Informatics

AR Visualisierung von Migrationsbewegungen in Europa

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Gabriel Walchhofer
Matrikelnummer 12024752

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Dipl.-Ing. BSc Daniel Pahr
Mitwirkung: Dr. Renata Raidou

Wien, 27. März 2025

Gabriel Walchhofer — Daniel Pahr

TU WIEN Informatics

AR Visualization of Migration Flows in Europe

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Gabriel Walchhofer
Registration Number 12024752

to the Faculty of Informatics

at the TU Wien

Advisor: Dipl.-Ing. BSc Daniel Pahr
Assistance: Dr. Renata Raidou

Vienna, March 27, 2025

Gabriel Walchhofer Daniel Pahr

Erklärung zur Verfassung der Arbeit

Gabriel Walchhofer

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 27. März 2025

Gabriel Walchhofer

Danksagung

An dieser Stelle möchte ich mich herzlich bei all denjenigen bedanken, die mich während der Entstehung dieser Bachelorarbeit unterstützt haben.

Mein besonderer Dank gilt meinem Betreuer Dipl.Ing. BSc Daniel Pahr, der mich mit kontinuierlichem und konstruktivem Feedback sowie mit fachlicher Unterstützung durch den gesamten Prozess begleitet hat.

Ein ebenso großer Dank geht an meine Familie und Freund*innen, die mich stets motiviert und unterstützt haben. Ihre Meinung und Feedback haben mir dabei geholfen diese Arbeit zu formen und zu erarbeiten.

Nicht zuletzt danke ich meinem Kollegen Marian Bramberger mit dessen hilfe Teile dieser Arbeit in Kollaboration entstanden sind. Diese Bachelorarbeit ist nicht nur das Ergebnis meiner eigenen Anstrengungen, sondern auch ein Produkt der wertvollen Unterstützung, die ich von vielen Seiten erhalten habe.

Acknowledgements

At this point, I would like to sincerely thank everyone who supported me during the creation of this bachelor's thesis.

My special thanks go to my supervisor, Dipl.-Ing. BSc Daniel Pahr, who guided me through the entire process with continuous and constructive feedback as well as professional support.

I am equally grateful to my family and friends, who constantly motivated and supported me. Their opinions and feedback helped me shape and develop this thesis.

Last but not least, I would like to thank my colleague Marian Bramberger, who is my collaborator on parts of this project. This bachelor's thesis is not only the result of my own efforts but also a product of the valuable support I have received from many sides.

Kurzfassung

Origin-Destination (OD) Flowmaps bieten eine Methode zur Visualisierung von Bewegungsmustern in Bereichen wie Transport, Handel oder Migration. OD Flowmaps stellen die Bewegung als Netzwerkvisualisierung mit gewichteten und gerichteten Kurven dar. Traditionelle zwei dimensionale OD Flowmaps leiden oft unter visuellem Durcheinander und Überlagerungen, was ihre Effektivität bei der Vermittlung komplexer Zusammenhänge einschränkt. Diese Arbeit untersucht den Einsatz von Augmented Reality (AR) zur Visualisierung von OD Flowmaps mit Migrationsdaten, um Interaktivität, Datenverständlichkeit und räumliches Bewusstsein zu verbessern. Durch die Implementierung von dreidimensionalen OD Flowmaps können Benutzer*innen dynamisch mit den Daten interagieren und die OD Flowmaps aus verschiedenen Perspektiven und über mehrere Jahre betrachten. Diese Arbeit stellt einen Ansatz zur Visualisierung und Kodierung der Zeitkomponente mithilfe eines Space-Time Cubes (STC) vor, welcher die Zeitkomponente als zusätzliche räumliche Dimension kodiert. Die Arbeit umfasst die Implementierung eines "force-directed"Layout-Algorithmus basierend auf der Arbeit von Jenny et al. [JSM+17] sowie die Entwicklung einer markerbasierten AR-Smartphone Applikation, welche Migrationsdaten für EU-Länder im Zeitraum von 2008 bis 2022 visualisieren kann. Diese Arbeit leistet einen Beitrag zu den Bereichen Datenvisualisierung, Computergrafik und Mensch-Computer-Interaktion und liefert Erkenntnisse darüber, wie immersive Technologien die Visualisierung raumzeitlicher Daten verbessern können.

Abstract

Origin-Destination (OD) flow maps are a tool for visualizing movement patterns in domains such as transportation, trade, or migration. OD flow maps visualize movement using a network of directional and weighted curves. Traditional 2D OD flow maps often suffer from visual clutter and occlusion, limiting their effectiveness in conveying complex spatial relationships. This thesis explores the use of Augmented Reality (AR) for OD flow map visualization of migration data to enhance interactivity, data comprehensibility, and spatial awareness. By extending OD flow maps to the third dimension of time, users can interact with the data dynamically and view OD connections from multiple perspectives across a span of years. This thesis develops an approach to visualize and encode the time component using a Space-Time Cube (STC), which encodes the time as an additional spatial dimension. The research involves the implementation of a force-directed layout algorithm based on work by Jenny et al. [JSM+17] and the development of a marker-based AR phone application prototype capable of visualizing migration data for EU countries spanning from 2008-2022. This thesis contributes to the fields of data visualization, computer graphics, and human-computer interaction, providing insights into how immersive technologies can enhance spatio-temporal data visualization.

Contents

CHAPTER 1

Introduction

The topic of migration is often extremely polarizing and emotionalized to the point where it is difficult to form objective thoughts on the underlying statistical data [KAG15]. Risam [Ris19] examines how the visualization of geo-spatial data in the migration domain can be used to challenge the framing of migrants as a "problem." In their research, they argue that the choice of language and the visual representation of the data can greatly influence the political and social narrative around migration and can either reinforce the negative stereotype or support a more open discourse.

The rise of technologies like augmented reality (AR) and virtual reality (VR) have opened new perspectives for interactive visualization [KK06]. AR is a subset of extended reality (XR) that combines physical and virtual objects by linking them based on algorithms. A notable benefit of using AR to visualize abstract data is the high level of interactivity without being too intrusive [HKSS17]. By embedding data into objects in the physical world, we can enhance our understanding of the information without being distracted from what is happening around us. Contemporary AR applications can be executed from various smart devices such as smartphones and laptops, making it more accessible to the masses without needing specialized hardware [RCEJR19].

Data visualization is an excellent way to create context for otherwise hard-to-understand connections [IJ19]. This holds especially true for abstract data that can be quite difficult to contextualize [Fri08]. In the context of migration in general, conveying movement over time and showing the different destinations of migrants poses unique challenges in handling large datasets. Especially putting them into a comprehensive but meaningful visualization, whether that may be a digital or physical model. Presenting the data within a digital approach embedded into an AR application allows users to grasp more detailed information and interact with the data in real time [YDJ+19].

With this thesis, I aim to develop an AR-based visualization of migration data within Europe over the last decade. The main goal is to process the data and represent them in a more accessible way by having an objective and intractable visualization without any political or emotional agenda attached to the empirical data. To achieve this, I use migration data from the Eurostat Institute of the European Union [Eur24]. I implement and compare two ways of visualizing Origin-Destination (OD) flow maps in an immersive environment using three-dimensional arcs and a space-time cube (STC). OD flow maps are a visualization tool to visualize geographical movement through a network of weighted and directional curves. The STC allows me to visualize the trends in migration over multiple years by encoding the time in the third dimension which allows me to stack multiple OD flow maps into a cube shape. Through this thesis, I aim to answer the question: "How can augmented reality enhance the visualization and analysis of migration flow maps to improve spatial and temporal understanding of migration patterns?"

Contribution Statement Parts of this project were done in collaboration with my colleague Marian Bramberger. Who worked independently on a related topic, "3D Physicalization of Migration Flows in Europe". The data pre-processing, as well as the implementation of a force-directed layout algorithm were created in collaboration. After the initial collaboration, we both developed a unique visualization model to represent the data.

CHAPTER 2

Related Work

2.1 Data Visualization

The visualization of time is especially relevant in the context of my thesis. Time-oriented data visualization is dependent on the scale and the scope of the time element. As described by Aigner et al. [AMST11], data can be ordinal, meaning only the relative relation of events like the sequence of events is relevant. In continuous time domains, time can be mapped to real numbers. This kind of time data is also known as "dense-time-data" and is especially important in synchronized computer systems. At last, the discrete-time domain allows the modeling of quantitative data by mapping the time steps to integers.

Additionally, the time scope of the data can greatly influence the visualization. In a point-based time scope, there is no information on what happens between two different points in time. This is in comparison to an interval-based time scope, where the change between two points in time is the main focus. Visualizing this class of data comes with its own methods and challenges. Research by Lee et al. [LAN21] proposes two main approaches to encoding time, including time-to-time encoding and time-to-space encoding. In time-to-time encoding, the time component is shown via animation or video playback. This approach works well when having multiple time points visible at the same time is not required. Otherwise, the preferred method is a time-to-space encoding where the domain of time is mapped to one of the spatial dimensions.

A way to visualize a time-to-space encoding is the STC. As described by Kristensson et al. [KDA+09], an STC is a visualization approach where spatiotemporal data is mapped using the height axis as an encoding for time, as shown in Figure 2.1. The data is then displayed in a cube, and each slice of the cube along its height (time) axis displays a discrete point in time over the data. They conclude that the STC is an effective way for novices to work and understand data displayed in this way. This helps users observe non-trivial relations and patterns in the data that require a comparison of multiple time points.

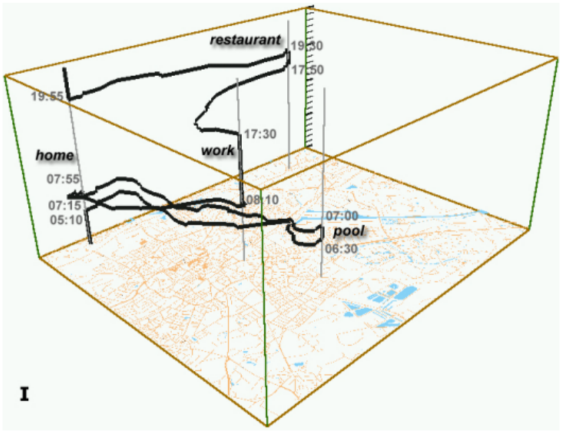


Figure 2.1: Space-Time Cube by Kraak [Kra03]

Furthermore, an STC-based approach to analyze conflicts in Africa and the resulting migration is also available. The work by Osman et al. [OOABA23] uses the STC to analyze these conflicts because previous conflict studies distinctly lack a focus on the time component. Previous methodologies for identifying conflict hotspots, such as nearest neighbor analysis and other clustering techniques, are effective spatial approaches but fall short when accounting for time. The research concludes that conflict studies benefit from the STC because of the way conflicts from different time periods and conflicts due to recent geopolitical events can be overlaid and analyzed. Additionally, the visualization helps the viewer understand how geographical factors can influence conflicts. One example is the change in season and climate change as well as access to education and drinking water over time in different regions of Africa.

2.2 Origin-Destination Flow Maps

OD flow maps are a way to visualize movement using weighted and directed arrows. The direction encodes the origin and destination of the flow, and the weight of the arrow is used to encode the quantity of the flow as [JSM+18], as shown in Figure 2.2. Recent trends show that organizations and foundations have started to migrate from tabular representations of data in the direction of interactive visualizations. Geographical movements like migration patterns, transportation, trade, digital traffic, tourism, and disease spread can effectively be visualized using OD flow maps[HNÅJ11].

Some key design principles have to be kept in mind when creating OD flow maps. Jenny et al. [JSM+18] define these principles and guidelines in their research. The main focus is on readability and reducing overlap as well as visual clutter on the maps. To verify these selected principles a user study was conducted. The three main findings of the study show that curved flows work better than straight flows or flows containing sharp edges and bends, arrowheads are the best way to indicate direction, and nodes at the origin and destination of the flow are better than using areas. The study hypothesizes that combining multiple encodings for the flow value, such as a mixture of width and brightness or saturation, can also increase readability.

Additionally, previous research by Yang et al. [YDJ+19] focuses on OD flow maps in immersive environments like augmented and virtual reality. When visualizing flows it is important to acknowledge the fact that the trajectory or curvature of the flow is not relevant as a data encoding. The work on immersive OD flow maps presents a few

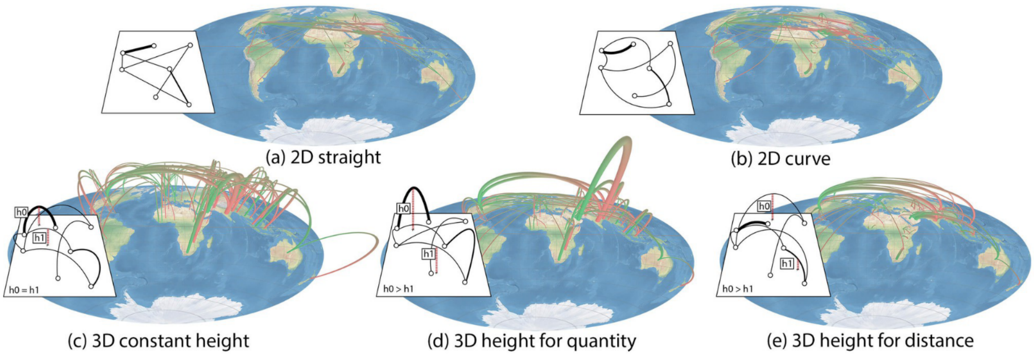


Figure 2.2: Different visualizations of OD flow maps in 2D and 3D [YDJ+19].

different ways of visualizing the flows. On one hand, the flows can still be displayed as two-dimensional lines or curves on a projected map. On the other hand, when working with immersive environments the flow maps can be elevated into the third dimension. In this case, the flows can be visualized as arches where the height allows for an extra way of encoding the flow value.

2.3 Augmented Reality

Carmigniani et al. [CFA+11] define AR as "a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer-generated information to it." The goal of AR is to enhance the experience of the user by integrating virtual information into the view of the user. It is important to note that AR is not restricted by particular display technologies like in the case of virtual reality. In VR, the user is completely immersed in the environment through the use of head-mounted displays (HMD) that completely obstruct the vision of the real environment around the user. In the case of AR, even overlays and enhancements on for example live sports broadcasts are considered to be a part of AR [GGD22].

The availability of software and easy hardware integration have helped AR become a reliable tool in data visualization [RCEJR19]. The ability to use a phone camera and its sensors to project virtual objects into the view of the camera makes AR available to anyone who owns a smartphone. AR on mobile devices comes with its own set of advantages and limitations. On one hand, the wide availability and built-in sensors like accelerometers and magnetometers make mobile AR a very viable option. On the other hand, limitations in computing power and performance can hinder more demanding and real-time applications [PT10]. There are multiple approaches and devices for tracking objects in space. Yi-Bo et al. [YBSPZHQ08] identify the general technology for tracking in AR to be optical, mechanical, and magnetic. Mechanical tracking refers to the use of physical sensors to track objects, whereas magnetic tracking utilizes electromagnetic

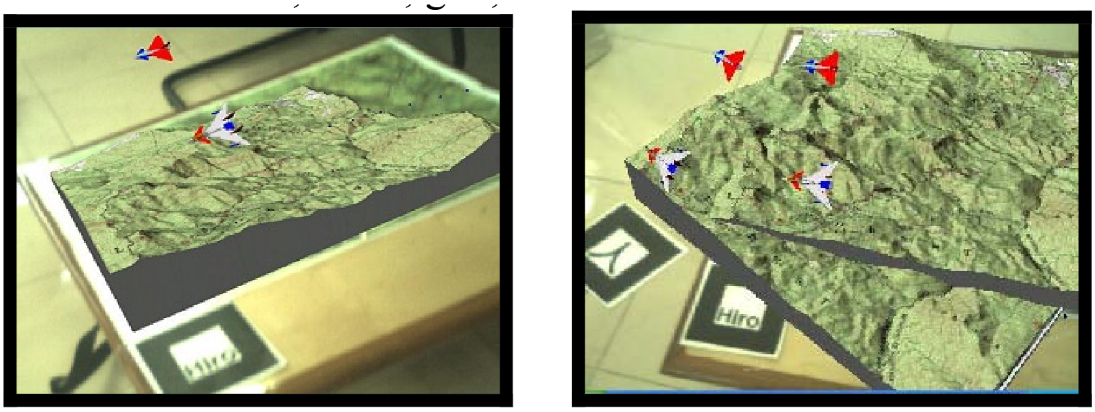


Figure 2.3: War planes projected onto a marker-based AR paper-based map visualization [AKN+10].

fields to find the position of objects. In my thesis, I use an optical approach, where markers are placed on a surface and tracked via a camera. These markers are then used to project the virtual information onto the surface.

Marker-based AR is a cheap and flexible implementation of AR that is especially advantageous for mobile phones, as described by Gherghina et al. [GOT13]. Virtually any image can be used as a marker, but the best practice is to use QR-Codes as marker images since QR-Codes are already well optimized for mobile use. Research by Adithya et al. [AKN+10] shows the viability of paper map visualization using marker-based AR. Even though their research mainly focuses on military training and learning, they have given a comprehensive solution for using AR to enhance a paper map by using the additional dimensions provided by the technology. As shown in Figure 2.3, they use the added third dimension to visualize digital elevation data and create a virtual terrain that is laid over the paper map. Additionally, virtual objects like planes are added to the visualization, allowing users to look at the flight path of, for example fighter jets interactively.

2.4 Research Gap

Even though there is prior research on migration data visualization using OD flow maps ([OOABA23]), there is a lack of immersive visualization approaches for research purposes. By combining a marker-based AR application with flow visualization, I aim to create an immersive solution for contextualizing and showing data. I propose to enhance the visualization with an encoding for time that allows the user to look at the data-specific points in time interactively by incorporating an interactive STC into the flow map visualization.

CHAPTER 3

Method

In the following chapter, I aim to give an overview of the methodology and main challenges of visualizing OD flow maps in the immersive environment of marker-based AR. I will outline the design principles and guidelines used to create a comprehensive and interactive visualization and briefly touch on specific technologies used to implement the visualization.

3.1 Design Guidelines

As the research by Jenny et al. [JSM+18] concludes, there are some important design principles and guidelines for OD flow maps that have to be kept in mind. In this section, I propose my design goals for a comprehensive AR flow map visualization based on recommendations and findings from related work [JSM+18, AKN+10]. Furthermore, there are also key design choices for marker-based AR that help me create a conclusive model for OD map visualization.

(G1) Accessibility Visualizing data such as migration movement can often lead to misleading conclusions [KAG15]. Users with little visualization literacy or computer experience as well as users who are uninformed on the topic, should be able to use interactive tools with ease to avoid such manipulation. The motivation for designing for accessibility in marker-based AR OD flow map visualization is to ensure that users can effectively interact with and understand the data. I want to make sure that users who are not well-versed in AR technologies can still use my visualization approach to further improve their knowledge on the topic of migration. OD flow maps represent complex movement patterns, and if the visualization is cluttered or unintuitive due to the overdrawing of flow curves, users may struggle to interpret the data. First and foremost, I propose to optimize the marker placement and scanning. The focus lies on having a very simple and intuitive system to project a virtual 2D paper map into the real-world environment. Additionally, I focus on intuitive interaction regarding the navigation of the visualization and flow exploration.

(G2) Clarity If the visualization is not concise, it can be overwhelming to inexperienced users [SES16]. Clarity is a design requirement that ensures that users can quickly and accurately interpret the movement data. If the flow paths, labels, or interactions are not designed clearly, users may misinterpret the data. The goal is to improve depth perception and spatial awareness as well as reduce visual clutter. AR integrates digital elements into the real world, and therefore, poor depth perception can make OD flows look distorted or misleading. If the visualization is cluttered, users may be overwhelmed or unable to comprehend the data. One of the main problems when working with OD flow maps is that as the number of flows increases, the flow visualization tends to get very cluttered. The research by Yang et al. [YDJ+19] proposes to solve this problem of visual clutter by introducing the third dimension provided by augmented or virtual reality.

(G3) Performance Even though there have been great advancements in processing power on mobile devices over the past years, AR visualization still has to be optimized for real-time rendering on a smartphone [PT10]. Poor performance can lead to lag, tracking instability, crashes, and user frustration, making the visualization ineffective. The marker tracking and detection can be greatly optimized by using high-contrast markers optimized for optical tracking like QR codes. To improve the performance of real-time flow rendering, some optimization steps are necessary. Another key aspect when implementing marker-based AR is to stabilize and maintain the visualization even when the camera moves or one of the markers gets temporarily occluded or moves out of frame. In addition to optimizing the marker tracking, the 3D models need to be improved for real-time rendering. When rendering flow maps, the flows can be visualized using simple splines with fewer polygons to increase rendering speeds. By pre-computing and caching flows, a lot of the redundant calculations can be avoided, which further optimizes the visualization for real-time rendering [WS09]. Furthermore, by designing with performance in mind I also aim to improve the accessibility of the visualization. Ensuring the visualization can run on a wide range of hardware like older smartphones the visualization is available to a wider range of users.

3.2 Origin-Destination Flow Maps

I propose two different time encodings in my visualization and therefore implement two separate OD flow map visualizations. On one hand, a static time encoding where the user can visualize the immigration for a single destination country in a specified year, and on the other hand an interactive time encoding using an STC. The STC allows the user to visualize the immigration for a single country over all years in the data.

In general, OD flow maps visualize geographic movement data like trade or migration as a network. Each node is an origin or destination, and the directional lines or curves connecting the nodes encode the flow value. The value for each flow can be encoded through a multitude of visualization techniques, such as flow thickness or saturation [JSM+18].

3.2.1 Arc Visualization

When visualizing a single year, I propose using the OD flow map visualization approach for immersive environments by Yang et al. [YDJ+19]. In their work, they use the third dimension to elevate the two-dimensional flows into three-dimensional arcs. When working with flow arcs there are multiple ways of using the arch height to either encode information or to help with visual clutter. Yang et al. [YDJ+19] propose three ways to visualize arcs in 3D space, as shown in Figure 2.2. The most straightforward approach is to use constant arc height, which can help with visual clutter but does not utilize the third dimension as an additional variable for encoding data. The other approaches both utilize the additional dimension as a form of data encoding. Using the arch height to encode flow distance is not relevant to my visualization since the distance between the countries does not encode relevant information when it comes to migration quantity. To comply with my design goal G2 (clarity), I choose the arc height as an additional encoding for flow quantity. This allows me to further highlight the flow quantity which is the most important aspect of the visualization.

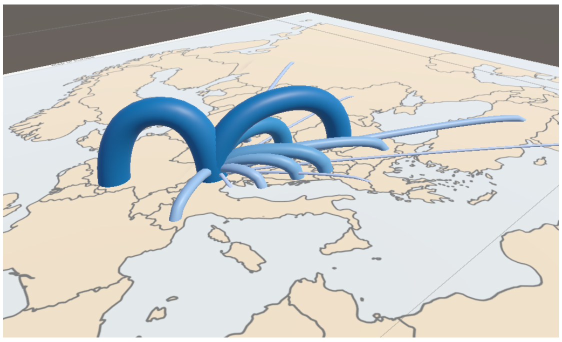


Figure 3.1: Render of arc flows for immigration into Austria in 2019. Each arced flow encodes the migration between an origin country and Austria

For the arc visualization, I combine flow thickness, saturation, and arc height to encode the flow value. This allows me to visualize the flows in a meaningful way. To retain G3 (performance), I propose to visualize the flows using spline objects which provide an intuitive and performant way to generate objects along curves and create trajectories as well as draw shapes. As shown in Figure 3.1, a flow is represented as a quadratic bezier curve where the control point is set as the midpoint between the origin and destination of the flow. I raise the control point depending on the flow value and introduce an upper bound to counteract extreme outliers. Analogously, I calculate the width of the flow by calculating the area of the circle depending on the flow value from which I extrude the spline into a tube shape. Lastly, I set the color of the flow to a blue tone determined by linearly interpolating a blue color gradient that goes from a de-saturated blue (#DCDDFF) for the smallest value to a saturated blue (#007BC2) for the largest observed flow value. I have chosen the blue gradient because, as suggested by research from Elliot et al. [EM14], blue is often associated with neutrality and trust, which is important for my visualization since this implies an impartial and neutral view of the data. It is important to note that the human eye is less sensitive to variations in blue compared to red and green. As a result, we perceive fewer shades of blue distinctly [NFBB22]. To mitigate these drawbacks and adhere to G1 (accessibility), I use multiple encodings for flow values simultaneously and users could easily change the color encoding in the future.

3.2.2 Space-Time Cube

Additionally, I want to be able to visualize immigration for a single country over multiple years. To achieve a visualization that allows the user to compare migration over multiple years, I propose an STC approach as presented by Bach et al. [BDA+17]. The STC provides a comprehensive visualization for datasets with multiple dimensions in addition to the time, such as geospatial data. As mentioned in my related work, the STC encodes time using one axis of the cube, where each slice along this axis represents a point in time. In my visualization, I use the third dimension of height supported by the AR approach to encode the time component of the data. The STC visualization utilizes similar technologies and encodings as my arc visualization approach. The flows are generated using splines and the flow value is encoded using the same thickness and color gradient as introduced in the arc flow visualization. Stacking multiple vertically arched OD flow maps does not work with the STC because the height of the arcs can overlap with other time slices of the STC. As shown in Figure 3.2, I propose using multiple flat OD flow maps, which can be stacked up along the time axis and allow the user to compare the differences in flows over multiple years. To preserve and ensure G2 (clarity), working with two-dimensional flow maps requires further optimization to address flow clustering. To combat flow clustering, I put forward a force-directed layout algorithm as introduced by Jenny et al. [JSM+17] to curve flows and to increase angular resolution.

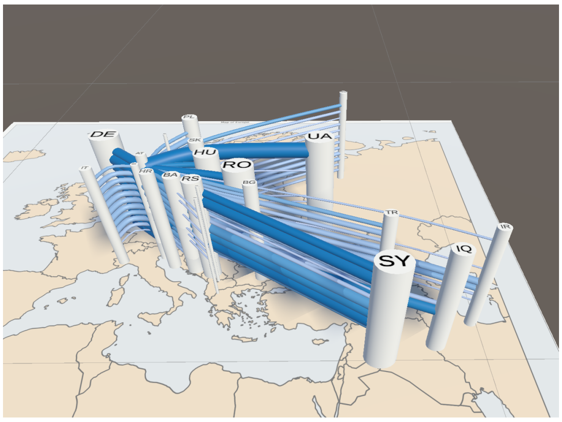


Figure 3.2: Render of the STC visualization for immigration into Austria between 2008 and 2022. Each pillar represents a Country, and every slice along the z-axis shows an OD flow map for a specific year of the data.

Force-Directed Layout The goal of a force-directed layout algorithm is to create a layout for a graph or, in my case, a flow network by moving nodes through simulated physical forces like attraction and repulsion. As shown in Figure 3.3, by continuously applying these forces over multiple iterations at decreasing weights, a layout for the flows

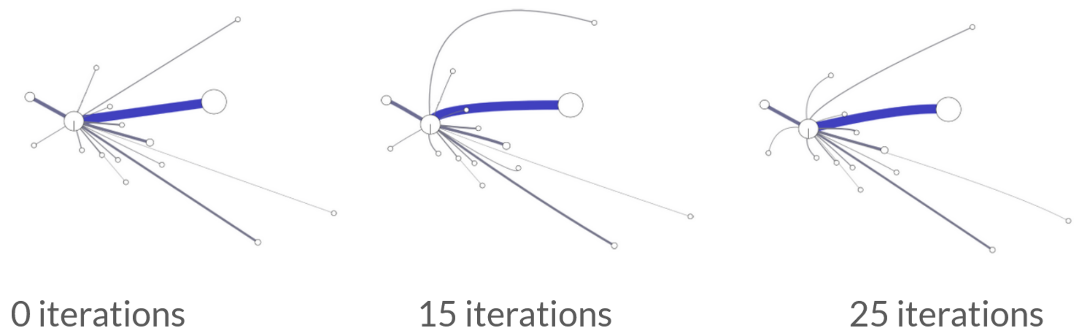


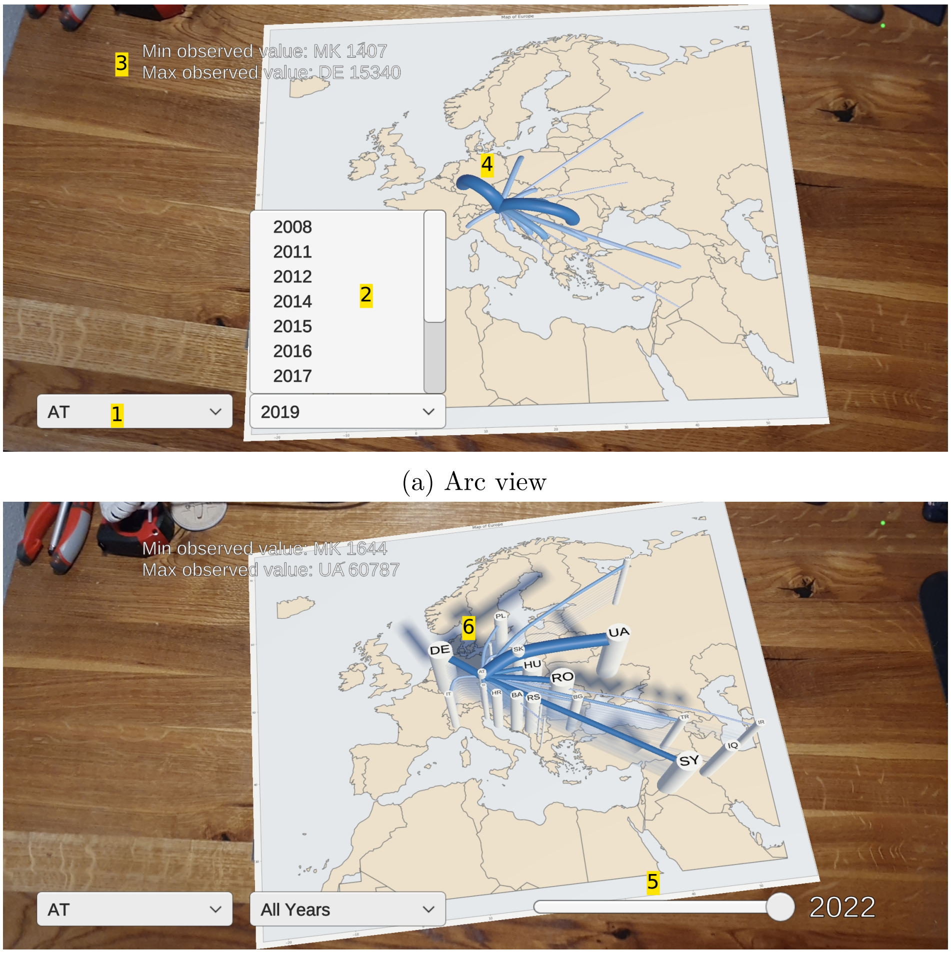
Figure 3.3: Force-Directed Layout after 0, 15, and 25 iterations for immigration into Austria in 2022 [Eur24]. The central destination Node displays Austria and all the other nodes form origin countries for flows.

is generated[HVW09]. The specific algorithm proposed by Jenny et al. [JSM+17] treats flows as quadratic bezier curves. The forces are repeatedly applied to bezier control points. Additionally, the algorithm implements a way to reduce flow intersections by checking the neighborhood of the control points. Control points are repulsed by other flows, and their start and end points and the angular resolution at the nodes is increased. This allows the flows to be spaced out throughout the visualization. Additionally, more forces have to be applied to counteract these repulsion forces. The goal is to prevent flows from curving away from each other uncontrolled and to increase symmetry. A spring and anti-torsion force attracts the control points back towards the midpoint of their respective start and end points. Lastly, the flows are bound to a rectangle determined by the flow length to further control the maximal curvature of the flow.

Because of my specific use case, I adjusted the proposed algorithm in how the forces interact with the flows. When visualizing immigration for a particular country, all flows share the same destination. The singular destination node is the only node in the system that requires the angular resolution force since it is the only node where all flows share a common node. The angular resolution force fans out the angles of the incoming flows to reduce clutter in the close vicinity of the destination node. Additionally, the destination node does not exert any repulsion forces to prevent strange flow behavior for flows with an origin close to the destination.

3.3 Application

Figure 3.4 shows the marker-based AR smartphone application I develop for my thesis. To help with G1 (accessibility), I choose intuitive controls to interact with the visualization. Users can use two simple drop-down menus to select a country and a year (see 1 and 2). As soon as both markers are tracked and a country and year are selected, the visualization will be displayed. If the user chooses a specific year from the drop-down menu, the migration for that year will be displayed using arc visualization (see 4). Additionally, the minimum and maximum observed value for the selected year is displayed in the top right corner of the phone screen (see 3). This allows the user to gain further information in addition to the rendered and encoded flows. Furthermore, if the user selects "All Years" from the year drop-down, the visualization switches to the STC view (see 6). The STC view introduces two additional controls. A slider in the bottom left corner (see 5) allows the user to highlight a specific year of the visualization, with all other years fading out of focus. The STC view introduces pillars at the origin and destination of each flow with the country code on top. The width of the pillar is determined by the maximum observed value over the data for said country and is used to cohere to G2 (clarity). To further highlight specific countries in the visualization the user can tap the pillar at an origin country to completely hide the attached flows from the visualization. This allows the user to focus on one or multiple origin countries.



(a) Arc view

(b) STC view

Figure 3.4: Marker-based AR smartphone application displaying immigration into Austria with country dropdown menu (1), year dropdown menu (2), min and max observed value (3), and the year slider (5), as well as an arc visualization (4) and the STC view (6)

3.4 Implementation

The pre-processing of the data, as well as the force-directed layout algorithm, was implemented using Python with the help of the packages Pandas, GeoPandas, and NumPy. Pandas is a data analysis library that I mainly used to filter the large dataset. After filtering the data, I used GeoPandas as an extension of Pandas to handle the geospatial data provided by the Natural Earth dataset. Finally, the numerical computing library NumPy was used to implement the force-directed layout algorithm introduced by Jenny et al. [JSM+17]. The resulting data, as well as the coordinates and the control points for the flows, were then exported as CSV files to be used in the AR visualization.

To visualize the flows I propose a mobile phone marker-based AR application that allows the user to intuitively control and interact with the data. Marker-based AR using a mobile phone camera solves some of the controls by design: As long as both markers are being tracked, the user can freely move the camera in the real world and consequently change the viewing angle of the visualization. Furthermore, physically moving the camera closer or farther away from the markers functions as a zoom feature. Users are supported by intuitive touchscreen controls like tapping the countries to hide or show flows as well as simple drop-down menus and a slider to highlight data entries or switch between years and data entries. The AR application is implemented in Unity using the Vuforia Engine and can be run on any Android device with ARMv7, ARM64, x86, or x86-64 architecture. The widely used computer vision-based AR framework Vuforia Engine solves performance and optimization issues by utilizing cloud technology, which helps me to adhere to G3 (performance). Using the cloud to offload the image detection reduces the load on local processing [XL14]. The specific versions of all technologies and tools used in my Project are listed in table 3.1.

Technology/Tool	**Version**	**Objective**
Python	3.13	Main programming language for data pre-processing and layout algorithm
Pandas	2.2.3	Data analysis and data pre-processing
GeoPandas	1.0.1	Handling geospatial data
NumPy	2.2.0	Numerical Computing
Matplotlib	3.9.1	Plotting two-dimensional OD flow maps
Unity Engine	2022.3.53f1	Implementation of the visualization
Vuforia Engine	10.28.4	Augmented reality software development kit
PyCharm Professional	2024.3.2	IDE for Python development
Android Studio	2024.2.1	Benchmarking application performance

Table 3.1: List of technologies and tools used in the project

Marker placement and tracking To ensure G1 (accessibility), the marker placement and tracking are designed to be as intuitive as possible. The two QR codes have to be printed on paper and labeled as bottom-left and top-right on their backside. The markers are generated using the Vuforia integrated VuMark database and are therefore optimized for use with Vuforia Engine. VuMark allows the markers to be saved to the cloud, which is helpful if the application needs a lot of markers, but because I only have exactly two markers, I chose to save them locally on the device. As visible in Figure 3.5 and Figure 3.6, if the markers are placed correctly, the application will automatically start projecting the virtual paper map. Vuforia manages all the tracking internally and no further implementation for tracking stability is required.

Figure 3.5: AR markers untracked

Figure 3.6: AR markers tracked with the projected paper map

CHAPTER 4

Result

In the following chapter, I will present the marker-based AR smartphone application that I developed for my thesis as well as explain the data used in the final visualization. Furthermore, I will show the AR Visualization with both time encodings through immigration into Austria. I chose to present the project through Austria as an example country because of its geographical location and extensive data availability.

4.1 Dataset

The first step of the visualization is to gather the underlying data and transform it as well as pre-process it. To achieve this visualization, I rely on a comprehensive dataset gathered by Eurostat. Eurostat is the statistics Institute of the European Union that provides high-quality statistics and data on Europe. Their data covers a wide variety of topics such as energy, environment, trade, transport, technology, population, and social conditions. The dataset I use in my visualization is called "Immigration by age group, sex and country of birth", which covers data on immigration starting in 2008 and includes close to nine million data entries. I use the version last updated on the 13th of November 2024 [Eur24].

On one hand, the dataset includes data for 301 countries or regions of birth. To ensure G2 (clarity), the visualization focuses on migration in Europe. Therefore, I reduced the countries of birth where flows can originate to all the nations in Eurasia as well as countries in northern Africa and the Near East, resulting in 76 possible origin countries, as seen in Figure 4.1. On the other hand, there are some data entries with very small quantities. To reduce visual clutter, remove all data entries where the observed value is below a certain threshold. I exclude all data entries where the observed value is below 90% of the mean in the data. This yields a sufficient threshold when outlier detection is not of importance [Dal09].

Additionally, as suggested by the name of the dataset, the data is further split into "sex" and "age," which is not important to my visualization. Therefore I combined the data to only show one combined observed value. Each entry in the processed data now corresponds to an origin-destination country pair in a specific year, and the observed value is then encoded in the flow.

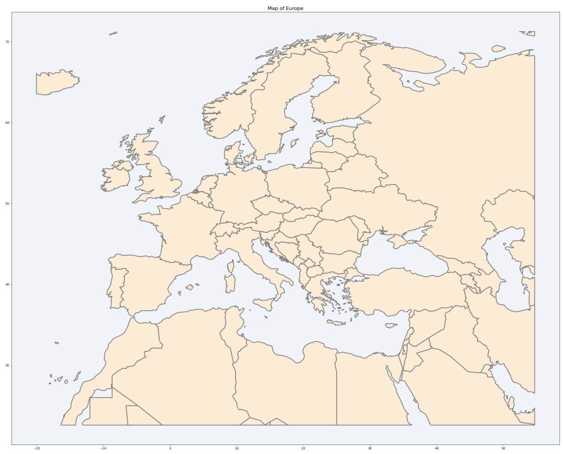


Figure 4.1: Geographical scope of the visualization including Eurasia, Northern Africa, and the Near East

Lastly, I need a way to handle country coordinates and map the data onto a paper map. I use a free public vector and raster map data set called Natural Earth. The Natural Earth dataset includes different scales and many different cultural and physical data themes such as countries, populated places, coastlines, and glaciers just to name a few. In my thesis, I use the 1:50m cultural vector country data set, which includes 258 countries and their boundaries according to *de facto* ("in fact") status rather than *de jure* ("by law"). The data uses the Geographic coordinate system WGS84, which is a standard in cartography, geodesy, and satellite navigation including GPS.

4.2 Example case: Austria

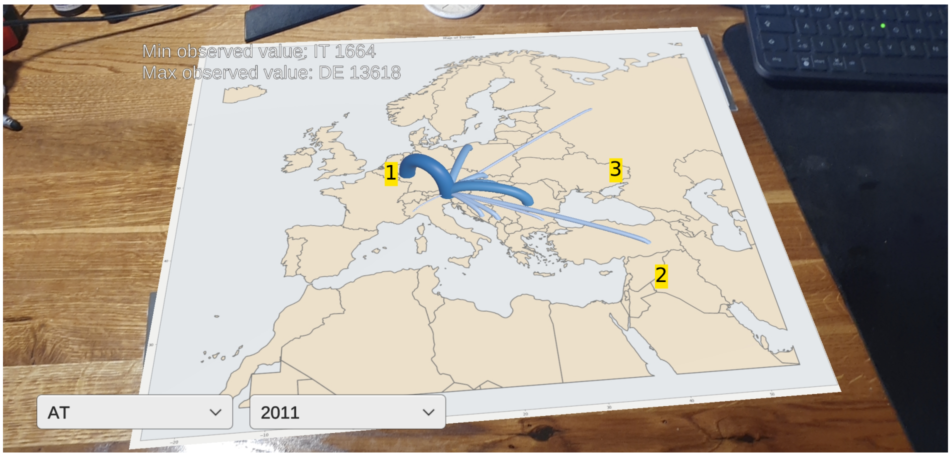
To show the marker-based AR application, I chose to present immigration into Austria. The data for Austria also spans from 2008 to 2022, with three years missing. The data for 2009, 2010, and 2013 is incomplete. This is a recurring theme in the dataset provided by Eurostat, where some countries do not display immigration on a country-by-country basis but rather abstract the data by combining multiple origin countries into groups like "Non-Eu 27 countries" or "Countries other than EU, EFTA, and Candidate countries". For example, Germany only began providing country-by-country immigration statistics in 2020, so the visualization I can create from the data is not yet fully comparable. The calculated threshold for the pre-processing of the Austria data is 1141, and in total 185 flows from 20 origin countries can be visualized.

4.2.1 Arc Visualization

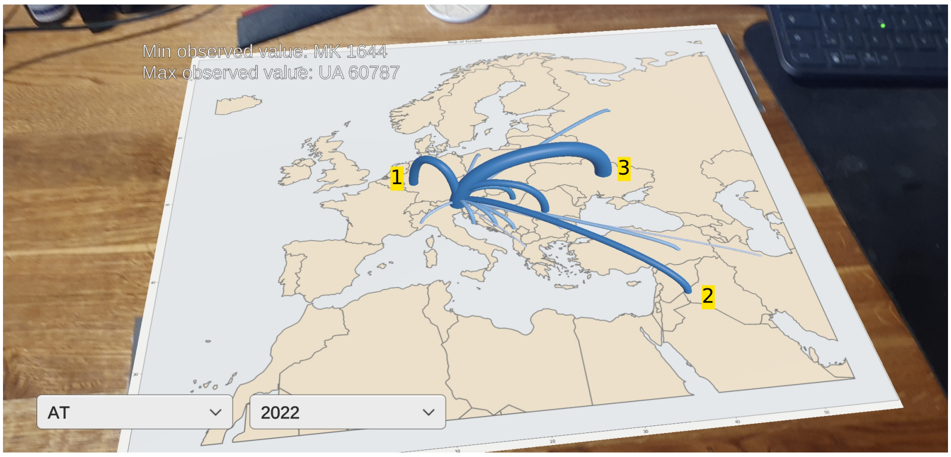
In Figure 4.2, I show two examples of arc visualized OD flow maps for immigration into Austria in 2011 and 2022. As soon as a country and year are selected, the arc visualization of the OD flow map is projected onto the previously tracked virtual paper map. For the year 2011, visible in Figure 4.2a, Germany (see 1) had the largest observed immigration into Austria with a value of 13618, and Italy had the smallest immigration with an observed value of 1664. As shown in the visualization, Germany is encoded with the most saturated color of the blue color gradient, has the highest possible arc trajectory, and has the thickest possible flow width.

Figure 4.2b shows the arc visualized OD flow map for immigration in 2022. Countries like Syria (see 2) and Ukraine (see 3), which did not pass the calculated threshold in 2011, are now visualized with observed values of 11142 for Syria and 60787 for Ukraine. One interesting observation when comparing immigration in 2011 and 2022 is the general increase in immigration in 2022. In 2011, only 10 countries surpassed the calculated threshold compared to 16 in 2022. Furthermore, the absolute numbers for each country increased as well with immigration from Ukraine having an observed value over four times higher than the maximum value recorded in 2011.

It is important to note that in the arc visualization approach, comparing two different years in a meaningful way is only limitedly possible. To adhere to G2 (clarity), the OD flow map for each year is encoded separately to further the clarity of the visualization. As visible in Figure 4.2, the flow for migration from Germany in the year 2022 is much slimmer than the flow in 2011. The absolute number of migrants for Germany in 2022 is 18422 and therefore bigger than the number recorded in 2011. If the flow width, height, and color were encoded over all the years simultaneously, OD flow maps for years with generally low migration would not be very expressive. By encoding each year separately, each OD flow map utilizes the full range of encoding, which increases the clarity of the visualization by focusing on one specific year at a time. So while it is possible to see trends in migration over the years by which countries are visualized, it is not possible to directly compare flows for a specific origin country over multiple years.



(a) 2011



(b) 2022

Figure 4.2: Marker-based AR arc visualization of Austria in 2011 and 2022. The relation of observed migration can be compared each year for countries like Germany (1), Syria (2), and Ukraine (3). Mind that there is no relationship between flow encodings across multiple years.

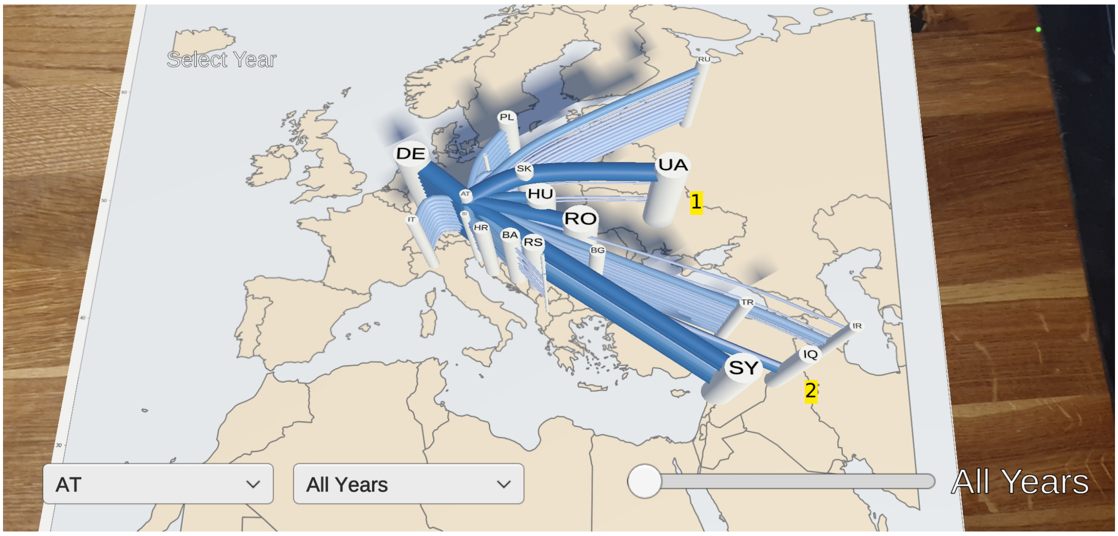


Figure 4.3: Overview of STC visualization for Austria with all flows visible. Irregularities in migration for countries like Ukraine (1) and Iraq (2) can be seen over the years.

4.2.2 Space-Time Cube visualization

To compare multiple years of immigration simultaneously and meaningfully, I implemented the STC visualization as presented in Figure 4.3. When a user selects "All Years" from the year drop-down menu, the visualization displays the STC view of the data. If no specific year is highlighted by the slider, the entire data is visualized. This allows the user to get an overview of the change in migration over multiple years. Years with more thick and saturated flows like the top STC slice 2022 show years with generally high migration. Additionally, users can deduct that countries like Ukraine (see 1) or Iraq (see 2) only surpass the visualized threshold in some years of the data. In contrast, countries like Italy have a constant migration over the entire data ranging between 1664 in 2011 and 3204 in 2022, with all flows being encoded in a similar color and thickness.

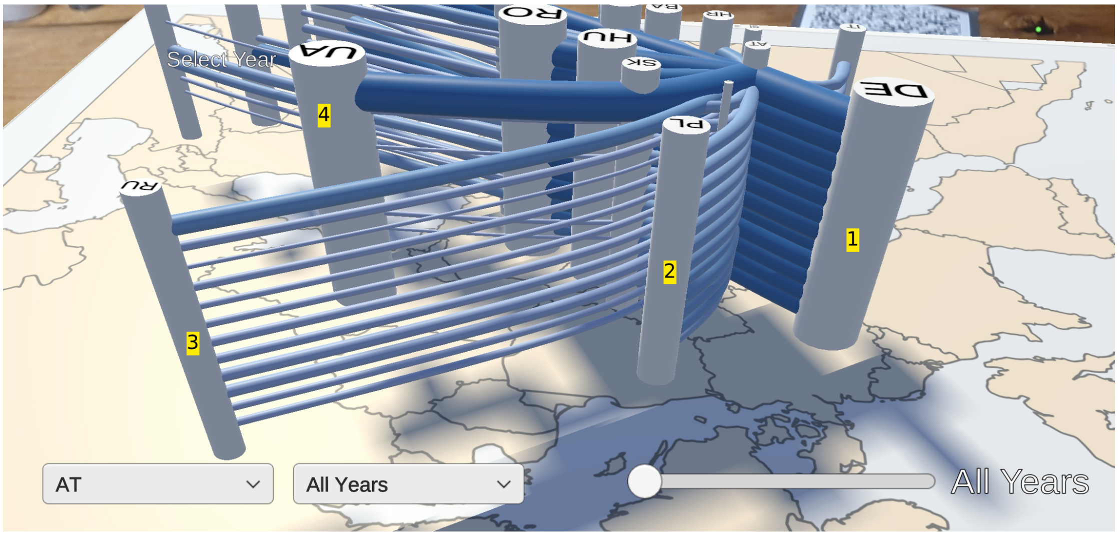
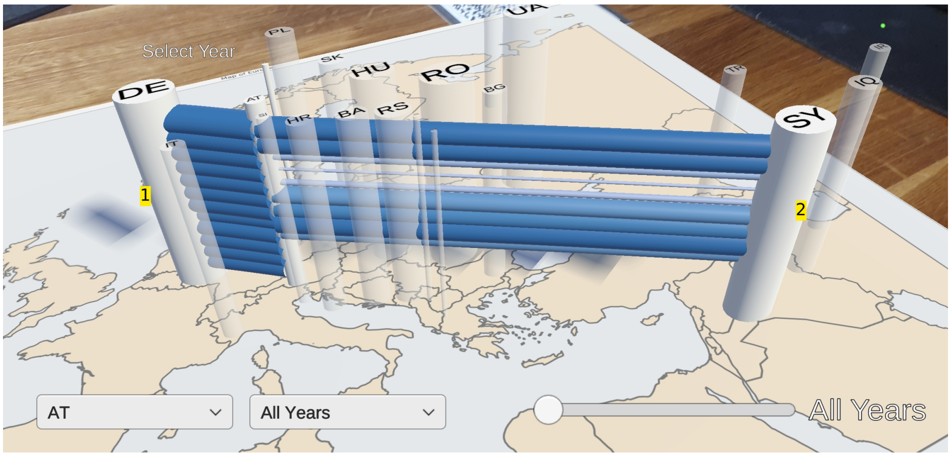


Figure 4.4: Close-up and angled view of the entire STC visualization of immigration into Austria with a focus on Germany (1), Poland (2), Russia (3), and Ukraine (4)

By physically moving the camera the user can view the STC visualization from different angles and scales intuitively. As shown in Figure 4.4, viewing the data at a different orientation and zoom can help users gain further insights into the data. Viewing at this angle makes it easier to compare the countries Germany (see 1), Poland (see 2), and Russia (see 3). It is interesting to see that all three countries have a rather consistent amount of migrants over the data with no real outliers over the years. This suggests that geopolitical factors do not have a noticeable impact on migration from these selected countries. In comparison, immigration from Ukraine (see 4) rarely crossed the calculated threshold up until 2022, when a huge spike in migration occurred, which coincides with the Russian invasion of Ukraine that started on February 24th, 2022.

The STC visualization provides the user with multiple ways of exploring the data like comparing two specific origin countries side by side. As shown in Figure 4.5, by deselecting all countries except Germany (see 1) and Syria (see 2), the user can focus the visualization on these specific countries. By keeping the slider at "All Years", as seen in Figure 4.5a, the entire data for Germany and Syria is displayed. The data shows constant high immigration from Germany with a yearly value between 12770 and 18422. In contrast, immigration from Syria only surpassed the calculated threshold after 2014 and had very low observed values in 2018, 2019, and 2020. To further compare the two countries, the user can select a specific year like 2017, as seen in Figure 4.5b using the slider. The data shows that for 2017 the observed value for Germany (13666) is more than twice the value for Syria (5798).



(a) All years

(b) 2017

Figure 4.5: STC view of migration from Germany (1) and Syria (2) into Austria

4.3 Performance

As mentioned in my design goal G3 (performance), I focused on making the application performant for real-time rendering and phone usage. To benchmark the performance of my application, I tested the application using Android Studio with two different emulated Android devices. The two phones I emulated ran on Android 11.0 (API 30) as well as Android 15.0(API 35), and the specific properties of both devices used are listed in Table 1. Both phone emulators used for testing have rather weak specs compared to real-world flagship Android devices like the Samsung Galaxy S24 Ultra or the Google Pixel 9 Pro. By testing the application on low to mid-range devices, I can verify that the application is available for a wide audience, which is necessary to fulfill G1 (accessibility). To measure performance, I logged the internal device timestamps from the time a new year is selected from the drop-down menu to the moment the flows are rendered and displayed in the visualization. I ran the test on both devices, switching between different arc and STC visualizations 50 times for each device. I then wrote a short Python script to evaluate rendering times for the arc and STC visualizations respectively.

An interesting observation when testing the application on both emulators is the similar performance on both devices. When switching between arc visualizations of different countries and years the average rendering time came out to 8.28ms. Modern phone screens have a maximal refresh rate of 120hz, which means the frame time is 8.33ms. The average time to render a new arc visualization is therefore shorter than the time interval between frames, which allows the user to seamlessly switch between countries and years. Switching between STC visualizations, on the other hand, is way more computationally expensive, with rendering times averaging at 508.67ms or about half a second. This leads to a noticeable wait time when switching between different STC visualizations, but once the flows are rendered navigating the data and viewing the visualization from different angles works without any noticeable lag or stutter. The main reason for these fast rendering times is the use of simple extruded spline objects with a rather small amount of segments per unit which increases rendering times by keeping the meshes rather simple. Additionally, all the flow data is shipped with the application and directly saved on the device. This reduces the time it takes to load the data compared to first having to retrieve the data from an online server. This comes at the cost of an increase in disk space required by the application which is negligible due to the small size of the CSV data files being just a couple of megabytes each.

CHAPTER 5

Conclusion & Future Work

To conclude my thesis, I want to address some limitations and further improvements of my visualization approach as well as give some personal thoughts and ideas on future work related to interactive data visualization of spatiotemporal data. Additionally, I will summarize my main findings, contributions, and insights on the research topic.

5.1 Improvements and Future Work

The visualization approach I developed for my thesis shows a promising and intuitive way to interact with spatiotemporal data in an immersive environment but has some space for improvement. The visualization is currently limited to a eurocentric view of the data and could be expanded to cover global migration trends and data. Countries such as Great Britain, Spain, and Portugal have a considerable amount of migration from continents that are not part of my visualization scope, which contains Eurasia, Northern Africa, and the Near East. Expanding the visualization to a global scope would allow users to view this data in a meaningful way.

The proposed expansion of the visualization scope comes with its own set of new challenges, like the additional visual clutter. It would be interesting to see some visualization approaches, like flow bundling to combat these problems. When visualizing migration globally, it would be intriguing to compare a projected paper map visualization with a globe visualization. Using mobile-based AR, it would be possible to project the data on an interactive globe that the user can rotate by swiping the projected sphere. The arc visualization approach presented in my thesis would be better suited for the globe projection, but it would be interesting to develop an STC approach for this kind of visualization.

My implementation allows users to interactively look at the data and compare multiple years of immigration for a selected country but could use some improvements to help users gain further insights. It is currently not possible to get numerical information for each flow but rather to gain insights into migration trends. It would be interesting to add additional controls where the user can tap a flow and get numerical information on the flow value in addition to all the encodings already in place.

5.2 Summary

To answer the research question: "How can augmented reality enhance the visualization and analysis of migration flow maps to improve spatial and temporal understanding of migration patterns?" I conclude that both the arc visualization and the STC give a beneficial approach to visualizing flow data in the immersive environment of AR. The arc visualization provides a meaningful visualization for highlighting specific years in the data, allowing users to view the migration distribution. The arc visualization is limited when it comes to viewing and understanding migration patterns over multiple years. These insights can be achieved through the STC visualization, which allows users to better understand migration trends over multiple years. Both visualizations are easy and intuitive to navigate and allow even novices in data analysis, visualization, and migration studies to gain further understanding of the topic of migration.

List of Figures

List of Tables

Bibliography

[AKN+10] C. Adithya, K. Kowsik, D. Namrata, V. S. Nageli, Sangeeta Shrivastava, and S. Rakshit. Augmented reality approach for paper map visualization. In *Proceedings of 2010 International Conference on Communication and Computational Intelligence, INCOCCI-2010*, 2010.

[AMST11] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data.* 2011.

[BDA+17] B. Bach, P. Dragicevic, D. Archambault, C. Hurter, and S. Carpendale. A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes. *Computer Graphics Forum*, 36(6):36–61, 9 2017.

[CFA+11] Julie Carmigniani, Borko Furht, Marco Anisetti, Paolo Ceravolo, Ernesto Damiani, and Misa Ivkovic. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1), 2011.

[Dal09] Christoph Dalitz. Reject options and confidence measures for knn classifiers. *Schriftenreihe des Fachbereichs Elektrotechnik und Informatik Hochschule Niederrhein*, 8(2009), 2009.

[EM14] Andrew J. Elliot and Markus A. Maier. Color psychology: effects of perceiving color on psychological functioning in humans. *Annual review of psychology*, 65:95–120, 2014.

[Eur24] Eurostat. Migration Statistics - Immigrants by Sex, Age, and Citizenship (Code: migr_imm3ctb). https://ec.europa.eu/eurostat/databrowser/product/page/migr_imm3ctb, 11 2024.

[Fri08] Michael Friendly. A Brief History of Data Visualization. In *Handbook of Data Visualization*, pages 15–56. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[GGD22] Chad Goebert, Gregory Greenhalgh, and Brendan Dwyer. A whole new ball game: Fan perceptions of augmented reality enhanced sport broadcasts. *Computers in Human Behavior*, 137:107388, 12 2022.

[GOT13] Alexandru Gherghina, Alexandru Corneliu Olteanu, and Nicolae Tapus. A marker-based augmented reality system for mobile devices. In *Proceedings - RoEduNet IEEE International Conference*, 2013.

[HKSS17] Sumit A. Hirve, Ajinkya Kunjir, Basil Shaikh, and Karan Shah. An approach towards data visualization based on AR principles. In *Proceedings of the 2017 International Conference On Big Data Analytics and Computational Intelligence, ICBDACI 2017*, 2017.

[HNÅJ11] Quan Ho, Phong H. Nguyen, Tobias Åström, and Mikael Jern. Implementation of a Flow Map Demonstrator for Analyzing Commuting and Migration Flow Statistics Data. *Procedia - Social and Behavioral Sciences*, 21:157–166, 1 2011.

[HVW09] Danny Holten and Jarke J. Van Wijk. Force-Directed edge bundling for graph visualization. *Computer Graphics Forum*, 28(3), 2009.

[IJ19] Mohaiminul Islam and Shangzhu Jin. An Overview of Data Visualization. In *International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2019*, 2019.

[JSM+17] Bernhard Jenny, Daniel M. Stephen, Ian Muehlenhaus, Brooke E. Marston, Ritesh Sharma, Eugene Zhang, and Helen Jenny. Force-directed layout of origin-destination flow maps. *International Journal of Geographical Information Science*, 31(8), 2017.

[JSM+18] Bernhard Jenny, Daniel M. Stephen, Ian Muehlenhaus, Brooke E. Marston, Ritesh Sharma, Eugene Zhang, and Helen Jenny. Design principles for origin-destination flow maps. *Cartography and Geographic Information Science*, 45(1), 2018.

[KAG15] Anna Kaisa Kuusisto-Arponen and Mary Gilmartin. The politics of migration, 2015.

[KDA+09] Per Ola Kristensson, Nils Dahlbck, Daniel Anundi, Marius Bjrnstad, Hanna Gillberg, Jonas Haraldsson, Ingrid Mrtensson, Mathias Nordvall, and Josefine Sthl. An evaluation of space time cube representation of spatiotemporal patterns. *IEEE Transactions on Visualization and Computer Graphics*, 15(4), 2009.

[KK06] Claudio Kirner and Tereza G. Kirner. A data visualization virtual environment supported by augmented reality. In *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, volume 1, 2006.

[Kra03] Menno-jan Kraak. The Space-Time Cube Revisited from a Geovisualization Perspective. *21st International Cartographic Conference (ICC)*, (August), 2003.

[LAN21] Alexandra Lee, Daniel Archambault, and Miguel A. Nacenta. The effectiveness of interactive visualization techniques for time navigation of dynamic graphs on large displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 2021.

[NFBB22] Lucia Natali, Alessandro Farini, Elisabetta Baldanzi, and John Barbur. Effects of tinted lenses on chromatic sensitivity: changes in colour vision assessed with the CAD test, a preliminary study. 8 2022.

[OOABA23] Adams Osman, Alex Barimah Owusu, Kofi Adu-Boahen, and Emmanuel Atamey. Space-time cube approach in analysing conflicts in Africa. *Social Sciences & Humanities Open*, 8(1):100557, 1 2023.

[PT10] Rémi Paucher and Matthew Turk. Location-based augmented reality on mobile phones. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010*, 2010.

[RCEJR19] Ananth N. Ramaseri Chandra, Fatima El Jamiy, and Hassan Reza. Augmented reality for big data visualization: A review. In *Proceedings - 6th Annual Conference on Computational Science and Computational Intelligence, CSCI 2019*, 2019.

[Ris19] Roopika Risam. Beyond the Migrant "Problem": Visualizing Global Migration. *Television and New Media*, 20(6), 2019.

[SES16] Bahador Saket, Alex Endert, and John Stasko. Beyond usability and performance: A review of user experience-focused evaluations in Visualization. In *ACM International Conference Proceeding Series*, volume 24-October-2016, 2016.

[WS09] Daniel Wagner and Dieter Schmalstieg. Making augmented reality practical on mobile phones, part 2. *IEEE Computer Graphics and Applications*, 29(4), 2009.

[XL14] Cheng Xiao and Zhang Lifeng. Implementation of mobile augmented reality based on Vuforia and Rawajali. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, pages 912–915, 10 2014.

[YBSPZHQ08] Li Yi-Bo, Kang Shao-Peng, Qiao Zhi-Hua, and Zhu Qiong. Development actuality and application of registration technology in augmented reality. In *Proceedings of the 2008 International Symposium on Computational Intelligence and Design, ISCID 2008*, volume 2, 2008.

[YDJ+19] Yalong Yang, Tim Dwyer, Bernhard Jenny, Kim Marriott, Maxime Cordeil, and Haohui Chen. Origin-Destination Flow Maps in Immersive Environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 2019.

Appendix

Property	**Device 1 (Medium Phone API 35)**	**Device 2 (Medium Phone API 30)**
Display Name	Medium Phone API 35	Medium Phone API 30
Encoding	UTF-8	UTF-8
AVD ID	Medium_Phone_API_35	Medium_Phone_API_30
Data Partition Size	6GB	6GB
Accelerometer Enabled	yes	yes
Arc Support	false	false
Audio Input	yes	yes
Battery Support	yes	yes
Rear Camera	virtualscene	virtualscene
Front Camera	emulated	emulated
CPU Cores	8	4
Device Manufacturer	Generic	Generic
Device Name	medium_phone	medium_phone
D-Pad Support	no	no
GPS Support	yes	yes
GPU Enabled	yes	yes
GPU Mode	auto	auto
Screen Density	420	420
Screen Height	2400	2400
Screen Width	1080	1080
RAM Size	2G	2G
SD Card Support	yes	yes
Orientation Sensor	yes	yes
Proximity Sensor	yes	yes
Android API Version	35	30
Network Latency	none	none
Network Speed	full	full
Show Device Frame	yes	no
Dynamic Skin Support	yes	yes
Tag Display	Google Play	Google Play
VM Heap Size	336	228

Table 1: Comparison of Properties for Device 1 and Device 2