte aus Experiment 3

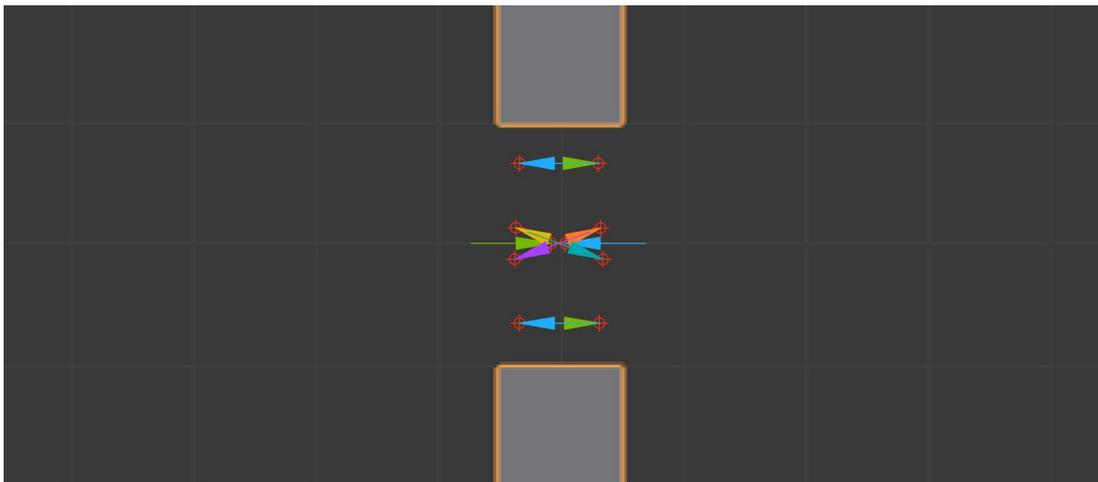


Abbildung A.109: Durchfahrt durch den türähnlichen Durchgang (Texturzusammenstellung 1) Die Methode war hier gegenüber der Texturzusammenstellung invariant (Abbildung 3.13).

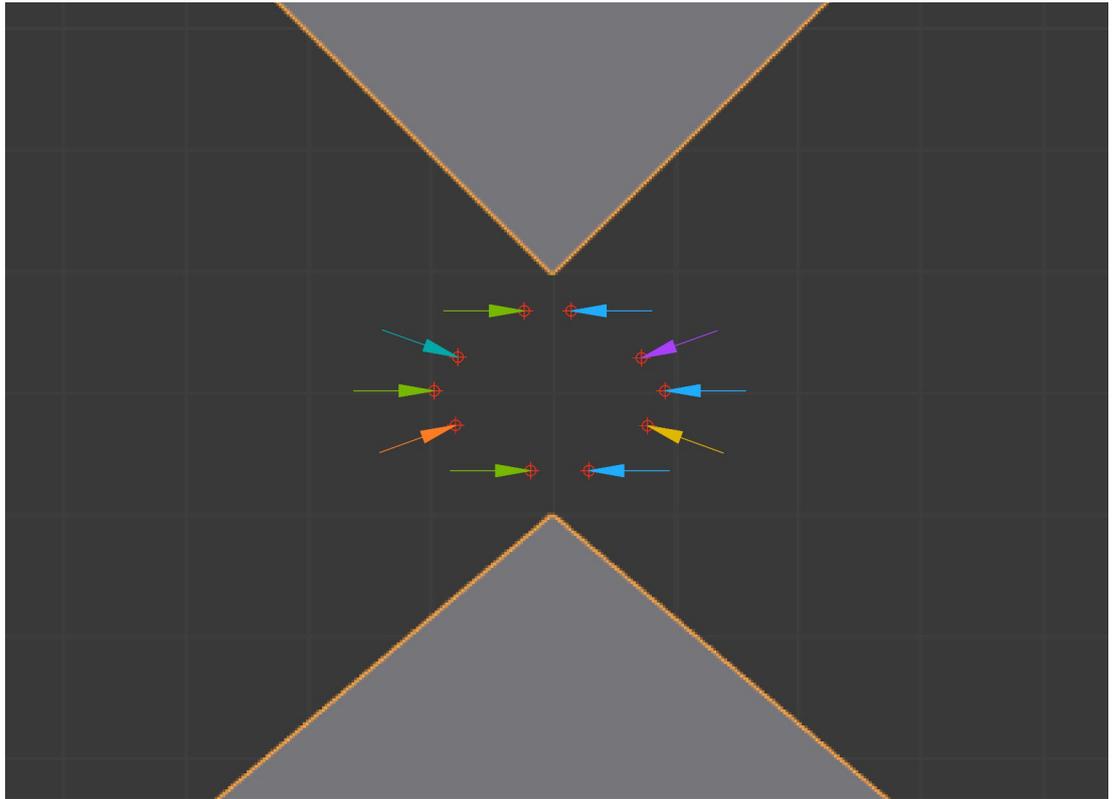


Abbildung A.110: Der Durchgang mit sich zuspitzenden Wänden zeigte für alle Bewegungsrichtung die markante, halbkreisförmige Anordnung der Ankerpunkte (Texturzusammenstellung 2). Die Methode war hier gegenüber der Texturzusammenstellungen invariant (Abbildung 3.14).

## B. Literaturverzeichnis

- [1] Bachelder, I. A., Waxman, A. M. (1994). Mobile Robot Visual Mapping and Localization: A View-Based Neurocomputational Architecture That Emulates Hippocampal Place Learning, *Neural Networks* 6/7, Vol. 7: 1083-1099
- [2] Cartwright B. A., Collett T. S. (1983). Landmark Learning in Bees, *Journal of Comparative Physiology A* 151: 521-543
- [3] Franz, M. O., Schölkopf, B., Mallot, H. A., Bühlhoff, H. H. (1998). Learning View Graphs for Robot Navigation, *Autonomous Robots* 5: 111-125
- [4] Franz, M. O., Schölkopf, B., Mallot, H. A., Bühlhoff, H. H. (1998). Where did I take that snapshot? Scene-based homing by image matching, *Biological Cybernetics* 79: 191-202
- [5] Gerstmayr-Hillen, L., Röben, F., Krzykawski, M., Kreft, S., Venjakob, D., Möller, R. (2013). Dense topological maps and partial pose estimation for visual control of an autonomous cleaning robot, *Robotics and Autonomous Systems* 61(5): 497-516
- [6] Harris, C., Stephens, M. (1988). A combined corner and edge detector, *Alvey Vision Conference*: 147-152
- [7] Koenderink, J. J., van Doorn, A. J. (1988). The internal representation of solid shape with respect to vision, *Biological Cybernetics* 32: 211-216
- [8] Kuipers, B. J., Levitt, T. S. (1988). Navigation and mapping in large scale space, *AI Magazine* 2, Vol. 9: 25-43
- [9] Laptev, I. (). On Space-Time Interest Points, *International Journal of Computer Vision* 64(2/3): 107-123
- [10] Moravec, H. (1980). Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, Tech Report CMU-RI-TR-3, Carnegie-Mellon University, Robotics Institute
- [11] Muller, R. U., Kubie, J. L. (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells, *J. Neurosci.* 7: 1951-1968
- [12] O'Keefe, J., Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely moving rat, *Brain Res.* 34: 171-175
- [13] Penna, M. A., Wu, J (1993). Models for Map Building and Navigation, *IEEE Transactions on Systems, Man, and Cybernetics* 5, Vol. 23: 1276-1301
- [14] Seibert, M., Waxman, A. M. (). Adaptive 3D object recognition from multiple views, *IEEE Transaction on Pattern Analysis and Machine Intelligence* 14: 107-124
- [15] Tinbergen N. (1932). Über die Orientierung des Bienenwolfes (*Philanthus triangulum* Fabr.), *Zeitschrift fuer Vergleichende Physiologie* Berlin 16: 305-335
- [16] Witkin, A.P. (1983). Scale-space filtering, *Proc. 8th Int. Joint Conf. Art. Intell.*, Karlsruhe, Germany: 1019-1022
- [17] Zeil, J., Hofman, M. I., Chahl, J. S. (2003). Catchment areas of panoramic snapshots in outdoor scenes, *J. Opt. Soc. Am. A* 20: 450-469