Visually Integrated Clinical Cooperation - Algorithmic Concepts, Implementation and Evaluation

----

Visuell Integrierte Klinische Kooperation - Algorithmische Konzepte, Implementierung und Evaluation

DISSERTATION

zur Erlangung des Grades eines Doktors
der Ingenieurwissenschaften (Dr.-Ing.)

vorgelegt von
M.Sc. Dmitri Presnov

eingereicht bei der
Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen

Siegen 2023

Betreuer und erster Gutachter
Prof. Dr. Andreas Kolb
Universität Siegen

Zweiter Gutachter
Prof. Dr. Volker Blanz
Universität Siegen

Tag der mündlichen Prüfung
13. Dezember 2023

Abstract

The workflow in modern hospitals entails that the medical treatment of a patient is distributed between several physicians and nurses. This leads to an intensive cooperation, which takes place under particular time pressure and requires an efficient conveyance of the relevant patient-related medical data to colleagues. This requirement is difficult to achieve with traditional data representation approaches, which usually do not consider the specifics of the cooperative work. This thesis introduces a concept of Visually Integrated Clinical Cooperation (VICC) to support the information transfer in cooperative tasks in a hospital. Developed in the context of an interdisciplinary research project, it exploits findings of a sociological field study in a hospital and feedback from the discussions with physicians. Accordingly, the VICC concept aims at an multivariate visualization on a mobile device that provides a synopsis of the relevant data and is intuitively comprehensible for medical personnel. The core components of the proposed concept are *anatomically integrated in-place visualization* and *iconic glyphs*.

The anatomically integrated visualization uses a 3D human avatar as spatial representation of visually encoded medical data *with* an (inherent) *anatomical reference*. This component comprises a set of formal requirements and procedures for this kind of visual encoding as well as a prototypical implementation for the diagnosis of spinal disc herniation, evaluated by neurosurgeons.

The VICC concept also includes a personalization option of the generic 3D human avatar, using a 3D reconstruction of a patient's body from range data. To allow it, a method for robust camera pose tracking for the spatially and temporally low-resolution range data, given in mobile applications, is proposed. It combines a geometry-based pose estimation by means of the iterative closest point (ICP) algorithm with inertial tracking, using an extended Kalman filter (EKF). In particular, it uses the extrapolated ICP pose estimates as virtual measurements and the output of the EKF as initial guess for the next ICP-based pose estimation.

The iconic glyph approach allows for representation of a patient's data *without anatomical reference*. It aims to combine visual metaphors, inherent for icons, and the glyph capability for multivariate visualization. Technically, it is based on a parametric representation that utilizes diffusion curves, enriched with new degrees of freedom in arc-length parametrization, which allows for automated, controllable manipulation of the icon contours' geometry and the related colour attributes. Besides the generic concept, an implementation for a specific design, based on periodic, wavelike contour modifications along with a perception and quantization model for these kinds of visual variables are proposed. The practicality of the approach is demonstrated by examples for visualization of weather forecast uncertainty, COVID-19 statistic trends and intracranial pressure.

Zusammenfassung

Die Arbeitsabläufe im modernen Krankenhaus sehen vor, dass die Patientenbehandlung zwischen mehreren Ärzten und Pflegekräften verteilt ist. Dies hat eine intensive Kooperation zur Folge, die unter einem besonderen Zeitdruck stattfindet und eine effiziente Übermittlung der relevanten medizinischen Patientendaten an Kollegen erfordert. Diese Anforderung kann mit den traditionellen Datendarstellungsansätzen, bei denen die Besonderheiten der kooperativen Arbeit in der Regel nicht berücksichtigt werden, kaum erfüllt werden. Die vorliegende Arbeit stellt ein Konzept der Visuell Integrierten Klinischen Kooperation (VIKK) zur Unterstützung vom Informationstransfer im Kontext der kooperativen Aufgaben im Krankenhaus vor. Es wurde im Rahmen eines interdisziplinären Forschungsprojekts entwickelt und basiert auf den Erkenntnissen einer sozialwissenschaftlichen Feldstudie im Krankenhaus sowie auf dem Feedback aus den Diskussionen mit den Ärzten. Daraus resultierend bezweckt das VIKK Konzept eine multivariate Visualisierung auf einem mobilen Endgerät, die eine Synopsis der relevanten Daten zur Verfügung stellt und für medizinisches Personal intuitiv verständlich ist. Die Kernkomponenten des vorgeschlagenen Konzepts sind *anatomisch integrierte Visualisierung* und *ikonische Glyphen*.

Die anatomisch integrierte Visualisierung nutzt einen 3D Menschenavatar als räumliche Darstellung von visuell kodierten medizinischen Daten *mit* einem (inhärenten) *anatomischen Bezug*. Diese Komponente schließt eine Reihe von formalen Anforderungen und Verfahren für diese Art von visuellen Kodierungen ein, sowie eine prototypische Implementierung für die Diagnose Bandscheibenvorfall, die von Neurochirurgen evaluiert wurde.

Das VIKK Konzept sieht auch eine Option der Personalisierung des 3D Avatars vor, wofür eine 3D Rekonstruktion des Patientenkörpers aus Tiefendaten benutzt wird. Um dies zu ermöglichen, wurde eine Methode zur robusten Verfolgung der Kameraposen anhand von Tiefendaten mit geringer räumlicher und zeitlicher Auflösung, typisch für mobile Anwendungen, vorgeschlagen. Sie kombiniert eine geometriebasierte Posenschätzung durch den Iterative Closest Point (ICP) Algorithmus mit inertialem Tracking, indem sie ein Erweitertes Kalman-Filter (EKF) anwendet. Insbesondere nutzt sie die Extrapolation der mit ICP geschätzten Posen als virtuelle Messungen und die Ausgabe von EKF als Initialwert für die nächste ICP-Posenschätzung.

Die ikonischen Glyphen ermöglichen Darstellung der Patientendaten *ohne anatomischen Bezug*. Das Ziel dieses Ansatzes ist eine Vereinigung der visuellen Metaphorik von Icons mit dem Potential der Glyphen für multivariate Visualisierung. Technisch gesehen, nutzt er eine parametrische Darstellung, die auf den mit zusätzlichen Freiheitsgraden und Bogenlängenparametrisierung erweiterten Diffusionskurven basiert und eine automatisierte, kontrollierte Manipulation der Geometrie der Iconkonturen sowie der dazugehörenden Farbattribute ermöglicht. Außer dem generischen Konzept wurde eine Implementierung für ein

spezifisches, auf den wellenartigen, periodischen Konturmodifikationen basierendes Design sowie ein Wahrnehmungs- und Quantisierungsmodell für diese Art von visuellen Variablen entwickelt. Die Praktikabilität des Ansatzes wurde durch Beispiele für Visualisierung von Wettervorhersageunsicherheit, Trends in COVID-19-Statistiken und intrakraniellem Druck gezeigt.

Contents

List of Figures

xii

List of Tables

Chapter 1

Introduction

1.1 Problem Statement

The diagnostic and treatment processes in a modern hospital are distributed be-
tween several specialists. Such a distribution happens on different levels and is
caused, for instance, by the expanding interdisciplinary collaboration, allocation of
responsibilities between professional groups such as physicians and nurses, or task
sharing within the same professional group due to shift rotations. To ensure the
patient care under these conditions, an intensive cooperation for transferring of the
collected data from colleague to colleague is required. In particular, in each spe-
cific cooperative workflow scenario a subset of the patient's data that describe his
medical status is crucial for decision making and initiation of the corresponding
procedures by physicians. Some of this information to be transferred is stored in a
systematised digital form as electronic health record (EHR) and thus can be shared
hospital-wide. However, the interface of a typical EHR system reflects an adminis-
trative logic rather than the requirements of collaborative work, which makes it less
suitable for supporting this kind of process. The consequences are a time-consuming
access to relevant data and a low EHR acceptance by physicians [Ema+17]. An-
other part of such diagnostic-cooperative data is not even digitized but is recorded
in several paper-based forms or hand-written scratches and often conveyed to the
colleagues orally.

   In summary, the cooperative tasks in a hospital are characterised by a low data
availability due to 1) heterogeneous and scattered data sources, 2) not coopera-
tion-oriented representation, 3) time-consuming data access, 4) reliance on the par-
tially volatile (i.e., oral) communication. This in turn often leads to a *deficient infor-
mation transfer*, impairing the efficiency of the clinical workflow.

1.2 Contributions

This thesis addresses the aforementioned problem with a concept of Visually Inte-
grated Clinical Cooperation (VICC). At its core is a visual integration of the contex-
tually relevant patient data on a mobile device. Particularly, the concept aims at an
increased data availability by bringing all required information together in one place
and by providing a concise, context-specific overview. It takes advantage of the abil-
ity of human visual perception to quickly recognise decisive features [Wij05], in this
way facilitating the information assessment "at a glance", which is particularly im-
portant by working under time pressure.

   Effectiveness and efficiency of visualization as well as its acceptance by users
strongly depends on its intuitiveness. Taking into account this issue, the proposed

concept builds novel visualization methods upon already familiar visual forms, achieving the target visual integration in two ways:

- as *anatomically integrated in-place visualization*, which exploits the existing visual tradition in the medicine, and

- as *iconic glyphs*, which use the generally comprehensible and intuitive symbolic visual objects.

In the following, this two concepts are discussed in more detail.

Anatomically integrated in-place visualization   The visual tradition in medicine is highly related to the human anatomy. Learning with anatomical atlases, evaluation of the medical imaging results or use of human body sketches for collecting the results of a clinical investigation are only few examples of daily dealings with visual representation of anatomy by physicians. At the same time, many medical data that are usually represented by means of an abstract visualization or as text, are still inherently related to anatomical structures. For instance, a clinical symptom refers to the affected organ or body part. Building upon these observations, the VICC concept uses an anatomical body model as spatial representation of such data. More precisely, the data in question are encoded by visual variables such as colour or texture, which are applied by rendering of the corresponding anatomical structures, changing their default appearance. This enables a *synoptical reading* of information directly 'from the body' – i.e., based on a principle that is familiar for physicians, e.g., from assessment of radiological images, – avoiding abstract and potentially less intuitive forms of representation.

    The benefits of such a linking of data to specific anatomical structures can be even enhanced by relating them visually to a specific patient. This relation can be achieved, e.g., through personalisation of a generic anatomical model, that is making the virtual human body look similar to the respective real person. The VICC concept comprises a solution approach towards a model personalisation using a 3D reconstruction from patient's scans, which are provided by a range camera, integrated in the same mobile device that serves for the visualization. The quality and robustness of a 3D reconstruction relies on the precision of camera tracking, which highly depends on the sensor resolution and computing capacity of CPU and GPU. On a mobile platform, the tracking issue becomes particularly challenging, as the respective hardware resources are rather limited. The proposed concept addresses this problem with a *sensor fusion algorithm*, integrating in the visual reconstruction pipeline pose-related data from inertial sensors.

Iconic glyphs   One of the main challenges in the context of synoptic visualization of patient's data is their commonly multivariate nature. The anatomical approach partially solves this problem by embedding the data in 3D space, that is using location as an additional visual variable. However, it does not cover all potentially problematic cases, e.g. if multiple data refer to the same human body location or for data that do not have any anatomical reference and, thus, the anatomical integration is not applicable. The standard solution for multivariate data in scientific visualization (SciVis) as well as in information visualization (InfoVis) are glyphs, i.e. visual objects that "represent different data variables by a set of visual channels" [Bor+13]. Most of the glyphs are abstract, e.g. build of geometric primitives, which potentially impede an intuitive information reading, making their direct application in the cooperative

clinical settings rather problematic. The glyph's intuitiveness can be enhanced using metaphoric associations, e.g. creating them from symbolic visual objects such as icons, which are related with the problem domain. However, the design of such icon-based glyphs is for the most part a manual process due to not existing control parameters. The VICC concept tackles this problem, proposing a parametric icon representation, based on the *diffusion curve* (DC) [Orz+08] vector primitives, along with the functionalities for controllable, quantifiable contour-based modifications of the base shape and colour in a (semi-)automated fashion.

In summary, this thesis proposes the following main contributions:

- the concept and a prototypical implementation of an anatomically integrated in-place visualization for clinical cooperation support [Pre+23b], comprising

- a robust camera tracking algorithm for 3D scene reconstruction on a mobile device [PLK18], as a component of the model personalization option;

- an approach for (semi-)automated generation of iconic glyphs, including design examples based on the glyph's periodic contour modifications [PBK23] and

- a respective perception and quantization model [PK22],

- as well as a prototype for visualization of intracranial pressure, developed together with neurosurgeons and social scientists applying the iconic glyphs approach [Pre+23a].

The VICC concept, presented in this thesis, was developed in the context of an interdisciplinary project A06, which was a part of the Special Collaborative Research Center (CRC) 1187 "Media of Cooperation" at the University of Siegen. Besides computer scientists, the project involved several neurosurgeons and sociologists. This joined expertise in different fields allowed to identify the communicational shortcomings in the hospital daily routines, to deduce the corresponding requirements, and to ensure the acceptance of the proposed solutions as well as their adequacy from the medical perspective. A high-frequency interaction between the respective groups of experts resulted in a gradual refinement of the initial visualization concept, followed by a prototypical implementation of some of its parts. Finally, the usability of these prototypes has been evaluated by means of user studies

1.3 Overview

Chapter 2 introduces the theoretical foundations, necessary for understanding of the specific approaches discussed in later chapters. Besides an exposition of the generic visualization principles, essential for the entire thesis, it comprises some fundamentals, necessary for the understanding of specific components of the VICC concept. In particular, the principles of scene reconstruction, inertial navigation and Kalman filter as a sensor fusion algorithm are related to the mobile camera pose tracking for the anatomically integrated visualization, whereby such topics as freeform curves and Diffusion Curve Images are relevant for the iconic glyph approach.

Chapter 3 provides an overview of the overall VICC concept, comprising a definition of requirements for the concept development and short description of its single components.

The subsequent chapters discuss specific VICC components in detail. Namely, chapter 4 describes the concept of an anatomically integrated visualization, its prototypical implementation and evaluation. Chapter 5 addresses the personalization option of the aforementioned visualization component, with a particular focus on a novel camera tracking approach for mobile 3D scene reconstruction. In chapter 6, the visualization by iconic glyphs is discussed, comprising a study of their perceptional aspects (6.1) and an approach for their (semi-)automated generation (6.2).

Finally, chapter 7 summarizes and concludes the thesis, delineating also the possible directions of future work.

Chapter 2

Foundations

2.1 Data visualization

2.1.1 Encoding by visual variables

The core of a visualization is the encoding of data that need to be represented by visual variables, i.e. graphic dimensions of a visible mark. The original concept, introduced by Bertin[Ber83], comprised seven visual variables: size, shape, value, colour, orientation, texture and position. Subsequently, some additional variables, e.g. motion [Car03], have been proposed.
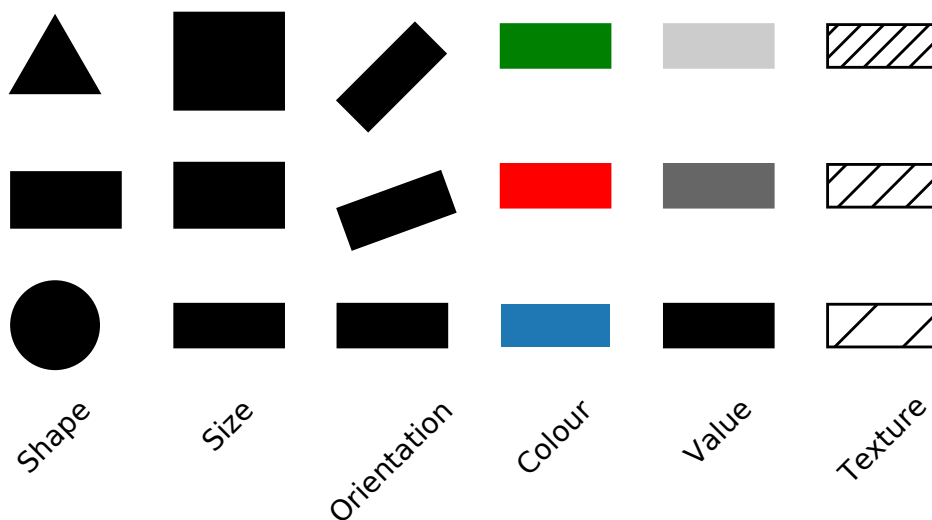
FIGURE 2.1: Bertin's (retinal) visual variables.

Depending on the nature of the visual variable "position", two main families of visualization approaches can be distinguished. Scientific Visualization (SciVis) deals with physically based data, which have a spatial reference, therefore, the respective mark position results from the mapped data variable, e.g. as measurement location. Information Visualization (InfoVis), on the contrary, represents abstract entities, which have no inherent spatial reference and their mapping into the visualization space is a design decision [Car12]. Nevertheless, the above distinction is not rigid, and there are several attempts to combine advantages of both perspectives, proposing hybrid approaches. These can be implemented as a juxtaposition, e.g. InfoVis methods such as scatterplots linked with 3D views of scientific data in the form of multiple coordinated views [Wei+06], or as a more in-deep integration, e.g. a graph, representing metabolic processes, embedded in 3D cell views [KS14].

The visual variables have properties, which determine their suitability to represent certain kinds of data. Following Bertin's definition, a visual variable can be

a) *selective*, i.e. it allows the observer to identify a group of visual signs with correspondences in this visual variables, e.g. a group of blue signs;

b) *associative*, i.e. it facilitates to identify groups of signs that differ in this variable and have correspondences in the others, e.g. different shapes of the same colour;

c) *ordered*, i.e. changes in this variable are perceived as a natural order, e.g. signs of increasing sizes;

d) *quantitative*, i.e. difference in this variable is perceived as a numeric ratio, e.g. different length or areas.

However, there is some controversy in the academic literature regarding the attribution of these properties to specific visual variables. For instance, while Bertin considers *shape* as not ordered, Chung et al. [Chu+16] demonstrate the orderability at least for star-like shapes.

2.1.2   Multidimensional visualization and glyphs

A challenge that a visualization method has to face is representation of *multidimensional* or *multivariate* data sets (the distinction between the both terms in the related literature is rather vague and they are often used as synonyms [Cha06]). A widely used technique to tackle this problem are *glyphs*, i.e. small visual objects, which visual variables encode the variables of the data set. Glyphs are particularly suitable for assessment of interdimensional and inter-record relations; on the other hand, they have their limitations regarding the number of representable data variables and values of each variable, due to their small size [War08].

The main aspects that have to be taken into account for an effective glyph design are described in the following. For an overview of the glyph design guidelines see, e.g., Borgo et al. [Bor+13].

Trade-off between complexity and accuracy   An increase of the number of visualized data variables potentially facilitates the recognition of dependences, patterns or trends but, at the same time, affects the representation accuracy of each single variable. The best compromise between complexity and accuracy depends on the goals of the specific visualization tasks.

Perceptual uniformity   Especially for a quantitative visualization it is important to guarantee that equal distances in data space lead to equal distances in perceptual space. Thus, it has to be taken into account that the relation between physical stimuli and perceived magnitudes of visual variables is in general not linear. In particular, Stevens [Ste57] showed that it follows the power law, i.e. a stimulus-to-perception transformation e is described by a power function

$$e(\mathbf{x}) = a \cdot \mathbf{x}^b, \tag{2.1}$$

where $\mathbf{x}$ is a stimulus.

Perceptual interrelations of visual variables   The visual variables can be classified as *separable* (or analysable), that is changes in one variable are perceived independently from others, e.g. hue and shape [HI72], and *integral* (or unanalysable), that is multiple variables are perceived as whole, e.g. value and chroma [Bur+78]. For separable visual variables, the dissimilarity is represented by city-block metric $d_a + d_b$ and for the integral ones by Euclidean metric $\sqrt{d_a^2 + d_b^2}$ [HI72]. Thus, the separable visual variables are preferable for mapping, in order to enhance the dissimilarity perception. At the same time, there are *augmentation* (reduction) effects in pairs of integral variables, i.e. an increase (decrease) in one variable augments (reduces) the perceived distance in the second one. For instance, a dissimilarity between hues increases by an increasing saturation [KT75]. Such an augmentation can be exploited to enhance the visualization distinguishability, applying a one-to-many mapping, i.e. a redundant encoding of one data variable by several visual variables [War08]. At the same time, if similar mutual effects arise in a case of a one-to-one mapping, i.e. where each visual variable represents a different data variable, a *normalization* of the visual variables against each other is needed. For example, changes in the shape of a glyph can influence the perception of its size and therefore require an appropriate adjustment of the latter [Bor+13].

Importance-based mapping   Visual variables differ in their *pop-out effect*, which describes how fast the respective visual sign can be identified. In the following list, the visual variables are arranged by this property in descending order: colour, size, shape, orientation [Mag+12]. As a consequence for glyph design, more important or informative data variables should be mapped to visual variables with a higher pop-out effect.

Metaphoric associations   The use of metaphoric associations between data and visual variables in encoding design facilitates learning, understanding and memorization of the resulting visualization [Mag+12].

2.2   3D Reconstruction

2.2.1   Range imaging

The term "Range imaging" comprises a number of approaches to capture distances to objects in a real-world environment, relative to the respective sensor pose (position and orientation). Such sensors are often arranged in a 2D matrix form, which results in 2.5D depth maps, i.e. 2D images containing per-pixel distance data.

The optical distance acquisition approaches can be classified in two main categories that use time-of-flight (ToF) or triangulation technique, respectively (for more details see, e.g., [Zan+16; KP15]). The latter makes use of the disparity effect and can be applied with a passive device, like a RGB camera, as well as with an active optical system. By a passive triangulation, the disparity is produced, for instance, by taking successive images with a single camera from varying poses (structure-from-motion) or capturing two simultaneous images by a pair of cameras with a known, fixed baseline (stereo-vision). The active triangulation uses a pair of a sensor and a light projector (illuminator); for instance, in structured light depth cameras the illuminator projects to the scene a light pattern that encodes the spatial directions. In this case, the disparity results form the displacement between the projector and the observing camera.

The ToF distance acquisition is based on the measurement of the time t, which elapses between the emission of a radiance (light/optical signal), its reflection by a point of the scene and its capturing in the sensor, co-positioned with the illuminator. With the speed of light in the air $c \approx 3 \times 10^8 m/s$, the resulting distance can be calculated as $d = \frac{ct}{2}$. An example of devices that apply this principle are LIDAR (Light Detection and Ranging) scanners, which provide point-by-point ranges. The ToF cameras, on the contrary, utilize arrays of illuminator-sensor pairs, allowing this way a simultaneous capturing of distances to multiple points in the scene (in a single frame) as depth maps.

There are different approaches for implementation of a ToF camera; from a practical point of view, however, the most relevant is the continuous wave (CW) amplitude modulation, which is realized in the most commercial devices, e.g. in PMD depth cameras [Zan+16]. The approach uses an active scene illumination with a periodic near infrared (NIR) signal , whereby the distance calculation is based on the fact that the time delay t leads to a measurable phase shift $\Delta\phi$ between the outgoing and incoming signal. In particular,

$$d = \frac{c\Delta\phi}{4\pi f_m},$$

where f_m is the modulation frequency.

2.2.2 Principles of scene reconstruction from range data

Having a 2.5D output of a range camera and its intrinsic parameters, the 3D geometric representation of the imaged scene part can be computed. Accordingly, moving the camera and merging depth information captured from different poses allows for 3D reconstruction of a complex scene. The entire process of scene reconstruction from range data can be considered as a pipeline [RHHL02], whose main structure is represented in Fig. 2.2. In the *depth map preprocessing* step, the incoming range data are filtered for smoothing and outlier removal as well as 3D information is extracted. Next, a *camera pose estimation* is performed to allow the transformation between the current and the common model coordinate system. Finally, the *depth map fusion* updates the so-far accumulated model with the current data.

From the perspective of internal data representation, there are two main families of scene reconstruction approaches. On the one hand, it is the volumetric approach. Introduced by Curless und Levoy [CL96], it has been for the first time adapted for a real-time application by Rusinkiewicz et al. [RHHL02] and serves as a basis for several further adaptation, e.g. as KinectFusion method [New+11; Iza+11]. In the volumetric approach, the reconstructed 3D surface is implicitly stored in a voxel grid, e.g. as level set of the Truncated Signed Distance Function (TSDF). The truncation allows to limit the stored data to a narrow region with reliable measurements around the surface. Facilitating high-quality reconstruction results, the approach, however, implies a certain computational overhead due to the necessity to convert between different representation forms, i.e. implicit volumetric representation and explicit 3D points, which is achieved by rendering using per-pixel raycasting. Furthermore, a grid with fixed resolution and sparse data leads to an inefficient memory usage, i.e. a memory overhead problem, whose solution requires special memory management strategies such as voxel hierarchies (e.g., [SSC14]) or hashing (e.g., [Nie+13]).

The point-based approach [Kel+13], on the other hand, allows for a direct storing of the explicit model as 3D points. This avoids representation conversions and also

provides more flexibility in the memory usage, since only the reconstructed surface and no free space is stored. Besides 3D positions, further point attributes, such as normal, radius, colour or curvature [Lef+17], can be stored. The rendering of the reconstructed surface is done by a splatting technique, similar to [Zwi+01], whereby the point radius is used by splat generation.

In the following, the single pipeline steps are described in more detail, with a special focus on the point-based approach of Keller et al. [Kel+13].
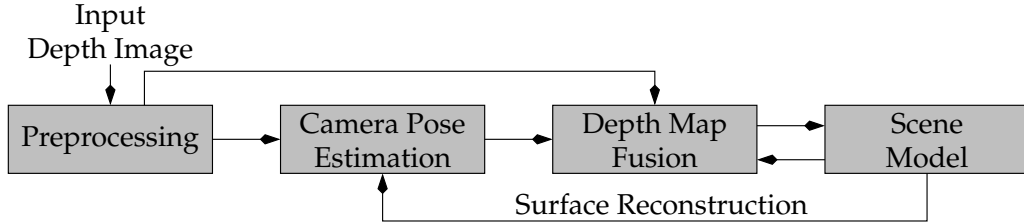
FIGURE 2.2: 3D reconstruction pipeline.

Depth Map Preprocessing: Considering a range camera that takes consecutive frames, the i-th input of the reconstruction pipeline is the depth map $\mathcal{D}_i$. After outlier removal and denoising with bilateral filter (for camera pose estimation), the depth map is further processed for extraction of the 3D point-related information. The scanning of a 3D scene point $(x, y, z)^T$ can be described, using the pinhole camera model, as projection into the image plane

$$z(\mathbf{u}, 1)^T = K(x, y, z)^T \tag{2.2}$$

where K is the intrinsic camera matrix and $\mathbf{u} = (u_x, u_y)^T$ are pixel coordinates, so that the respective depth value, i.e. camera-to-point distance, is stored as $z = \mathcal{D}(\mathbf{u})$. The parameters of the matrix

$$K = \begin{Bmatrix} \frac{f_x}{pix_w} & 0 & x'_c & 0 \\ 0 & \frac{f_y}{pix_h} & y'_c & 0 \\ 0 & 0 & 1 & 0 \end{Bmatrix} \tag{2.3}$$

are the focal length in x and y direction f_x, f_y, the pixel width and height pix_w, pix_h, and the coordinates of the sensor's centre x'_c, y'_c. The standard approach for the calculation of 3D positions from depth maps is the inverse projection [New+11]

$$\mathcal{V}_i(\mathbf{u}) = \mathcal{D}_i(\mathbf{u})K^{-1}(\mathbf{u}^T, 1)^T \tag{2.4}$$

where $\mathcal{V}_i$ is the vertex map, which associates the resulting 3D data with the source pixel. In addition to the positions, other vertex attributes such as normals and radii are calculated and stored in the normal and radius map $\mathcal{N}_i$ and $\mathcal{R}_i$, respectively.

Camera Pose Estimation: The vertex map $\mathcal{V}_i$ created in the first pipeline step contains 3D positions in camera coordinates. However, the fusion of the incoming points with the already reconstructed part of the scene requires a common coordinate system, i.e. the current camera pose in a reference coordinate frame must be known. Such a pose is defined by the transformation matrix

$$T = \begin{Bmatrix} R & & & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{Bmatrix}, \tag{2.5}$$

with the rotational part $R \in \mathbb{R}^{3\times3}$ and translational part $\mathbf{t} \in \mathbb{R}^3$.

A widely used solution for camera pose estimation is the Iterative Closest Point (ICP) algorithm [BM92], which performs a featureless alignment of two (partially) overlapping 3D geometries, iteratively refining the rigid transformation between them. More precisely, it contains the following steps:

1. transform the input geometry to the target coordinate system using the currently estimated transformation,

2. for each transformed input point seek for a corresponding (closest) point in the target geometry, and

3. for the set of corresponding points minimize the determined alignment error function by adjusting the transformation.

The above procedure is iterated until the alignment error is small enough.

In the context of 3D scene reconstruction, ICP is commonly used to align the incoming depth map $\mathcal{D}_i$ to the so-far accumulated scene model, i.e. in a *frame-to-model* approach, estimating the relative transformation from the current to the previous view $T_{i\rightarrow(i-1)}$ [New+11]. Assuming a relatively small camera motion between two consecutive frames, the initialization of this estimate with the identity matrix [New+11] is sufficient. Then, in each iteration k it is updated with an increment as $\tilde{T}_{i\rightarrow(i-1),k} = \tilde{T}_k^{inc}\tilde{T}_{i\rightarrow(i-1),k-1}$. Finally, after the convergence of the algorithm, the current camera pose can be computed from the previous pose and the ICP output as $T_i = T_{i-1}T_{i\rightarrow(i-1)}$.

The pairs of corresponding points can be efficiently found by means of projective data association [BL95; RHHL02]. For this purpose, the accumulated model is represented as a vertex map in the camera coordinates of the last frame $\mathcal{V}_{i-1}^{\mathcal{M}}$. Thus, for a point in the input vertex map at the pixel $\mathbf{u}$, $\mathbf{p} = \mathcal{V}(\mathbf{u})$, the transformation into the previous camera coordinates and projection into the image plain provides the estimated pixel coordinates of its corresponding model point, i.e.

$$\hat{\mathbf{u}} = K\tilde{T}_{i\rightarrow(i-1),l}\mathcal{V}_i(\mathbf{u}) \tag{2.6}$$

with the resulting correspondences $(\mathcal{V}_i(\mathbf{u}), \mathcal{V}_{i-1}^M(\hat{\mathbf{u}}))$.

Commonly, the *point-to-plane* alignment error function [CM92] is used in order to estimate $T_{i\rightarrow(i-1)}$:

$$E(T_{i\rightarrow(i-1)}) = \tag{2.7}$$
$$\sum_{\mathbf{u}\in S} \langle (T_{i\rightarrow(i-1)}\mathcal{V}_i(\mathbf{u}) - \mathcal{V}_{i-1}^{\mathcal{M}}(\hat{\mathbf{u}})) , \mathcal{N}_{i-1}^{\mathcal{M}}(\hat{\mathbf{u}})\rangle^2,$$

where $\mathcal{N}_{i-1}^{\mathcal{M}}$ is the model normal map in the camera coordinates of the frame $i-1$ and S a pixel subset of the incoming vertex map, so that $\mathcal{V}_i(\mathbf{u} \mid_{\mathbf{u}\in S})$ has a valid corresponding model point. This error function measures the distance between the incoming point and the tangential plane of the model at the corresponding model point. In this way, it only constrains one direction, which leads to a faster convergence comparing with the point-to-point metric.

Using a linearised version of Eq. 2.7, which is based on the assumption of a small angle between iterative steps, the minimization problem results in solving a 6×6 linear system.

The successful convergence heavily depends on a good initial guess for the camera pose $\tilde{T}_{i\rightarrow(i-1),0}$: if the algorithm does not start close enough to the actual pose, it can be trapped in a local minimum, which yields significant geometric artefacts (e.g., ghost geometries). In order to cope with faster camera motions still maintaining a simple initialization, the ICP is commonly applied in a hierarchical manner [New+11]. By setting up an image pyramid, coarser version of the underlying matching problem are deduced. Solving this hierarchical optimization in a coarse-to-fine approach leads to a refinement of the camera pose which can overcome local minima to some extend.

Model Update: After camera pose estimation, the currently incoming range data are fused with the so-far accumulated model. The appropriate correspondence search is performed in image space, based on the high-resolution index map, which is created by projection of the model points with a 4×4 supersampling relative to the resolution of $\mathcal{D}_i$. For the established correspondences, the geometric attributes, i.e. position, normal and point radius, are merged into the model as their weighted average. In particular, in order to cope with outliers and isolated points, the model points are weighted on the basis of their confidence attributes, which count the number of merges, i.e. observations. Points are tagged as "unstable" or "stable", if their confidence counter is below or above a given threshold, respectively, and the unstable points are deleted after a specified number of the model update cycles. The incoming points are weighted in accordance to the respective measurement accuracy, assuming that it decreases with increasing distance to the sensor centre.

2.3 Inertial Navigation

The inertial navigation describes tracking of orientation and position relative to a reference coordinate system, based on the measurements of inertial sensors, such as gyroscopes and accelerometers [TW04; GWA07]. For this purpose, multiple inertial sensors are often used as a cluster with a common base, building an Inertial Measurement Unit (IMU). In particular, the modern mobile devices, like smartphones or tablets, are equipped with IMUs, containing a 3-axes gyroscope and accelerometer in the form of a micro-electromechanical system (MEMS) [GP16], allowing this way for a 3D navigation. The section below describes the basic principles concerning a strapdown system, that is an Inertial Navigation System (INS), whose sensors are fixed with respect to device and, thus, provide measurements in the local coordinate system. For further details and regarding different sensor types and other systems see, e.g., [GWA07; Jek12; TW04].

Let be $\mathbf{q}$ the INS orientation in the global coordinate system, i.e. the rotational part of the local-to-global transformation, expressed in quaternion form. In particular, the rotation around a unit vector $\hat{\mathbf{u}}$ by an angle θ yields the corresponding quaternion expression

$$\mathbf{q} = \begin{bmatrix} \cos(\frac{\theta}{2}) \\ u_x \sin(\frac{\theta}{2}) \\ u_y \sin(\frac{\theta}{2}) \\ u_z \sin(\frac{\theta}{2}). \end{bmatrix}$$

Furthermore, let be $\mathbf{s} = (x, y, z)^T$ the INS position and $\mathbf{v} = (v_x, v_y, v_z)^T$ its velocity. Then, the following measurements and computational steps allows for a straightforward tracking implementation.

Orientation   The gyroscope senses the rotational motion around the local axes as angular velocity $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^T$, in radian/s, which relates to the change of orientation in the time, i.e. its time derivatives, namely:

$$\dot{\mathbf{q}} = \frac{1}{2}\mathbf{q} \otimes (0\ \boldsymbol{\omega}^T)^T, \tag{2.8}$$

where the operator $\otimes$ describes quaternion multiplication. Thus, having a known initial orientation, the current orientation can be computed by integration of $\dot{\mathbf{q}}$.

Position   The accelerometer provides acceleration measurements $\mathbf{a}^m = (a_x^m, a_y^m, a_z^m)^T$ in the local coordinates, which also contains the acceleration due to gravitational attraction. The actual device acceleration with respect to the global coordinate system can be derived by transformation of $\mathbf{a}^m$ using the current orientation and subtraction of the gravitational part $\mathbf{g}$, defined in the global coordinates according to the gravitational field model, i.e.

$$\mathbf{a}^g = \mathbf{q} \otimes (0\ \mathbf{a}^{mT})^T \otimes \overline{\mathbf{q}} - \mathbf{g}, \tag{2.9}$$

where $\overline{\mathbf{q}}$ is the conjugate of $\mathbf{q}$. This implies that the tilt of the reference coordinate system has to be determined during the initialization, i.e. the y-axis has to be aligned with the gravity vector, which can be accomplished without external sources as gyrocompass alignment, that is by sensing acceleration with the device in a stationary position [GWA07]. Then, taking into account the following dependencies:

$$\dot{\mathbf{v}} = \mathbf{a},$$
$$\dot{\mathbf{s}} = \mathbf{v},$$

and having a known initial velocity and position, the current position can be calculated by double integration of $\mathbf{a}$.

Integration and error treatment   A widely used solution for integration of pose changes are Runge-Kutta methods. Depending on the tracking range and precision requirements, a low order integration scheme can be sufficient [TW04]; for instance, using the first-order integration (forward Euler) results in the following equation for motion propagation:

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \frac{1}{2}\Delta t \mathbf{q}_k \otimes (0\ \boldsymbol{\omega}_k^T)^T, \tag{2.10}$$

$$\mathbf{v}_{k+1} = \mathbf{v}_k + (\mathbf{a}_k^g - \mathbf{g})\Delta t, \tag{2.11}$$

$$\mathbf{s}_{k+1} = \mathbf{s}_k + \mathbf{v}_k \Delta t, \tag{2.12}$$

Furthermore, under certain assumptions regarding the behaviour of $\boldsymbol{\omega}$, the integration of rotational changes has a close form solution. This can be shown by expressing Eq. 2.8 in matrix form:

$$\dot{\mathbf{q}} = \frac{1}{2}\Omega(\boldsymbol{\omega})\mathbf{q}, \tag{2.13}$$

with

$$\Omega(\boldsymbol{\omega}) = \begin{Bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{Bmatrix}$$

Thus, an assumption that, e.g., $\boldsymbol{\omega}$ is constant over the time interval $\Delta t = t_{k+1} - t_k$, i.e. $\dot{\boldsymbol{\omega}} = 0$, yields the solution of Eq. 2.13 as

$$\mathbf{q}_{k+1} = \exp(\frac{1}{2}\Delta t \Omega(\boldsymbol{\omega_k})) \mathbf{q}_k \tag{2.14}$$

(see, e.g., [TW04; TR05]). Using the Taylor expansion, this can be re-written in quaternion form as

$$\mathbf{q}_{k+1} = \mathbf{q}_k \otimes (\cos(\frac{|\boldsymbol{\omega_k}| \Delta t}{2}) \ \sin(\frac{|\boldsymbol{\omega_k}| \Delta t}{2})(\frac{\boldsymbol{\omega_k}}{|\boldsymbol{\omega_k}|})^T)^T. \tag{2.15}$$

Still, the sensors do not provide exact real values, since their measurements are affected by errors, which have to be taken into account by tracking. The main IMU error sources are bias as well as errors due to scaling and axes misalignment. The bias can be further subdivided into the fixed term, the bias stability and the bias drift, usually modelled as a random walk [GWA07]. For a more detailed error model see also [TW04]. In general, the fixed errors can be measured and compensated accordingly. However, not for all kinds of errors a deterministic solution is possible [TW04]. Some approaches dynamically estimate the respective biases as a part of a filtering method, e.g. an Extended Kalman Filter (EKF) [Sab11].

The accumulation of errors due to noisy sensor measurements leads to a time-dependent drift in INS's [TW04]. Especially the positional tracking is affected by this effect because of the double integration. Thus, for a reliable tracking over a longer time period, it is common to fuse an INS with other navigation sources, e.g. Global Positioning System (GPS) or magnetic sensors. A conventional practice is to combine sources with complementary features, i.e. a high update rate and a sufficient short-term precision, on the one side, and a high long-term stability (by a potentially lower availability), on the other [TW04]. The aim of an integrated navigation system is a mutual error compensation. A well-established solution for such a sensor fusion is, e.g., Kalman filter (see also Sec. 2.4).

2.4 Kalman filter

The Kalman filter is an iterative algorithm for state estimation of a system, described by its dynamics with a known model and unknown perturbations (process noise), and state measurements with unknown errors (measurement noise), whereby both, the process and measurement noise, are assumed to be Gaussian and white. It computes the mean of the state probability distribution, conditioned by measurements, i.e. provides the optimal estimate, minimizing the estimation error in the least-square or maximum likelihood sense. (For more details see, e.g., [May79; May82; GA14].)

The algorithm comprises two main steps (see Fig. 2.3). First, in the *prediction* step, the estimates of the state and uncertainty (the latter given as error covariance matrix) from the last time step are projected to the next one, using the system model and (optional) input parameters. This yields *a priori* estimates (commonly marked with $\cdot^-$ superscript). Then, in the *update* step, they are corrected by incorporation of

the measurement-based information, resulting in *a posteriori* estimates (the common notation is a $\cdot^+$ superscript).
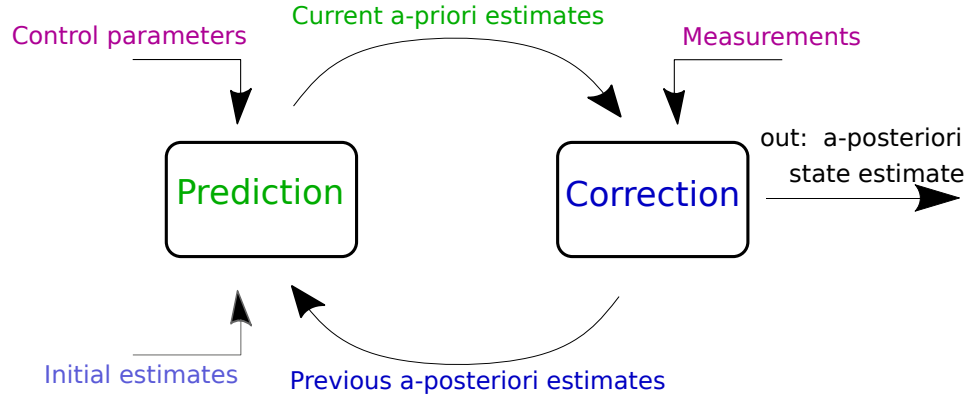
FIGURE 2.3: Kalman filter diagram.

The original Kalman filter addresses the linear system models. To tackle non-linear processes, some algorithm adaptations that use a local linearisation, have been proposed. A well-established non-linear variation, widely used, e.g., in the inertial navigation [GWA07] is Extended Kalman Filter (EKF). The equations that define the corresponding steps of a time-discrete EKF are described in the following.

General definitions   Let $\mathbf{x}_i$ be the vector of the system state variables at the $i - th$ sampling step, $\mathbf{u}_i$ the respective vector of control parameters and $\mathbf{z}_i$ the measurements. Furthermore, let $\mathbf{w}_i$ and $\mathbf{v}_i$ be the process and measurement Gaussian zero-mean noise, respectively. Then, the system *process model* is a function

$$\mathbf{x}_{i+1} = \mathbf{f}(\mathbf{x}_i, \mathbf{u}_i, \mathbf{w}_i), \tag{2.16}$$

which describes the transition to the next state, and the *measurement model* is a function

$$\mathbf{z}_i = \mathbf{h}(\mathbf{x}_i, \mathbf{v}_i), \tag{2.17}$$

which describes the relation between the system state and the measurements. Moreover, let us define the estimation error covariance matrix as

$$P_i = E[(\mathbf{x}_i - \hat{\mathbf{x}}_i)(\mathbf{x}_i - \hat{\mathbf{x}}_i)^T], \tag{2.18}$$

where $\hat{\mathbf{x}}$ is the respective state estimate.

Prediction   Having the a-posteriori estimates $\hat{\mathbf{x}}_k^+$, P_k^+ at the last time step k, their propagation to the next step can be estimated as

$$\hat{\mathbf{x}}_{k+1}^{-} = \mathbf{f}(\hat{\mathbf{x}}_k^{+}, \mathbf{u}_k) \text{ and} \tag{2.19}$$

$$P_{k+1}^{-} = J_f(\hat{\mathbf{x}}_k^{+})P_k^{+}J_f(\hat{\mathbf{x}}_k^{+})^T + Q_k, \tag{2.20}$$

respectively, where

$$J_f(\hat{\mathbf{x}}_k^{+}) = \left. \frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}_k^{+}} \right|_{\mathbf{u}_k} \tag{2.21}$$

is the Jacobian of $\mathbf{f}(\hat{\mathbf{x}}_k^{+})$ and Q_k the process noise covariance matrix.

Update   For the estimation update, first, the Kalman gain is computed as

$$K_{k+1} = P_{k+1}^{-}J_h(\hat{\mathbf{x}}_{k+1}^{-})^T(J_h(\hat{\mathbf{x}}_{k+1}^{-})P_{k+1}^{-}J_h(\hat{\mathbf{x}}_{k+1}^{-})^T + R_{k+1})^{-1}, \tag{2.22}$$

where

$$J_h(\hat{\mathbf{x}}_{k+1}^{-}) = \frac{\partial \mathbf{h}}{\partial \hat{\mathbf{x}}_{k+1}^{-}} \tag{2.23}$$

is the Jacobian of $\mathbf{h}(\hat{\mathbf{x}}_{k+1}^{-})$ and R_{k+1} the measurement noise covariance matrix.

Then, the respective a-posteriori estimates are calculated, using the following update equations:

$$\hat{\mathbf{x}}_{k+1}^{+} = \hat{\mathbf{x}}_{k+1}^{-} + K_{k+1}(\mathbf{z}_{k+1} - h(\hat{\mathbf{x}}_{k+1}^{-})) \tag{2.24}$$

$$P_{k+1}^{+} = (I - K_{k+1}J_h(\hat{\mathbf{x}}_{k+1}^{-}))P_{k+1}^{-}. \tag{2.25}$$

2.5   Curves

2.5.1   Bézier curves

A Bézier curve $\mathbf{b}(u)$ of degree n with parameter $u \in [0, 1]$ is represented as

$$\mathbf{b}(u) = \sum_{i=0}^{n} \mathbf{b}_i B_i^n(u), \tag{2.26}$$

where $\mathbf{b}_i$ is a Bézier control point and

$$B_i^n(u) = \binom{n}{i}(1-u)^{n-i}u^i \tag{2.27}$$

is a Bernstein polynomial of degree n.

Bernstein polynomials have the following important properties [BFK84; PBP02]:

a) partition of unity:

$$\sum_{i=0}^{n} B_i^n(u) \equiv 1 \tag{2.28}$$

b) positivity:

$$B_i^n(u) \geq 0 \text{ for } u \in [0,1] \tag{2.29}$$

c) recursion:

$$B_i^k(u) = (1-u)B_i^{k-1}(u) + uB_{i-1}^{k-1}(u) \tag{2.30}$$

d) endpoint interpolation, since

$$B_i^n(0) = \delta_{i,0} \tag{2.31}$$
$$B_i^n(1) = \delta_{i,n}$$

Moreover, 2.28-2.30 ensure some further characteristics. In particular, from 2.28 results that the Bézier representation is *invariant under affine transformations*. Also, 2.29 guarantees that every curve $\mathbf{b}(u)$ lies in the *convex hull* of its Bézier polygon, formed by $\mathbf{b}_i$. Due to 2.30, a Bézier curve can be constructed recursively, applying the de Casteljau algorithm:

$$\mathbf{b}_i^{k+1} = (1-u)\mathbf{b}_i^k + u\mathbf{b}_{i+1}^k, \tag{2.32}$$

with initialization $\mathbf{b}_i^0 = \mathbf{b}_i$ and the evaluated curve point $\mathbf{b}(u) = \mathbf{b}_0^n$.
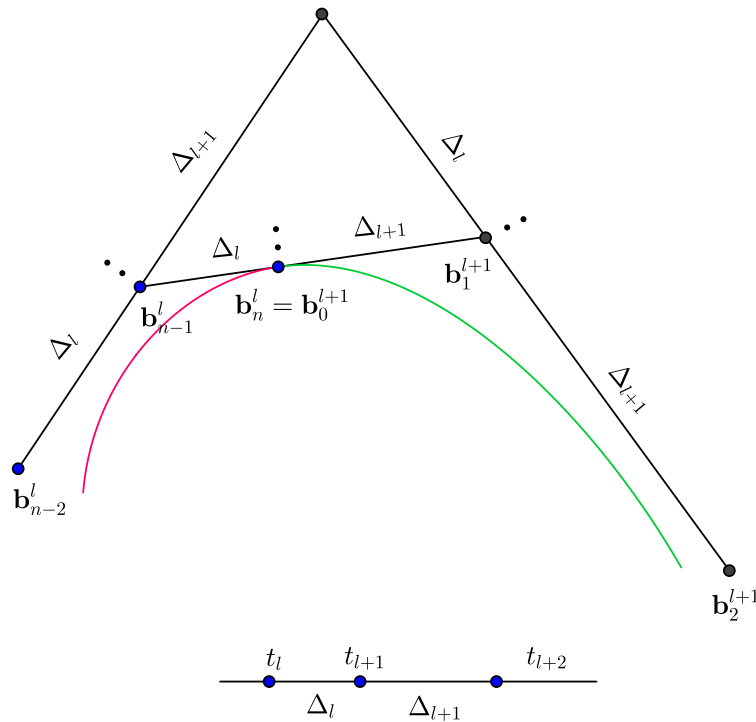
FIGURE 2.4: C^2 joint by A-frame construction.

To increase the degrees of freedom, single Bézier curves can be connected to form a composite curve. Such a composite curve of p segments has a domain, subdivided by $p+1$ knots $t_0 < t_1 < \ldots < t_p$, where the $l-th$ curve $\mathbf{b}^l(u)$ is defined on the respective interval, i.e. with its local parameter $u \in [0,1]$ and the corresponding global parameter $t_l(1-u) + t_{l+1}u = t \in [t_l, t_{l+1}]$. The resulting curve is C^r continuous at the joint points, if the left and right curves have the same derivatives

$(\mathbf{b}^l)^p(t_l) = (\mathbf{b}^{l+1})^p(t_l)$, $p = 0, \ldots, r$. To meet this condition, the respective next control polygon can be constructed from the previous one, using, e.g., Stärk's construction [PBP02]. In particular, the C^2 continuity condition is ensured by the so called A-frame construction (see Fig. 2.4).

The p-th derivative of a Bézier curve $\mathbf{b}^l(t)$ of degree n can be computed as

$$(\mathbf{b}^l)^p(t) = \frac{n!}{(n-p)!(t_{l+1} - t_l)^p} \sum_{i=0}^{n-p} \Delta^p \mathbf{b}_i B_i^{n-p}(u), \tag{2.33}$$

where $\Delta^p \mathbf{b}_i = \Delta^{p-1}\mathbf{b}_{i+1} - \Delta^{p-1}\mathbf{b}_i$ is the p-th forward difference $\mathbf{b}_i$.

2.5.2 B-spline curves

B-spline curves are piecewise polynomial curves with "built-in" continuity constraints. A B-spline curve of degree n

$$\mathbf{s}(u) = \sum_{i=0}^{m} \mathbf{d}_i N_i^n(u), \tag{2.34}$$

where $\mathbf{d}_i$ are de Boor or control points and $N_i^n(u)$ are basis functions, is defined over a knot vector $t_0, \ldots, t_{n+m+1}$, with $t_i \leq t_{i+1}$ and $u \in [t_n, t_{m+1}]$.

B-splines have the following properties [BFK84; PBP02]:

Similarly to Bézier curves:

a) partition of unity:

$$\sum_{i}^{m} N_i^n(u) \equiv 1 \tag{2.35}$$

b) positivity:

$$N_i^n(u) \geq 0 \tag{2.36}$$

c) recursion:

$$N_i^k(u) = (u - t_i)\frac{N_i^{k-1}(u)}{t_{i+k} - t_i} + (t_{i+k+1} - u)\frac{N_{i+1}^{k-1}(u)}{t_{k+i+1} - t_{i+1}} \tag{2.37}$$

$$N_i^0(u) = \begin{cases} 1, & u \in [t_i, t_{i+1}[\\ 0, & else \end{cases} \tag{2.38}$$

Additionally:

e) local support:

$$N_i^n(u) = 0 \text{ if } u \notin [t_i, t_{i+n+1}], \text{ i.e. supp } N_i^n =]t_i, t_{i+n+1}[\tag{2.39}$$

Thus, analogously to Bézier curves, it can be deduced that a B-spline curve is *affine invariant* and its de Boor points form its *convex hull*, more precisely, any curve point $\mathbf{s}(u)$ lies in the hull of those de Boor points whose corresponding basis function $N_i^n(u) \neq 0$ [BFK84].

Based on 2.38, the de Boor algorithm describes the recursive construction of a B-spline as follows: a curve point $\mathbf{s}(u)$ with $u \in [t_l, t_{l+1}[$, i.e. $N_i^n(u) \neq 0$ only for $i \in \{l-n, \dots, l\}$, can be evaluated, recursively applying

$$\mathbf{d}_i^k = (1 - \alpha_i^k)\mathbf{d}_{i-1}^{k-1} + \alpha_i^k\mathbf{d}_i^{k-1}, \qquad\qquad \alpha_i^k = \frac{u - t_i}{t_{i+n+1-k} - t_i}, \qquad (2.40)$$

with $\mathbf{d}_i^0 = \mathbf{d}_i$ and $\mathbf{d}_n^n = \mathbf{s}(u)$. This is a generalization of the de Casteljau algorithm [PBP02].

The degree of freedom of a B-spline curve can be increased by knot insertion, without changing its shape. For instance, according to Böhm's algorithm [Boe80], the insertion of a r-fold knot $t^* \in [t_l, t_{l+1}[$ leads to the replacement of $n-1$ old control points by $n-1+r$ new ones, which result from the recursive de Boor scheme. The extended control point sequence is

$$\dots, \mathbf{d}_{l-n}^0, \mathbf{d}_{l-n+1}^1, \dots, \mathbf{d}_l^r, \dots, \mathbf{d}_l^1, \mathbf{d}_l^0, \dots \qquad (2.41)$$

with $\mathbf{d}_{l-n}^0 = \mathbf{d}_{l-n}$ and $\mathbf{d}_l^0 = \mathbf{d}_l$. For other insertion approaches, such as the Oslo algorithm, see, e.g., [PBP02]. Multiple knots reduce the curve continuity, i.e. it is $n - r$ times continuous at any r-fold knot, with $r \leq n+1$ to guarantee a non-vanished support (see also Fig. 2.5). Increasing the multiplicity of all knots to n produces the Bézier representation of the curve, i.e. the de Boor points form the Bézier polygon and the only not vanished B-Spline bases are Bernstein polynomial. Thus, B-splines are a generalization of the Bézier representation.
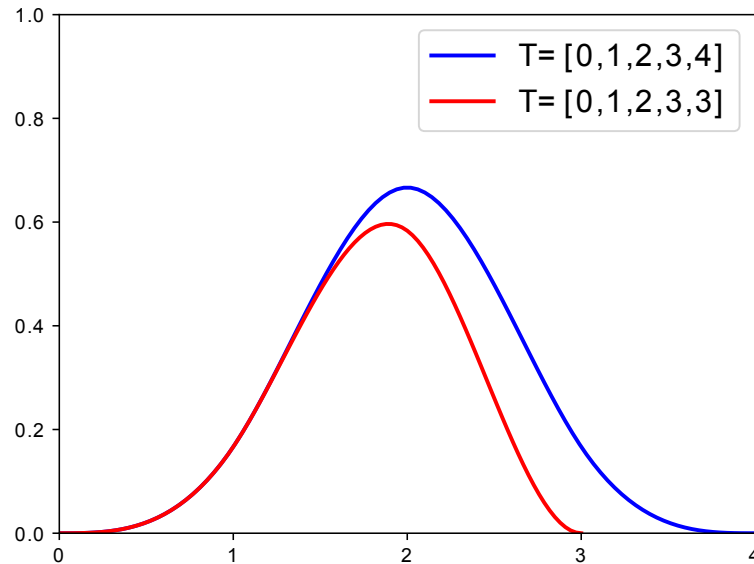
FIGURE 2.5: Cubic B-spline N_0^3: reduced support by a 2-fold knot (red).

The derivative of a B-Spline curve $\mathbf{s}$ of degree n is a B-Spline curve of degree $n-1$

$$\mathbf{s}'(u) = \sum_{i=1}^m \frac{n\nabla\mathbf{d}_i}{t_{i+n} - t_i} N_i^{n-1}(u), \qquad (2.42)$$

where $\nabla \mathbf{d}_i = \mathbf{d}_i - \mathbf{d}_{i-1}$ is the first backward difference. To compute higher derivatives this formula has to be applied recursively.

2.5.3 B-spline interpolation and approximation

Let $\mathbf{s}$ be the interpolating curve, $\mathbf{p}_0, \ldots, \mathbf{p}_k$ the points that have to be interpolated by this curve, and $u_0, \ldots, u_k$ the corresponding parameters, i.e. interpolation nodes, so that $\mathbf{p}_i = \mathbf{s}(u_i)$. Then, the curve de Boor points can be calculated as solution of the following linear system:

$$
\begin{bmatrix} N_0^n(u_0) & \ldots & N_m^n(u_0) \\ \vdots & \ddots & \vdots \\ N_0^n(u_k) & \ldots & N_m^n(u_k) \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \vdots \\ \mathbf{d}_m \end{bmatrix} = \begin{bmatrix} \mathbf{p}_0 \\ \vdots \\ \mathbf{p}_k \end{bmatrix} \tag{2.43}
$$

or, in an abbreviated form, $ND = P$. The matrix N is invertible if $k = m$ and $\forall i : N_i^n(u_i) \neq 0$ [PBP02], i.e. no more than $n + 1$ conditions per a knot vector segment can be specified.

The interpolation conditions can also include derivatives for a better shape control, e.g. tangents as end conditions. That is, more generally, the element of N is $n_{i,j} = N_j^n(u_i)^{(r)}$, where $r = 0, \ldots, n - 1$ describes the derivative order.

The interpolation results, especially the curve smoothness, are determined to a large extent by the choice of the interpolation nodes. There are different choice strategies, depending, e.g., on the curve degree or order. In particular, for a B-spline curve with an even order $h = n + 1$ and the parameter $q = h/2$, the interpolation at knots provides smooth results [Boo78], i.e. the nodes $\{\mathbf{u}_i\}_{i=0}^k \in [a, b]$ are placed to satisfy the following conditions:

$$
t_{h-1} = a \leq u_0 < \ldots < u_{q-1} < t_h,
$$
$$
u_{q+j} = t_{h+j} \text{ for } j = 0, \ldots, k - h,
$$
$$
t_k < u_{k-q+1} < \ldots < u_k \leq b = t_{k+1}
$$

Moreover, it is good practice to take into account the geometry of $\{\mathbf{p}_i\}_{i=0}^k$, using the arc-length parametrization [Boo78; BFK84]. The arc length of a curve segment corresponding to the interval $[0, u]$ is given by the equation

$$
l(u) = \int_0^u \|\mathbf{s}'\| dt. \tag{2.44}
$$

Thus, a B-spline curve $\mathbf{s}$ can be parametrized with respect to its arc length as $\mathbf{s}(u(l))$, where $u(l)$ is the inverse of $l(u)$. Such a parametrization has a useful property that the curve's first derivative with respect to l has unit Euclidean length, i.e.

$$
\|\mathbf{s}'\| = 1, \tag{2.45}
$$
$$
\mathbf{s}' = \frac{d}{dl}\mathbf{s}(u(l)).
$$

For applications, a chord-length approximation

$$
L = \sum_i \|\Delta \mathbf{p}_i\|, \tag{2.46}
$$

where $\mathbf{p}_i = \mathbf{s}(u_i)$ and $\Delta\mathbf{p}_i = \mathbf{p}_{i+1} - \mathbf{p}_i$, is commonly used instead [Far02]. It sufficiently precise approximates the arc length by choosing a small parameter step, as $L \to l$ if $\Delta u \to 0$.

In some cases, for more shape control, in specific, to minimize the curve oscillation, it may be necessary to set more constraints than available degrees of freedom, i.e. $k > m$, which results in an overdetermined linear system 2.43. The solution can be approximated by the least-square method as

$$N^T N D = N^T P. \tag{2.47}$$

Besides curve construction, the same approach can also be applied to produce local changes to an existing curve. In particular, Fowler and Bartels [FB93] use the least-square approximation to solve a constrained-based direct curve manipulation problem. The constraints ΔP result from positional or derivative changes, applied to some selected curve points, and the curve is accordingly modified, computing offsets $\Delta\tilde{D}$ for those de Boor points that contribute to the evaluation of the affected points, i.e.

$$\Delta\tilde{D}^T = N^T (NN^T)^{-1} \Delta P^T, \tag{2.48}$$

Alternatively to the direct solution, an iterative least-square approximation is used to avoid a possible numerical instability [He+15]. In case specific shape properties are required, the respective loss function L is extended by an energy term, i.e.

$$L(D) = \sum_{i=0}^{k} (\mathbf{s}(u_i) - \mathbf{p}_i)^2 + \omega E,$$

where ω is a constant weight. The curve energy can comprise several components, e.g. stretching and bending energy E_s and E_b [VW95]:

$$E_s(\mathbf{s}) = \int \|\mathbf{s}'(u)\| \, du$$

$$E_b(\mathbf{s}) = \int \kappa^2(u) \|\mathbf{s}'(u)\| \, du,$$

where κ is the curve curvature. In practice, the following approximations are used to reduce computational cost:

$$\hat{E}_s(\mathbf{s}) = \int \|\mathbf{s}'(u)\|^2 \, du$$

$$\hat{E}_b(\mathbf{s}) = \int \|\mathbf{s}''(u)\|^2 \, du,$$

whereby $\int \kappa^2(u)\|\mathbf{s}'(u)\| \, du = \int \|\mathbf{s}''(u)\|^2 \, du$ for an arc-length curve parametrization (see also Seq. 2.5.4). For instance, He et al. [He+15] proposed an Iterative Approximation for Least Squares, incorporating the stretching Energy (ELSPIA), which minimizes the loss function with the steepest descent method.

2.5.4 Local curve properties and Frenet frame

Local properties of a curve $\mathbf{s}$ in a point $\mathbf{p} = \mathbf{s}(u)$ can be described by means of the Frenet frame, which is an orthonormal coordinate system with origin at $\mathbf{p}$. The Frenet frame can be constructed from the derivatives of $\mathbf{s}(u)$, using the Gram-Schmidt orthonormalization and assuming for $\mathbb{R}^d$ that the first d derivatives are linearly independent. In particular, for $\mathbb{R}^3$ the axes of the Frenet frame are *tangent*

vector **t**, *normal* vector **n** and *binormal* vector **b**, which are calculated as

$$\mathbf{t} = \frac{\mathbf{s}'}{\|\mathbf{s}'\|}, \quad \mathbf{b} = \frac{\mathbf{s}' \times \mathbf{s}''}{\|\mathbf{s}' \times \mathbf{s}''\|}, \quad \mathbf{n} = \mathbf{b} \times \mathbf{t}, \tag{2.49}$$

respectively (the parameter u is omitted for simplicity).

The variation of the Frenet frame by moving along a curve is related to the curve's behaviour regarding its smoothness. This relation is described by the Frenet-Serret formulae:

$$\begin{aligned} \mathbf{t}' &= & \kappa\mathbf{n} & \\ \mathbf{n}' &= -\kappa\mathbf{t} & +\tau\mathbf{b} & \\ \mathbf{b}' &= & -\tau\mathbf{n} & , \end{aligned} \tag{2.50}$$

where κ is curvature and τ is torsion. The use of arc-length parametrization yields some simplifications in the corresponding calculations, namely

$$\mathbf{t} = \mathbf{s}',$$

with the resulting property that the curve tracing speed is always 1 (see also Eq. 2.45),

$$\kappa = \|\mathbf{s}''\|$$

and

$$\tau = \frac{1}{\kappa^2} \det \begin{bmatrix} \mathbf{s}' & \mathbf{s}'' & \mathbf{s}''' \end{bmatrix}.$$

More details can be found in [Far02; PBP02].

2.6 Diffusion Curves

In raster graphics a 2D picture is represented as a rectangular grid, whose cells are denominated as pixels. This essentially corresponds to the graphic representation in the image acquisition hardware (e.g., sensor of a RGB camera) as well as in the image processing (e.g., a graphic card) and display (e.g., a computer monitor) devices, which facilitates a wide usage of raster graphics. A further important graphic representation are vector graphics. Compared to raster graphics, this form offers such benefits as resolution independency, compact representation and a straightforward geometry manipulation. These properties make it especially suitable for design and artistic work. However, its most significant drawback is the limitation in representing arbitrary colour gradients.

To combine the benefits of both of these graphic representations, Orzan et al. [Orz+08] introduced *Diffusion Curves* (DC). A DC stores the geometry, the colour and the blur of a picture region in parametric form, and a set of DCs describes the respective entire picture as a *Diffusion Curve Image* (DCI). Mathematically, a DC is a spline of K consecutive cubic Bézier curves, defined via Bézier control points $\{\mathbf{p}_i\}_{i=0}^{4 \cdot K} \subset \mathbb{R}^2$, in conjunction with *colour parameters*,

i.e. colour values and blur attributes located at parametric position u along the respective Bézier curves and defined separately for their left and right side: $\left\{\mathbf{c}_i^l(u)\right\}_{i=0}^M$, $\left\{\mathbf{c}_i^r(u)\right\}_{i=0}^N$, $\left\{\mathbf{b}_i^l(u)\right\}_{i=0}^O$, $\left\{\mathbf{b}_i^r(u)\right\}_{i=0}^P$, $u \in [0, K]$. This allows for a controlled and independent manipulation of the image geometry and colour.

For rendering, the colour and blur values are linearly interpolated along the respective curve. Furthermore, a gradient field $\mathbf{w}$ to control the colour diffusion is calculated as $w_{x,y} = (\mathbf{cl} - \mathbf{cr})\mathbf{n}_{x,y}$, where $\mathbf{cl}$, $\mathbf{cr}$ and $\mathbf{n}_{x,y}$ are, respectively, the left and right colour and the curve normal at the given position. Then, the colour parameters and the gradient filed are rasterized. In the original approach, the colour is diffused by solving the Poisson equation

$$\Delta I = \operatorname{div} \mathbf{w},$$

where $I(x, y) = C(x, y)$ if (x, y) stores a colour. Subsequently, several adaptations have been proposed in order to enhance the rendering precision and performance (e.g., [JCW09; Jes16]).

Chapter 3

The Concept of Visually Integrated Clinical Cooperation

This chapter provides an overview of the concept of Visually Integrated Clinical Cooperation (VICC), comprising a definition of goals and a description of its main components.

3.1   Definition of Goals

The goal definition builds upon interdisciplinary collaboration in the context of the project A06, in the Special Collaborative Research Center 1187 at the University of Siegen. Having the Neurosurgery Department of the Jung-Stilling Hospital in Siegen as a cooperation partner, allowed for the routine work observations on the ward and the subsequent interviews with neurosurgeons. This field study and the first evaluation of its results were mainly performed by the social science project members: Julia Kurz, Judith Willkomm and Cornelius Schubert. Then, the respective findings served as basis for the goals to achieve designing the VICC concept, specified and discussed during recurrent interdisciplinary meetings.

First of all, the field study identified the main settings, in which a visual support would have the most potential to enhance the cooperation between healthcare workers. In particular, it is *the transfer of the most relevant patient's diagnostic data* to colleagues, e.g. during a shift handover. Currently, this information is passed on mostly in verbal form and is documented in scattered, heterogeneous sources, which makes it hard to retrieve. The main findings regarding this cooperation settings are summarized below.

- **Conservatism of the medical personnel.** Due to a high responsibility and potentially severe consequences of errors in their profession, the physicians adhere to the proven traditions in medicine and have a rather conservative attitude to innovations.

- **Time pressure.** All routines on the ward are subject to strict timing constraints, which implies that the information transfer takes place in very restricted time slots.

- **Focus on abnormality.** The physicians are primarily interested in the data that indicate pathological changes and thus are decisive for diagnosis and treatment.

- **Crucial role of assessment of the patient's status.** Getting the patient-related information from colleagues, the neurosurgeons are mainly interested in a specific subset of diagnostic data that, considered *as whole*, define the current neurological *patient's status* and are relevant for determination of further diagnostic or therapeutic measures.

- **Temporal changes and trends as deciding factors.** In particular, the main attention is paid to potential changes and trends in the patient's status as criteria for decision-making.

An interdisciplinary analysis and discussion of these findings allowed to deduce the design goals for the VICC concept, which are listed in the following.

G1 Intuitive comprehensibility. Due to a particular time pressure mentioned above, a substantial requirement to facilitate the acceptance of a visual support is a notable speed-up in information transfer in comparison with the current situation. Furthermore, the physicians have no possibility to invest a lot of time in the learning of new tools. Therefore, the goal is an intuitively comprehensible visualization, relying on the existing tradition.

G2 Multivariate visualization. Most of the patient's data are multivariate, which requires an appropriate visual representation.

G3 Integrated visualization. The crucial role of the patient's status, i.e. a synopsis of multiple parameters, in the information flow implies the necessity of an integrated visualization.

G4 Relevance-dependent filtering. Since the patient's status comprises a subset of all patient-related data that is relevant in the given context, an appropriate pre-filtering is required.

G5 Visualization of chronological changes. Because assessment of the disease progression is decisive for determination of an appropriate treatment, the visualization has to provide an option for representation of chronological changes in a patient's status.

3.2 Main Components

The following section describes the concept components as well as their interrelation. An overview is provided in Fig. 3.1.

3.2.1 Visualization

The concept aims at providing a visual support for information transfer in the context of clinical cooperation, in consideration of the design goals defined above. Accordingly, it focuses on an intuitive synoptical visualization of the patient's status. The major challenges, arising within this scope, i.e. visual integration (**G3**) and representation of multivariate data (**G2**), are addressed by the concept's core components: anatomically integrated visualization and iconic glyphs.

The anatomically integrated visualization uses a 3D model of a human body for a spacial integration (**G3**) of medical data with an (inherent) anatomical reference. In particular, the 3D object that represents the anatomical structure the data relate to,
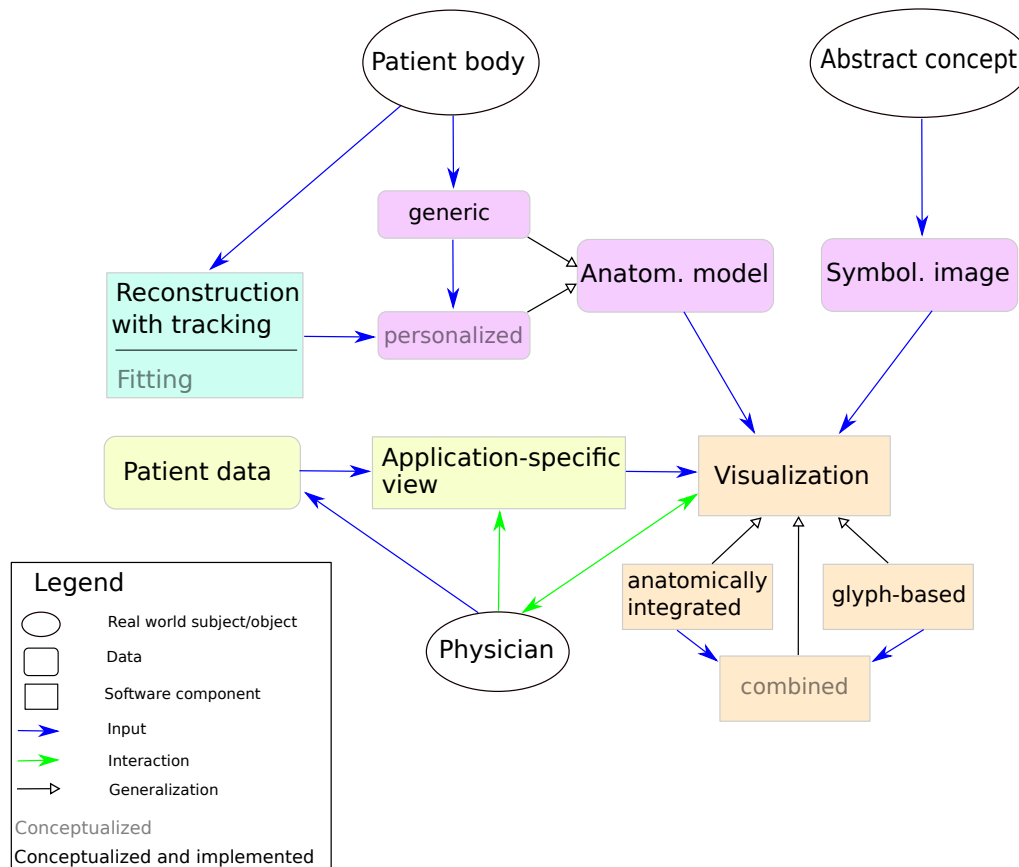
FIGURE 3.1: The main components of the VICC concept.

serves as 3D location for visualization. The latter is created by encoding of data variables by visual variables of the object, such as colour or texture, which are applied during the model rendering. This direct visualization in the anatomical space facilitates a "CT-like reading" of the represented information, familiar to the physicians, and, therefore, a more intuitive visualization (**G1**) of abstract data like symptoms.

Besides, this approach supports multivariate visualization (**G2**), allowing simultaneous representation of multiple data records that refer to distinct anatomical structures and a moderate number of data variables relating to same anatomical structure, encoding them with different visual variables, i.e. hue, saturation, value and texture.

The representation of temporal changes inside an anatomically integrated visualization is realised as scrolling in a time sequence of patient's statuses (**G5**).

The realization of this VICC component is described in Chapter 4, based on [Pre+23b], and comprises the following main contributions:

- definition of the corresponding visualization concept including requirements and visual encoding procedures;
- prototypical concept implementation for the diagnosis spinal disc herniation;
- prototype evaluation by physicians.

The iconic glyph visualization approach allows for a (semi-)automated creation of glyphs from icons, providing a parametric control over appearance and combining this way the intuitiveness of an icon, related to data topic (**G1**), and the glyph's capability of the multivariate data visualization (**G2**). Moreover, it augments the

standard glyph's visual variables such as size or colour with the contour-based visual channels, which are particularly suitable for encoding of gradient-like data, e.g. temporal changes in the respective variable (G5).

The implementation of this VICC component is described in Chapter 6 and is based on [PK22; PBK23; Pre+23a]. It can be subdivided in two main parts. First, a perception and quantization model for periodic contour modifications, including

- an online user study about perception of these kinds of modifications, whose results served as a basis for
- modelling of a corresponding stimulus-to-perception transformation function and
- definition of distinguishable quantization levels.

Second, an approach for controllable geometric and colour modifications of an icon, comprising

- a diffusion-curve based concept of parametric, contour-based modifications;
- the implementation of this concept with a focus on the automated, wave-like modifications;
- application examples along with an online user study to evaluate the usability of the approach.

Conceptually, the anatomical approach and the iconic glyphs can be used complementary as well as in an combined form. In the former case, the glyphs can be utilized to augment the visualization with patient's data without anatomical reference, which are not covered by the anatomical approach. For a combined use, an iconic glyph can serve as quantifiable texture, mapped to the respective anatomical structure, increasing the capability and controllability of the texture-based data encoding. The implementation of such a combined visualization is not part of this thesis.

3.2.2   Data

The VICC concept comprises two kinds of data. On the one hand, there are medical data to be visualized and in this way conveyed to the physicians. For this purpose, the individual data variables need to be visually encoded. On the other hand, there are visual objects that serve as a basis for visualization and whose visual variables are used for encoding of medical data.

Patients' data are the input of visual encoding. In the first line, these are diagnostic data, which play a key role in the information transfer on the ward and can be available in digital, paper-based or even oral form.

Anatomical model is a hierarchical collection of 3D objects that represent organs and other structures of human body. It serves as a basis for the anatomically integrated visualization.

Icons, i.e. rather minimalistic, symbolic images, function as source for glyph creation.

3.2.3   Data processing

The visualization of patient's data according to the principles described above, requires some preprocessing steps.

An application-specific view of a patient's record (see Secs. 4.3.2.1 and 4.4.2.1) is aimed at pre-selecting of data that are relevant in the given context, avoiding an information overload, accordingly to **G4**. It serves as an intermediate data structure, optimized for a visual representation without altering the underlying information, which implies an interdisciplinary categorization and hierarchization work, involving the domain experts.

This component has been implemented in the context of the anatomically integrated visualization (see Chapters 4.3.2.1, 4.4.2.1). The implementation comprises

- formal definitions for organization of medical data that have to be visualized as well as
- data categories for clinical neurological symptoms related to the diagnosis spinal disc herniation, determined in collaboration with neurosurgeons.

Personalization of the 3D body model, i.e. adaption of the generic anatomical model to the appearance of a specific person can be optionally performed to strengthen the mental linking between the visualized data and the respective patient. This aims to support a more intuitive and faster data reading and understanding (**G1**).

The main steps of the personalization process are a) reconstruction of a 3D model from (partial) patient's scans, registered with a mobile range camera, and b) fitting of the generic model to this (potentially incomplete) reconstruction, to achieve a complete and personalized anatomical model. The latter step is not implemented in the presented thesis. The implementation of the 3D reconstruction step is described in Chapter 5, based on [PLK18], and comprises the following contributions:

- an adaptation of the point-based framework, as proposed by Keller et al. [Kel+13], for mobile platforms;
- a robust, multi-sensor tracking approach for initialization of the ICP-based registration, which combines inertial tracking and extrapolated ICP pose estimations using an EKF;
- demonstration of the approach suitability due to an enhanced tracking robustness.

Chapter 4

Anatomically Integrated In-Place Visualization

This chapter describes the components of the VICC concept that are related to the anatomically integrated visualization (see Sec. 3.2). Having the main focus on the respective visualization concept and its implementation, it also addresses the creation of an application-specific view for patient's data with anatomical reference (see Secs. 4.3.2.1 and 4.4.2.1). The presented approach has been published in [Pre+23b].

4.1 Introduction and Prior Work

Patient-related medical data are stored and maintained in order to support different diagnostic and treatment processes in a hospital and are shared within and between professional groups, such as physicians, nurses and sometimes technicians. Here, the increasing complexity of these processes due to, for instance, new diagnostic methods and the expanding interdisciplinary as well as inter-hospital collaboration, is amplified by the challenges in organizing clinical work due to shift rotations and understaffing. Therefore, the timely access of the medical personnel to the currently relevant patient data is crucial for treatment quality. This chapter presents a visualization approach to support clinical cooperation by building on the visual tradition in medicine. Compared to prior work, the proposed approach focuses on the specific requirements of such cooperative tasks, that is the need of a synoptic-like representation of the patient's current status for a fast information transfer. In particular, it allows to visualize *multiple* clinical parameters in an intuitive, *anatomically integrated* fashion. In the following, the main contributions and the delimitation to other methods are described in more detail.

Various visualization approaches have been proposed for intuitively accessing patient's medical data in order to improve clinical support and quality assurance [WBH14]. In general, one can distinguish two main categories, depending on the aim of such visualizations. The first category is characterized by the exploratory nature of the approaches and commonly refers to electronic health records (EHRs), that is, the visualization provides an overview of a patient's complete set of medical and potentially administrative data and offers functionality for data selection including refinement, reconfiguration of layout and visual data encoding, and detection of data correlations or outliers [Rin+13]. The second category comprises task-oriented visualization approaches that address medical tasks with specific requirements and with a focus on efficient information access and transfer. Examples are, for instance, communication to patient of the prostate cancer health risk [Hak+17] or reading

of imaging diagnostic records [Shi+18]. For the purpose of such visualization systems, a relatively small subset of patient data, in some cases not documented in the EHR [Jaa+19], is relevant.

Visualization tools for patient data predominantly utilize concepts from information visualization, for example, encoding time-related medical data as colourized geometric primitives that are placed at a timeline [Pla+98; BAK07; Cra11]. Some newer works intend to enhance the comprehensibility and navigability of the timeline-based EHR visualization, using it in combination with a data abstraction layer [Mar+20] or a generic data model such as the Observational Medical Outcomes Partnership Common Data Model [Gli+19]. Still, the physicians are in general not very familiar with abstract visual representations, which can impair their understanding [MS13]. Therefore, some visualization approaches additionally utilize a virtual human body in order to provide a more intuitive data access. For instance, Kirby and Rector [KR96] presented one of the first systems that uses a visualization of the human body as the central entry point for an efficient access of medical data items that are presented as text. Sundvall et al. [Sun+07] presented a prototype of a 2D visualization framework for patient's medical data based on Google Earth. It supports so-called placemarks, the standard Google Earth approach, to position body-referenced information as icons and text. An et al. [An+10] developed a 3D navigation and visualization method for medical data that uses different levels of detail (LOD) in combination with different visualization strategies applied to data at different levels. Particularly, on the two topmost levels a virtual human body is used to refer to affected organ systems or organs, which serve as access points to respective disease descriptions. Shi et al. [Shi+18] proposed a visualization tool that uses a 3D anatomical human model in combination with a three state colour encoding of pathological severity in order to facilitate for radiologists the access to the current and historical medical status of a patient. The user can access the original data, such as DICOM images or imaging diagnostic reports, by selecting the highlighted organs. Jaatun et al. [Jaa+19] developed a tablet-based digital pain body map that allowed patients with cancer to visually express their pain by drawing its location on a 2D virtual body and applying a colour encoded scale for its intensity. The tool is aimed at an improvement of the communication between patients and medical personnel and supports the information exchange between different physicians and nurses during a patient handover. The existing anatomy-based approaches have in common the focus on the on-body visualization of few, mainly a single parameter, while further data are presented using abstract visualization methods or text.

The overall goal of the interdisciplinary research project this thesis is based on, is the analysis of cooperative clinical work practices on a neurosurgical ward and the development, implementation and evaluation of appropriate visualization concepts to improve these practices by providing efficient and task-oriented access to the health status, i.e. the relevant medical data, of a patient. The focus is on pre- and post-operative medical tasks that are distributed among several doctors and nurses due to shift rotations and occasional understaffing and therefore rely on their cooperation in the form of circulation of patient's status. As this cooperation takes place in the strict time slots of the main medical tasks, e.g. ward rounds, it is characterized by the particular time pressure under which the patient's status must be conveyed from one physician or nurse to the next. In the initial analysis, the field studies, mainly conducted by the social scientists (see also Sec. 3.1), revealed major findings that are relevant for the target visualization concept (further details are given in Sec. 4.2.1). (1) Only a specific and relatively small subset of data in a concise

and objectified form has to be conveyed, depending on the usage context. (2) A significant part of such subsets are diagnostic data related to the anatomy, e.g. clinical symptoms. (3) Neurosurgical personnel strongly rely on their visual tradition and can therefore easily get used to anatomical visualization metaphors.

These findings about the cooperative clinical work practices on a neurosurgical ward served as a basis for developing of a new visualization approach, presented in this chapter, which focuses on transferring diagnostic medical data with anatomical reference. Inspired by the prior work that partially utilizes a virtual human body for visualizing medical data [KR96; Sun+07; An+10; Shi+18; Jaa+19] and rested on the positive feedback from the neurosurgical personnel regarding anatomy related visualization techniques, it exploits the visual tradition in medicine as much as possible in order to achieve an *anatomically integrated in-place visualization*. This means that several, task specific data are encoded with appropriate visual variables, which are then applied by displaying the respective anatomical structures of a 3D avatar.

Following the basic visualization principle of the previous body-based solutions, the present anatomically integrated visualization approach, however, especially focuses on providing a *visual synopsis* of a patient's task-specific status, i.e. a consistent, simultaneous representation of multiple key data. As observed during the aforementioned field studies and stressed in the interviews by the project's medical partners, in the considered cooperative constellations only relatively small subsets of patient diagnostic data are relevant for the physicians (see the major finding 1 above): this fact is exploited here to minimize the necessity of manual online data selection and filtering. Still, the challenging task of mapping multiple quantified medical data to visual variables in a way that they can be effectively perceived in a simultaneous visualization has to be addressed appropriately. A further challenge in this context is that the visual variable "spatial location" cannot be freely selected, being pre-defined by the anatomy.

The main contributions of the proposed visualization approach are as follows:

1. The concept of an *anatomically integrated in-place visualization* of medical data with anatomical reference including the *definition of requirements and visual encoding procedures* for this kind of visualization.

2. The *design goals* and *data categories* for clinical neurological symptoms related to the differential diagnosis for patients with spinal disc herniation. This serves as case study for applying the visualization concept and the formal requirements.

3. A *prototypical implementation* of the visualization concept and its *evaluation* by ten neurosurgeons.

The evaluation shows a strong consensus among the assessing physicians about the usefulness of applying the proposed visualization concept to the diagnostic data of patients with spinal disc herniation.

It should be emphasized that this concept is designed for visualization of a specific kind of medical data that refer to anatomical structures. Therefore, for a more advanced use it may need to be combined with iconic glyphs (see Sec. 6) or other information visualization techniques to incorporate medical data without anatomical reference.

4.2 Design Method and Goals

The interdisciplinary long-term project on a neurosurgical ward that this thesis is a part of, investigated fundamental questions of analysing and in part reconfiguring

cooperative medical work practices. From a sociological perspective, articulation work [Str+85], i.e. the ongoing work of integrating distributed tasks and maintaining a coherent treatment trajectory by collection, processing and organizing patient data, is at the core of clinical cooperation. However, this information flow is hampered by gaps and distortions due to the limited availability or time-consuming access to patient-related medical data, which in turn is at least partly related to their inadequate visual presentation.

One of the project's objectives was the design and development of *an integrated in-place visualization* of patient-related medical data in order to support and potentially even modify specific cooperative workflows. By the development of the anatomically integrated visualization, the fundamental problem of the acceptance, effectiveness and efficiency of new technologies in working practices has been addressed using a participatory design approach, similar to Belden et al. [Bel+18] and Sedlmair et al. [SMM12]. This requires the set-up of a design and implementation process (see Fig. 4.1) that involves visualization researchers, sociologists and neurosurgeons in order to analyse specific cooperative real-world situations on the neurosurgical ward. Its substantial part is a high-frequency interaction between these groups of experts, in particular during the initial definition of the overall goals of the intended visualization tool and the *field study* that comprises observations of cooperative workflow situations and interviews. This interaction is based on the mutual engagement in the design process following a symmetrical approach (Schubert and Kolb [SK21]). Chapter 3 provides a summarized overview of the field study findings and resulting goals, related to the overall VICC concept. In the following, this information is taken up and re-focused according to the perspective of a specific VICC component, described in this chapter. In particular, the insights of the field study are depicted in Sec. 4.2.1, whereas the deduced design goals are presented in Sec. 4.2.2.

Moreover, the last phase represents a *participatory refinement of design & implementation*, similar to the principle of agile software development (see, e.g., Martin [Mar03]), in which visualization researchers, sociologists and neurosurgeons jointly advanced and refined the visualization design on the basis of the formulated design goals (see Sec. 4.2.2) by utilizing prototypical implementations of a visualization tool. For reasons of efficiency and in order to not bias the evaluation, three physicians (*expert group*) were involved in this stage, whereas the evaluation involved a distinct set of ten physicians (*test group*; see Sec. 4.5).
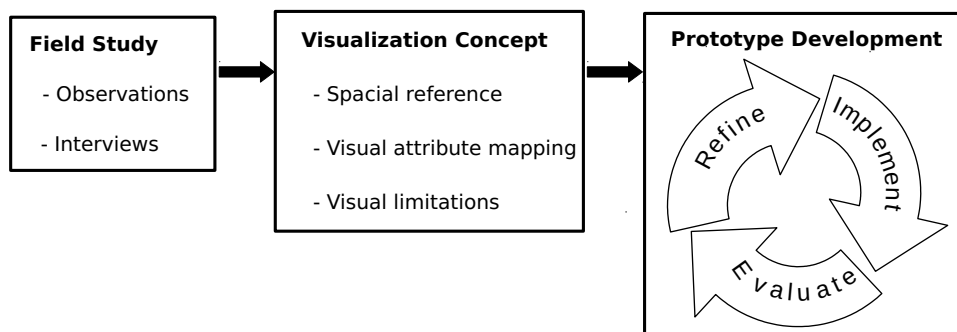

FIGURE 4.1: The design and development process for the visualization prototype. The rectangles represent the three main stages. The last stage is an iterative prototype development, which leads to a gradual concept refinement on the basis of experts' feedback.

4.2.1 Field Study

In this first phase of data collection and processing, field observations and interviews were conducted on the neurosurgical ward of a local, 480 bed hospital. Here, the sociologists, and partially visualization researchers, observed a total of 14 neurosurgeons for 22 days during their routine ward work to become acquainted with the workflows and relevance structures of their daily work. An important finding was the fact that the physicians rely on the patient's anatomy as spatial reference of the diagnostic results not only using medical imaging for instance, but also handwriting information, e.g., on anatomical sketches in examination forms. Furthermore, during these observations, specific cooperative work constellations have been identified. Due to the distributed character of the clinical work, the patient-related medical data collected by one physician, has to be conveyed to the colleague who continues with the patient's treatment. In this context, the task-related subset of medical data that describes the *patient's current status* and its *change over time* is of particular importance. A significant part of this subset constitute diagnostic data related to the anatomy, e.g. clinical symptoms. Moreover, the *timely efficient assessment* of the relevant medical data and its temporal change plays a crucial role for decision making and treatment quality. An effective conveyance of the patient's status under these particular conditions requires a data representation in an *objectified* and *concise* form.

The cooperative work constellations can occur in synchronous or asynchronous mode. An example of synchronous cooperation is the *shift handover* or, more precisely, the presentation of patients during the morning ward meetings. In these cooperative constellations the information is passed on personally, mostly in verbal, i.e. volatile, form. However, a personal shift handover is not always possible due to the inherent irregularity of the clinical workflow that results from possible emergencies. In these cases, the required information has to be passed on to colleagues asynchronously. This kind of asynchronous cooperation frequently occurs when dealing with unfamiliar patients, e.g. during *ward rounds* or when *writing of discharge reports*. Currently, the physicians have to recover the patient's status from different data sources. Furthermore, even when treating a patient they are familiar with, the physicians often need to recollect the patient-related medical data they acquired themselves.

4.2.2 Design Goals

Following the project motivation and the findings of the field studies, the overarching goal of the intended visualization system is to facilitate *efficient transfer of the patient's status in the context of distributed medical tasks*. The resulting design goals listed below have been jointly deduced by the visualization researchers, sociologists and neurosurgeons from findings of the field studies, using a symmetrical approach (see above Sec. 4.2), and represent a specific implementation of the overall VICC goals (see Sec. 3.1), related to the anatomical visualization component:

DG1 Familiarity in novelty. The goal is to exploit the existing visual tradition, i.e. the usage of the anatomy as spatial reference, as far as possible and to prevent abstract visualizations in order to achieve a high degree of intuitiveness.

DG2 Visual discriminability. A data mapping that allows the user to clearly discriminate between corresponding visual variables in a given clinical usage context is crucial for the correct understanding of a visualization, and thus for

an effective information transfer. The preference for anatomical visualization concepts defines the spatial location as a visual variable a-priori, which makes such a mapping even more challenging to achieve.

DG3 Context-related synopsis. In cooperative settings, it is important to provide a simultaneous visual access to all the data that are relevant in the given clinical usage context for conveying the patient's current status at a glance.

DG4 Intuitive comprehensibility. The visualization design has to focus on an intuitive and comprehensive interpretability in order to address the overall goal of efficient and effective transfer of the patient's status.

DG5 Concise visualization. In order to avoid clutter and distraction by secondary details, the aim is to provide only the data that are relevant in the current clinical usage context and to focus on medically relevant, i.e. abnormal, value constellations.

DG6 Visualization of chronological changes. The visualization should provide access to the evolution of the patient's data.

DG7 Mobility. The goal is to make the visualization available in all places using a mobile device.

4.3 The Visualization Concept

4.3.1 General Considerations

This section describes the proposed visualization concept, which aims at the achievement of the design goals (see Sec. 4.2.2) under consideration of practical insights into the hospital workflow gathered during the field studies (see Sec. 4.2.1).

At the core of the concept is an *anatomically integrated in-place visualization* of medical data relevant to specific cooperative tasks. According to this visualization principle, an anatomical model serves as spatial representation of medical data that inherently refer to its structures, e.g. the clinical symptom *paresis* referring the affected *muscle*. Particularly, the visual variables that appropriately encode the medical data to be visualized are applied by rendering of the corresponding anatomical structures, changing their default or "natural" appearance.

The use of the human body as spatial reference directly addresses the design goal **DG1** of exploiting the existing visual tradition, which is tightly bound to the human anatomy. On the one hand, this approach reduces the freedom in assigning medical data to visual variables compared to Information Visualization (InfoVis) (see the respective explanation in Sec. 2.1.1), since the attributes 'spatial position' and 'form' are defined by the model and are no longer free parameters. On the other hand, the focusing on a specific cooperative context and on data relevant therein (cf. **DG5**), significantly reduces the amount of data that have to be visualized and, consequently, the number of required visual variables. This data preselection in combination with in-place visualization provides a context-related synopsis at a glance (cf. **DG3**). Moreover, it is to be expected that the spatial embedding of medical data in the anatomical context makes its relation to the underlying real phenomena more evident and facilitates its interpretability (cf. **DG4**).

However, the anatomically integrated in-place visualization poses a very fundamental challenge: How to *design the mapping* from abstract medical data to visual variables under the restriction of a preset location/geometry, such that visual

discriminability (**DG2**), context-related synopsis (**DG3**) and intuitive comprehensibility (**DG4**) can be achieved in practice? To cope with this challenge, the current section provides the formal requirements that any specific realization of the anatomically integrated in-place visualization has to fulfil and the corresponding procedure of visual encoding (Sec. 4.3.2) as well as a discussion regarding further visualization concepts (Sec. 4.3.3).

4.3.2 Formal Definition and Requirements

4.3.2.1 Medical Data

The developing of an anatomically integrated in-place approach to visualize a synopsis of task-specific patient data requires a formal definition of the medical data and visual encoding attributes. The data organization proposed here links the formalism of the Unified Medical Language Service (UMLS)[1] to the data required in the visualization process. As UMLS itself is organized as ontology, it covers all notions and terms, as well as spatial and diagnostic relations. During the implementation process, UMLS serves as knowledge background in the discussion with the physicians to extract task-specific medical data (see Sec. 4.4.2.1). This information has to be augmented with *data types* and *domains*, as they play a central role in the visual encoding of the patient data. The resulting medical data to be visualized are organized in categories, each of them can have one or several properties; the categories are grouped into usage-dependent views. A precise definition of these concepts and their relations is given below.

- *Property*. A property p is defined as triple

$$p = (\text{propType, dataType, domain}),$$

  where p.propType refers to the underlying medical concept, p.dataType is a nominal, ordinal or numerical data type and p.domain is the domain set of elements of the corresponding p.dataType.

- *Category*. A category c is a pair

$$c = (\text{spatialRef, props}),$$

  where c.spatialRef is the reference to the corresponding anatomical structure and c.props $= \{x | x \text{ is a } Property\}$ are all properties of c. The set of all categories is denoted as C.

- *View*. A view $V \subseteq C$ is defined as the categories relevant to the given usage context. $\mathcal{V}$ is the set of all views.

  Note: Views automatically reduce the amount of information to be visualized, and thus the necessity of user navigation, by preselecting the context relevant data.

The following examples illustrate the structural concept introduced above.

The category 'muscle strength' refers to the anatomical structure 'muscle', comprising a single property with the type 'intensity' that documents the strength a

---

[1] https://www.nlm.nih.gov/research/umls/index.html

patient can create in a specific muscle, quantified in 6 numerical values.

$$\text{muscleStrength} = (\text{muscle},$$
$$\{(\text{intensity}, \text{numerical}, [0, \ldots, 5])\}).$$

The second example describes the category 'radicular pain', i.e. pain caused by irritation of a nerve root and related to the skin region that is associated with the latter, i.e. *dermatome*. This category comprises one numerical and one nominal property.

$$\text{radicularPain} = (\text{dermatome},$$
$$\{(\text{intensity}, \text{numerical}, [0, \ldots, 10]),$$
$$(\text{trigger}, \text{nominal}, \{\text{constant}, \text{stress})\}\}).$$

There are several fundamental relationships between views, categories and properties:

- A property type is unique within a category. Property types obtain their medical meaning only in combination with the category they are used in.

- A property type can be shared between properties of several categories. This expresses similarity of medical concepts, such as 'intensity' of clinical symptoms in the prior examples, even if the respective property domains can be distinct.

- A category is unique within a view.

- A category can be shared between several views, as it can be relevant in different usage contexts.

Moreover, in order to properly specify the mapping of medical data to visual variables, let be

$$T = \{p.\text{propType} \mid p \in c.\text{props}, c \in C\}.$$

the set of all property types.

4.3.2.2   Visual Variables

Shape and position are mainly predefined as the anatomical structures that are used for visualization in our concept, i.e. the *geometric* visual variables cannot be used for data encoding.[2] The remaining visual variables are colour components, namely *hue*, *brightness* and *saturation*, *textures* and *transparency*, as well as *time* using, e.g., animations.

Let a denote a visual variable, A the set of all visual variables, and $a.\text{range}$ the discrete and finite set of distinctively perceivable values of $a \in A$. Note, that $|a.\text{range}|$ is commonly smaller than the number of displayable visual variables. For example, the visual variable 'hue' is a floating point value, however, the human vision is able to distinguish only up to eight hue values w/o external reference (see, e.g., Kuehni and Schwarz [KS08]).

---

[2]In principle, specific transformations or deformations that do not degenerate the anatomical structure could be used.

4.3.2.3 Visual Encoding / Mapping

The mapping of medical data to visual variables is represented in Fig. 4.2 and comprises three levels, i.e. (a) encoding of categories (step 1-2), (b) encoding of property types (step 3), (c) encoding of property values for all properties of each category (step 4).
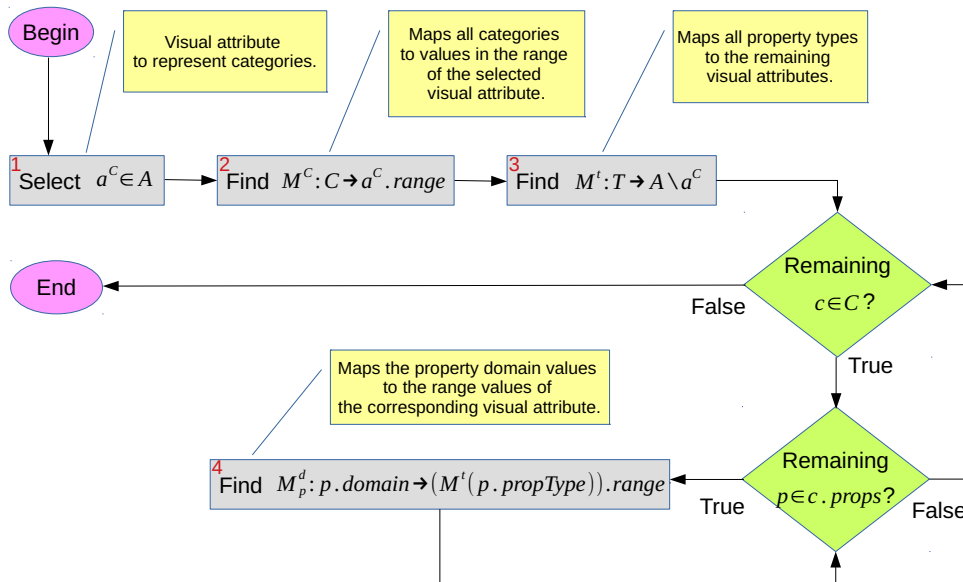
FIGURE 4.2: The three level visual encoding of medical data.

4.3.2.4 Injectivity Requirement

In general, *injectivity* is a pre-requisite for any visual encoding / mapping in order to lead to an unambiguously comprehensible representation (see, e.g., Ziemkiewicz and Kosara [ZK10]). The proposed visualization concept distinguishes the following situations where either the mapping injectivity is strongly required or its violation has to be recognized and appropriately tackled.

Local Injectivity. In the case the mapping $M^c\colon C \to a^C.\text{range}$ or $M^t\colon T \to A \setminus \{a^C\}$ (see Fig. 4.2) is not injective within a given view V or category c, respectively, it is impossible to trace back the categories or property types from their visual representation. Local injectivity can be guaranteed if (1) all categories of a view are mapped to distinct values of a^C, and (2) all property types of a category are mapped to distinct visual variables.

Global Injectivity. While local injectivity guarantees the visual distinctiveness of categories and properties *inside* each single view and category, respectively, the global injectivity ensures the uniqueness of visual encoding *across* respective contexts, i.e. categories and views for property types and views for categories. Due to the restricted distinctiveness of visual variables, global injectivity is hardly achievable, still it should be pursued as far as possible.

Spatial Injectivity. Obviously, several categories of a view can refer to the same anatomical structure. The violation of spatial injectivity rules out simultaneous visualization of the respective medical data and needs to be handled explicitly (see Sec. 4.3.3).

Property Domain Injectivity A non-injective mapping of a property domain to the respective visual variable range causes quantization and, consequently, leads to a loss of information. In some cases a quantized visualization can be acceptable, in particular in the context related synopsis (cf. **DG3**), in which a qualitative overview is sufficient. In any case, quantization needs to be detected and reported. In general, quantization cannot be prevented and needs to be handled explicitly (see Sec. 4.3.3).

4.3.2.5   Visibility Restrictions

A 3D visualization with free camera motion may intrinsically affect the visibility of geometric objects representing anatomical structures. In the context of the presented approach, there are two causes for restricted visibility. (1) An anatomical structure that relates to relevant medical data (target) may be *occluded* by another anatomical structure. For example, a muscle affected by paresis is hidden under the skin. (2) The *spatial extension* of an anatomical structure is too small in relation to the entire human body, so that the visualization cannot be clearly recognized. An example would be a tendon. Both aspects are handled in Sec. 4.3.3.

4.3.3   Further Visualization Concepts

Independently of the specific mapping, introduced in Sec. 4.4, the anatomically integrated in-place visualization concept handles some further aspects, necessary to achieve the design goals postulated in Sec. 4.2.2.

Data selection. Besides two already described data selection mechanisms, which apply automatically, i.e. focusing on the abnormal and usage-dependent views, the user has the possibility to additionally *filter the data by their categories*. Note that the usage-dependent preselection reduces the available categories to a manageable amount. This filter allows to tackle, inter alia, the visualization of multiple medical data for the same anatomical structure, i.e. spatial non-injectivity (see Sec. 4.3.2.4).

Alternating visualization. In the constellations where a simultaneous visualization of multiple medical data on the same anatomical structure is not possible, i.e. the spatial injectivity is not fulfilled (see Sec. 4.3.2.4), an alternating visualization with additional user control to select one of the alternatives, e.g. through the data category filter, can be applied (see Fig. 4.6A-4.6B).

Textual overlays. In some situations, the physician may want to access the underlying information explicitly, i.e. in textual form. This option is enabled by textual overlays on top of the corresponding anatomical structure in order to, for example, resolve the quantization problem (see Sec. 4.3.2.4) that results in a simplified visual representation or to support the physicians in the getting acquainted with the visualization tool.

Proxies. The small object extension problem (see Sec. 4.3.2.5) can be tackled by means of an appropriately scaled proxy that is displayed on the body surface over the location of the target anatomical structure and rendered with the corresponding visual variables.

Transparencies. View dependent transparency is used to handle depth occlusions of anatomical structures carrying relevant information by other anatomical structures (see Sec. 4.3.2.5). In particular, the occluder, in case itself is not carrying relevant information, is rendered as (semi-)transparent, while the surrounding anatomical context is preserved.

The specific technical approaches taken to implement textual overlays, proxies and transparencies are described in Sec. 4.4.2.2.

4.4 Prototype Implementation

Based on the design goals stated in Sec. 4.2.2 and the visualization concept introduced in Sec. 4.3, the following section presents a prototype implementation on a mobile device. As a proof of concept, the focus has been set on patients with *spinal disc herniation* and the relevant medical data for assessing a patient's neurological status.

4.4.1 Development Environment

As geometric model for the anatomically integrated visualization is used *plasticboy's* 3D human avatar[3], which fulfils the substantial requirement of an exact and faithful representation of the human anatomy. Even though the model already includes the main organ systems subdivided into the corresponding anatomical structures, this subdivision is not sufficiently detailed for the use with the proposed visualization approach. Thus, the initial plasticboy anatomical human model was mapped to the hierarchy of the Foundational Model of Anatomy Ontology (FMA)[4]. Furthermore, dermatomes, which constitute very important anatomical structures in the context of neurosurgery, are not reflected in the model. Hence, an anatomy refinement procedure, based on indexed texture maps, is applied, allowing to define substructures on the existing geometries.

The prototype was developed in C++, using Qt 5.12 for Android for the platform integration and a rendering framework, which is based on the Vulkan 1.1 API. The latter allows an efficient resource management, which is important especially on mobile devices.

4.4.2 Prototype Features

4.4.2.1 Mapping of Spinal Disc Herniation Data

Raw Data Categories. In collaboration with the expert group of neurosurgeons (see Sec. 4.2), seven data categories and their properties (see Tab. 4.1) that are of high relevance for cooperative tasks with respect to the representation of neurological status of a patient with spinal disc herniation, i.e. in the respective view (cf. Sec. 4.3.2.1), have been determined. The type of the main property to be visualized, common for all these categories, is the *intensity* with which the respective symptom manifests, whereas its domain is individual for each symptom, i.e. category.

The category *radicular pain* has a further property with the type *trigger*, which states if the pain is constant or only occurs under stress, e.g. during movements, whereby the former is assumed to be the normal, i.e. default situation for pain that does not need any visual indication.

Data Category Refinement. The above raw categories mainly reflect concepts from the Unified Medical Language System (UMLS) vocabularies, related to neurological

---

[3]www.plasticboy.co.uk/store/Human_Male_Female_Anatomy_Complete_V05.html
[4]http://sig.biostr.washington.edu/projects/fm/AboutFM.html

findings, which then have been partially refined with the participation of domain experts, in order to better map the application logic to the given context of cooperative tasks.

During the first trials with the prototype and discussions with neurosurgeons, two changes to the initial raw data categories have been applied. First, the three categories related to *sensory disorder*, i.e. paresthesia, hypoesthesia and anaesthesia have rather complex interrelations. For example, hypoesthesia and anaesthesia can be considered as different stages of sense decrease, whereas paresthesia is in a certain sense orthogonal because it does not describe a decrease of sensation but rather its abnormality, e.g. tingling, and, thus, it can occur in combination with hypoesthesia. Therefore, the new category sensory disorder represents anaesthesia and hypoesthesia as a joint property with the type 'intensity' and has an additional property with the type 'paresthesia'. Second, insufficient *muscle strength*, measured in the Medical Research Council scale $0, \dots, 5$, is used in daily clinical practice as indication for *paresis* that triggers for potential urgent actions such as emergency surgery. Therefore, this practice has been adopted in the data model by using the category *paresis* with the ordinal data type comprising the values 'mild', 'moderate', 'severe'.

Mapping of Data to Visual Attributes. The final medical data categories in Tab. 4.1 are visually encoded in consideration of the rules deduced in Secs. 4.3.2.3 and 4.3.2.4, that is, the mapping functions M^c and M^t are at least locally injective. *Hue* has been selected as the visual variable a^C to encode *category* (cf. step 1 in Fig. 4.2). The category mapping M^c (cf. step 2 in Fig. 4.2) takes into account the distinctiveness of the resulting five hues with regard to each other as well as to their context in the anatomical model, e.g. a red colour is not suitable for visualization on muscles because it highly coincides with their natural, i.e. healthy, appearance. The shared property type *intensity* is mapped in all categories to a composite visual variable *saturation-brightness* in the HSV colour space, i.e. M^t (cf. step 3 in Fig. 4.2) is also globally injective. The decision to combine two respective visual variables in a single one aims to increase the visual discriminability (**DG2**) by means of an augmentation effect (see Sec. 2.1.1). *Additional properties* are encoded by means of *textures*, which have been selected such that they are visually as complementary to the visual variable *colour* as possible in order to allow for a simultaneous visualization. In case the additional property is non-binary, e.g. the property with the paresthesia type in sensory disorder (cf. Tab. 4.1), the required range of visual values corresponds to the texture's frequency and amplitude (cf. step 4 in Fig. 4.2); see also Fig. 4.8.

There are two additional features of the mapping to visual variables that need to be mentioned (cf. Table 4.1): (1) The 10 intensities of the pain category result in a visual quantization, i.e. the corresponding property domain mapping is not injective (see Sec. 4.3.2.4). In this case, the technical reason for the quantization coincides with the physician's suggestion for a reduction of the number of intensity levels for this symptom to account for the subjectivity of the patient's sensation (cf. Jaatun et al. [Jaa+19], who encoded pain intensity using three colours instead of the 1-to-10 scale, which was evaluated as suitable by most users; cf. also the 0-to-3 Digital Palpation pain scale in Masulo et al. [Mas+19]). (2) the anatomical reference for the categories *T-reflex* and *excretion disorder* are too small and require a proxy geometry (see also Sec. 4.4.2.2 and, e.g., Fig. 4.6).

TABLE 4.1: Data categories: The raw data categories (left block), the final categories after discussion with the physicians (centre block), and the visual variables incl. anatomical reference (right block). Only the abnormal states are listed. Specific aspects are indicated as [1]: quantization, [2]: usage of proxy geometry.

Raw Data Categories			Final Data Categories			Visual Attribute	Anatom. Reference
Category	**Prop. Type**	**Domain**	**Category**	**Prop. Type**	**Domain**	**(⊢range⊣)**	
Radicular Pain			Radicular Pain			Red	
	Intensity	$\{1,\dots,10\}$		Intensity	$\{1,\dots,10\}$	Saturat.-Brightn. (3)[1]	Dermatome
	Trigger	binary		Trigger	binary	Texture *Normal Pert.*(1)	
Muscle Strength			Paresis			Purple	
	Intensity	$\{1,\dots,5\}$		Intensity	{mild, moderate, severe}	Saturat.-Brightn. (3)	Muscle
T-Reflex			T-Reflex			Green	
	Intensity	$\{1,\dots,5\}$		Intensity	$\{1,\dots,5\}$	Saturat.-Brightn. (5)	Tendon[2]
Excretion Disorder			Excretion Disorder			Orange	
	Intensity	binary		Intensity	binary	Saturat.-Brightn. (1)	Urethra or anus[2]
Paresthesia			Sensory Disorder			Cyan	
	Intensity	$\{1,\dots,3\}$		Intensity	$\{1,\dots,4\}$	Saturat.-Brightn. (4)	Dermatome
Hypoesthesia				Paresthesia	$\{1,\dots,3\}$	Texture *Noise* (3)	
	Intensity	$\{1,\dots,3\}$					
Anaesthesia							
	Intensity	binary					

4.4.2.2 Rendering Implementation Details

This section briefly describes some implementation details related to the realization of specific visualization features, partially mentioned in Sec. 4.3.3.

Anatomical Proxies. The main idea in generating anatomical proxies is to utilize projective textures [Seg+92], e.g. appropriately scaled circles, on the skin surface above the anatomical structure that is too small for a direct visualization (cf. Sec 4.3.2.5 and the patellar reflex, e.g., in Fig. 4.6).

Occlusion Handling In order to visualize hidden anatomical structures in an integrated overview of the most relevant information without requiring specific selection, navigation or zooming efforts, the opacity of areas above the occluded target structure is dynamically decreased, depending on the current camera transformation (cf. Sec. 4.3.2.5). Similar to Viola et al. [VKG04] and Burns and Finkelstein [BF08], an image space approach that allows an efficient detection of occluding fragments in real time on mobile hardware is used (see Fig. 4.3).

Context Menu The prototype includes a context menu for accessing the advanced features such as the *data category filter* (see Fig. 4.7A) and the *overlays* with textual data (see Fig. 4.7C). The filter provides for the user the possibility to hide/show data visualizations by their category as described in Sec. 4.3.3, e.g. for handling of the spatially non-injective cases (see Sec. 4.3.2.4). The overlays display textual data corresponding to the visualization on a anatomical structure selected by the user.

Visualization of Temporal Changes For assessment of the healing progress a slider with dates of available clinical examinations is integrated below the main 3D view (see Fig. 4.6). Moving the slider, the user can navigate to the date of interest or scroll
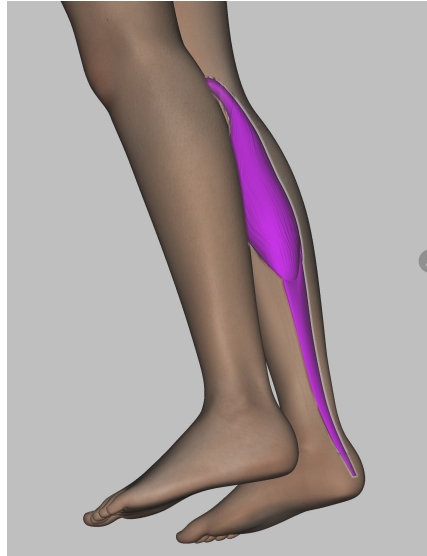
FIGURE 4.3: Dynamic transparency area with hierarchical information: the right gastrocnemius muscle with paresis data (purple) is visible through the skin of the same body region and partially occluded by the left leg.

through consecutive examination results. By selecting a date, the entries with the corresponding timestamp are retrieved from the database and the virtual body is rendered with updated visual variables.

4.5 Prototype Evaluation

4.5.1 Objectives

With the goal to evaluate the potential of the visualization concept and its limitations, a user study was conducted, addressing the following objectives.

EO1 Perceived usefulness of the anatomically integrated in-place visualization concept in comparison with the conventional data representation, i.e. hospital information system, examination forms etc. The following core concepts was evaluated separately:

   (a) Use of anatomy for spatial data representation (cf. **DG2**);

   (b) Synopsis-like visualization of multiple data (cf. **DG3**);

   (c) Narrowing the amount of data to be visualized according to the given usage-dependent view (cf. **DG5**);

   (d) Use of a mobile platform (cf. **DG7**).

EO2 Effectiveness of the information transfer. In particular, the evaluation addressed the visual encoding of the medical data, i.e. that the medical data can be read from the visualization unambiguously regarding the following aspects:

   (a) Reading of a neurological patient status;

   (b) Reading of the healing progress by comparison of different status snapshots (cf. **DG6**);

(c) Usage of interactive mechanisms, for instance data category filter or zooming (see Sec. 4.4.2.2), where the direct reading is impeded.

EO3 Perceived efficiency in comparison with the conventional data representation. Since currently the relevant data are scattered in different sources, a direct quantitative comparison is impossible. Instead, users should estimate the expected speed-up due to the proposed visualization approach on the background of their long-term experience in using the traditional data sources.

EO4 Completeness / meaningfulness of the implemented data selection for the view of *neurological status* (see Tab. 4.1). In particular, the aim is to assess the convergence between the usage-dependent view defined in collaboration with the expert group (see Sec. 4.2) and the data considered relevant in the given context by the test group.

4.5.2 Setup

	Description	Objectives
Task 1	Interpretation without prior knowledge of the prototype	EO4
Learning	Introduction of the visual encoding legend	
Task 2	Assigning the visualization to a textual description of symptoms	EO2a, EO2c
Task 3	Reading the healing progress	EO2b, EO2c
QNR 1	Assessment of relevance of the data selected for the diagnosis to be visualized	EO4
QNR 2	Assessment of usefulness	EO1
QNR 3	Assessment of effectiveness and efficiency	EO2, EO3

FIGURE 4.4: The evaluation process of the visualization prototype. The left and middle column show the numbers of practical tasks or questionnaires and their short descriptions, respectively. The right column displays the evaluation objectives, as described in **EO1-EO4**, targeted by the corresponding task or questionnaire.

The evaluation comprises three practical tasks and a questionnaire. Fig. 4.4 gives an overview of the evaluation process, including the objectives to achieve in each specific stage.

The practical tasks are designed as follows.

Task 1. In the first task, the physicians need to give their intuitive interpretation of two visualizations (see Fig. 4.5) without previous knowledge about the meaning of the visual variables. The assumption to be tested is that embedding in the anatomy (see **EO1**a) and usage-dependent preselection of medical data (see **EO1**c) increase the intuitiveness of the visualization.

Task 2. The visualization of a neurological patient status (see Fig. 4.6A-4.6B) has to be assigned to the corresponding textual description of clinical symptoms. The participants need to choose the correct status out of seven real-life descriptions. The task can be seen as succession of two sub-tasks, in which different evaluation objectives are addressed. (a) The direct reading of visual variables (see **EO2**a) narrows

(A)   (B)
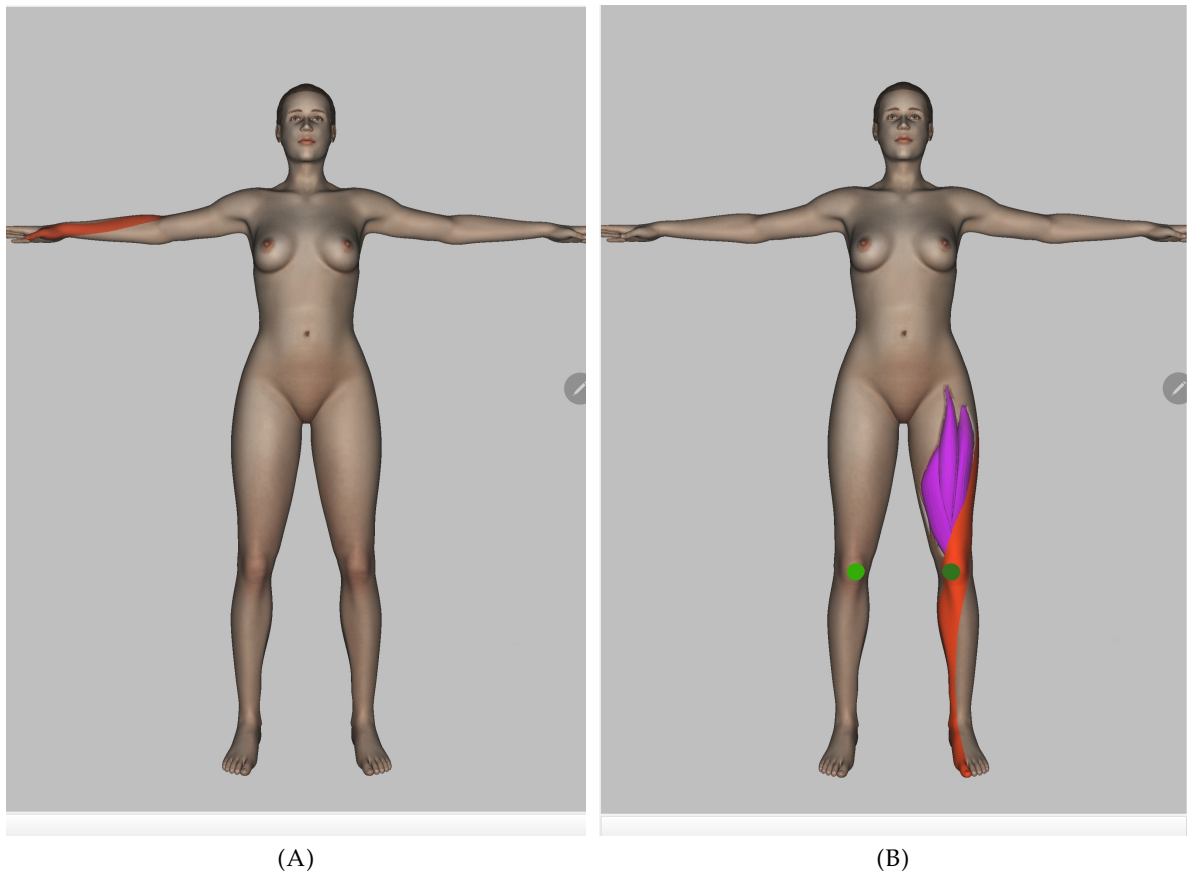
FIGURE 4.5: Two visualization samples used for spontaneous inter-
pretation in **Task 1**. (A) shows the right C6 dermatome with radicular
pain. (B) shows several symptoms visualized in parallel: radicular
pain in the left L4 dermatome, paresis of the left quadriceps muscle,
asymmetric patellar reflex by means of the proxies.

down the possible choices to the subset of two very similar descriptions[5]. (b) Sep-
arating the remaining two similar cases requires some user interaction (see **EO2**c),
namely applying data category filter in order to uncover the muscles, and, poten-
tially, zooming and textual overlays (see Fig. 4.7A-4.7C).

Task 3. The last task is to read the healing progress (see **EO2**b) from two visual-
izations (see Fig. 4.6A-4.6B and Fig. 4.6C) that show successive patient states. The
first one equals **Task 2** and corresponds to the preoperative status of a patient. The
second one represents the postoperative neurological status of the same patient.
Switching between both visualizations by means of the timeline, the physicians
have to recognize (cf. **DG2**) and to describe the respective changes. Analogously
to **Task 2**, there is an additional level of difficulty due to the partially remaining
paresis that is still occluded in the postoperative status (see Fig. 4.6C), which again
requires the use of the data category filter (see **EO2**c).

The questionnaire comprises three main sections:

---

[5]The correct and the similar status differ only in location of the paresis. The corresponding muscles
are relatively small and anatomically adjacent, and they are partially occluded by the L5 dermatome.

FIGURE 4.6: Visualization of temporal changes. (A)-(B) show the visualization of the preoperative neurological status of a patient with an asymmetric patellar reflex, a moderate paresis in the right extensor hallucis longus muscle, a mild paresis in the tibialis anterior muscle, a severe pain and a hypoesthesia in the left L5 dermatome, whereby the pain (s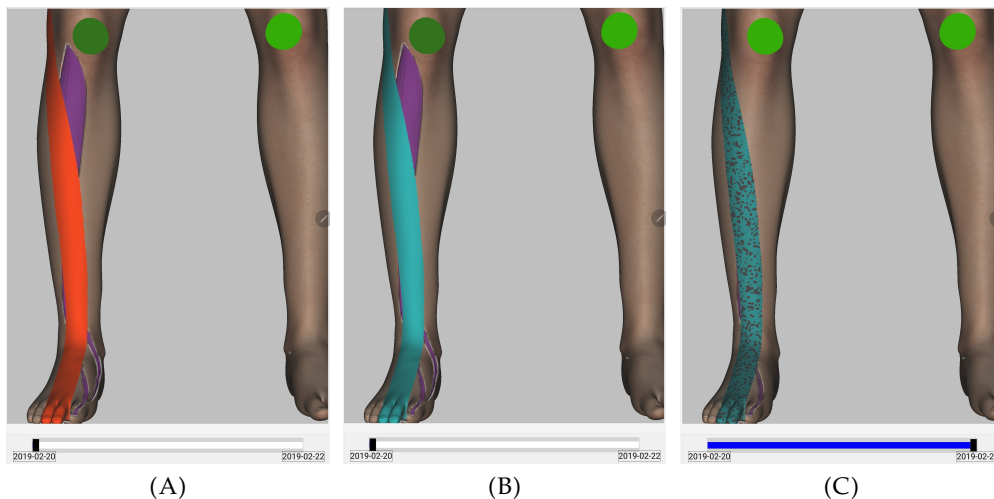ee Fig. (A)) and the hypoesthesia (see Fig. (B)) are visualized alternately; this visualization is used in **Task 2**. (C) shows the postoperative status of the same patient with symmetry in patellar reflex, a decreased hypoesthesia in combination with paresthesia and a remaining mild paresis in the right extensor hallucis longus muscle (see **Task 3**)

QNR 1. The first section is dedicated to the evaluation of data in the usage-dependent view related to patients with spinal disc herniations (see **EO4**). Particularly, the physicians assess the relevance of the pre-selected data categories in different cooperative constellations as described in Sec. 4.2.1 as well as their completeness (see Fig. 4.9A). The considered cooperative constellations are 1) shift handover; 2) visiting a familiar patient during ward rounds, i.e. a patient who already has been examined by the same physician; 3) visiting an unfamiliar patient during ward rounds, i.e. a patient who has been admitted and examined by a colleague; 4) writing a discharge report, especially for patients who were treated by another, currently absent colleague.

QNR 2. In the second section the physicians assess the usefulness of the anatomically integrated in-place visualization concept, including the simultaneous visualization of multiple data and the availability on a mobile device (see **EO1**).

QNR 3. The third section includes questions regarding the effectiveness (cf. **EO2**) and efficiency (cf. **EO3**) of the prototype that implements the anatomically integrated in-place visualization concept (see Fig. 4.9C).

4.5.3 Procedure

The exact evaluation procedure is structured as follows.

(1) After a brief introduction to the overall context, the participants are directly confronted with **Task 1**.
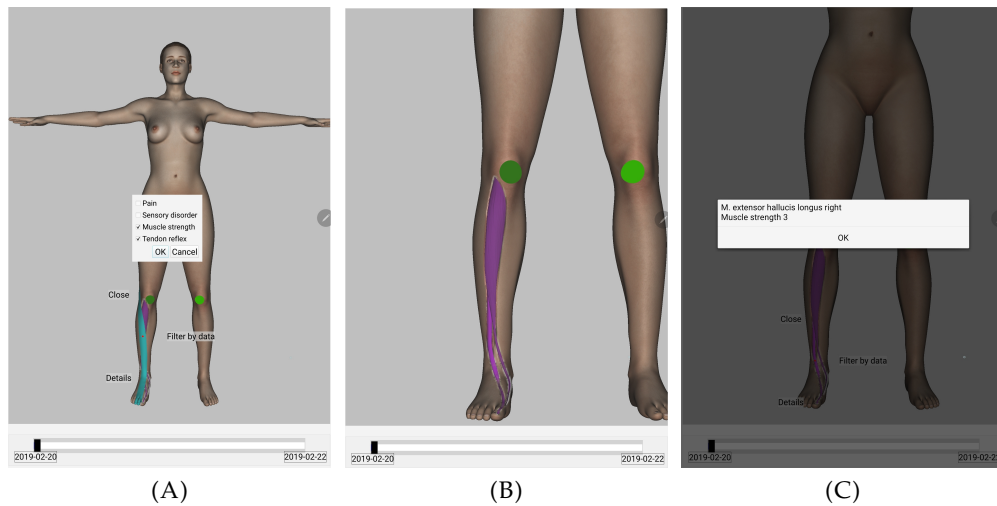
FIGURE 4.7: Usage of further prototype features in **Task 2**. (A) illustrates the use of the data category filter. (B) shows a close-up of paresis after filtering, i.e. without the occluding dermatome. (C) exemplifies the textual overlay for one of the affected muscles comprising its anatomical name and associated raw muscle strength data.

(2) Afterwards, during a learning phase of approx. 15 minutes, the visual encoding legend (see Fig. 4.8) is introduced and the participants are familiarized with the most important interaction features of the prototype.

(3) Next, the physicians have to perform **Task 2** and **Task 3**, explaining their choices.

(4) Finally, the participants have to reflect on the practical experience answering the questionnaire, i.e. **QNR 1**, **QNR 2** and **QNR 3**.

Ten neurosurgeons from the same department took part in the evaluation, none of which was involved in the participatory refinement process described in Sec. 4.2, i.e. this test group had only a very basic understanding of the overall aim of the visualization prototype. The group covers very different professional experiences and hierarchical positions, namely it consists of three assistant physicians, one specialist, five senior physicians and one chief physician (for more details see Tab. 4.2). The evaluation study was conducted in the hospital, in groups of two persons or, partially, individually, and has been filmed for documentation and analysis. The complete procedure for a single group took about an hour. A total of nine (out of ten) completed questionnaires was submitted, because one of the interviews was interrupted due to a medical emergency. The total number of answers varies, since a few participants skipped some questions.

TABLE 4.2: Professional experiences as well as age and gender structure of the test group.

Prof. experience		Age		Gender	
years	# persons	years	# persons		# persons
<10	3	30-40	4	female	4
11-20	3	41-50	2		
21-30	2	51-60	2	male	6
>30	2	>60	2		
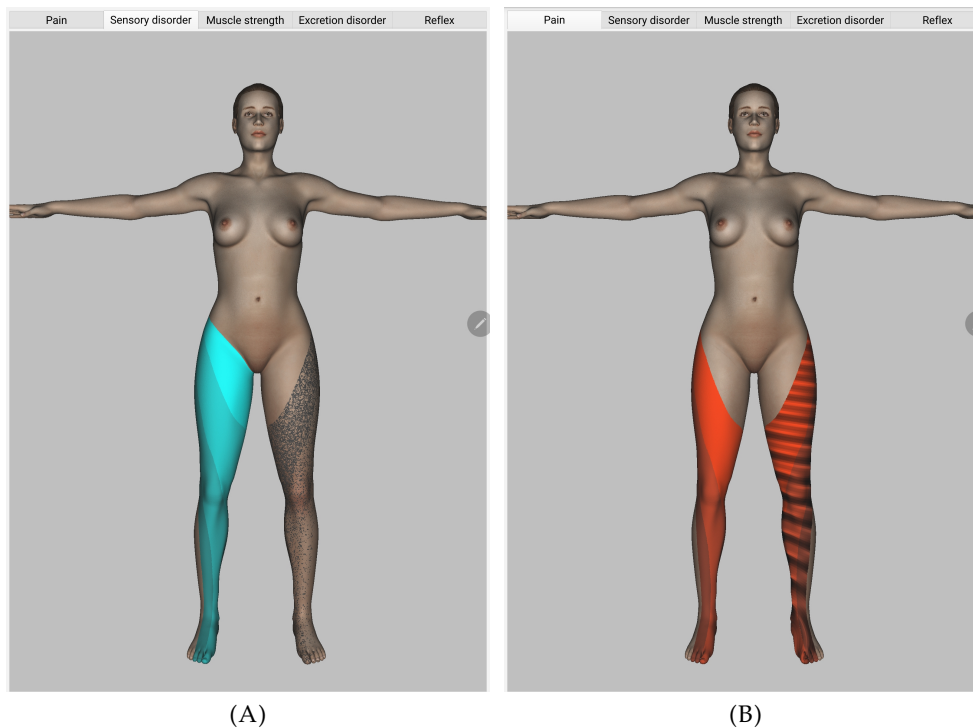
FIGURE 4.8: The legend of the prototype visual encoding. (A) depicts the active tab of the sensory disorder category: the right L2-L5 dermatomes visualize four intensity levels, the left L3-L5 dermatomes show three levels of the paresthesia, i.e. Noise texture; (B) shows the active radicular pain tab: the right L3-L5 dermatomes visualize three intensity levels, the left L3-L5 dermatomes show same levels in combination with the stress trigger, i.e. Normal Pert. texture (cf. Tab. 4.1); the exact property values can be looked up by means of overlays similarly to Fig. 4.7C.

4.5.4 Quantitative Results

In the following, the reader finds an overview of the quantitative evaluation results, structured according to the aforementioned tasks and questions. The qualitative results and lessons learned are discussed in Sec. 4.6.

In **Task 1** (i.e., before the learning phase), reading the visualization without any specific pre-knowledge, most physicians (8/10) correctly interpreted the *red dermatome* (see Fig. 4.5A and 4.5B) as pain and the rest of the participants (2/10) as sensory disorder. The *purple colour in the muscle* (see Fig. 4.5B) was consistently associated with paresis (10/10), which is the correct meaning of this visual variable. The *green proxy geometries at the knees* (see Fig. 4.5B) have often been misinterpreted (6/10), e.g. as "problems with knee joints" or "hematoma". Nevertheless, four participants (4/10) recognized them as visualization of patellar reflex.

In **Task 2** (i.e., after the learning phase), assigning the visualization to a textual description, all physicians (10/10) could easily solve step 1, i.e. they narrowed down the conceivable variants to two similar cases. By step 2, three of the participants (3/10) independently found the correct case and the other seven of them could solve the task after a suggestion of the test supervisors to use data category filter and zooming function.

The healing progress in **Task 3** was interpreted correctly by all participants
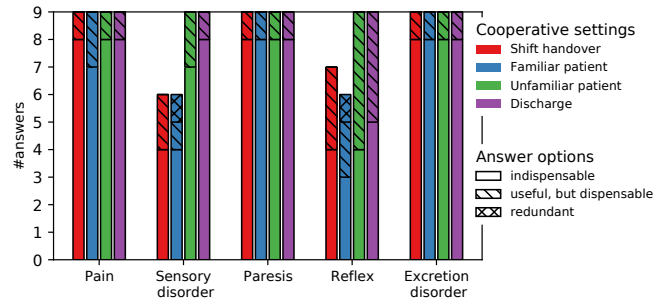
(A) **QNR 1**: *data category relevance*.
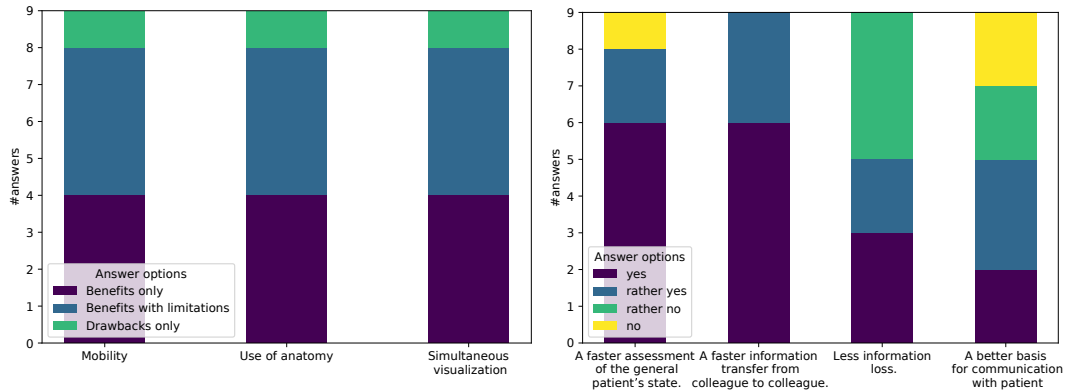
(B) **QNR 2**: *assessment of potential benefits and drawbacks*.

(C) **QNR 3**: *efficiency and effectiveness of the visualization*.

FIGURE 4.9: Quantitative results of the questionnaire. (A): the x-axis represents the symptom categories. (B): the x-axis shows the prototype features to be assessed. (C): the x-axis shows possible answers to the question: *Would the proposed visualization concept yield the following benefits in your work routine?*

(10/10). Several of them (6/10) also could independently find the occluded remaining paresis in the postoperative status (see Fig. 4.6C), hiding the occluding dermatome by means of the category filter without external suggestions.

The quantitative results regarding the questionnaires **QNR 1-QNR 3** are summarized in Figs. 4.9A-4.9C, and are discussed in Sec. 4.6.

4.6   Discussions

4.6.1   Evaluation findings

In the following, the evaluation results are discussed according to the objectives determined in Sec. 4.5.1

Usefulness   The results of **QNR 2** (see Fig. 4.9B) and the physicians' comments during the evaluation study show that the proposed visualization concept was perceived as very beneficial. In particular, the use of anatomy (cf. **EO1a**) was well received. Several physicians pointed out that, as a positive collateral effect, the 3D body model also serves as anatomical quick reference guide. The results of **Task 1** underline the increased intuitivity of the used visualization, due to the anatomically integrated visual variables in combination with a usage-dependent view (cf. **EO1c**).

Both allow for the comprehension of the visualized symptom even without legend (cf. interpreting muscle colour as paresis) or significantly reduce the possible interpretations (cf. colour in a dermatome as pain or sensory disorder). However, the interpretation of visual variables corresponding to the patellar reflex in the same task was perceived as less intuitive. According to the above claim, this can be explained by a lower importance of the data category "T-reflex" (see Fig. 4.9A and the discussion regarding **EO4** below), and the use of a proxy geometry, i.e. a weaker anatomical integration.

The physicians appreciated the availability of "a lot of information at a glance" thanks to the synopsis-like visualization (cf. **EO1**b). At the same time, some of them expressed concern about a potentially too complex appearance due to such an approach and the required user interaction, e.g. to hide an occluding anatomical structure. Also **Task 2** revealed difficulties of the users to utilize the prototype's interaction functionalities. However, this can be at least partially explained by an extremely short learning phase of about twenty minutes and the lack of experience with the corresponding prototype features, e.g. the data category filter. This conjecture is confirmed by the observation that in the next task (**Task 3**) the physicians used the interaction mechanisms for reading (partially) occluded information much more reliably.

Most physicians see a clear benefit in the mobility of the prototype (cf. **EO1**d), since it allows to save time in making the information available immediately in any location. However, it was mentioned that the current device is still too large and heavy to be carried in a coat pocket, and one participant also commented that a medical doctor with a mobile may be perceived by patients as "less professional".

Effectiveness The evaluation tasks demonstrate that the proposed visualization allows to effectively convey the neurological status of the patient (cf. **EO2**a) and its development over time (cf. **EO2**b). The observed limitations refer to complex situation with impaired visibility, which require user interaction (cf. **EO2**c), as discussed above.

The prototype's capability to reduce the loss of information in the hospital workflow (see Fig. 4.9C) has been rated controversially. The participants who remain rather sceptical, explained their attitude by the fact that the information still can be lost before reaching the prototype, due to a deficient collection and input of raw data by the examining physician. Surprisingly, only few physicians assessed the prototype as advantageous for communication with patients (see Fig. 4.9C). This, however, is due to the characteristics of the workflow itself, i.e. all physicians admitted that the proposed visualization would allow to explain diagnosis to the patient more easily, but this would stimulate further discussions with the patients, which is in conflict with the strongly limited time budgets.

Efficiency Comparing with the currently used textual representation of the respective medical data, the physicians attest to the prototype's potential for accelerating the transfer and evaluation of the patient's status (see **EO3**), which is due to the visualization of "all symptoms at a glance", i.e. synopsis, and to the fact that the information is "always available", i.e. mobility (see also Fig. 4.9C). Only one participant claimed that simple lists are as good as visualizations.

Data selection As follows from the rating in Fig. 4.9A, all defined data categories are of high importance for assessment of neurological status of a patient with spinal

disc herniation (cf. **EO4**), i.e. there is a high coincidence in the selection of relevant data within the test group as well as between the test group and the expert one.

Considerable variations occur in the category T-reflex only. The latter is caused by differences in the individual diagnostic methods: several participants actively apply reflex test during a clinical examination, whereas other physicians consider this test less relevant (cf. also the above discussion about effectiveness). In relation to the completeness of the view, six (6/9) physicians proposed further data categories to be added. While some of these proposals reflect individual variations in the diagnostic practices, referring to, for instance, additional pathological reflexes or pain qualities, other suggestions address the extended use of the prototype by integrating, for example, prior diseases or instability signs, required for decision making or surgery planning.

4.6.2   Limitations

Primarily, the prototype supports the evaluation of the proposed core concept, i.e. anatomically integrated in-place visualization, and, thus, focuses on the visualization of usage-dependent medical data with anatomical reference, more precisely, neurological symptoms of spinal disc herniation, in the form of a synopsis, i.e. a consistent, simultaneous representation of multiple key data. The preference for using anatomy as a spatial reference and for the synoptic form of data visualization are major findings from the field studies, and their effectiveness, efficiency and acceptance by the physicians have been verified in the aforementioned evaluation. Nevertheless, these visualization principles are potentially less useful in other medical application scopes, such as trend analysis of extensive medical data records, visualization of a single data category with a high domain cardinality, or the handling of medical data without anatomical reference. These kinds of data are handled by another component of the VICC concept, i.e. iconic glyphs (see Sec. 6). Furthermore, there is the option to integrate other visualization approaches such as abstract timeline-based techniques into the presented concept.

The use of human anatomy as spatial representation of abstract medical data, facilitating a more intuitive visualization semantic, at the same time implies some additional challenges with regard to visualization design. The general problem of mapping of a large amount of data to the limited number of visual variables becomes in this case even more significant, because the spatial position and the form of anatomical structures are fixed. As discussed in Sec. 4.3, this problem is addressed with such concepts as data selection at different levels and alternating visualization. Nevertheless, by further increasing the number of patient data associated with same anatomical structure, these solutions are not sufficient. To address this challenge, the overall VICC concept (see Sec. 3.2) includes the possibility to integrate iconic glyphs in the anatomically-based visualization, using them as quantifiable, multidimensional textures mapped to the respective anatomical structures. This technique, however, is not implemented yet.

Moreover, the on-body visualization can lead to configurations that are challenging in terms of visibility (cf. 4.3.2.2). The proposed solutions are described in Sec. 4.3.3 and 4.4.2.2. However, they potentially require additional user interaction, which can impair the information reading, as demonstrated in the practical tasks during the evaluation. In this respect, the evaluation results should be deemed intermediate. A more substantial evaluation of the efficiency of user interaction mechanisms requires a greater familiarity of physicians with the prototype.

Finally, the current prototype does not allow for a personal adaptation to individual variations in diagnostic practices (see Sec. 4.6.1). This suggests to introduce some personalization options by the view definition, on top of the core data categories.

4.7 Conclusion

In this chapter, a novel concept for anatomically integrated in-place visualization of medical data has been presented. The concept is designed in accordance with the requirements arising from specific tasks in cooperative clinical workflow, namely conveying of cooperation relevant patient data between colleagues, and it aims at closing the existing gap in current visualization approaches. The presented approach allows for a spatially integrated comprehensive visualization of medical data, such as clinical symptoms, on a 3D human avatar using their inherent references to affected anatomical structures and an appropriate visual encoding. Preselecting patient data as a function of their relevance in the given clinical usage context, i.e. view, provides an at-a-glance synopsis of relevant information to physicians.

The evaluation of the prototypical implementation of the visualization concept by a group of neurosurgeons revealed positive feedback, in particular concerning the use of anatomy as spatial representation of data and the potential speed-up of information assessment. The current prototype covers the application scenarios in the context of distributed clinical cooperation related to the diagnosis of spinal disc herniation. Still, having a sufficient level of abstraction, the visualization concept is also transferable to other cooperative tasks and diagnoses, where a) data with anatomical reference are used, and b) a quick patient status overview is beneficial. For a better integration in the existing workflow, each specific implementation would require field studies and participation of medical experts in the design process, particularly in regard to selection of task-specific medical data and intuitive visual encoding.

Chapter 5

Personalization of Anatomical Model

This chapter describes a personalization option of the anatomically integrated visualization, presented in Sec. 4. As outlined in Sec. 3.2, the generic anatomical model can be adapted in order to look similar to the respective real person and thus to make the relation of a data record to the associated patient more intuitive. The personalization process comprises 3D reconstruction of a patient body model and model fitting, as described in Sec. 3.2.3.

One of the main overall requirements to the VICC concept is the capability of being smoothly integrated in the current hospital workflow without the necessity of an additional time investment on the physicians' side, which implies two specific requirements to the model personalization:

R1 Range data acquisition using *mobile device*, in order to allow the patient scanning during an usual workflow step, e.g., on the ward (cf. **DG7** in Sec. 4.2.1);

R2 *Online* scene reconstruction, in order to provide real-time feedback to the medical personnel, who executes the scanning.

The combination of **R1** and **R2** results in a challenging task of performing computationally expensive scene reconstruction steps with limited mobile hardware resources. This chapter presents an approach to mobile scene reconstruction, published in [PLK18], which addresses the aforementioned problem, particularly focusing on camera tracking enhancement. The second component of the personalization procedure can be performed, using a state-of-the-art human body model fitting method (e.g., [KYW14], an overview can be found, e.g., in [Bar+21]). However, its implementation is not a part of this thesis.

5.1   Introduction

Real-time 3D scene reconstruction from depth data is a well-established research area where several approaches have been proposed [RHHL02; New+11; Kel+13]. However, because of high computational requirements and due to the absence of highly integrated range cameras, their implementation was limited for a long time to high-end platforms, including robots, that involve PCs or laptops and rather energy intensive and bulky range cameras such as the Microsoft Kinect. Alternative approaches utilize the standard colour camera of a mobile device, i.e. smartphone or tablet, in order to extract 3D information from RGB image streams [Tan+13; OKI15; Sch+15; Mur+16]. In the last decade, however, the availability of highly integrated Time-of-Flight (ToF) depth cameras such as the Real3™ area sensor

from infineon [Inf15] integrated in ToF cameras such as the picoflex from pmd-technologies [pmd15] or in mobile devices such as Lenovo's PHAB2 Pro, makes mobile 3D reconstruction from depth data possible.

As discussed in Sec. 2.2.2, the 3D scene reconstruction problem requires the estimation of the camera's pose (position and orientation) and the reconstruction of the scene in parallel. Thus, it is structurally very similar to simultaneous localization and mapping (SLAM), which also can involve ToF cameras [May+08]. 3D reconstruction and SLAM, however, have different foci: while 3D scene reconstruction aims for high geometric quality of the reconstructed scene, in SLAM trajectory and pose estimation have the highest priority. This section mainly addresses 3D scene reconstruction, a survey of visual SLAM approaches can be found, e.g., in [FPRARM15].

Commonly, in 3D scene reconstruction the camera pose estimation is formulated as a registration problem. Given a sequence of overlapping depth maps delivered by the range camera, the camera pose is estimated by finding the best alignment between two successive depth maps (frame-to-frame registration) or between the current frame with the reconstructed model of all preceding frames (frame-to-model alignment). There are various registration approaches, e.g. ICP, RANSAC, PCA, or cross-correlation [BH95; LCD12]; see Salvi et al. [Sal+07] for an overview. In online scene reconstruction, the ICP algorithm is commonly applied.

Camera tracking based on, e.g., ICP, is a task that is computationally demanding and data intensive even in the standard, desktop 3D scene reconstruction pipelines. Since the accuracy of 3D map registration heavily depends on the size of overlapping areas and on the magnitude of relative transformation that has to be estimated, it requires both, significant CPU and GPU capabilities and high spatial and temporal resolution of the depth data. However, mobile devices with highly integrated ToF cameras comprise restricted computational resources and significantly lower temporal and spatial resolution (see Tab. 5.1), making high quality online 3D scene reconstruction hard to achieve.

On the other hand, online scene reconstruction can benefit from additional sensory information, e.g. from an IMU (see Sec. 2.3). A common approach to embed these motion data is to provide them as initial guess for geometric registration, i.e. to ICP. Corresponding pose estimates can be either obtained by mere integration of the IMU data [NDF14], or by sensor fusion, e.g. with an EKF (see Sec. 2.4), that delivers more robust predictions. However, to get a complementary data source, fusion algorithms require either a third sensory input from, e.g., a camera-based feature tracker [Dry+17], or they are restricted to the rotational component only, combining gyroscope and accelerometer measurements [Käh+15]. Some approaches apply a final fusion of the ICP results with other sensory information, which, however, is only beneficial, if the ICP pose is less reliable, leading otherwise to unwanted geometric reconstruction errors.

This chapter describes a lightweight solution for real-time 3D reconstruction on mobile devices. In order to address the challenges arising from a low temporal and spatial depth resolution and the limited computational resolution, a novel multi-sensor tracking, solely using IMU as additional sensory input, is proposed. In particular, the proposed approach comprises the following contributions:

- An adaptation of the point-based framework, as proposed by Keller et al. [Kel+13], for mobile online scene reconstruction.

- A novel EKF-based fusion approach for initialization of the ICP-based registration, which robustly estimates rotational and positional pose information,

using the IMU sensor in combination with extrapolated ICP pose estimations as virtual measurements.

- Demonstration of the approach suitability due to risk minimization of an ICP failure, particularly in case of a fast camera motion.

5.2 Related Work

The integration of the inertial tracking and 2D imagery for 3D reconstruction on mobile devices has been investigated in several prior works. In the context of scene reconstruction based on motion stereo, inertial tracking has been used for scene scale estimation and bridging the temporal gap between the two camera frames which are asynchronously acquired [Tan+13] or as initial guess of the current pose within the scope of a photo-consistency-based tracking [OKI15]. In [Mur+16] an online estimation of the relative rotation between two camera frames by integration of only gyroscope data, combined with the information from a 2D feature point tracker, is used for an offline structure-from-motion algorithm. Other approaches use visual inertial odometry (VIO) as filter-based fusion of inertial and feature point tracking data, e.g. VIO from the Tango library [Sch+15].

In the context of scene reconstruction from depth data, several approaches use additional sensory information, such as inertial tracking and visual tracking, in order to improve camera pose estimation.

Kähler et al. [Käh+15] estimate rotational and translational components of the camera pose separately from different tracking sources. The IMU data in connection with a fusion algorithm are used for orientation estimation whereby ICP and the colour-based tracker only optimize position. This way the rotational drift is reduced.

Klingensmith et al. [Kli+15] incorporate VIO from a Tango mobile device, which fuses inertial data with visual odometry, based on feature tracking using a fish-eye camera. The low depth frame rate of the range camera of $\approx 3 - 6$ Hz does not allow for standard dense ICP-based pose estimation. Thus, the ICP is initialized with the pose from Tango VIO and refined incrementally. The approach is designed for large scale scene reconstructions with low to moderate reconstruction resolution of $2 - 3$ cm and yields camera drift of $\approx 5m$ in the case of long trajectories (≈ 175 m). In [Dry+17] the authors further improved their approach by applying a corrective transformation to the Tango-based pose estimation for ICP initialization in order to compensate the VIO drift. The correction term is updated with each depth frame using the difference between the initial guess and the final ICP pose.

Huai et al. [HZY15] combine inertial data with ICP and SIFT odometry on a heavy-weight mobile platform consisting of a notebook and a Microsoft Kinect range camera. The integration of IMU readings provides a predicted camera orientation, which is used for initialization and validity check of the ICP. In case ICP fails, SIFT odometry is used. If both ICP and SIFT odometry fail, the incremental motion from the inertial tracking is used as final pose estimation. An EKF is used to correct the IMU-based predictions, where ICP or SIFT odometry serve as measurements for position correction, whereas the states with a small acceleration are utilized for roll and pitch correction by means of the gravity direction from the accelerometer.

In the context of SLAM, Chow et al. [Cho+14] describe a 3D terrestrial LiDAR system that integrates a MEMS IMU and two Microsoft Kinect sensors to map indoor urban environments in a stop-and-go fashion. The pose estimation is achieved by an implicit iterative extended Kalman filter (IEKF) that predicts poses integrating IMU

data and gets measurements from visual tracking. The latter uses alternatively ICP-based point cloud matching initialized by sparse 3D feature point matching (during the "going" state), LiDAR data (during the "stopped" state) or 5-point monocular visual odometry (VO) in highly textured areas or regions with few depth features. The IMU-based pose predictions are also used for initialization of visual tracking.

At the time of the conception and implementation of this VICC component, the approaches solely relying on IMU data as additional sensory input were restricted to non-lightweight platforms and non-integrated range cameras such as the Microsoft Kinect. Niessner et al. [NDF14] combine the IMU of a mobile device with a Kinect depth sensor on a laptop platform. The inertial tracking integrates IMU data without sensor fusion and provides an estimate of the transformation between last and current camera pose that is used for ICP initialization. The authors demonstrate that their method reduces the number of ICP iterations and makes the tracking more robust in scenarios such as fast camera motion or scanning of planar surfaces. Hervier et al. [HBG12] propose a general framework for the fusion of ICP-based pose tracking with data from motion sensors by means of an invariant EKF. The approach uses ICP noise covariances estimates on the basis of the Fisher matrix which allows detection of unobservable directions and prevents that information along these directions flows from ICP "measurements" in a-posteriori pose estimates. The pose prediction from motion data is used to initialize the ICP. Based on this framework, an implementation with a Kinect camera and a tri-axial gyroscope is described and experimentally tested. Camurri et al. [Cam+15] combine ICP-based camera tracking and inertial tracking in order to solve the SLAM problem for legged robots in a three-folded manner: the pose predicted by means of IMU data is used for ICP initialization, for validity check of the ICP result and, finally, for correction of ICP pose estimates by replacing roll and pitch.

Refocusing on online 3D scene reconstruction, i.e. using additional sensory information for the initialization of the geometric registration, it can be concluded that the then actual approaches either use integrated IMU data as ICP initialization (which requires high temporal depth resolution) or use Kalman filtered IMU for orientation estimation only. More sophisticated approaches use a second sensory input, e.g. a fisheye-based feature tracker, in order to obtain robust position and orientation estimates via Kalman filtering. In contrast to this prior work, the present approach requires only IMU as additional sensor information in conjunction with a mobile ToF camera with low spatial and temporal resolution and applies a fusion algorithm including extrapolated ICP poses as virtual EKF measurements in order to obtain a robust ICP initialization.

5.3 Implementation of Online Scene Reconstruction on a Mobile Device

The 3D scene reconstruction, implemented in the VICC personalization component, follows the generic reconstruction pipeline, as described in Sec. 2.2.2. The use of mobile devices, however, implies an additional challenge regarding the camera tracking. In particular, since the ICP algorithm solves a highly non-linear optimization, its successful convergence heavily depends on a good initial guess for the camera pose. In the case of high temporal and spatial resolution range cameras such as the Microsoft Kinect, a simple initialization, such as the identity matrix [New+11], is fully sufficient. In the context of the VICC implementation, i.e. for a low temporal and spatial range camera resolution (cf. Tab. 5.1), this leads to misalignments, i.e. the

optimization results in a local minimum, which yields significant geometric artefacts (e.g. ghost geometries). A solution for the tracking problem is presented in the next section.

The implementation basically uses the point-based online scene reconstruction framework proposed by [Kel+13; Lef+17]. Due to its memory efficiency, this approach does not impose specific modifications with respect to memory or algorithmic layout when porting to a mobile platform. Still, there are some adaptions required in order to use it in the case of range cameras with low temporal and spatial resolution, such as the picoflexx [pmd15]. Highly integrated ToF cameras are designed for near-range and wide-field-of-view types of applications. The wide field of view leads to severe camera distortions which need to be accounted for by camera calibration and in the pre-processing of the range maps. Furthermore, wide field of view in combination with low spatial resolution leads to points with larger sizes and higher noise even at moderate camera-to-object distances and, in average, to fewer model point observations. This effect needs to be taken care of by adapting the system thresholds for, e.g., the confidence counter and the ICP convergence.

5.4 Multi-sensor Camera Tracking with EKF

ICP is known to produce accurate estimations on the condition that transformations between successive camera poses are relatively small. However, larger transformations can lead to false results due to convergence in a local minimum, or even to a failure due to false correspondences. On a mobile system the probability of such critically large frame-to-frame transformations increases because of a lower range camera frame rate and longer processing times due to limited computational resources. The latter may lead to the situation that not all camera frames can be processed in real time, which results in frame dropping and, consequently, reduces the effective frame rate even more. The proposed camera tracking approach improves robustness by incorporating inertial sensor data in order to predict the camera pose that is used as ICP initialization. As a consequence, it speeds up convergence as it allows reducing the amount of image pyramid levels.

Since integrating noisy IMU data quickly leads to error accumulation, the results of inertial tracking are commonly fused with data from another tracking source, e.g. from a feature point-based visual odometry. An often used sensor fusion algorithm is EKF. The present approach, however, deliberately does not incorporate any additional tracking device in order to keep hardware efforts as low as possible. Alternatively, it takes advantage of the available ICP camera pose estimations under the assumption of moderate change rates in human motion. In particular, future camera poses are predicted by extrapolating ICP pose estimates from previous frames and then are taken as "virtual measurements", which are fused with the IMU integration results using an EKF with the objective to compensate the errors due to the high-frequency IMU noise.

Consequently, the proposed EKF approach aims at estimating ICP results, which are considered as true states. Since the true states corresponding to the camera frame i are known, it is preferable to reset the filter to these values and estimate the next frame-to-frame transformation $T_{(i+1)\rightarrow i}$ only, thus keeping the error accumulation as small as possible. However, due to a processing delay in the reconstruction pipeline (see Fig. 2.2) the last ICP result T_i^{ICP} is not available at t_i, i.e. at the beginning of the next inter-frame interval $i \rightarrow (i+1)$. This issue is addressed by splitting the

rotational part of the camera pose R_{i+1} in the prior frame orientation R_i (base orientation) and the current frame-to-frame rotation $R_{(i+1)\to i}$, i.e. $R_{i+1} = R_i \cdot R_{(i+1)\to i}$. The EKF starts for each interval $i \to (i+1)$ with zero frame-to-frame rotation and translation and with an estimate $\widetilde{R}_i$ without waiting for the ICP result. These states are propagated by integration of incoming IMU data within the interval. At t_{i+1}, i.e. at the end of the interval, the true previous orientation R_i^{ICP} is always available and can be used in the measurement vector that is passed to the EKF for correction of the predicted states. Due to the split representation of the orientation, it is possible to weight each of its components separately by setting appropriate values in the noise covariance matrices and, thus, to control their influence on the corrector. A generic scheme of the tracking workflow is shown in Fig. 5.1.
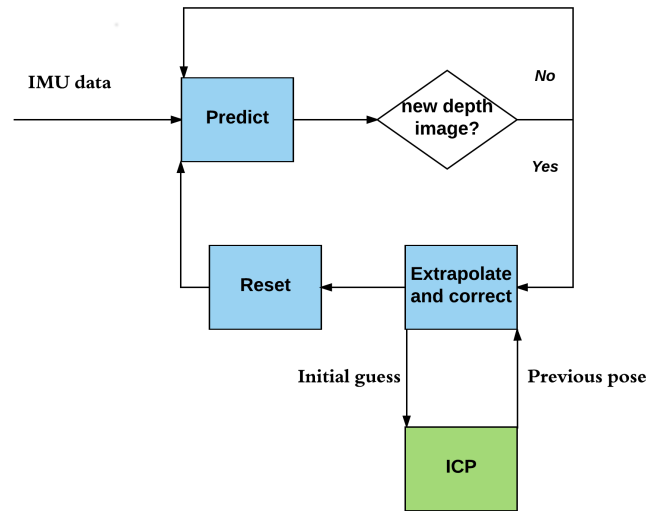
FIGURE 5.1: Scheme of camera tracking. The components of the initial guess estimation module are highlighted in blue, the ICP-based pose estimation module is highlighted in green.

The (partial) filter reset after each inter-frame interval and the weighting of the orientation components according to their confidences reduces errors in the a-posteriori estimates despite multiple prediction steps without corrections. Furthermore, executing the more computationally expensive correction step only once per frame supports a high tracking performance. In the following, the EKF design is described in more detail.

5.4.1 System State and Prediction

In the following orientations are represented as unit quaternions $\mathbf{q} = (s, \mathbf{v}^T)^T$, where s is the scalar and $\mathbf{v}$ the imagery vectorial component. Below, all quaternions are assumed to be normalized although the normalization step is omitted for simplicity. Regarding the time step indexing, k corresponds to the IMU steps whereas the i refers to ICP frame index.

The state vector is defined as

$$\mathbf{x} = [\mathbf{q}\ \mathbf{v}\ \mathbf{p}\ \mathbf{q_b}] \tag{5.1}$$

where $\mathbf{q}$ is the current frame-to-frame rotation, $\mathbf{v}$ is the velocity, $\mathbf{p}$ is the translation relating to the last camera position (both in global coordinates) and $\mathbf{q_b}$ is the estimated last orientation (base orientation).

Let

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \tag{5.2}$$

be the discrete-time non-linear process function where $\mathbf{u}_k$ is the control parameter vector and $\mathbf{w}_k$ is the process white noise with normal distribution. $\mathbf{u}_k$ consists of angular velocity $\boldsymbol{\omega}_k$, acceleration $\mathbf{a}_k$, gyroscope bias $\hat{\mathbf{b}}_g$ and accelerometer bias $\hat{\mathbf{b}}_a$, all expressed in local coordinates. The prediction, i.e. state propagation

$$\hat{\mathbf{x}}_{k+1}^- = \mathbf{f}(\hat{\mathbf{x}}_k^+, \mathbf{u}_k, \mathbf{0}) \tag{5.3}$$

is specified with the following difference equations:

$$\hat{\mathbf{q}}_{k+1}^- = \hat{\mathbf{q}}_k^+ + \frac{1}{2}\Delta t_k \hat{\mathbf{q}}_k^+ \otimes (0, (\boldsymbol{\omega}_k - \hat{\mathbf{b}}_g)^T)^{T1} \tag{5.4}$$

$$\hat{\mathbf{v}}_{k+1}^- = \hat{\mathbf{v}}_k^+ + \Delta t_k(\mathrm{qrot}(\hat{\mathbf{q}}_{\mathbf{b},k}^+ \otimes \hat{\mathbf{q}}_k^+, \mathbf{a}_k - \hat{\mathbf{b}}_a) - \mathbf{g}) \tag{5.5}$$

$$\hat{\mathbf{p}}_{k+1}^- = \hat{\mathbf{p}}_k^+ + \Delta t_k \hat{\mathbf{v}}_k^+ \tag{5.6}$$

$$\hat{\mathbf{q}}_{\mathbf{b},k+1}^- = \hat{\mathbf{q}}_{\mathbf{b},k}^+, \tag{5.7}$$

where $\cdot^-$ refers to the a-priori and $\cdot^+$ to the a-posteriori estimate, Δt_k is the current sample step, $\mathbf{g}$ is the acceleration due to gravity, $\otimes$ represents quaternion multiplication, and $\mathrm{qrot}(\mathbf{r}, \mathbf{s})$ describes rotation of vector $\mathbf{s}$ by unit quaternion $\mathbf{r}$ (cf. also Sec. 2.3, Eq. 2.10-2.12).

In order to reduce the filter complexity and to increase performance, it is assumed that the gyroscope and accelerometer biases are constant over time and thus are not included in the state vector. Periodical filter resets allow avoiding large error accumulation despite this simplifying assumption. During a short steady phase at the beginning of each experiment, $\hat{\mathbf{b}}_g$ is estimated by averaging angular velocity measurements over this time interval. Analogously, the averaging across accelerometer readings produces an estimate of the acceleration due to gravity affected by the accelerometer bias in the reference (start) frame $\tilde{\mathbf{g}} = \mathbf{g} + \hat{\mathbf{b}}_a^{\mathrm{ref}}$. Thus, the object acceleration in global coordinates can be approximated rotating the measured local acceleration $\mathbf{a}_k$ by means of the current estimated device orientation $\hat{\mathbf{q}}_{\mathbf{b},k}^+ \otimes \hat{\mathbf{q}}_k^+$ and subtracting $\tilde{\mathbf{g}}$. Thus, the right part of the Eq. 5.5 can be rewritten as:

$$\mathrm{qrot}(\hat{\mathbf{q}}_{\mathbf{b},k}^+ \otimes \hat{\mathbf{q}}_k^+, \mathbf{a}_k - \hat{\mathbf{b}}_a) - \mathbf{g} =$$
$$= \mathrm{qrot}(\hat{\mathbf{q}}_{\mathbf{b},k}^+ \otimes \hat{\mathbf{q}}_k^+, \mathbf{a}_k) - \hat{\mathbf{b}}_a^{ref} - \mathbf{g}$$
$$= \mathrm{qrot}(\hat{\mathbf{q}}_{\mathbf{b},k}^+ \otimes \hat{\mathbf{q}}_k^+, \mathbf{a}_k) - \tilde{\mathbf{g}}.$$

The propagation of the error covariance matrix P_k is done according to the standard EKF formulation, as described in Sec. 2.4, Eq. 2.20-2.21.

At t_0 (the beginning of scene reconstruction) the filter is initialized with zero velocity, translation and orientations. In the performed experiments, the error covariance matrix has been set to

$$P_0 = 10^{-4} \cdot Id_{14}. \tag{5.8}$$

---

[1]Note that this equation had an error in [PLK18], missing the first term on the right-hand side, i.e. $\hat{\mathbf{q}}_k^+ +$.

As soon as new IMU values are available, the system state vector (Eq. 5.4-5.7) and the error covariance matrix (Eq. 2.20) are propagated. Since the filter correction is executed only for the last sample of each estimated interval, the intermediate steps are given as:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- \text{ and } P_k^+ = P_k^-.$$

5.4.2   Measurement model

The measurement vector is defined as:

$$\mathbf{z} = [\mathbf{q}^e \ \mathbf{v}^e \ \mathbf{p}^e \ \mathbf{q}_\mathbf{b}^{\text{ICP}}], \tag{5.9}$$

where $\cdot^e$ represents extrapolated equivalents of the state vector components, while $\mathbf{q}_\mathbf{b}^{\text{ICP}}$ is the previous ICP camera orientation.

In order to calculate $\mathbf{z}_i$, first the velocity $\tilde{\mathbf{v}}_i$, the acceleration $\tilde{\mathbf{a}}_i$ and the angular velocity $\tilde{\omega}_i$ (all in global coordinates) are estimated with a backward Euler method.

$$\tilde{\mathbf{v}}_i = \begin{cases} (\mathbf{p}_i^{\text{ICP}} - \mathbf{p}_{i-1}^{\text{ICP}})/\Delta t_{i-1}, & \text{if } i > 0 \\ \mathbf{0}, & \text{else} \end{cases} \tag{5.10}$$

$$\tilde{\mathbf{a}}_i = \begin{cases} (\tilde{\mathbf{v}}_i - \tilde{\mathbf{v}}_{i-1})/\Delta t_{i-1}, & \text{if } i > 1 \\ \mathbf{0}, & \text{else} \end{cases} \tag{5.11}$$

$$\tilde{\mathbf{q}}_{\omega,i} = \begin{cases} \frac{2}{\Delta t_{i-1}}(\mathbf{q}_i^{\text{ICP}} - \mathbf{q}_{i-1}^{\text{ICP}}) \otimes \mathbf{q}_{i-1}^{*\text{ICP}}, & \text{if } i > 0 \\ (1, \mathbf{0}^T), & \text{else} \end{cases}, \tag{5.12}$$

where $\mathbf{p}^{\text{ICP}}, \mathbf{q}^{\text{ICP}}$ are the translational and rotational part of the respective ICP camera pose estimates, Δt_i is the time between ICP frames i and $i + 1$, and $\tilde{\mathbf{q}}_\omega = (0, \tilde{\omega})$ contains the angular velocity $\tilde{\omega}$ as its imagery vector component.

After that, the measurements are calculated with an explicit Euler method:

$$\mathbf{q}_{i+1}^e = \mathbf{q}_i^{*ICP} \otimes (\mathbf{q}_i^{ICP} + \frac{1}{2}\Delta t_i \tilde{\mathbf{q}}_{\omega,i} \otimes \mathbf{q}_i^{ICP})^2 \tag{5.13}$$

$$\mathbf{v}_{i+1}^e = \tilde{\mathbf{v}}_i + \Delta t_i \tilde{\mathbf{a}}_i \tag{5.14}$$

$$\mathbf{p}_{i+1}^e = \Delta t_i \tilde{\mathbf{v}}_i. \tag{5.15}$$

The measurement function $\mathbf{h}(\cdot)$ is linear and only copies the a-priori estimates:

$$\mathbf{h}(\hat{\mathbf{x}}_{k+1}^-) = \hat{\mathbf{x}}_{k+1}^-. \tag{5.16}$$

5.4.3   Noise Modelling

Besides bias, whose treatment is explained in 5.4.1, the IMU measurements are perturbed by white noise $\mathbf{w}$, which is commonly modelled as Gaussian [TW04]. $\mathbf{w}$ is considered in the process noise covariance matrix Q.

The noise in the virtual measurements arises from uncertainties in the estimation of the motion parameters (Eq. 5.10-5.12) used by the extrapolation. More precisely, the linear and angular velocities, estimated over an interval $(i - 1) \rightarrow i$, and the acceleration over $(i - 2) \rightarrow i$, are assumed to be constant over the next interval

---

[2]Note that this equation had an error in [PLK18], missing the second term on the right-hand side, i.e. $\mathbf{q}_i^{ICP}+$.

$i \rightarrow (i + 1)$, which generally is an approximation and not the true state. Having the camera pose T_{i+1}^{ICP}, the true values $\mathbf{v}_i$, $\mathbf{a}_i$ and $\boldsymbol{\omega}_i$ can be calculated retrospectively, substituting the left parts of the Eq. 5.13-5.15 with ICP results and solving them for the respective motion parameters. Then, the current deviations from the true values, normalized by the time intervals, can be obtained as

$$\Delta \mathbf{v}_i = (\tilde{\mathbf{v}}_i - \mathbf{v}_i)/\Delta t_i \tag{5.17}$$

$$\Delta \mathbf{a}_i = (\tilde{\mathbf{a}}_i - \mathbf{a}_i)/\Delta t_i \tag{5.18}$$

$$\Delta \boldsymbol{\omega}_i = (\tilde{\boldsymbol{\omega}}_i - \boldsymbol{\omega}_i)/\Delta t_i. \tag{5.19}$$

The above calculations have been applied to the outcomes of the three scene reconstruction experiments in order to estimate the distributions of $\Delta \mathbf{v}$, $\Delta \mathbf{a}$ and $\Delta \boldsymbol{\omega}$ (see Fig. 5.2A-5.2B for two examples). The results demonstrate that they follow a Gaussian distribution with zero mean, which justifies modelling them as white Gaussian noise.

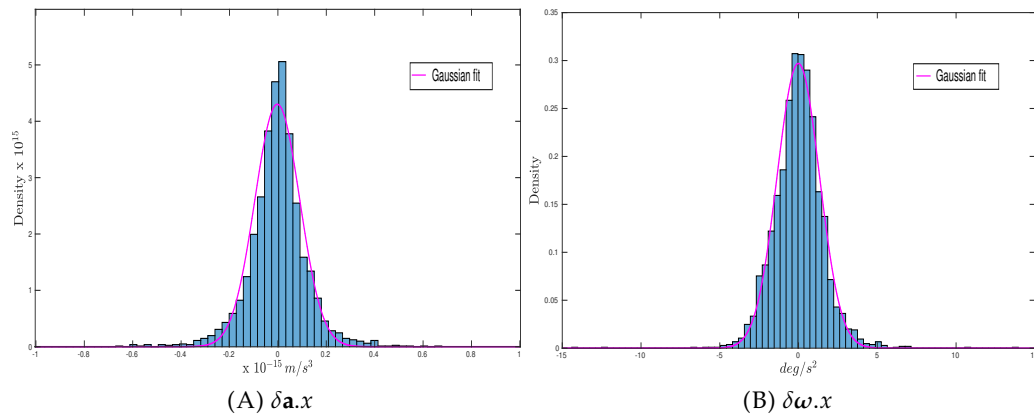(A) $\delta \mathbf{a}.x$ (B) $\delta \boldsymbol{\omega}.x$

FIGURE 5.2: Examples of estimated pdf of virtual noise.

Even though both noise covariance matrices are conceptually related to real-world physical noise measures, in the present approach they are treated as filter tuning parameters by setting them as constant, diagonal matrices. The matrix values are determined experimentally by manual optimization with respect to filter response and noise rejection as described hereafter. This kind of simplification is repeatedly applied in the context of inertial navigation, e.g. [Cav+14], and it yields satisfactory precision at minimal computational costs, i.e. minimal time delays for the overall system (see Sec. 5.5.2).

The values of the noise covariance matrix Q are determined taking the following considerations into account. The integration of angular velocity (after bias subtraction) delivers rather precise orientation estimations. The velocity estimation, and, even worse, the position estimates, are less precise due to the high-frequency noise in the acceleration values and the double integration. After manual optimization, the Q's diagonal matrix elements that correspond to $\mathbf{q}$, $\mathbf{v}$, $\mathbf{p}$ and $\mathbf{q_b}$ are set to 10^{-5}, 10^{-2}, 10^{-1} and 10^{-5}, respectively. The measurement noise covariance matrix R is set in an analogous manner. Initially, some simulation experiments were conducted in order to get more insight into the precision of the ICP extrapolation values. They showed a higher precision for positional and a lower precision for rotational values. The base orientation from ICP, interpreted as true state, gets the highest confidence, i.e. the smallest noise value. After manual optimization, the R's values that correspond to $\mathbf{q}^e$, $\mathbf{v}^e$, $\mathbf{p}^e$ and $\mathbf{q_b^{\mathrm{ICP}}}$ are set to 10^{-3}, 10^{-6}, 10^{-5} and 10^{-13}, respectively.

From the perspective of accuracy and precision, more sophisticated approaches to calculate Q and R incorporating, e.g., measured IMU noise values and dynamic adaptation, are promising. However, besides the additional work load and resulting loss in temporal performance, the asymmetrical design of the proposed EKF, i.e. the unequal number of prediction and correction steps, is a major challenge for conceptual design and implementation of these kinds of filters. For instance, a commonly used adaptive approach applies noise covariance matching by means of innovation or residual covariances during each correction step [Meh70; Wan99]. These covariances are estimated within a sliding window of the n latest measurement updates, and it is assumed that the error covariance matrix P (Eq. 2.18) is constant within this temporal window [AS16]. This assumption, however, is unjustifiable in the present case, as ≈ 10 system propagations that update P (see Eq. 2.20) occur between two consecutive measurement updates.

5.4.4   Correction and Filter Reset

Once a new camera frame $i + 1$ initiates a new cycle of the reconstruction pipeline (see Fig. 2.2), the estimation of the pose transformation for the interval $i \to (i+1)$ is terminated and the result $\hat{T}^{\text{EKF}}_{(i+1)\to i}$ is available for the ICP pipeline module. At t_{i+1}, i.e. at the end of the interval, the virtual measurements are computed via ICP extrapolation (Eq. 5.13-5.15) and the EKF correction is executed. Due to the linearity of $\mathbf{h}(\cdot)$ the Kalman gain equation (cf. Eq. 2.22) can be written in a simplified form:

$$K_{k+1} = P^-_{k+1}(P^-_{k+1} + R)^{-1}. \tag{5.20}$$

Knowing the Kalman gain, the a-posteriori estimates of the system states and of the error covariance matrix (see also Sec. 2.4, Eq. 2.24-2.25) are computed as

$$\hat{\mathbf{x}}^+_{m_i} = \hat{\mathbf{x}}^-_{m_i} + K_{m_i}(\mathbf{z}_{m_i} - \hat{\mathbf{x}}^-_{m_i}) \tag{5.21}$$

$$P^+_{m_i} = (I - K_{m_i})P^-_{m_i}, \tag{5.22}$$

where m_i is the last IMU sample index in the current inter-frame interval $i \to (i+1)$.

Than the current camera pose estimate is determined as

$$\hat{\mathbf{p}}^{\text{Init}}_{i+1} = \mathbf{p}^{\text{ICP}}_i + \hat{\mathbf{p}}^+_{m_i} \tag{5.23}$$

$$\hat{\mathbf{q}}^{\text{Init}}_{i+1} = \mathbf{q}^{\text{ICP}}_i \otimes \hat{\mathbf{q}}^+_{m_i}. \tag{5.24}$$

After the submission of the pose frame-to-frame component to the ICP module, the EKF is reset for the next inter-frame interval $(i+1) \to (i+2)$ as

$$\hat{\mathbf{x}}_{0_{i+1}} = \mathbf{g}(\hat{\mathbf{x}}^+_{m_i}, \mathbf{q}^{\text{ICP}}_i), \tag{5.25}$$

where $\mathbf{g}$ is the reset function that re-initializes the state vector as follows:

$$\hat{\mathbf{q}}_{0_{i+1}} = (1, \mathbf{0}^T) \tag{5.26}$$

$$\hat{\mathbf{v}}_{0_{i+1}} = \hat{\mathbf{v}}^+_{m_i} \tag{5.27}$$

$$\hat{\mathbf{p}}_{0_{i+1}} = \mathbf{0} \tag{5.28}$$

$$\hat{\mathbf{q}}_{\mathbf{b},0_{i+1}} = \mathbf{q}^{\text{ICP}}_i \otimes \hat{\mathbf{q}}^+_{m_i}. \tag{5.29}$$

Consequently, the errors in the frame-to-frame translation and rotation are not propagated to the next interval and the base orientation includes only an error from the

last frame-to-frame rotation estimate. Finally, the error covariance matrix P is re-initialized:

$$P_0 = J_g(\hat{\mathbf{x}}_{m_i}^+) P_{m_i}^+ J_g(\hat{\mathbf{x}}_{m_i}^+)^T, \tag{5.30}$$

where

$$J_g(\hat{\mathbf{x}}_{m_i}^+) = \left. \frac{\partial \mathbf{g}}{\partial \hat{\mathbf{x}}_{m_i}^+} \right|_{\mathbf{q}_i^{\text{ICP}}} \tag{5.31}$$

is the Jacobian of $\mathbf{g}(\hat{\mathbf{x}}_{m_i}^+)$. Since the part that relates to the frame-to-frame translation and rotation is cleared to zero, the corresponding diagonal values are reinitialized like in Eq. 5.8.

5.4.5 Synchronization

Although incoming depth images and IMU data are provided with timestamps, the respective clocks are unsynchronized. In the present implementation a simple synchronization approach considering relative times between samples is used. The EKF is initialized with the first incoming depth image. The arrivals of subsequent depth images serve as synchronization events. At each such synchronization event, the time Δt_D between two last depth image timestamps (camera inter-frame time) and the integration time Δt_{EKF} of the EKF, accumulated between the events provoked by the corresponding images, are compared.

If $\Delta t_{\text{EKF}} > \Delta t_D$, the number of integration steps is cut off to match Δt_D. If $\Delta t_{\text{EKF}} < \Delta t_D$, we cut off the corresponding IMU raw data at the beginning of the next interval.

5.5 Results

5.5.1 Experimental setup and evaluation criteria

The system described above is implemented on a Tango Yellowstone Tablet with a NVIDIA Tegra K1 processor. The range data are provided by an external CamBoard picoflexx fixed on the tablet and connected via USB. The picoflexx provides ≈ 15 FPS at a resolution of 224×171 px. For a comparison of principle characteristics of the equipments used in the handheld system in this project and in a recent desktop solution [Lef+17], they are summarized in the Tab. 5.1.

TABLE 5.1: Key measures of a typical desktop platform used in recent publications, e.g. [Lef+17], and mobile devices (Tegra K1). Note that the performance and fill-rate figures are imperfect, as there are no standardized acquisition processes.

|	Desktop Platform	Mobile Platform
CPU Capacity: • CPU/#cores	Intel i7-4790 / 4	A15 ARM / 4
GPU Capacity: • GPU	NV GeForce GTX 980 Ti	NV Kepler
• # Cores / RAM	2816 cores / 6 GB	192 cores / 4 GB
• Perform./Fillrate	5.6 TFLOPS / 176 GTex-el/s	0.365 TFLOPS / 7.6 GTex-el/s
3D Camera: • Type	Kinect One (ToF)	picoflexx
• Spatial@temp.	512x424 px @ 30 Hz	224x171 px @ 15Hz
• Point Diameter@1.2 m dist.	2.7 mm	6.2 mm
• Noise/Variance@1.2m dist.	1.2 mm	2.8mm

In order to evaluate the achievable accuracy with the given system, a geometric resolution target consisting of bars and spaces of different widths was designed (see

Fig. 5.3). The width decreases from left to right: 22 mm, 14.5 mm, 9 mm, 6 mm, 4 mm, 2.5 mm and 1.5 mm. The target was captured from a distance of 0.5 m, which corresponds to a theoretical point diameter of ≈ 2.5 mm (see Tab. 5.1). Fig. 5.3 shows the result of the target reconstruction. The two largest spaces are clearly recognisable, the next two are partially closed and the space of 4 mm is only distinguishable as a groove without holes. The space smaller than 4 mm cannot be recognized. This result is expected, as the imaging system's optics, e.g. its point-spread function, limits the system resolution.

FIGURE 5.3: Reconstruction of the geometric resolution target. Space widths left to right: 22 mm, 14.5 mm, 9 mm, 6 mm, 4 mm, 2.5 mm and 1.5 mm.

The depth camera is calibrated against the integrated RGB camera of the tablet using standard calibration methods [Zha00]. In order to calculate the complete calibration matrix, the transformation between RGB camera and IMU from the Tango library is used.
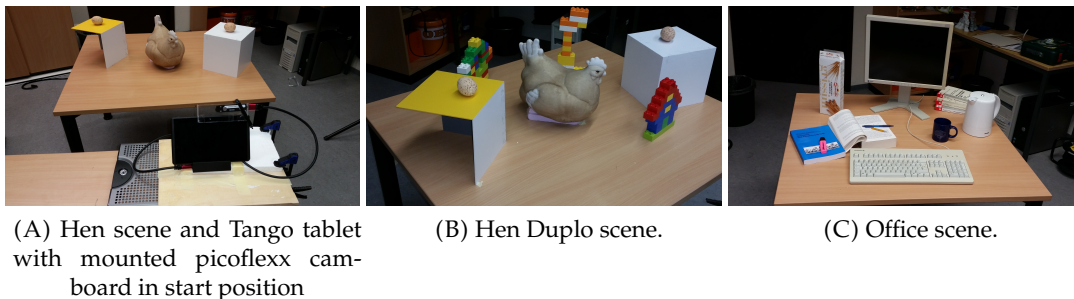
(A) Hen scene and Tango tablet with mounted picoflexx cam-board in start position

(B) Hen Duplo scene.

(C) Office scene.

FIGURE 5.4: The test scenes used for evaluation.

To demonstrate an improvement of tracking robustness by the proposed method, extreme scenarios with fast handheld camera motions and abrupt directional changes were applied. In the experiments, a data set consisting of range and IMU data was acquired for each of three different scenes used for this purpose (see Figs. 5.4A-5.4C). The first scene, with relatively sparse geometric details, is composed of a gypsum hen figure and two cubes with gypsum eggs thereon (**Hen**). For the second data set **Hen** was extended with some Lego Duplo figures (**Hen_Duplo**). The third scene shows an office desk (**Office**). The data was acquired by moving the camera around the respective scene with loop closure. In addition, the (**Hen**) scene was recorded with slower, smoother camera movements (**Hen_Slow**) in order to evaluate the system under moderate motion conditions.

The result of the proposed tracking method (**ICP+IMU**) is compared with the ICP initialized with identity matrix (**ICP**) in consideration of the following aspects.

Levels of Image Pyramid: Executing the ICP on several levels of the input image stabilizes the camera pose estimation (see Sec. 2.2.2). However, this hierarchical approach requires additional computations and memory. The results demonstrate that due to the enhanced initial guess by **ICP+IMU**, good correspondence pairs already are obtained at the full image resolution, thus decreasing the overall computation and memory requirements. In the described experiments, **l=3, l=2** and **l=1** pyramid levels are used.

Frame Dropping: The throughput of the reconstruction pipeline on the given mobile hardware is ≈ 8 FPS while the picoflexx delivers ≈ 15 depth frames per second. Thus, in real time the pipeline is unable to process all data and has to drop some frames. Such an online frame dropping depends on different internal factors of the operating system and is therefore unpredictable. In order to achieve reproducible results, a controlled frame dropping approach was used in the respective experiments. This approach processes a predefined set of depth images, independently of the processing time of the mobile device, leading to test sequences **All**, **1/2** and **1/4** that contain all, each second and each forth depth frame, respectively. The frame dropping enlarges pose transformation between consecutive processed frames and leads to similar inputs as capturing the same motion at a higher speed.

In order to evaluate the overall tracking precision and global drift errors, the camera path is set up as loop-closures, i.e. with the same start and end pose. Thus, the positional and rotational drift for each test sequence and pose estimation method can be calculated as absolute differences between the first and final estimated position and orientation, respectively (see Tab. 5.2-5.5). Large drift errors indicate unreliable pose estimations: in cases where they exceed the specified thresholds (0.15 m for positional and $10°$ for rotational error), the respective results are placed in parentheses, even though ICP itself did not fail, i.e. no matrix singularity occurred.

A reduction of the image pyramid does not automatically lead to a smaller total number of ICP iterations since a bad initial guess can slow down the convergence on lower pyramid levels. In order to demonstrate the impact of pyramid reduction on the computational effort, the mean and the standard deviation of the total iteration number are considered for each experiment.

Finally, the enhancement of the initial guess by the proposed method is evaluated, considering the difference in position and orientation between the EKF estimate, i.e. the ICP's initial pose, and the final pose after ICP (**Init_EKF**) and comparing it with the respective difference between the last and current ICP pose (**Init_Id**), which corresponds to the common ICP initialization with the identity matrix.

Besides the demonstration of an improvement over the standard ICP initialization, the present method is also compared with the Tango VIO-based initialization approach proposed in [Dry+17]. For this purpose, the respective pose from Tango tracking for each incoming depth image was saved during the **Office** scene capturing. After the scene is reconstructed, the pose predictions are calculated offline, using the stored Tango data and the ICP results. Again, differences between Tango-based predictions and ICP poses (**Init_Tango**) are obtained.

5.5.2 Evaluation

The experimental results in Tab. 5.2-5.4 show a considerable improvement of tracking robustness by **ICP+IMU**. Considering fast motion scenarios, **ICP** delivers reliable results only for one configuration, i.e. the **Hen_Duplo** scenario (Tab. 5.3) with
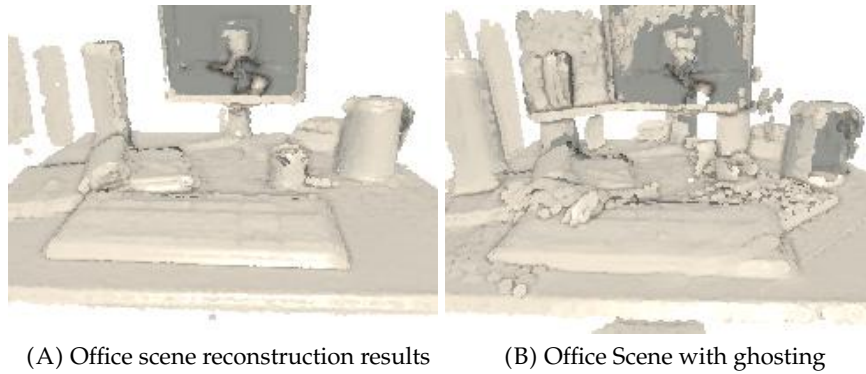
(A) Office scene reconstruction results   (B) Office Scene with ghosting

FIGURE 5.5: The **Office** scene reconstruction results. (A) reconstructed with **ICP+IMU** method, **1/2** frames, **l=1**. (B) shows ghost artefacts by the reconstruction with **ICP** method, **All** frames, **l=3**.

all pyramid levels and all frames (which is de-facto not achievable in online mode), whereas **ICP+IMU** yields valid reconstruction in almost all scenarios and configurations. Fig. 5.5B shows visual artefacts due to a large drift error in the **Office** scene reconstructed by **ICP** using **All** frames and **l=3**. The comparison Fig. 5.5A shows the more precise reconstruction results achieved with **ICP+IMU** using fewer frames (**1/2**) and pyramid level (**l=1**). Note that the artefacts on the computer screen are attributable to the general weakness of ToF cameras in capturing strongly light absorbing, i.e. dark, surfaces.

For the frame dropping rate **1/2** that is close to the online frame dropping due to the hardware limitations, **ICP+IMU** works successfully in all experiments, which demonstrate real-time capability of the proposed method. Furthermore, **ICP+IMU** preserves tracking stability with fewer pyramid levels in most cases. In particular, in all tests with the "online-like" frame dropping **1/2** the pose estimation on the basis of original image resolution (**l=1**) was possible.

On the other hand, in the experiment with a slow camera motion (Tab. 5.5), suitable for scene reconstruction with **ICP**, the application of **ICP+IMU** does not deteriorate tracking results and exhibits higher robustness under the highest frame dropping (**1/4**) with a reduced image pyramid (**l=2, l=1**).

As can be seen in Tab. 5.3 and 5.5, in those experiments where both methods have delivered reliable results, the mean iteration number in **ICP+IMU** is only marginally below the one of **ICP**. In general, however, **ICP+IMU** achieves a robust result with considerable fewer overall iterations when fewer pyramid levels are used. As expected, the drift errors are comparable in cases where **ICP** and **ICP+IMU** deliver reliable results.

TABLE 5.2: **Hen** scenario: The rotational and positional drift errors, as well as the total number of iterations for the **ICP** and the **ICP+IMU** method for different frame dropping (**All, 1/2, 1/4**) and different pyramid level (**l=3, l=2, l=1**). ∅ indicates ICP failure.

Frames	Pyr. lvl.	ICP l=3	ICP l=2	ICP l=1	ICP+IMU l=3	ICP+IMU l=2	ICP+IMU l=1
All	Rot. drift, °	(176.389)	(37.729)	∅	2.689	2.724	∅
	Pos. drift, m	(1.3352)	(0.6740)	∅	0.0122	0.0126	∅
	Iterations	(9.14±2.76)	(6.56±2.21)	∅	8.34±1.74	5.79±1.52	∅
1/2	Rot. drift, °	∅	∅	∅	2.847	2.836	2.826
	Pos. drift, m	∅	∅	∅	0.0129	0.0127	0.0126
	Iterations	∅	∅	∅	8.59±1.75	6.13±1.68	3.72±1.10
1/4	Rot. drift, °	∅	∅	∅	2.719	∅	∅
	Pos. drift, m	∅	∅	∅	0.0127	∅	∅
	Iterations	∅	∅	∅	9.24±2.16	∅	∅

TABLE 5.3: **Hen_Duplo** scenario: The rotational and positional drift errors, as well as the total number of iterations for the **ICP** and the **ICP+IMU** method for different frame dropping (**All, 1/2**) and different pyramid level (**l=3, l=2, l=1**). ∅ indicates ICP failure.

Frames	Pyr. lvl.	ICP l=3	ICP l=2	ICP l=1	ICP+IMU l=3	ICP+IMU l=2	ICP+IMU l=1
All	Rot. drift, °	8.948	∅	∅	8.944	8.954	8.963
	Pos. drift, m	0.1106	∅	∅	0.1106	0.1107	0.1107
	Iterations	8.61±1.74	∅	∅	8.46±1.74	5.74±1.39	3.34±1.01
1/2	Rot. drift, °	∅	∅	∅	8.744	8.742	8.737
	Pos. drift, m	∅	∅	∅	0.1088	0.1088	0.1087
	Iterations	∅	∅	∅	8.96±2.00	6.10±1.69	3.62±1.70

TABLE 5.4: **Office** scenario: The rotational and positional drift errors, as well as the total number of iterations for the **ICP** and the **ICP+IMU** method for different frame dropping (**All, 1/2**) and different pyramid level (**l=3, l=2, l=1**). ∅ indicates ICP failure.

Frames	Pyr. lvl.	ICP l=3	ICP l=2	ICP l=1	ICP+IMU l=3	ICP+IMU l=2	ICP+IMU l=1
All	Rot. drift, °	(44.835)	(36.385)	(38.623)	1.089	1.097	0.814
	Pos. drift, m	(0.6259)	(0.4438)	(0.4921)	0.0179	0.0178	0.0114
	Iterations	(9.63±2.33)	(6.85±2.13)	(4.13±1.6)	8.93±1.61	5.95±1.55	3.31±1.05
1/2	Rot. drift, °	∅	∅	∅	1.268	1.270	1.260
	Pos. drift, m	∅	∅	∅	0.0185	0.0186	0.0185
	Iterations	∅	∅	∅	8.97±1.50	6±1.35	3.44±0.89

TABLE 5.5: **Hen_Slow** scenario: The rotational and positional drift errors, as well as the total number of iterations for the **ICP** and the **ICP+IMU** method for different frame dropping (**All, 1/2, 1/4**) and different pyramid level (**l=3, l=2, l=1**). $\varnothing$ indicates ICP failure.

Frames	Pyr. lvl.	ICP l=3	ICP l=2	ICP l=1	ICP+IMU l=3	ICP+IMU l=2	ICP+IMU l=1
All	Rot. drift, °	2.286	2.408	2.301	2.348	2.361	2.309
	Pos. drift, m	0.0162	0.0178	0.0166	0.0169	0.0173	0.0167
	Iterations	8.44±1.60	5.82±1.38	3.44±0.99	8.35±1.70	5.55±1.30	3.23±1.47
1/2	Rot. drift, °	2.255	2.289	2.275	2.262	2.254	2.268
	Pos. drift, m	0.0146	0.0150	0.0151	0.0147	0.0147	0.0148
	Iterations	8.63±1.62	6.19±1.52	3.87±1.15	8.58±1.80	5.71±1.24	3.42±1.00
1/4	Rot. drift, °	2.323	(37.044)	(36.609)	2.262	2.254	2.313
	Pos. drift, m	0.0165	(0.4694)	(0.4721)	0.0147	0.0147	0.0160
	Iterations	9.47±2.49	(7.51±2.38)	(5.12±2.06)	8.57±1.80	5.71±1.24	4.13±1.76

Concerning the difference between the initial guess and final pose, Fig. 5.6 and 5.7 show the absolute errors for a sequence of ≈ 130 frames of the **Office** scenario with **l=1** for frame dropping **All** and **1/2**. In the vast majority of the frames there is a significant improvement by the EKF prediction. The overall relative error **Init_EKF** / **Init_Id** for the range image sequence of **All** frames is 0.5392 ± 0.6965 for the mean relative positional and 0.4556 ± 0.3697 for the mean relative rotational error. The dropping of one half of the frames (**1/2**) results in a mean relative position error of 0.4348 ± 0.4546 and a mean relative rotational error of 0.3283 ± 0.2876. Thus, for faster motion relatively better initializations could be obtained using the present method.
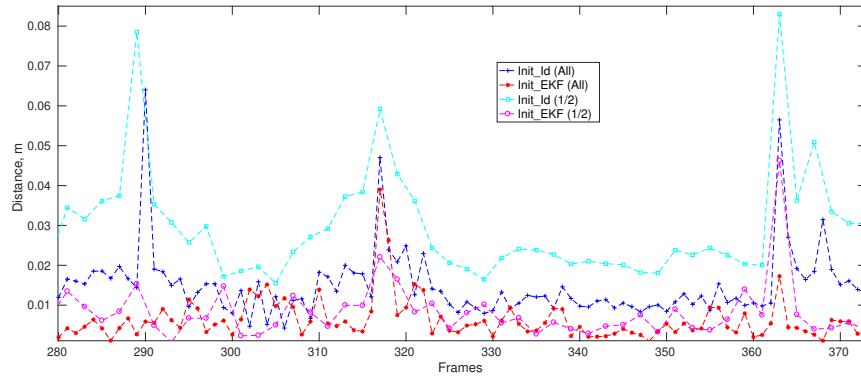

FIGURE 5.6: Distance quality of EKF estimate **Init_EKF** comparing with the standard initialization approach **Init_Id** for the **Office** scenario with frame dropping **All, 1/2**.

As a comparison, the Tango-based initialization according to [Dry+17] produces with the frame sequence **1/2** a mean relative error **Init_Tango** / **Init_Id** of 0.2926 ± 0.3147 in position and 0.4349 ± 0.4659 in orientation. The results demonstrate that the method proposed here is only slightly less precise in position prediction, and at the same time has a slightly higher precision in orientation prediction. In summary, it allows to obtain an initial guess precision comparable with the above VIO method without requiring a fisheye camera with large opening angle and high temporal and spatial resolution.
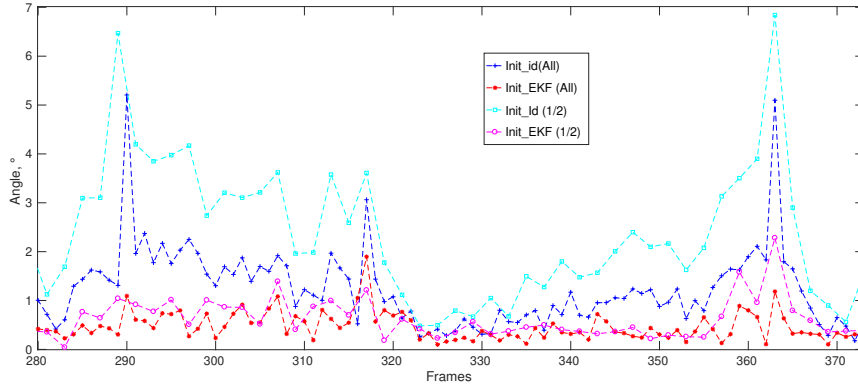
FIGURE 5.7: Angular quality of EKF estimate **Init_EKF** comparing with the standard initialization approach **Init_Id** for the **Office** scenario with frame dropping **All, 1/2**.

The processing of an individual input depth frame according to Fig. 2.2 requires ≈ 23 ms for the preprocessing stage, $20 - 25$ ms and ≈ 11 ms for the pose estimation stage in the **l=3** and the **l=1** case, respectively, $41 - 51$ ms for the depth map fusion, and $35 - 57$ ms for the surface reconstruction.

5.5.3 Limitations

Gap in sensor sampling rate A higher sampling frequency of IMU allows to bridge longer intervals between depth images. However, the integration of raw IMU data leads to a considerable sensor drift due to error accumulation (see discussion in Sec. 5.4). Although the proposed method partially corrects the integration errors by the EKF at the end of each inter-frame interval, the error increases for longer temporal gaps between successive processed depth images, causing the a-posteriori estimates of the EKF to be less precise. Furthermore, larger inter-frame intervals lead to a decreased reliability of the ICP pose extrapolations, particularly in the case of abrupt changes of motion direction. Thus, the difference between initial guess and final pose in the present method increases by higher frame dropping rates albeit, as a general tendency, remaining below this parameter in the standard approach. For example, only one "fast" data set (**Hen**) can be processed with the dropping rate **1/4** (see Tab. 5.2).

The only way to counteract this problem is to reduce the inter-frame gap by faster depth frame processing, which, in turn, requires novel reconstruction approaches.

Noise estimation A further aspect that can limit the prediction precision is low accuracy of the noise estimation. As described in Sec. 5.4, the process noise covariance matrix was tuned experimentally, however, a measurement-based noise estimation, for instance using Allan variance [ESHN08], might provide more accurate results. Concerning the noise of ICP extrapolations, an online calculation depending on the length of the inter-frame interval may lead to a more appropriate noise estimation model.

ICP failure handling A good initial guess of the current camera pose can prevent ICP errors, such as convergence in local minima or false correspondence association, related to large frame-to-frame transformations. However, a lack of geometric features, e.g. when sliding along planar geometries, also results in ICP stability

problems. In these cases the ICP and EKF estimations diverge. If this constellation appears in several successive frames, tracking failure may occur. The experiment **Hen, All, 1/2, l=1** (Tab. 5.2) shows that in this particular case even discarding frame prevents tracking failure.

A possible solution for the problem above is to discard ICP results for such frames, continuing inertial-based estimation until new reliable ICP estimates are available. This requires sophisticated ICP error detection [NDF14; HBG12]. This improvement is orthogonal to the proposed method and can be integrated in the system.

5.6   Conclusion

TABLE 5.6: Overview of the fundamental characteristics of mobile approaches for scene reconstruction from range data: required sensors, application of sensor fusion, internal model representation and system performance (value is n/a if no data are available).

	Ours	**Dryanovski et al. [Dry+17]**	**Kähler et al. [Käh+15]**
Sensors:			
• Range camera resolution	224x171	n/a	320x240
• FPS	15	5	60
• IMU	+	+	+
• Additional sensors	-	RGB Cam	-
Sensor fusion:			
• Estimated parameter	Position+orientation	Position+orientation	Orientation
• Input sources	IMU+ICP	IMU+RGB Cam	IMU
• Target	ICP initial guess	ICP initial guess	Rotational part of end pose
Scene representation / performance:			
• Type of representation	Sparse points	Spatially hashed grid	Spatially hashed grid
• Object space resolution	2.5 mm	2 cm	n/a
• Model extent	Unlimited	Unlimited	Unlimited
• FPS	9	≥ 5	47

In this chapter, a novel method for real-time 3D reconstruction on mobile devices, which is a part of the VICC personalization component, was presented. This lightweight solution uses IMU data as the only additional sensory input. Its pose estimation incorporates a novel EKF-based fusion of inertial tracking data with extrapolated ICP camera poses in order to initialize the ICP pose estimation. The experimental results demonstrate considerable enhancement of the tracking robustness in comparison with the common initialization approach. On the one hand, this method shows a higher stability in the case of a fast camera motion. On the other hand, it reliably deals with the low temporal resolution of highly integrated depth cameras such as the picoflexx. In addition, the robust ICP initialization allows to reduce the number of ICP image pyramid levels and, consequently, to achieve a higher depth frame throughput. All in all, the proposed approach minimizes the negative impact of hardware limitations existing on a mobile system and allows 3D point-based scene reconstruction for fast camera motion. Tab. 5.6 summarizes some distinctive characteristics of the approaches for scene reconstruction from range data on mobile devices, illustrating the most important conceptual differences between the present method and existing solutions.[3] Having demonstrated the feasibility of a sufficiently

---

[3]The parameter "Model extent" in this table refers to the extent limitations due to the system design. Note that the real model size is also limited by available memory.

precise 3D reconstruction using mobile devices only, this approach can be complemented with a model fitting method to produce personalized patient body models for anatomically integrated visualization.

Chapter 6

Iconic Glyphs

This chapter describes the iconic glyph approach as a visualization component of the VICC concept (see Sec. 3.2.1). This approach combines the principles of the icon-based and glyph-based visualization, at the same time, augmenting the traditional design space of glyphs with periodic contour-based visual variables. To allow the use of these kinds of visual variables for quantitative visualization, first, a suitable perception and quantization model is required. Sec. 6.1 addresses this problem, basing on the study published in [PK22]. Then, Sec. 6.2 describes the technical aspects of the approach for glyph generation from icons, published in [PBK23], comprising some application examples and their evaluation.

6.1 Perception and Quantization Model for Periodic Contour Modifications

6.1.1 Introduction

In visualization, the visual augmentation of contours is often used to extend the amount of available visual channels to represent further attributes of multidimensional data in parallel. This can be implemented, for instance, by variation of a contour's *colour*, *width* and/or *fuzziness*, i.e. degree of blurring, or by means of a periodic, wave-like contour modification, which results in additional visual variables such as *frequency*, *amplitude*, and *waveform* (see Fig. 6.1).

In Scientific Visualization, this kind of contour augmentation is applied, for example, to enhance the encoding capacity of isolines [AOB08; ZMG21]. In Information Visualization, there are a number of approaches, which create glyphs or their parts from circular shapes, modulating their contours by a (mostly sinusoidal) wave. Such applications include the visualization of uncertainty (encoded by frequency [Hol+19; Gör+17] or amplitude [Gör+17]) and sport event data (encoded by frequency [Chu+13]) as well as the generation of more complex glyphs such as Rose-Shapes [Cai+15]. Besides circles, the contour waves have been also applied to iconic shapes such as leaf icons, to represent environmental data by means of frequency and amplitude [Fuc+15].

Even if such contour waves prove to be a promising design solution, the perceptual aspects of the respective visual variables have not been thoroughly investigated. While there are several studies that address the problem of perception and discrimination of more "classical" visual variables, e.g. colour [Wij+08; SSG14], size [LMW10], or the interdependencies of both [SSS14], the work related to periodic contour modifications, to the best of my knowledge, is limited to the demonstration of the shapes' orderability by frequency [Hol+19; Chu+16], based on user

experiments, and to a rather abstract discussion about dependencies between wave amplitude and frequency without user studies [Gör+17].
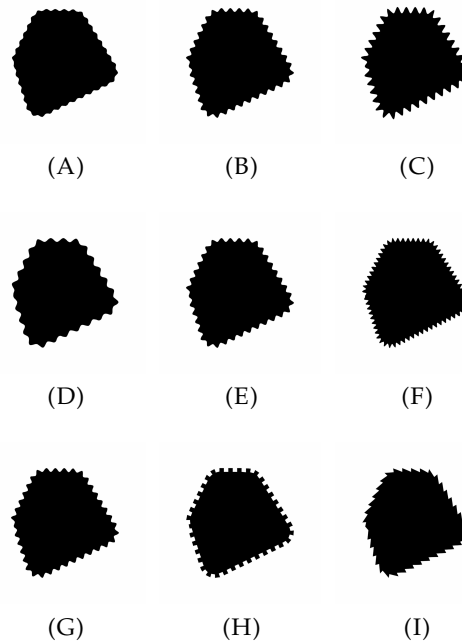

FIGURE 6.1: Examples of periodic contour modifications. (A)–(C): visual encoding by varying wave amplitude; (D)–(F): visual encoding by varying wave frequency; (G)–(I): visual encoding by varying waveform.

Motivated by this fact, in this section, a *perception and quantization model* for periodic contour modifications is proposed, which provides a basis of glyph design for visualization of *multivariate quantitative* data in general and for the iconic glyph design, described in Sec. 6.2.4, in particular. The proposed model addresses the essential properties of a quantitative visualization such as *perceptual uniformity* [Bor+13] and *accurate legibility* [War09], performing a purposefully created user study and evaluating its results. *Perceptual uniformity* signifies that the representation of equally sampled data values have to be perceived in visual space as equidistant. However, mapping data to equally distributed discrete stimulus levels does not guarantee perceptual uniformity, since the relation between stimuli and sensation is generally not linear [Ste57]. Therefore, the present model comprises an estimation of an appropriate transformation function between stimulus magnitudes and perceived magnitudes. *Accurate legibility* corresponds to the distinguishability of the levels of a visual variable that encode quantitative data [War09], which, in turn, implies a quantization, where the distances between single levels are equal or greater than the just noticeable difference (JND). To satisfy this requirement, a quantization scheme, which aims to an optimal balance between the legibility and the visual capacity of the respective visual variable, i.e., the number of values it can encode [Chu+13], has been derived.

Since the goal is a generic approach that covers a large number of potential visualization tasks in the context of VICC, different waveforms and iconic shapes instead of a single, fixed geometric primitive such as a circle are considered by derivation of the respective model. Moreover, it is taken into account that the visual effect of

contours modulated by a geometric wave can also be achieved by periodic modifications of the contours' colour components (cf., e.g., the "null-case glyph" in [Hol+19]). In particular, this study focuses on the colour modifications, created by alternation of the contours' segments with different intensity levels (see Figure 6.2). Considering both, geometric modifications and colour modifications by varying intensity, allows to obtain a more generic model of periodic contour modifications, which is evaluated in an appropriate user study.

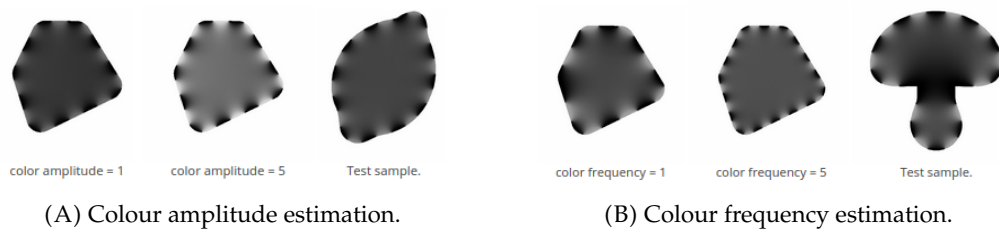(A) Colour amplitude estimation.   (B) Colour frequency estimation.

FIGURE 6.2: Design of different colour experiment types. (A),(B) show examples for direct estimation of colour amplitude magnitude (with fixed colour frequency) and colour frequency magnitude (with fixed colour amplitude), respectively; the participants need to assess the corresponding magnitude in the "Test sample" (right) on the basis of the reference shapes (left and middle).

To sum it up, this section comprises the following contributions:

- An online user study about perception of periodic geometric and colour contour modifications.

- Modelling of a stimulus-to-perception transformation function for sinusoidal and colour contour modifications.

- Analysis of perceptual dependencies between amplitude and frequency for geometry and colour, respectively.

- Evaluation of the waveform influence on the amplitude and frequency perception, including a calibration model for sinusoidal, rectangular, and sawtooth waves.

- Definition of distinguishable quantization levels for geometric and colour contour modifications.

- A method for transferring the quantization model to shapes with different sizes.

6.1.2   Related Work

6.1.2.1   Periodical Contour Modifications in Visualization

There are several recent glyph visualization approaches that use periodic contour modifications, with application in different fields. For instance, Holliman et al. [Hol+19] used modified contours of a circular glyph, with wave frequency serving as a measure of visual entropy to encode uncertainty, while the inner colour of the circle visualizes the respective mean value. Similarly, Görtler et al. [Gör+17] proposed a contour-based design space for hierarchical uncertainty visualization by means of Bubble Treemaps, which includes, among possible alternative visual variables, sine

wave frequency and amplitude as well as dashed frequency, whereby the latter can be considered as a kind of discontinuous rectangular wave. Here, mean values are encoded by circle size. Chung et al. [Chu+13] used contour wave frequency and radius of a circular silhouette as visual variables in a composite glyph for visualization of sport event analysis data.

Cai et al. in [Cai+15] followed a different approach, in which frequency, amplitude, and form of contour modifications do not serve as separable visual variables but as control parameters for construction of *unique* shapes, so-called RoseShapes, resulting from periodic functions plotted in polar coordinates.

On the contrary, in the glyph design for visualization of environmental data, developed by Fuchs et al. [Fuc+15], the original leaf shapes maintain their recognizability and meaning after contour modifications, while frequency and amplitude of the resulting serrated boundaries can be used as additional visual channels.

6.1.2.2   Studies of Perception of the Contour Modifications

Early research on visual perception stresses the role of contours as regions of a high information concentration, especially in the peaks of curvature [Att54], i.e. wave's peaks. Nevertheless, since the visualization based on periodic contour modifications is a relatively new approach, its perception and discriminative capacity have not yet been investigated in detail and existing research is limited to contours of circular shapes. Besides psychophysical work that demonstrates the ability of the human visual system to discriminate shapes on the basis of radial frequencies (e.g., [WWH98]), there are two recent visualization studies dedicated to the orderabilty issue. In particular, Chung et al. [Chu+16] investigated the suitability of specific visual channels to represent an ordinal scale. The results of their user study show that star shapes, which can be considered as circles modulated by a triangle wave with different frequencies, have an ordering. Furthermore, Holliman et al. [Hol+19] performed a user experiment to evaluate their uncertainty visualization approach and could demonstrate the orderability of circular glyphs with sinusoidal contours. Both studies used predefined frequency levels.

6.1.3   Materials and Methods

6.1.3.1   Components of the Perceptually Uniform Quantization Model

To develop a perceptually uniform quantization model of wave-like contour modifications, the following aspects are evaluated:

Stimulus-to-perception transformation function, i.e. transformation between stimulus magnitudes and perceived magnitudes. Assuming that this function follows Stevens's power law [Ste57], the corresponding parameters are estimated statistically.

Perceptual dependencies between amplitude and frequency. Considering a pair of arbitrary geometric or colour amplitude and frequency values, the goal is to investigate how changes in one parameter influence the perception of the other.

Perceptual influence of waveform for geometric amplitude and frequency. It is assumed that the waveform of a geometric contour modification influences the

perception of the respective amplitude and frequency. Thus, taken the sinusoidal shape as reference, the stimulus magnitudes for other shapes that produce the same sensation need to be acquired.

Quantization of visual variables, i.e. definition of clearly distinguishable and perceptually equidistant magnitude levels. The aim is to achieve a balance between the number of available levels and their distinguishability.

Size-dependent adaptation. Some rules for transferring the corresponding quantization to shapes with different sizes are proposed.

The expected model's outcome are perceptually equidistant levels of each visual variable for data encoding, and their transformations to stimuli magnitudes and, where appropriate, to other geometric waveforms for glyph generation. The model is mainly derived from the results of an online user survey. However, it must be mentioned that not all aspects could be addressed equally in a single user study. Firstly, the number of possible dependencies is in exponential relationship to the number of visual variables, and thus testing all of them in one study leads to an excessive experiment complexity and time exposure. Secondly, various advanced experiments imply a previous academic validation of primary test results, which are provided in this work.

Therefore, several investigations that are undoubtedly of high scientific interest could not be addressed in-depth in this study. This mainly applies to the following: (1) The interferences between amplitude and frequency in geometry as well as in colour for which, however, an initial insight is provided in Section 6.1.4.5. (2) The experimental validation of the proposed rules to transfer the derived model to shapes of different sizes (Sec. 6.1.5). (3) The dependencies resulting from a combination of geometric and colour modifications. All these aspects have to be addressed in future work.

6.1.3.2   Design of the Experiment

The test samples used in the survey are created from four monochrome base shapes with white background and black foreground (see Fig. 6.3). The geometric modifications are produced by modulation of the shape contours according to the given *geometric frequency, amplitude, and waveform*, namely sinusoidal, rectangular, and sawtooth-like (see Figs. 6.1 and 6.4). The amplitude and frequency range, used in the experiments, is accordingly narrowed as strong perceptual interferences are to be expected for extremely low and extremely high magnitudes (see [Gör+17]).
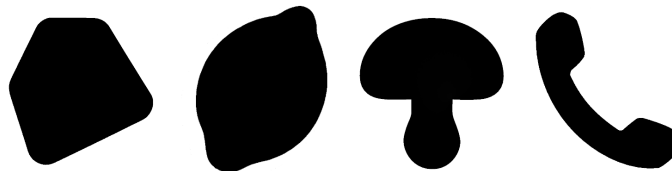

FIGURE 6.3: The four base shapes used for generation of modified contours in the experiments.

The intensity modifications result from alternating intervals of given length (i.e., inverse *colour frequency*) along the contour (see Fig. 6.2). The intensity inside each next interval changes between the current modified value (i.e., *colour amplitude*) and the original black foreground (i.e., colour amplitude = 0). Since all images used in

the experiments have white background, the maximal colour amplitude is limited to a value resulting in a light grey colour, in order to maintain contrast.

(A) Amplitude estimation.

(B) Frequency estimation.
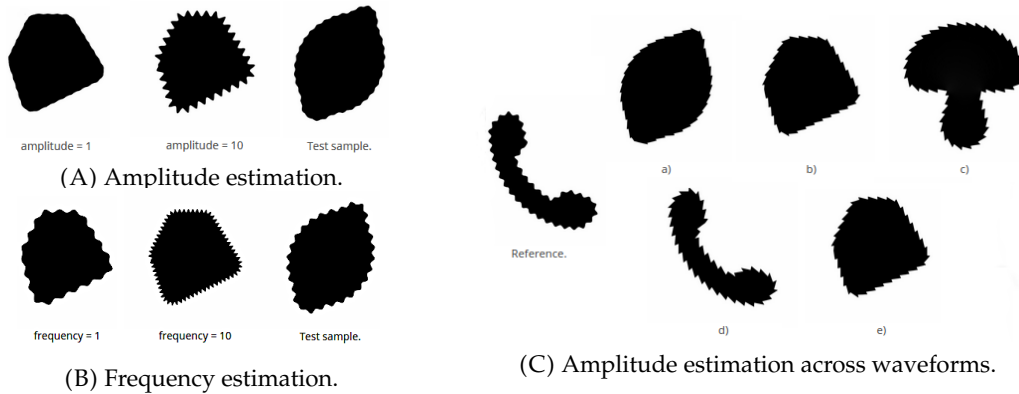
(C) Amplitude estimation across waveforms.

FIGURE 6.4: Screenshots of different geometric experiment types. (A),(B) are examples for direct estimation of amplitude (with fixed frequency) and frequency (with fixed amplitude) magnitude, respectively; the participants need to assess the corresponding magnitude in the "Test sample" (right) on the basis of the reference shapes (left and middle). (C) shows an example of a sawtooth amplitude calibration against a sinusoidal reference; the task is to select the test shape (a–e) whose amplitude is perceived as the closest to the "Reference" shape (left).

All shape images have size 512×512 px and are displayed at size 50×50 mm. Table 6.1 summarizes the metric values used to generate the stimuli and gives the mapping to the stimulus parameter values used for communication in the experiment (see also Fig. 6.4A,B). In each question, the base shape was selected randomly. The survey comprises two main categories of experiments.

TABLE 6.1: Metric values for the experiment with the glyph size 50 mm. The number in {}-brackets are the corresponding visual variable values (arbitrary digital unit, adu) used for communication in the experiment. Note that the frequency is proportional to the inverse period length. The colour amplitude values are given as value/brightness V in $[0, 1]$, in HSV colour space.

Visual Variable	min [mm]	max [mm]	step [mm]
Geom. amplitude	0.1 {1}	1.2 {12}	0.1
Geom. period length	0.8 {12}	5.1 {1}	0.4
Colour period length	5.9 {5}	12.1 {1}	1.6
	min [V]	max [V]	step [V]
Colour amplitude	0.425 {1}	0.85 {5}	0.10625

Magnitude estimation. Several magnitude estimation experiments [Ste57] were performed to determine a proper quantization of the visual variables as well as the transformation function between the stimuli and perception parameter space.

For each visual variable to estimate, the participants got displayed the available magnitude range by presenting a minimum and maximum reference shape with the corresponding stimulus parameter values in arbitrary digital unit (adu; for mapping of the metric or intensity values to the respective adu, see Table 6.1). Fig. 6.4A,B shows the design of the magnitude estimation experiments for geometry, and Fig. 6.2A,B for colour. The test shape with randomly selected magnitude

was hidden by default and was uncovered for eight seconds by clicking the corresponding button, and the participants had to assign the perceived magnitude from a drop-down list (the respective magnitude ranges available for selection are displayed in Table 6.2). The stepsize for generating the visual stimuli for the test shapes (see Table 6.1) was selected to be below a conservatively estimated JND, i.e. significantly smaller than the distance distinguishable for the experiment designers, to be able to derive a suitable quantization from a statistical evaluation.

There are two subtypes of the magnitude estimation experiments in the conducted survey (see Table 6.2):

1. Fixed second stimulus, e.g. geometric amplitude estimation with a fixed geometric frequency.

2. Randomly selected second stimulus.

TABLE 6.2: Overview of survey experiments. Each row is one experiment type, where the visual variables, i.e. the perceptional parameters to be assessed, are plotted **bold-faced**. Other parameters might be fixed values, indicated by F, or randomly selected, indicated as R. The numbers given in []-brackets are the stimuli values defined in Table 6.1. The individual experiments are of two kinds: $[\mathbf{2}..\mathbf{10}] \rightarrow [\mathbf{0}..\mathbf{12}]$, for example, generates stimuli in the value range $[\mathbf{2}..\mathbf{10}]$ and asks for assessing the perceptional values in $[\mathbf{0}..\mathbf{12}]$, while $\mathbf{x}_{|\in[\mathbf{2}..\mathbf{10}]} \rightarrow [\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$ generates stimuli values x in the range $[\mathbf{2}..\mathbf{10}]$ and asks for assessing the perceptional values in the dependent range $[\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$.

Experim.	Waveform	Geometric Amplitude	Geometric Frequency	Colour Amplitude	Colour Frequency	# exp.
Ampl1	sin.	$[\mathbf{2}..\mathbf{10}] \rightarrow [\mathbf{0}..\mathbf{12}]$	F [6]	n.a.	n.a.	6
Ampl2	sin.	$[\mathbf{2}..\mathbf{10}] \rightarrow [\mathbf{0}..\mathbf{12}]$	R [2..10]	n.a.	n.a.	20
Freq1	sin.	F [6]	$[\mathbf{2}..\mathbf{10}] \rightarrow [\mathbf{0}..\mathbf{12}]$	n.a.	n.a.	6
Freq2	sin.	R [2..10]	$[\mathbf{2}..\mathbf{10}] \rightarrow [\mathbf{0}..\mathbf{12}]$	n.a.	n.a.	20
SawtAmpl	sin.→sawt.	$\mathbf{x}_{|\in[\mathbf{2}..\mathbf{10}]} \rightarrow [\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$	F [6]	n.a.	n.a.	5
RectAmpl	sin.→rect.	$\mathbf{x}_{|\in[\mathbf{2}..\mathbf{10}]} \rightarrow [\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$	F [6]	n.a.	n.a.	5
SawtFreq	sin.→sawt.	F [6]	$\mathbf{x}_{|\in[\mathbf{2}..\mathbf{10}]} \rightarrow [\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$	n.a.	n.a.	5
RectFreq	sin.→rect.	F [6]	$\mathbf{x}_{|\in[\mathbf{2}..\mathbf{10}]} \rightarrow [\mathbf{x}-\mathbf{2}..\mathbf{x}+\mathbf{2}]$	n.a.	n.a.	5
ColAmpl1	n.a.	n.a.	n.a.	$[\mathbf{1}..\mathbf{5}] \rightarrow [\mathbf{0}..\mathbf{5}]$	F [3]	3
ColAmpl2	n.a.	n.a.	n.a.	$[\mathbf{1}..\mathbf{5}] \rightarrow [\mathbf{0}..\mathbf{5}]$	R [1..5]	6
ColFreq1	n.a.	n.a.	n.a.	F [3]	$[\mathbf{1}..\mathbf{5}] \rightarrow [\mathbf{0}..\mathbf{5}]$	3
ColFreq2	n.a.	n.a.	n.a.	R [1..5]	$[\mathbf{1}..\mathbf{5}] \rightarrow [\mathbf{0}..\mathbf{5}]$	6

The experiments with fixed second stimulus had been placed at the beginning of the specific experiment section to make the participants acquainted with the experimental setting, as the experiments with randomly selected second stimulus are more challenging.

Waveform-dependent calibration has been performed by selecting the modified shapes with the closest magnitude. To reduce the number of questions, all magnitude estimation experiments for geometric visual variables are done with the sinusoidal waveform. To estimate a waveform calibration function, the participants had to select one out of five glyphs with the perceptually most similar magnitude to a presented sinusoidal reference (see Fig. 6.4C). The modified shapes offered for selection were created with the magnitude levels $l \in [l^{\text{ref}}-2 .. l^{\text{ref}}+2]$, where l^{ref} denotes the visual variable values (adu) used for the reference shape, and have been arranged randomly. These experiments were done separately for each waveform, i.e. rectangular or sawtooth-like, and for each visual variable (see Table 6.2).

Additionally, one experiment, not listed in Table 6.2, was performed to verify the visual distinguishability between the three waveform types—sinusoidal, rectangular and sawtooth-like—for combinations of low frequencies/low amplitudes and high frequencies/high amplitudes. The recognition rates were approximately 92%, 99%, and 99% for sinusoidal, rectangular, and sawtooth-like, respectively.

The design of the experiment assumes to have "cooperative" participants, i.e. participants that will not "cheat their way through" the experiment, and that the time limit for the ability to concentrate is at most 20–25 min. Table 6.2 states the number of experiments taken per experiment type. Each participant was asked to go through 90 experiments in total.

6.1.3.3   Survey Evaluation

The persons, invited to participate in the online survey, were students and researchers mainly from the University of Siegen, from the fields of computer science and sociology, and an anonymous group of 73 persons participated. The average time to take the survey was ≈26 min. Given the raw results from the survey experiments conducted by the participants, the required stimulus-to-perception transformation, quantization, and calibration parameters were determined after having applied an outlier removal.

Outlier Removal   First, the "senseless" answers are filtered out, i.e. answers which deviate from the expected value to an extend not explainable by the subjective character of perception alone. These outliers are caused, e.g., by a misunderstanding of the respective experimental setting or by an external distraction of the participant while conducting the experiment. Such cases are handled with the two-step Chebyshev outlier detection method of Amidan et al. [AFC05], applying the filtering parameters $p_1 = 0.375$ and $p_2 = 0.175$ for all visual variables.

Modelling the Stimulus-to-Perception Transformation Function   Following Stevens [Ste57], it is assumed that the stimulus-to-perception transformation has the form of a power function (Eq. 2.1). Furthermore, to compensate the lack of a proper zero-origin due to the use of a adu-scale, the power function was extended with an additive constant, i.e. $e(\mathbf{x}) = a \cdot \mathbf{x}^b + c$. Thus, having the perceived magnitudes, stated by the participants, as data points $e(\mathbf{x})$ and the stimulus magnitudes $\mathbf{x}$ as the independent parameter, a, b, and c are estimated using non-linear least-square fitting.

Quantization   The aim is to find a quantization step Δv in perceptual space such that all resulting magnitude levels do not overlap with neighbouring confidence intervals for a given confidence level. This is analogous to the principle applied by estimation of just noticeable difference (JND), which is also defined regarding the probability of correct assignments, usually 50%: the same threshold also is applied in the experiment, described here. Table 6.3 gives an overview of the quantization steps and the resulting number of discrete levels for each visual variable.
   More precisely, the quantization step is calculated as follows:

   1. For each discrete stimulus magnitude level, observe and model the distribution of perceived magnitudes.

TABLE 6.3: Quantization results derived from the user experiment
with the image size 50×50 mm.

Visual Variable	Geom. Ampl.	Geom. Freq.	Col. Ampl.	Col Freq.
Quant. step Δv (adu)	≈ 2.91	≈ 2.01	≈ 1.23	≈ 1.14
# levels	4	5	4	4

2. Compute the 50% confidence interval, symmetrically placed about the respective mean.

3. Use the largest confidence interval as Δv.

Waveform-Dependent Calibration   The aim is to define a function that calibrates the stimulus amplitude and frequency of a rectangular or sawtooth-like waveform against the respective parameters of a sinusoidal shape. This calibration is modelled as a linear function $h(\mathbf{x}) = k \cdot \mathbf{x} + l$, with $\mathbf{x}$ being the reference sinusoidal magnitudes from the calibration experiments (see Sec. 6.1.3.2) and the perceived magnitudes regarding the rectangular or sawtooth-like waveform as data points. The fitting of h is done with a least-square method.

Evaluation of Perceptual Dependencies between Amplitude and Frequency   To evaluate the perceptual dependencies between geometric amplitude and frequency, stimulus-to-perception transformation functions of the investigated visual variable v_1 in dependence on the magnitude of the respective second visual variable v_2 are additionally estimated. To get a sufficient number of samples and taking advantage of the fact that the differences between neighbouring magnitudes are sufficiently small, v_2 is pooled in intervals. The resulting extended transformation functions for amplitude and frequency, respectively, are

$$\cdot a_p = e(\cdot a_s, \{\cdot f_s\}),$$
$$\cdot f_p = e(\cdot f_s, \{\cdot a_s\}),$$

where $\cdot$ is a placeholder for g (geometry) and c (colour), and the second function parameter is a set of adu levels of v_2, which serves as a mask, i.e. only samples where v_2 has a corresponding value are considered. For instance, $e(ga_s; \{2, 3, 4\})$ gives perception of stimulus geometric amplitude ga_s with geometric frequency $gf_s \in \{2, 3, 4\}$. The function fitting is performed in the same way as described in Sec. 6.1.3.3, "Modelling the Stimulus-to-Perception Transformation Function".

For geometry, v_2 is pooled in three intervals of three levels each. Since colour visual variables have been tested with five adu levels in the user study, the following weighted pooling scheme is applied to get three intervals again. Exemplarily, the colour frequencies are pooled as

$$e(ca_s, \{1, 1, 2\}), \quad e(ca_s, \{2, 3, 3, 4\}), \quad e(ca_s, \{4, 5, 5\}),$$

where a double occurrence of a level in a set signifies that it is considered twice, i.e. weighted with factor 2. Colour amplitudes are pooled analogously.
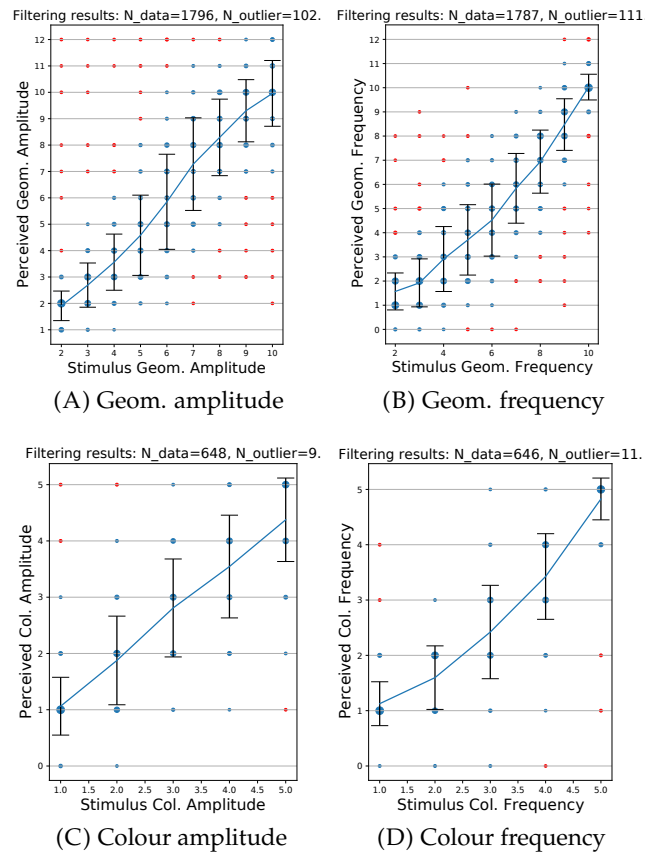
FIGURE 6.5: Results of the two-step Chebyshev outlier detection: outliers are marked in red; the point size encodes the number of occurrences.

6.1.4 Results

6.1.4.1 Outlier Removal

Table 6.4 gives an overview for all visual variables. The overall amount of removed outliers is 4.8%. The detailed statistics of the outlier removal are represented in Fig. 6.5.

TABLE 6.4: Outlier removal results.

Visual Variable	Geom. Ampl.	Geom. Freq.	Col. Ampl.	Col Freq.
# data	1796	1787	648	646
# outliers	102	111	9	11

6.1.4.2 Stimulus-to-Perception Transformation Function

The modelled functions for transformation of stimuli into perceived magnitudes are presented in Fig. 6.6 (for the estimation method, see Sec. 6.1.3.3, "Modelling the Stimulus-to-Perception Transformation Function"). For both, geometry and colour amplitude, a mainly linear and positive power dependency of the perceived magnitudes on the stimulus magnitudes can be observed, with the exponent $b = 1.0604$ and $b = 0.928$, respectively. The corresponding relationship for frequency is in both cases close to quadratic, $b = 1.7918$ and $b = 1.9463$.
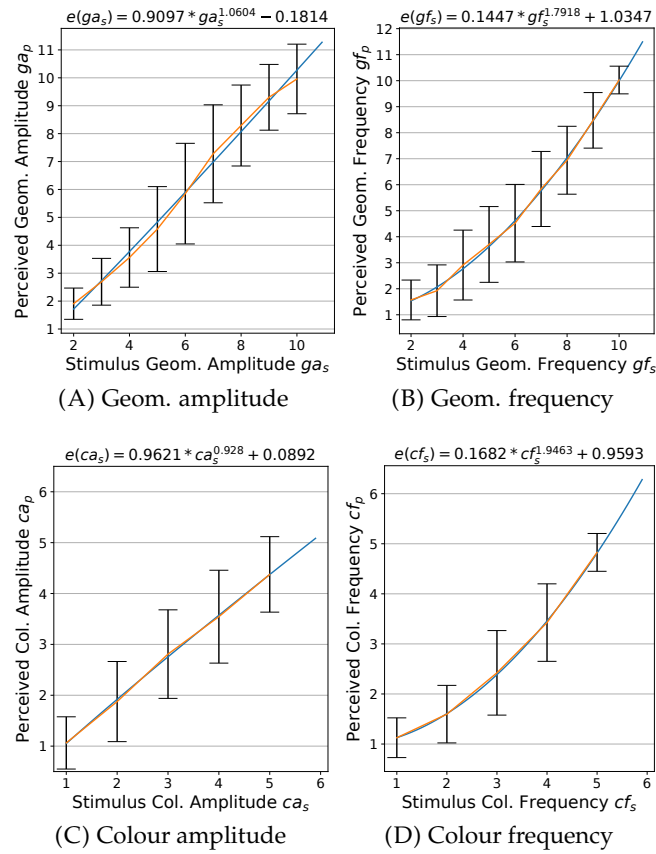
FIGURE 6.6: Modelling stimulus-to-perception transformation: blue: fitted transformation functions; orange: lines connecting perceptual means.

6.1.4.3 Quantization

The distributions of the perceived magnitudes reveal a mono-modal Gaussian nature for most magnitude levels and a bi-modal Gaussian behaviour for medium values of the geometric amplitude and frequency. The latter can be explained by a larger distance to the min. and max. references, which can be seen as a design-related phenomenon. Fig. 6.7 shows the distribution of the perceived magnitudes with the respective 50% confidence intervals for each visual variable. The 50% confidence for a bi-modal Gaussian distribution is computed by identifying the 25% and 75% limits of the cumulative distribution of the superposition of both Gaussians. For the calculation methods, see also Sec. 6.1.3.3, "Quantization".

(A) Geometry amplitude

(B) Geometry frequency

(C) Colour amplitude

(D) Colour frequency
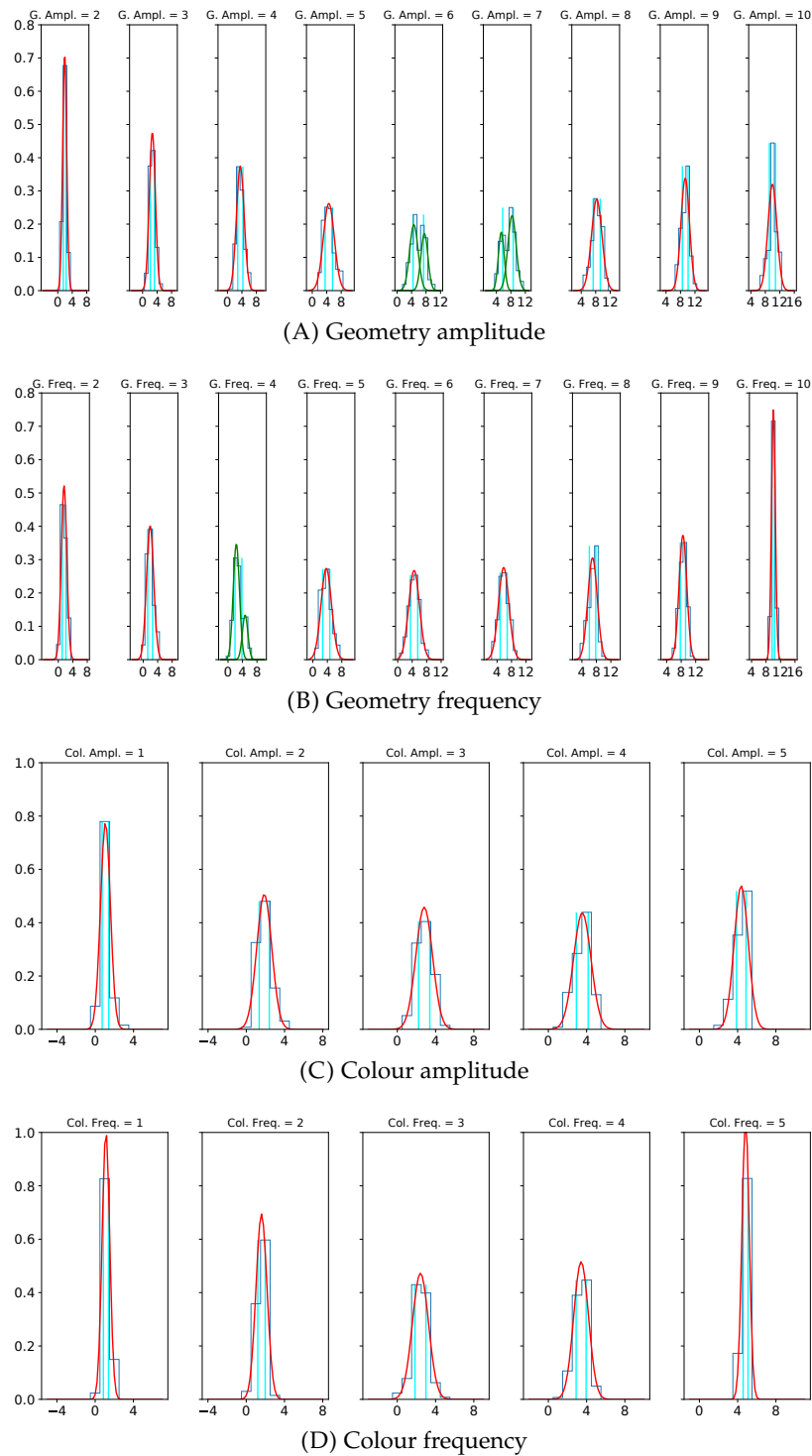
FIGURE 6.7: Perceived magnitudes as normal distributions. Red: simple Gaussian; green: two-component Gaussian mixture; cyan: borders of 50% confidence intervals.

Table 6.3 gives an overview of the quantization steps and the resulting number of discrete levels for each visual variable. It can be observed that the distinguishability of both geometric and colour frequency is slightly better, i.e. their quantization steps are smaller than in the case of the respective amplitudes.

6.1.4.4   Waveform Calibration

Fig. 6.8 shows the modelled linear calibration functions (for the estimation method, see Sec. 6.1.3.3, "Waveform-Dependent Calibration). The data demonstrate that the influence of a specific waveform on the perception of amplitude and frequency magnitudes is rather marginal, i.e. $k \approx 1$ for all four parameters.
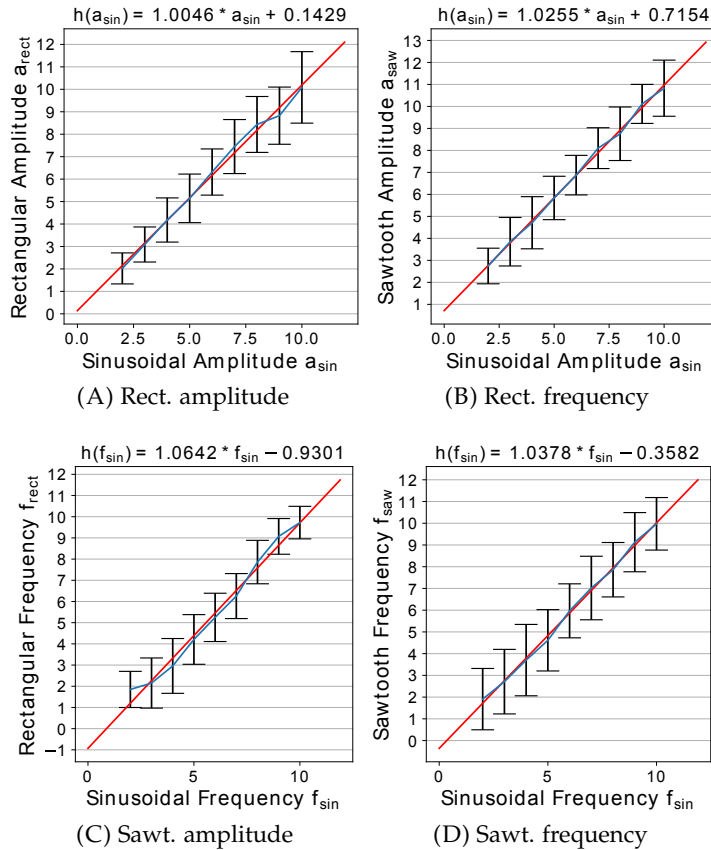
FIGURE 6.8: Calibration of rectangular and sawtooth-like waveforms against sinusoidal waveform: red: fitted calibration functions; blue: lines connecting perceptual means.

6.1.4.5   Evaluation of Perceptual Dependencies between Geometric Amplitude and Frequency

The modelled functions for transformation of stimuli to perceived magnitudes in dependence on the second visual variable are presented in Fig. 6.9 (see Sec. 6.1.3.3, "Evaluation of Perceptual Dependencies between Amplitude and Frequency" for the evaluation method). For the geometric amplitude (Fig. 6.9A), slightly higher perceived magnitudes for medium frequencies can be observed, which corresponds to an approximately linear stimulus–perception relationship with the exponent $b_{medium} = 0.903$ (see Section 6.1.4.2), while for low and high frequencies, the transformation function is close to the square root form with $b_{low} = 1.2257$ and $b_{high} = 1.3093$, respectively. For the geometric frequency (Fig. 6.9B), the perceived magnitudes increase with increasing amplitudes, i.e. the transformation function varies from a weaker to a more pronounced exponential function with $b_{low} = 1.4345 < b_{medium} = 1.8575 < b_{high} = 2.0297$. A similar trend can be observed for the colour

frequency (Fig. 6.9D). However, in this case, it can be explained by the contrast at the colour interval borders: while high colour amplitudes lead to hard transitions, making the interval alternation more salient, low amplitudes produce a kind of blurry border, making the intervals appear larger. Finally, colour amplitude perception does not show any apparent pattern in its dependency on colour frequency (Fig. 6.9C). In general, for all four visual variables, the respective deviations are rather marginal.

(A) Geom. amplitude

(B) Geom. frequency

(C) Colour amplitude

(D) Colour frequency
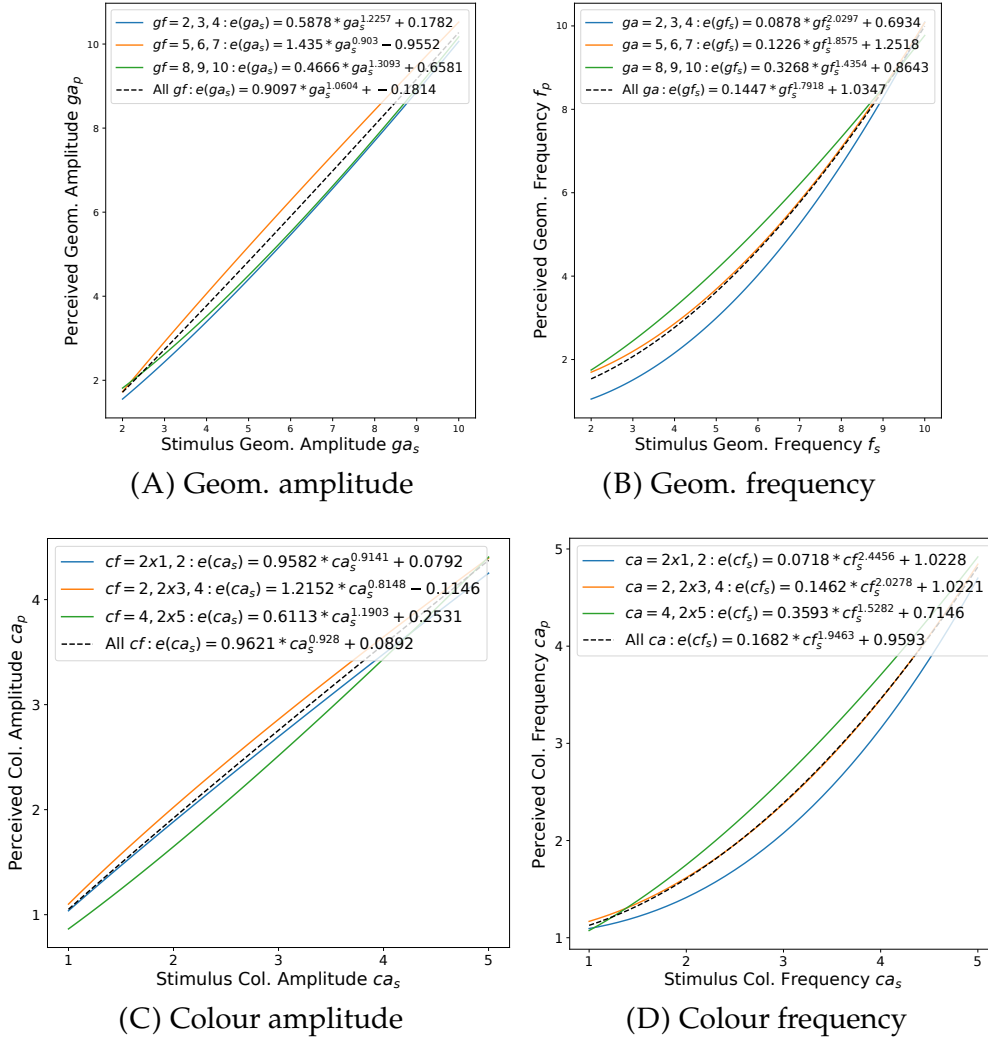
FIGURE 6.9: Modelling stimulus-to-perception transformation in dependence on the magnitude of the second visual variable.

6.1.5 Transfer to Different Shape Sizes

The quantization levels of visual variables derived in Sec. 6.1.4 are based on the experiments with a fixed image size 50×50 mm. To allow a flexible application in different visualization contexts, a scheme of how to apply the quantization results to glyphs with different sizes is proposed in the following, even though the evaluation of this scheme must be deferred for future research. Note that the proposed scheme does *not generate visual variables that are comparable across different scales.*

Before starting with the definition of the transfer rules, first, it is necessary to sum up the quantization process for the fixed size, introducing some necessary notation. More precisely, for a perceived visual variable v, initially, the minimum and the maximum stimulus values s_{min}, s_{max} (in mm) are fixed and the quantization size Δv (in adu) is deduced, whereby Δv is applied to the range v_{min}, v_{max} (in adu), which corresponds to the perception of s_{min} and s_{max}, according to the user experiment (see Sec. 6.1.3.3). Note that Δv corresponds to $\Delta s(v)$ (in mm), which is in general not constant (see Fig. 6.6).

To transfer the aforementioned quantization parameters to shapes with a scale $\omega > 0$ relative to the original shape of 50×50 mm, the following rules are proposed, where it is assumed that $0 < \omega < 1$, since glyphs are rather used at smaller scales:

- Colour amplitude should not be scaled, as intensity perception is independent of size.

- The "perceptual" stepsize Δv (and the corresponding stimulus stepsizes $\Delta s(v)$) should not be reduced to preserve the absolute variation (in mm), and thus, the visual distinguishability.

- The minimum and the maximum stimulus and visual variable values s_{min}, s_{max} and v_{min}, v_{max}, respectively, are scaled according to the following rules:

  - The minimum values s_{min} and v_{min} can only be scaled moderately, i.e. reduced using $\omega_{min} > \omega$, potentially even $\omega_{min} = 1$, to prevent, for example, visually vanishing amplitudes.
  - The maximum values s_{max} and v_{max} should be scaled by $\omega_{max} = \omega$, to prevent, for example, extreme distortions for small shapes.

  Consequently, the number of levels gets potentially reduced for $\omega < 1$ as the "usable" range $[\omega_{min} \cdot s_{min}, \omega_{max} \cdot s_{max}]$ gets smaller while the stepsize $\Delta s(v)$ remains unchanged. This problem can be counteracted by reduction of the scaling effect for the maximum values $\omega_{max} = \omega + \epsilon \leq 1$ with a user-defined parameter ϵ that also depends on the shape's complexity.

Fig. 6.10 shows some exemplary results of the quantization transfer to shapes with $\omega = 0.4$.

6.1.6 Conclusion

6.1.6.1 Summary

In this section, based on the results of a user study, the definition of a perceptually uniform quantization model of periodic contour modifications for a glyph-based visualization design, comprising the visual variables such as geometric amplitude and frequency, waveform as well as colour amplitude and frequency, was presented. The main model components are stimulus-to-perception transformation, waveform-dependent calibration, and definition of the quantified levels for the corresponding visual channels. Moreover, an evaluation of the potential impact of the perceptual dependencies between specific visual variables was provided.

Below, the core findings are summed up:

1. Following [Ste57], the *relation between stimuli and their perception* for all four quantitative visual variables, considered in the model, can be modelled as a
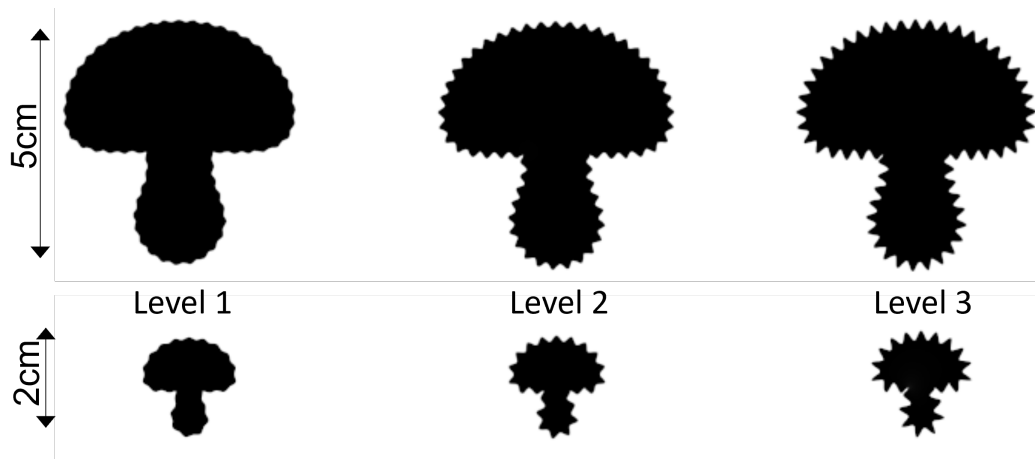
FIGURE 6.10: Transfer of quantization results to different shape's sizes: geometric amplitude levels $l_1, \ldots, l_3$. Top row: original size. Bottom row: scaled with $\omega = 0.4$. This figure needs to be displayed according to the given scale. Note that the quantization levels are *not intended to be comparable across scales*.

power function (see Sec. 6.1.4.2). Since the adu-scale used in the experiments does not have a proper zero-origin, the power function was extended with an additive term to compensate this fact.

2. The *influence of waveform* on the perception of geometric amplitude and frequency is marginal (see Sec. 6.1.4.4). As a consequence, the corresponding calibration step in a visualization design can be skipped.

3. The user study shows that the geometric as well as the colour frequency have a better *discriminative capacity* than the respective amplitudes (see Sec. 6.1.4.3). Overall, the geometric frequency has the highest number of quantified levels in the tested range.

6.1.6.2   Limitations

Additionally to these results, the presented study also allows the following assumptions regarding further perceptual aspects of the contour modifications, which still require an in-depth investigation or validation in future work:

1. A first insight into *amplitude–frequency dependencies*, provided in this work, shows certain perceptual trends as a function of the respective second parameter, but the resulting deviations are rather marginal and thus can be likely neglected by a visualization design (see Sec. 6.1.4.5).

2. A method to transfer the model, estimated for shapes with a fixed size 50×50 mm, to *arbitrary sizes* was proposed (see Sec. 6.1.5). The respective rules were derived heuristically. Also, first exemplary results created with this method were shown in this section. However, the evaluation of this method requires a separate user study.

6.1.6.3   Future Work

Finally, the presented results can serve as inspiration for some related topics, which, however, are beyond the scope of this work:

1. The current quantization has been statistically estimated on the basis of perceptual data. It may be of interest to compare the presented results with other estimation methods, for instance, direct JND tests.

2. The corresponding ranges of the geometric visual variables were consciously narrowed in this user study to avoid the expected strong interferences for low and high magnitudes [Gör+17], as mentioned in Sec. 6.1.3.2. At the same time, it can be assumed that the current maximum is still relatively far away from critical magnitudes. Consequently, the current *limits* needs further investigation in a separate experiment. Furthermore, it is to assume that the *amplitudes and frequency limits* depend to some degree on the respective base shape, especially on its local curvature.

3. Colour contour modifications, limited in the current user study to black-white images, can be transferred to shapes with *other foreground colours*, but a potential reduction of the number of colour amplitude levels, depending on the base shape intensity and the resulting shift of zero amplitude, has to be taken into account.

4. A *combination of* two main modification types—*geometry and colour*—is also conceivable. According to a specific visualization design, it can be implemented as four independent quantitative visual variables as well as in a coupled form, e.g. with colour frequency equal to geometric frequency and colour amplitude linked to geometric amplitude. Such combinations potentially entail dependencies between colour and geometry perception.

6.2   PACEMOD: Parametric Contour-based Modifications for Glyph Generation

6.2.1   Introduction

Icon-based or iconographic techniques, according to the taxonomy of Keim and Kriegel [KK96], form one fundamental class of methods for visual exploration of multivariate, multidimensional data. Their characteristic property is the mapping of data dimensions to varying visual features [DOL03], i.e. visual variables or channels, such as size, colour hue, luminance (or colour value), grain, orientation and shape [Ber83]. In literature, there is no clear consensus on differences between icons and glyphs. For the sake of consistency, this thesis follows the definition that *icons* "represent a sign that itself resembles the qualities of the object it stands for", while *glyphs* "represent different data variables by a set of visual channels" [Bor+13]. In other words, icons are considered to be static visual objects that transfer to glyphs by controllable variations of their visual variables.

Glyph design involves two main tasks that affect the intuitive mapping of data variables to application-related visual variables, namely (1) the selection of an appropriate initial visual object, or icon, including shape and colour, and (2) the modification of the visual variables of the initial icon, i.e. its geometry and related colour attributes. To enhance the intuitivity, metaphoric glyphs make heavy use of familiar and well-understood visuo-spatial phenomena related to the underlying problem domain [Ris08] in both tasks. The main advantages of metaphoric glyphs relate to their potential of increased readability in case of realistic glyphs [FR81], and improved data understanding [Fuc+16], e.g., by mapping data to corresponding glyph parts [Sur05].

On the technical level, the design of metaphoric glyphs is empowered and, at the same time, limited by the given means of manipulating the glyph's base icon and its visual variables, which often involves a large amount of manual graphics design. Examples are glyphs to visualize the health state of corn cobs [NSS05], environmental data related to forest fires using leaf-like glyphs [Fuc+15], or car glyphs that map car-related data to corresponding parts of the base icon [Sur05]. More automated approaches include procedural and data-driven methods. Procedural methods, such as RoseShape glyphs [Cai+15; LLZ15], which are sinusoids plotted in polar coordinates applied to circles, have a very restrict set of base icons. Data driven methods, such as the automatic generation of emoji-like metaphoric glyphs [Cun+18], are hardly controllable regarding the variation of their visual variables.

This section introduces the *PArametric Contour-basEd MODification (PACEMOD)* concept, a novel approach that allows for controllable geometric and colour modifications of an icon provided by the user. The proposed approach enhances the automated generation of metaphoric glyphs, providing two main technical features, particularly it (1) allows almost any base icon as input and (2) supports the design of various techniques for modification of the icon contour's geometry and the related colour attributes. Technically, it utilizes *diffusion curves* [Orz+08] (see also Sec. 2.6) as parametric representation of a user given icon, which are re-parametrized to add new degrees of freedom (DOF) in an arc-length encoded manner. Using these DOFs allows to modify the geometry and colour of icon's contours in a consistent way, creating visual variables for data encoding but also retaining the icon's overall shape,

and thus its recognizability. The PACEMOD concept is exemplified by implementing periodic, wave-like geometry and colour modifications and size variations, optionally applied to selected icon parts. This implementation is far from being exhaustive, i.e. the present concept is open to allow other modification approaches for the visual variables.

In summary, this section comprises the following contributions.

- *PACEMOD*, a new, diffusion-curve based concept of parametric, contour-based modification of icons, comprising

  - the parametrization of the icon contours as diffusion curves using B-splines,

  - the methodology to insert the DOFs in an arc-length encoded manner, required to control the visual variables, and

  - the utilization of distance transforms (DT) to allow further, geometry related contour manipulations and the prevention of self-intersection.

- The implementation of the PACEMOD concept with a focus on the automated, wave-like modification of the icon contour's geometry and colour.

- Two application examples, namely uncertainty visualization for rain forecast and gradient glyphs applied to COVID 19 data.

6.2.2   Prior Work

Prior work, conceptually related to the PACEMOD concept, is briefly discussed below.

Metaphoric Glyphs.   Metaphoric glyphs form a specific sub-group of glyphs that try to enhance the underlying communication process by additionally utilizing visual analogies from the related application domain and ultimately strive for "the picture becomes the thing it represents" [Ris08]. Several works underline their potential to improve readability [Fuc+16]. For a general overview of glyph design and application, the reader is referred to the surveys from Ward [War08], Borgo et al. [Bor+13] and Fuchs et al. [Fuc+16].

Technically, approaches to generate metaphoric glyphs are largely dominated by manual design processes or rely on structured databases. Nocke et al. [NSS05], for example, propose a mosaic paradigm that decomposes the icon into tiles, alters the tiles in size, shape or colour according to the data values, and recombines them to achieve the final corn glyph. Fuchs et al. [Fuc+15] propose a manual design of leave-shaped glyphs to utilize the humans ability to visually discriminate natural shapes by modifying, e.g., the leaf's morphology, venation and boundary for visualizing multidimensional data related to environmental events such as forest fires. Often, icons are combined with abstract glyph components. Legg et al. [Leg+12] and Chung et al. [Chu+13], for example, propose wave-like shape deformations and colour modifications of a circle to encode a sportsman's performance, which surrounds the icon describing the specific sports event to be assessed. Li et al. [Liu+16] propose a dashboard-like glyph to summarize the key features of locations for billboard selection, which combines abstract glyphs and feature-related icons.

Besides glyphs based on parametric shape representations, such as RoseShape, i.e. polar sinusoidal plots resulting in a flower-like glyph [Cai+15], different data driven approaches have been developed. Cunha et al. [Cun+18], for example,

present a data-driven strategy for the automatic generation of emoji-like metaphoric glyphs utilizing a structured emojinating database [CMM18]. Ying et al. [Yin+21] presented GlyphCreator, a tool for deep learning based decomposition of circle-like abstract glyphs into several visual elements, which are then manually bound to the input data attributes.

Contour-based Methods. Contour-based uncertainty visualization methods [Bon+14] relate to the general question of how to modify shape contours, also addressed by the present approach. Some of them have already been mentioned in Sec. 6.1.2 in relation to the perception of their visual variables. In this section, instead, the focus is on the generative technical aspects.

Contour-based approaches to encode uncertainty in the underlying data include the variation of the contour lines' width [AOB08], the usage of sets of contours in regions with high uncertainty in segmentation [PRH10], and the modification of graph nodes and radial colour gradients to encode the relation uncertainties in graphs. Görtler et al. [Gör+17] propose Bubble Treemaps as an extension of circular treemaps that encode uncertainty using wave-like modifications and blur-effects applied to the circular-arc spline contours during the treemap generation process. Holliman et al. [Hol+19] use abstract circle-shaped glyphs for uncertainty visualization with wave-like modified contours, manually modelled in Blender. Both techniques are intrinsically linked to a specific primitive as the glyph's base shape, i.e. a circle.

Diffusion Curves. Diffusion curves images (DCIs), proposed by Orzan et al. [Orz+08], are a widely used vector graphics form that combines benefits of raster and vector-based representations, as described in Sec. 2.6. Various functionalities and improvements have been developed over the years. The main focus lies on the conversion of raster images to DCI [Lu+19], enhanced DCI representation [Xie14], DCI editing [JCW11; Jes16; Lu+20], and efficient rendering of DCIs [JCW09; STZ14]. The editing and manipulation approaches proposed so-far for DCIs are all based on manual intervention. Examples include the methods of Jeschke and colleagues [Jes16; JCW11], who propose a click-and-drag metaphor for manipulating diffusion curve properties, and Lu et al. [Lu+20], who present a combination of global and local deformations for manipulating coarse and fine image content, respectively. In summary, there exists a rich tool set for utilizing DCIs, but there is no adequate approach to control an icon's visual variables as required for automated glyph generation.

PACEMOD in the context of prior work. The PACEMOD approach is conceptually related to the aforementioned contour-based techniques, especially to [Gör+17; Hol+19] in using wave-like contour modifications. However, the prior approaches focus on a specific visualization tasks, e.g. uncertainty visualization, and use a specific base shape, yielding an appropriate, but relatively limited design space. Moreover, automated glyph generation is mainly realized for simple base geometries like circles [Gör+17; Hol+19], while glyphs with more complex base shapes are commonly designed manually [Fuc+15]. In contrast, PACEMOD aims at a *generic glyph generation concept* comprising technical functionalities for a *flexible and automated* appearance modification of an *arbitrary iconic visual object*, used as base shape. To achieve these goals, the present approach uses a parametric glyph representation, based on DCIs.

6.2.3 PACEMOD Concept

This section describes the technical foundations of the proposed approach for PArametric Contour-basEd MODifications (PACEMOD) of an icon. The discussion of the basic principles comprises the representation of the initial icon (Sec. 6.2.3.1) and the pre-processing (Sec. 6.2.3.2), the general approach to contour-based modifications (Sec. 6.2.3.3), and the final post-processing (Sec. 6.2.3.4). A specific approach implementation for periodic, wave-like modifications of the contour's geometry and colour is explained in Sec. 6.2.4. An overview of the PACEMOD generation process is represented in Fig. 6.11. The main notations, used for the PACEMOD description, are provided in Tab. 6.5.
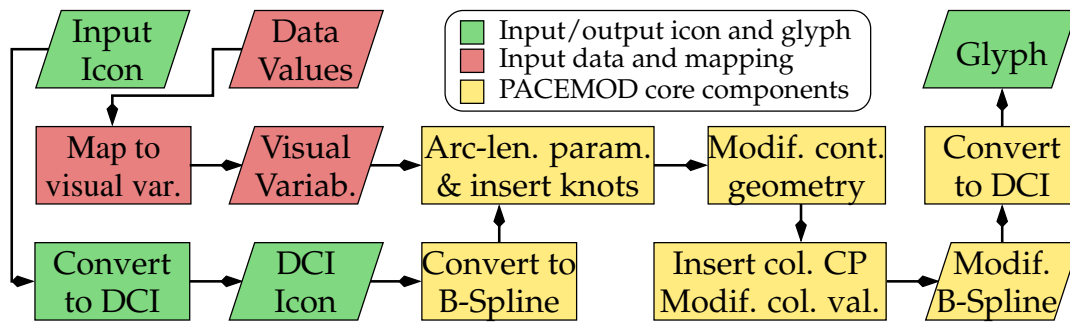
FIGURE 6.11: The PACEMOD concept applied to glyph design.

TABLE 6.5: Table of symbols (cp = control point).

$\mathbf{p}_i$	Orig. Bezier cp		K	# Bezier curves
$\mathbf{c}_i^{l/r}$	Orig. le./ri. col. cp		$\tilde{\mathbf{c}}_i^{l/r}$	Modif. le./ri. col. cp
$\mathbf{d}_i$	Orig. B-spline cp		$\tilde{\mathbf{d}}_i$	Modif. B-spline cp
$t, \mathbf{T}$	Orig. knot (vector)		$\tilde{t}, \tilde{\mathbf{T}}$	Modif. knot (vector)
u	Orig. colour param.		$\tilde{u}$	Modif. colour param.

6.2.3.1 Icon Representation

To achieve controllable and automatable contour modifications, the present approach uses the *parametric* diffusion curve image (DCI) representation [Orz+08] of the given base icon. A DCI represents an image as a set of K cubic *Bézier curves* in conjunction with *colour parameters*. For more details and mathematical definitions see Sec. 2.6 and also Fig. 6.12:2a-b. Note that the current PACEMOD concept omits the blur parameter due to its visual dependency on colour attributes.

It is assumed that a given base icon is composed of one or several closed, non-overlapping contours and that the regions defined by these contours have a homogeneous inner colour, since a colour gradient strongly restricts the options of an automated colour modification.

6.2.3.2 Pre-Processing

For conversion of the given raster icon into a DCI, the diffusion curves drawing tool from Orzan et al. [Orz+08] is used. As modifications of the individual Bézier segments would easily lead to unwanted cracks, the DCI Bézier representation is
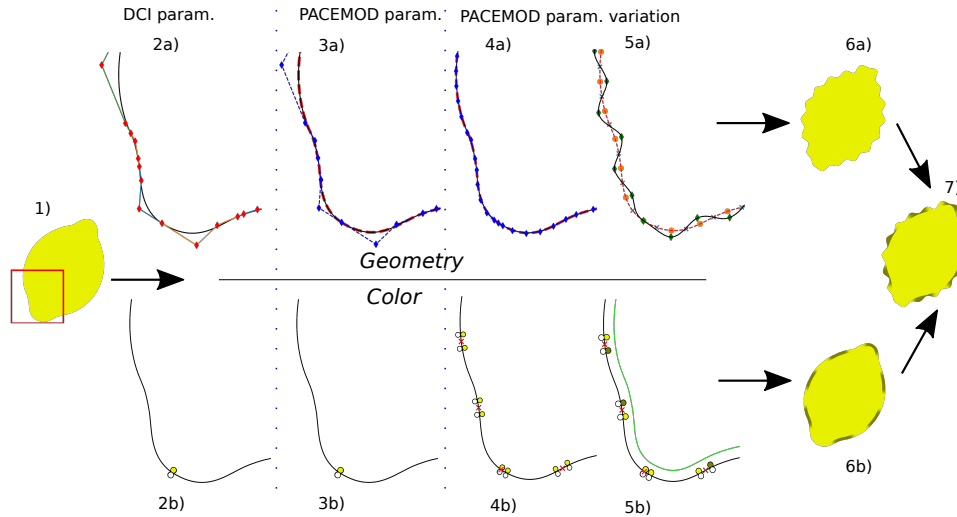
FIGURE 6.12: The PACEMOD modifications: The input raster image (1) is converted into a DCI, comprising Bézier control points for geometry (2a) and colour control points (2b). Afterwards, the geometry is converted into B-splines and a C^1 approximation is applied (3a) w/o changing the colour control points (3b). In the next step, DOFs are added according to the target knot point and colour border positions, resulting in an arc-length re-parametrization (4a) and new colour points at the virtual borders of colour intervals (4b: red crosses), respectively. The final shape modification transforms subgroups of knot points (5a: orange points are moved to green points). The colour modification changes the colour points of each second interval (5b: darker circles) at the icon's inside, while the outer colour is masked out and remains; a diffusion barrier (light green) is added to limit the diffusion and maintain the inner colour unchanged. The final glyph is obtained by re-conversion into standard DCI and rendering (6a, b; 7).

converted into cubic B-spline curves. To increase the curve continuity wherever possible, C^1-transitions are constructed if consecutive Bézier segments have first derivatives with (approximately) the same direction. Deviations in the length of the curves' end tangents are corrected by an appropriate interval scaling, while slight directional deviations below a threshold α ($\alpha = 3.5°$ is used) are ignored.

More precisely, given the K (consecutive) Bézier curves $b^i(u)$, $i = 0, \ldots, K-1$ with control points $P^i = \{\mathbf{p}_{3i}, \cdots, \mathbf{p}_{3i+3}\}$, the i-th Bézier segment is converted by adapting the parameter interval to fit the tangent length of the end point of the prior segment. In case the directional deviation is below the threshold α, the transition is assumed to be C^1, thus and a double knot is appended to the initial B-spline knot vector $\mathbf{T}$ and the last de Boor point is dropped from the control point list D. In case of a C^0 transition, a 3-folded knot is appended (see Fig. 6.12:3a).

Subsequently, the parametric positions of colour attributes are transformed by mapping them into interval defined by $\mathbf{T}$.

The resulting PACEMOD icon representation comprises de Boor points $\{\mathbf{d}_i\}_{i=0}^{L}$ with the corresponding knot vector $\mathbf{T} = \{t_i\}_{i=0}^{L+4}$, and colour control points $\{\mathbf{c}_i^l(u')\}_{i=0}^{M}$ and $\{\mathbf{c}_i^r(u')\}_{i=0}^{N}$, where u' refers to the colour parameter after the conversion to B-splines (see also Fig. 6.12:3a-b).

6.2.3.3   General Contour-Based Modification

According to the two different types of the PACEMOD control parameters, the icon's geometry and colour can be modified separately.

Geometry   Geometric modifications are accomplished by a direct manipulation, i.e. translation, of the contour points Q belonging to the corresponding B-Spline curve and the subsequent interpolation. They are controlled by source position on the curve and the translation vectors. In the following, the prerequisites and basic functionalities, allowing such modifications, are considered.

Arc-length parametrization and knot vector adjustment.   In general, defining the source position in terms of arc length $l(u)$ instead of using the original parametric space provides a more intuitive control (cf. Sec. 2.5.3). An arc-length parametrization of a B-Spline curve $\mathbf{s}(u(l))$ is realized by means of a numerically calculated lookup table, i.e. it is approximated by the chord length.

Moreover, based on $u(l)$, the knot vector $\mathbf{T}$ is converted to $\tilde{\mathbf{T}}$ by shifting the original knots or inserting new ones, according to the following requirements. For a better modification control, the curve points Q to be translated are converted to *knot points* [Boo78], i.e. $\mathbf{q} = \mathbf{s}(\tilde{t})$ with $\mathbf{q} \in Q$, $\tilde{t} \in \tilde{\mathbf{T}}$. The B-spline curves created in the pre-processing step (see Sec. 6.2.3.2) do not always have enough DOFs to interpolate the translated curve points. In such cases, additional DOFs are created inserting further knots by means of Boehm's algorithm [Boe80]. Furthermore, additional knots can be required to control the shape of a curve modification. For instance, inserting knots before and after the parametric position of a point to be translated limits the influence of the corresponding de-Boor points [FB93], and thus the width of the modified segment (more details in Sec. 6.2.4). After $\tilde{\mathbf{T}}$ has been defined, the respective de Boor points D are adjusted by means of the least-squares progressive iterative approximation algorithm with energy term (ELSPIA) [He+15], which minimizes the least-square distance to the original curve taking into account a stretching term (see also Sec. 2.5.3). The result is a cubic B-spline with single inner knots and de Boor points $\tilde{D}$, which approximates the original curve sufficiently accurate. Depending on the target shape, the multiplicity of some knots may be increased to three, to produce sharp corners (for specific examples see also Sec. 6.2.4).

Local reference frame.   To achieve visually appealing and consistent modifications, it is helpful to apply the translation of knot points in relation to the intrinsic properties of the respective curve, especially in the context of an automated process. In particular, the curve normal vectors provide information about inward/outward direction: by conversion into DC (see Sec. 6.2.3.2) all curves are defined in the way that the normals point inward into the region defined by the respective closed contour, i.e. there is a consistent Frenet–Serret frame in which the translation of the knot points is applied. In the specific implementation, presented see Sec. 6.2.4, only translations in normal directions are applied, but other approaches are possible.

Prevention of intersections and Distance Transform.   Depending on the geometry of the base icon, the translation can lead to self-intersections that have to be prevented, e.g., by locally reducing the length of the translation vector. The present approach uses a conservative intersection prevention scheme, based on Distance Transform (DT) [MQR03] of the base icon. A subpixel precision DT allows

to construct a (virtual) skeleton between icon's curves, which, in turn, is used to constrain the modified curves. More specifically, the transformed knot points are restricted to move inside the skeleton's half-space of the initial curve point. An implementation for a specific translation along normals is given in Sec. 6.2.4.

After curve point translation, applied on an adjusted curve $(\tilde{D}, \tilde{\mathbf{T}})$, and exploiting the aforementioned functionalities, the resulting new B-Spline curve is constructed by adjusting the de Boor point positions. Particularly, the corresponding offset vector $\Delta\tilde{D}$ is computed according to Fowler and Bartels [FB93] (see Sec. 2.5.3, Eq. 2.48).

Colour Controllable colour modifications are realized by setting colour control points $\tilde{C}^l$, $\tilde{C}^r$, i.e. defining respective locations and colour values. Analogously to geometric modifications, the arc-length parametrization is used to facilitate a more intuitive control over colour point positions.

Changing the colour of the original colour control points C^l, C^r modifies the entire icon subregion, influenced by the respective curve. Inserting new colour control points, in turn, creates areas with different colouration. The distance between $\tilde{\mathbf{c}}_j^*$ and $\tilde{\mathbf{c}}_{j+1}^*$ that serve as areas boundaries, controls the smoothness of the transition between neighbour colours.

Moreover, the diffusion range of colours on each curve side, that is the area affected by emitted colour values, can be limited by insertion of a *diffusion barrier* (see [Bez+10]), i.e. a DC that has no own colour at least on one side. In the present approach, diffusion barriers are created as DT isolines, where the iso-value limits the impact of the colour diffusion (see Sec. 6.2.3.3, "Geometry").

Shrinking and Inflating of Icon Parts The PACEMOD approach allows to modify the size of a specific icon part by shrinking or inflating the respective contour. This type of modification is controlled by offsets to the original curve. Instead of using curve point translations, the new curve is defined as isolines with a given offset, analogously to diffusion barriers described above, as this approach is more robust against local curve distortions.

6.2.3.4   Post-Processing

After a glyph has been constructed by applying the corresponding modifications, the PACEMOD representation is back-converted into a DCI. The de Boor points $\tilde{D}$ of the modified B-Splines are transformed to Bézier control points using a basis transformation matrix [CR04]. Accordingly, the parametric positions of the colour attributes are remapped onto the intervals of the resulting Bézier splines. For rendering the DCIs as raster images, an extended rendering tool of Jeschke [JCW09] is used.

6.2.4   Periodic Modifications

The PACEMOD concept, presented in Sec. 6.2.3, performs morphological modifications, i.e. it is basically agnostic to the semantic structure of the input icon. Conceptually, it allows for arbitrary modifications, but providing automated means to achieve meaningful visual effects, e.g. visual variables having metaphoric associations, is not trivial. Here, an implementation that comprises a generic set of automated contour modifications, namely the generation of periodic, wave-like patterns, is proposed.

The suitability of the periodic contour modifications for visual encoding is motivated by the findings of the existing psychophysical studies, described in Sec. 6.1.1–6.1.2, and specially by the results of the user study, presented in Sec. 6.1. The PACEMOD concept builds on this study, in particular, utilizing the proposed quantization model (see also Sec. 6.2.5).

6.2.4.1 Periodic Geometric Modifications

The wave-like modification of the curve geometry is performed by a periodic translation of the knot points along the curve's normals in alternating directions and is defined by frequency, amplitude and waveform (see Fig. 6.13). In this case, $\tilde{\mathbf{T}}$ is composed of knots that are equidistant in terms of arc length, placed at a distance controlled by the given frequency.

(A)   (B)   (C)

(D)   (E)   (F)
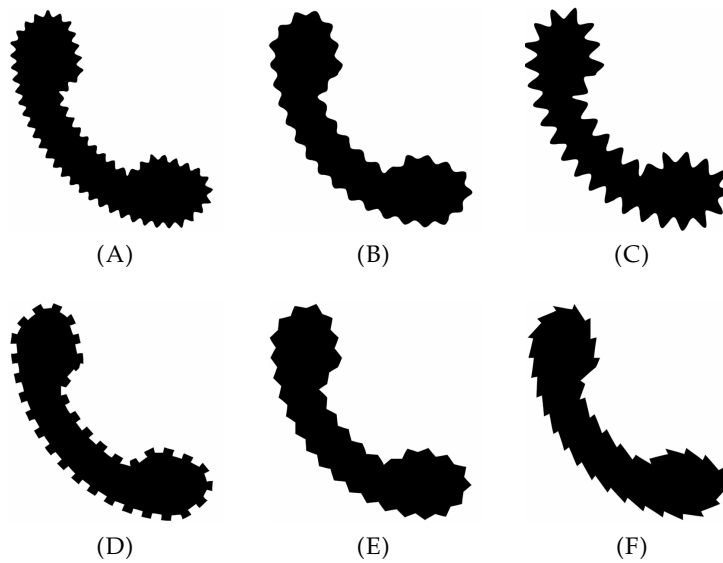
FIGURE 6.13: Examples of periodic geometric modifications. 1st row: varying wave amplitude and frequency; (B): middle frequency and amplitude, (A): a higher frequency, (C): a higher amplitude. 2nd row: different waveforms; (D): rectangular, (E): triangular, (F): sawtooth-like.

The implementation of the intersections prevention mechanism (see Sec. 6.2.3.3, "Geometry") checks whether a translated knot point $\mathbf{k} + a \cdot \hat{\mathbf{n}}^k$ is closer to another curve point $\mathbf{p}$ than to original knot point $\mathbf{k}$ itself, i.e. whether it is located outside its skeleton half-space. In this case, a *skeleton point* $\mathbf{s} = \mathbf{k} + a' \cdot \hat{\mathbf{n}}^k$ that is equidistant to $\mathbf{k}$ and $\mathbf{p}$ is locally reconstructed along the translation vector, whose amplitude a' is given as the length of a leg in the isosceles triangle with base $\mathbf{p} - \mathbf{k}$, i.e. $a' = \frac{0.5\|\mathbf{p}-\mathbf{k}\|^2}{\langle\hat{\mathbf{n}}^k,(\mathbf{p}-\mathbf{k})\rangle}$ (see Fig. 6.14A). The final amplitude is $a' - \epsilon$, to preserve a free space between the curves. This procedure is applied recursively, as several intersections may occur.

This intersection test is applied to the translated knot points, i.e. the wave peaks, which represent the most protruding parts of the modified curves. While the proposed strategy successfully prevents intersections in almost all cases observed during the experiments, extreme cases, e.g. highly curved contours build by multiple curves, which are less suited for PACEMOD, may require an increased skeleton pixel offset ϵ.
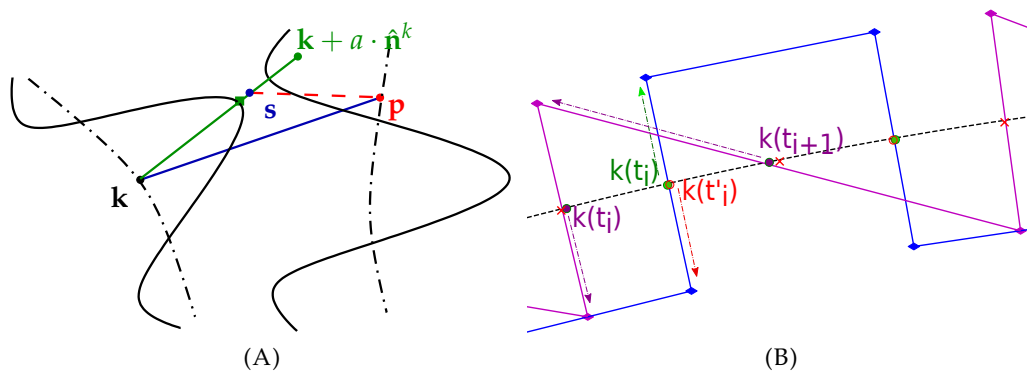
FIGURE 6.14: (A): Prevention of curve intersection. Dashed lines
show the original curves and solid lines their sinusoidal
modification. The skeleton point **s** at the translation vector is
equidistant to both curves. The sin peak (green square) has a pixel
offset $\epsilon = 3$ from **s**.
(B): Construction of rectangular and sawtooth shapes of a B-Spline
$\mathbf{k}(u)$; arrows represent the respective translation vectors.
Rectangular: additional knot points with a small offset (red) are
added to the initial knot point for the given frequency (green), and
are translated in the opposite direction. Sawtooth: the target position
for each second knot point is shifted.

Different waveforms can be created by means of additional knots and by shifting
the translated positions along the curve (see Figs. 6.14B and 6.13). For instance, in
the case of a sinusoidal wave one intermediate knot is inserted between each two
'peak knot points'. These additional knots do not function as curve modification
constraints, but restrict the influence of the control points to one period and guaran-
tee a smooth shape (see also Sec. 6.2.3.3, "Geometry"). For the triangular, sawtooth-
like and rectangular waveform, the knot multiplicity is increased to 3 to achieve the
C^0-continuity and thus to produce sharp corners. After a curve modification, each
third control point, starting with the first one, is located in a corner, while the inner
two control points are placed on the lines between the corners to get straight line
strips.

6.2.4.2 Periodic Colour Modifications

A periodic colour modification along the corresponding contour is created by al-
ternating equidistant intervals of two different colours. It is controlled by frequency,
i.e. inverse interval arc length, and respective RGB values. To construct the intervals,
initially their virtual boundaries are computed as positions along the curve that are
equidistant in arc length, according to the target frequency. This is done analogously
to the knot vector calculation for periodic geometric modifications, described above
(see Sec. 6.2.4.1). Afterwards, two new colour point locations are placed before and
after each interval boundary with a small arc length offset to get a hard transition
between the both colours (see Fig. 6.12:4b and Sec. 6.2.3.3, "Colour"). For instance,
alternating the original curve colour with a new one and limiting the colour point in-
fluence with a diffusion barrier at a small offset, creates the appearance of a dashed
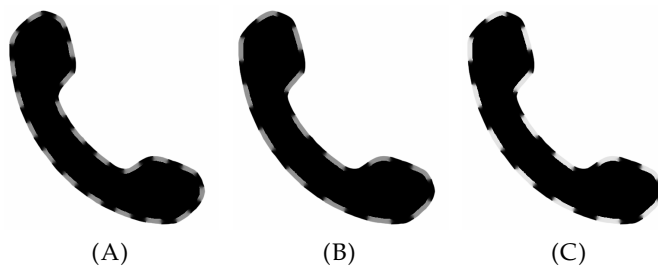line, as shown in Fig. 6.15.

FIGURE 6.15: Periodic colour modifications. (B): low frequency and dark grey dashes; (A): a higher frequency; (C): a lighter dash colour.

6.2.5 Application Examples

In the following, the usefulness of the PACEMOD concept and its implementation using periodic, wave-like contour modifications is demonstrated by giving two application examples. The first application shows how the PACEMOD framework allows to apply existing contour-based uncertainty visualization approaches [Gör+17; Hol+19] to iconic glyphs, facilitating a more intuitive relation to the application domain. The corresponding examples show visualization of the rain forecast uncertainty (Sec. 6.2.5.1). In the second application, a design of *gradient glyph* is proposed, i.e. a glyph that besides main parameters also encodes their changes, e.g., with respect to time. The visualization of COVID-19 related statistics has been chosen as specific gradient glyph example.

To guarantee the readability and perceptual linearity of the visual encoding, the stimulus-to-perception transformation and the quantization levels, estimated in Sec. 6.1, are applied for contour-based visual variables. The size levels have been generated according to Stevens and Guirao [SG63]. The maps in all examples are created with the Maputnik open source tool[1].

6.2.5.1 Uncertainty Visualization with Iconic Glyphs: Rain Forecast

Two commonly used rain forecast parameters are visualized here: amount of precipitation and forecast uncertainty or rain probability. The visualized data represent a weather forecast for Europe, September 8th, 2022, collected from WetterOnline[2] on September 6th, 2022. A cloud with drops, which is an image often used in forecast websites and apps, serves as the glyph's base icon. To demonstrate the flexibility of the presented framework, two different design variants were created.

In the first one (see Fig. 6.16), the uncertainty is encoded by frequency of a sinusoidal wave, as proposed in [Gör+17] and [Hol+19]. To generate the corresponding glyphs from the base icon, a periodic geometric modification with fixed amplitude, varying frequency and sinusoidal waveform is applied to DCs that represent drop contours (see Sec. 6.2.4.1). Since the drops in the image symbolize the rainfall, the amount of precipitation can be intuitively encoded with their size. However, the fact that the drops are relatively small glyph's details can impair the size discrimination. Hence, to enhance the distinguishability using the augmentation effect (cf. Sec. 2.1.1), drop size is coupled with saturation and luminance of their colour. Technically, the curves representing drops of different sizes are generated from DT isolines (see Sec. 6.2.3.3, "Shrinking and Inflating of Icon Parts") and new colours are

---

[1]https://maputnik.github.io/
[2]www.wetteronline.de

created by setting colour control points for the inner part of the respective DCs. The mapping of the forecast data to the respective visual variables is represented in Tabs. 6.6 and 6.7.
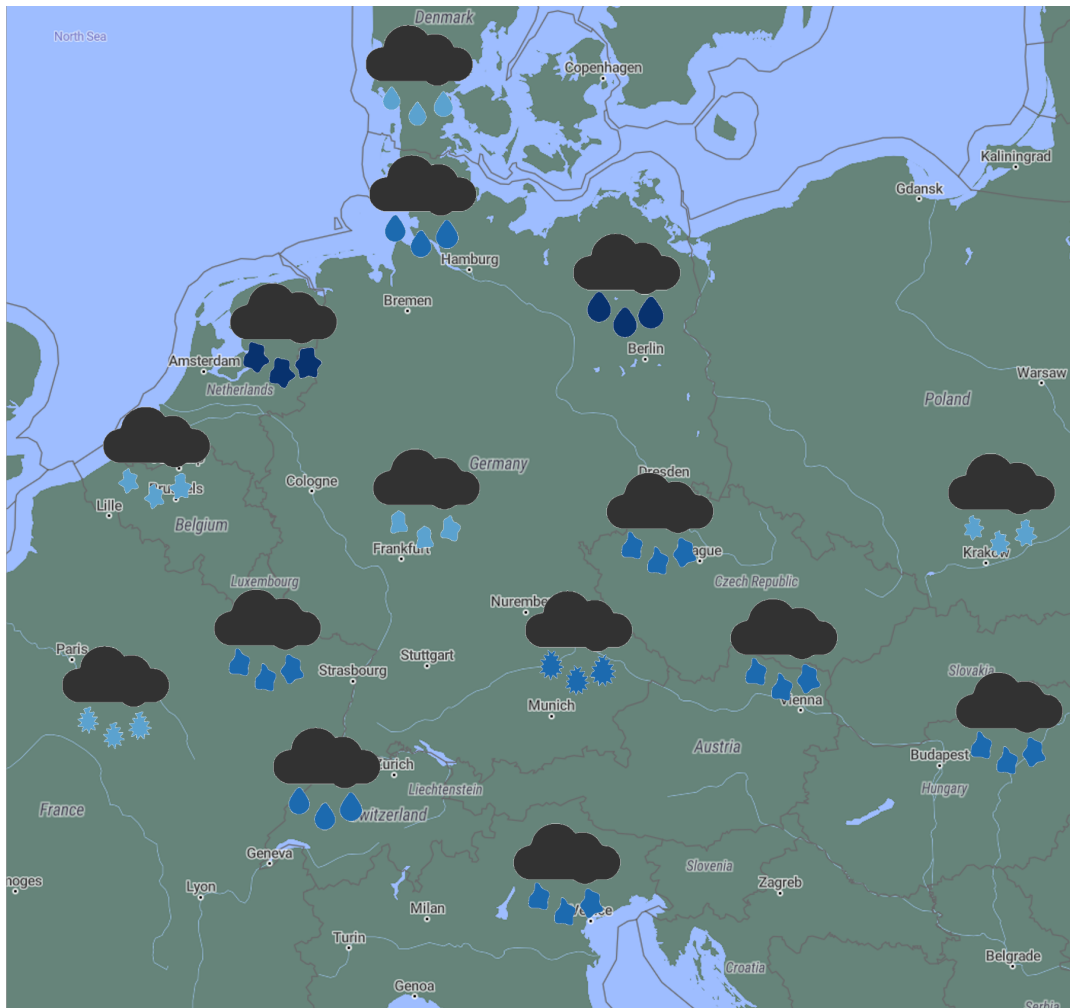

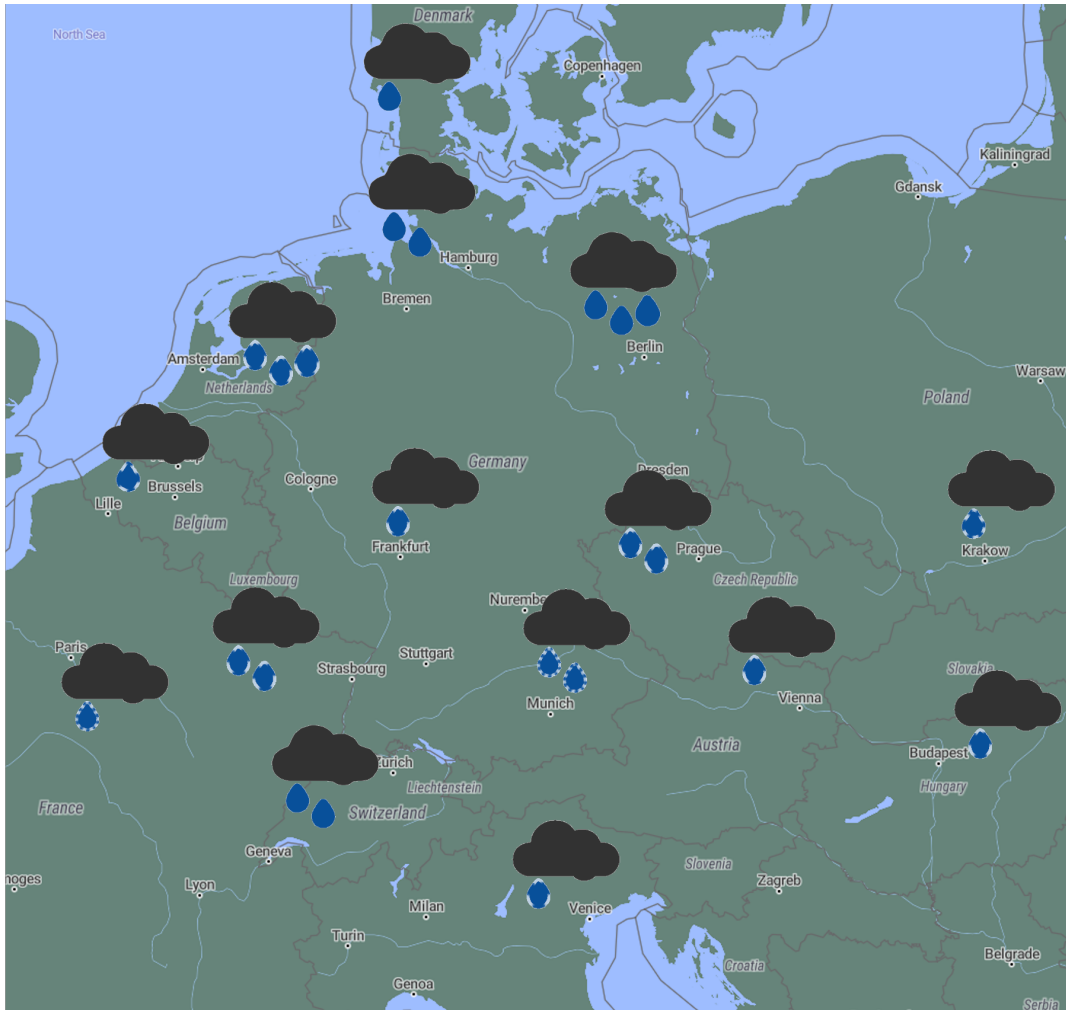FIGURE 6.16: Rain forecast visualization: the uncertainty encoded with sine wave frequency.

FIGURE 6.17: Rain forecast visualization: the uncertainty encoded
with dashed line frequency.

TABLE 6.6: Visual encoding of amount of rain: by size and colour (1st
row) and by number (2nd row) of drops.

Data	Mapping		
Amount of rain	light	moderate	heavy

TABLE 6.7: Visual encoding of rain forecast uncertainty: by sine wave
frequency (1st row) and dashed line frequency (2nd row).

Data	Mapping				
Rain probability, %	≥ 90	$[80, 90[$	$[70, 80[$	$[60, 70[$	< 60

In the second variant (see Fig. 6.17), the uncertainty is encoded by dash frequency like in [Gör+17]. To create dashed contours of the drops, a periodic colour transformation is applied as described in Sec. 6.2.4.2, alternating intervals of the original drop colour and a lighter colour. The dash frequency is controlled by interval arc length. To avoid a possible perceptual interference due to the multiple use of colour-based encoding, in this example the amount of precipitation is encoded by number of drops: depending on the parameter value, some drops are hidden by cancelling the respective DCs. For the data mapping details see Tabs. 6.6 and 6.7.

Since one of the goals is to enhance the visualization intuitivity by facilitation of metaphoric associations with the related real world phenomenon, the natural water colour, i.e. blue, is used as drop colour, appropriately selecting the background. In particular, it is a greyish green map, which allows for a sufficient contrast with the original dark blue drop colour as well as with the light shades of the scaled drops or dash line strokes in the first and second variant, respectively.

6.2.5.2   Gradient Glyphs

Multivariate data are often given as time series, while a glyph-based visualization usually captures steady data values. Accordingly, for a better understanding of the underlying *dynamics* in the multivariate data, it is also desirable to additionally visualize the temporal changes or trends. There are several visualization approaches, based on abstract glyphs, which visually mimic a function's slope to represent the data changes over time [GRT17; Wic+12; Car+99]. However, finding an intuitive gradient representation for an icon-based visualization, e.g. exploiting a metaphoric association between a visual variable and the data (see Sec. 2.1.1), is a challenging problem, as the main shape is already predefined by the base image.

The below section describes a glyph design proposal, based on the PACEMOD framework, that allows for visualization of up to two data parameters with the respective derivatives, using periodic contour modifications as additional visual variables. In particular, it exploits the fact that geometric contour modifications are closely related with other geometric visual variables of the glyph, such as size, while colour contour modifications are intuitively linked to the glyph's inner colour. Furthermore, it has to be assumed that the contour-based visual variables have a weaker pop-up effect. Due to the use of these perceptual features, the primary data become visually linked to the derived data, i.e. their gradient. In the following, the rules and considerations applied for design of a target glyph from an appropriate icon are explained in detail.

For the sake of brevity, in the following the two dimensional data parameter and the corresponding gradient components are described as p_i, p_i', $i = 1, 2$, respectively.

1. p_1 is mapped to the glyph's size and p_2 to its inner colour, as the visual variables with the highest pop-up effect (cf. Sec. 2.1.1).

2. The goal is to reflect the relation between p_i and p_i' on the visual level, that is the visual variables that encode p_i' need to be intuitively interpretable as *change of size* or *change of colour*, respectively.

3. To visualize p_1', a periodic, wave-like geometric modification of the glyph's contours with triangular waveform (see Sec. 6.2.4.1) is applied. The triangles symbolically represent arrowheads, which indicate the "movement" of the contour, i.e. the glyph's growing or shrinking. The wave amplitude encodes the amount of changes, i.e. the gradient magnitude. The gradient direction, i.e.

positive or negative, is visualized by varying the contrast between the triangular contour and the inner part of the glyph, which influences the perception of the "arrow" direction. Particularly, a narrow band along the contour is separated from the rest of the glyph by insertion of a new diffusion curve, which is created from a distance field isoline (see Fig. 6.18A). In the case of a positive change, the luminance inside of the band decreases, while the saturation increases, leading to a higher contrast in outward direction (see Figs. 6.18B and 6.19). In the same way, for negative gradients the band becomes brighter and more desaturated, creating a blur effect and a higher contrast in the outward and inward direction, respectively (see Fig. 6.19).

4. Analogously, p_2' is visualized by means of a periodic colour modification (see Sec. 6.2.4.2). For this purpose, a new region in the central part of the glyph is created. It is separated by an outer diffusion curve, where a corresponding colour modification is applied, and an inner diffusion barrier (see Fig. 6.18A). Both are created automatically from the distance field. The diffusion curve is subdivided in alternating segments of the glyph's inner colour and a "sign colour", which represents the change direction. The length of the sign colour segments encodes the gradient magnitude, while its direction is encoded by setting the sign colour as one of the both extremes of the colour map in use.

—	Original contour modified by a triangular wave.
—	Border of the arrow sign band.
—	Outer border of the colour gradient region.
—	Inner border of the colour gradient region.
•	Main colour points.
•	Colour gradient sign colour points.

(A)   (B)
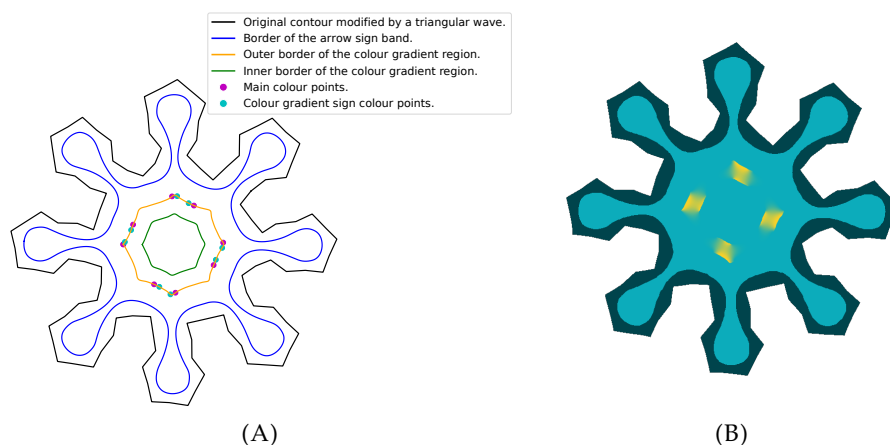
FIGURE 6.18: Gradient glyph. (A): the glyph structure; (B): the resulting glyph after rendering.

Based on these rules, a visualization of two of the most conclusive and widely used COVID-19 parameters, 7-days cases incidence and 7-days hospitalization incidence, were created (see Fig. 6.19). The data set is taken by the RKI [Rob92] and refers to the status on June 7, 2022. As the glyph's base icon serves a corona virus icon with the typical spikes. The 7-days cases incidence can be interpreted as a kind of "amount of virus", and thus intuitively encoded by size of glyph. The hospitalization incidence is mapped to glyph colour using parula colourmap, whereby the yellowish hues in the right part, which, correspondingly, represent higher incidence values, are intuitively interpretable as more dangerous. Since in this example the glyph colour is used as visual variable, a neutral grey was selected as background to minimize a possible interference of the colour perception. The visual encoding of theses data is shown in Tabs. 6.8 and 6.9, respectively. The definition of data intervals is similar to the RKI COVID-19 visualization [Rob22]. Besides the current incidence values, the difference to the previous week is an important parameter,
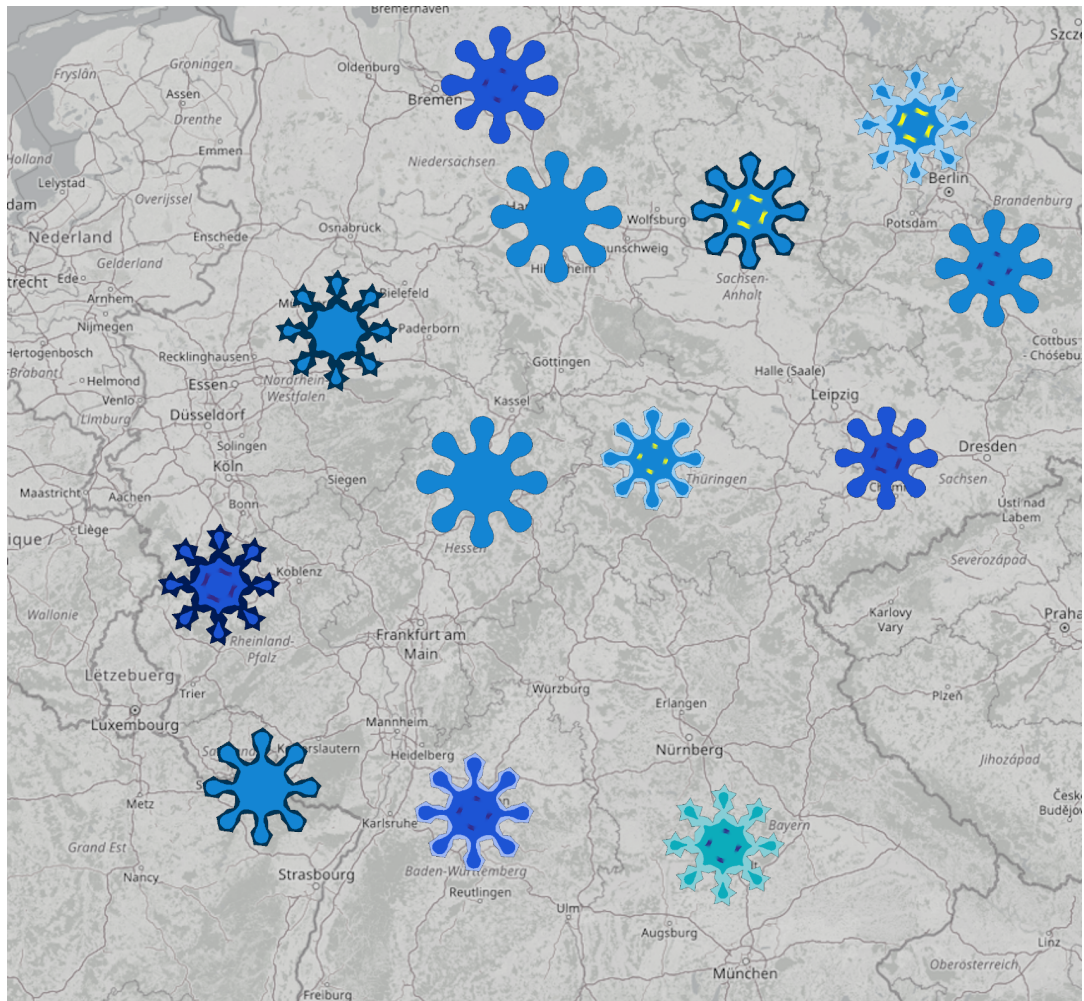
FIGURE 6.19: COVID-19 data visualization: 7-days cases incidence and hospitalization incidence along with the respective changes to previous week. Accordingly to the epidemiological situation as of June 7, 2022, the visualization only uses levels 1, 2 and 3 from Tab. 6.8.

which captures the dynamics of the disease. However, such differences are usually represented as a graph in a separate view, impairing a comprehensive overview of the status and the dynamics of the pandemic. The proposed glyph design allows for visualization of these differences simultaneously with the current values, by means of contour-based modifications. The corresponding visual encoding is summarized in Tab. 6.10. Note that to achieve an uniform mapping, the visual "zero gradient", i.e. no geometric modifications, includes low levels of data change. In total, the present approach aggregates in one single visualization the data that are usually spread in four separated views.

TABLE 6.8: Visual encoding of COVID-19 7-day cases incidence.

Data	Mapping					
7 d. cases incid, n./100K popul.	< 50	[50, 100[	[100, 250[	]250, 500]	]500, 1000]	> 1000

TABLE 6.9: Visual encoding colour bar of COVID-19 7-day hospitalization incidence.

Data	Mapping					
7 d. hosp. incid, n./100K popul.	≤ 1	]1, 2]	]2, 3]	]3, 4]	]4, 5]	> 5

TABLE 6.10: Visual encoding of COVID-19 data gradient.

Data	Mapping				
7 d. cases incid, last week diff., %	< −9	[−9, −3[	[−3, 3]	]3, 9]	> 9
7 d. hosp. incid, last week diff., %	< −24	[−24, −8[	[−8, 8]	]8, 24]	> 24

To evaluate the usability of the gradient glyph design, an online study with 22 participants, whose task was to assess the COVID-19 data visualization, given in Fig. 6.19, was performed. This test was mostly distributed among students and university employees. The study was anonymous and no personal data, such as age and gender, have been collected. After a brief explanation of the visual encoding, the participants were confronted with three main blocks of questions. Tab. 6.11 provides an overview of the applied questions (in an abbreviated form) and the respective test statistics.

The first block consists of five question (see Tab. 6.11, Q1-Q5), whose primary goal is to evaluate the visualization readability and comprehensibility. Each question presents a state characteristic regarding one or two COVID-19 parameters and a list of five states, for each of them the attendees needed to indicate whether the characteristic applies or not. In the second block (see Tab. 6.11, "Trend assessment block"), the goal is to assess the potential of the proposed visualization to facilitate the recognition of trends. In particular, it comprises five statements about statistical trends of the COVID-19 in Germany, and the task is to specify whether they are true or false. While in the first two blocks the map with glyphs was presented along with a legend (analogous to Tabs. 6.8-6.10), the legend was dropped in the third block (see Tab. 6.11, Q6-Q8) with the aim to evaluate the learnability, i.e. whether the participants were able to read the visualization without aid after a very short learning phase. The questions here are of the same type as in first block. Besides the aforementioned task classification, it has to be taken into account that the question difficulty level varies depending on the number of parameters to assess. In particular, questions Q1-Q3 and Q5-Q6 as well as statements 2 and 5 in the trend assessment

TABLE 6.11: Evaluation of the COVID-19 data visualization using gradient glyphs. The results represent percentage of the correct answers; bold numbers indicate the answer options where the respective statement does apply. We use the following abbreviations of the German state names: NRW: North Rhine-Westphalia, BR: Bremen, ST: Saxony-Anhalt, TH: Thuringia, SN: Saxony, BE: Berlin, RP: Rhineland-Palatinate, BW: Baden-Württemberg, BV: Bavaria, BB: Brandenburg, SL: Saarland, HE: Hesse, LS: Lower Saxony.

	Ans. a)	Ans. b)	Ans. c)	Ans. d)	Ans. e)
Min. 7-d. cases incid.	NRW: 82%	BR: 95%	ST: 100%	TH: **95**%	SN: **59**%
Max. 7-d. hosp. incid.	NRW: 100%	BE: 86%	RP: 100%	BW: 100%	BV: **91**%
Incr. 7-d. cases incid.	NRW: **86**%	TH: 86%	BB: 95%	SL: **73**%	BV: 100%
Decr. 7-d. cases AND hosp. incid.	ST: 95%	BB: 86%	RP: 91%	HE: 100%	BW: **82**%
Max. decr. 7-d. hosp. incid.	BR: 91%	ST: 100%	BB: 95%	RP: **95**%	SN: **95**%
Max. incr. 7-d. cases. incid.	BR: 100%	BE: 91%	NRW: **91**%	RP: **91**%	SL: 91%
Decr. 7-d. cases BUT incr. hosp. incid.	LS: 95%	BE: **91**%	TH: **77**%	SN: 91%	BV: 95%
Incr. 7-d. cases BUT decr. hosp. incid.	ST: 91%	BE: 100%	RP: **86**%	HE: 100%	BV: 95%
The states with the highest 7-d. cases incid. also have the highest 7-d. hosp. incid.					False: 82%
The 7-d. cases incid. tends to increase in the western regions faster than in the eastern ones.					True: 86%
A decreasing 7-d. cases incid. always correlates with a decrease in 7-d. hosp. incid.					False: 91%
A higher 7-d. hosp. incid. always implies a more rapid increase in this parameter.					False: 95%
In South Germany, the 7-d. hosp. incid. is stable or decreasing.					True: 91%

block require assessment of one single data parameter, while the remaining questions are two-dimensional.

Tab. 6.11 shows the test results as percentage of correct answers per question. In summary, this percentage lies in 42 of 45 cases between 82% and 100%, which demonstrates the suitability of the visualization to convey the encoded information and to serve as basis for trend recognition, despite a very short learning time. In the following, two cases that are below 75% are discussed.

In Q1, answer option e), the percentage of correct answers is only 59%. In this question, the task is to identify the listed states with the lowest 7-day cases incidence, and the correct answers are options d) Thuringia (the glyph close to "Thüringen" in Fig. 6.19) and e) Saxony (the glyph close to "Sachsen" in Fig. 6.19). Both glyphs have the same size, encoding the 7-day cases incidence interval $[50, 100[$. However, the Thuringia-glyph also has a brighter band along its contour, which encodes a negative last week difference for this COVID-19 parameter, which results in a visual "shrinkage". Due to this visual effect, Thuringia-glyph appears smaller than the Saxony-glyph for 41% participants. In general, the "visual shrinkage" is an intentional effect, aiming at an intuitive encoding of a negative gradient of the main parameter, visualized by size. The critical aspect here is the superposition of the "explicit" geometric scaling and "implicit" visual shrinking that may lead to the misinterpretation as the next smallest size. An option to compensate this effect of implicit visual shrinking, which has to be investigated in future work, is application of a slight glyph scaling. Moreover, it has to be taken into account that, in this study, the respective problem appears in the very first question when the participants are still not acquainted with the visualization concept.

The second case of low performance corresponds to Q3, option d), where 73% of the participants gave the correct answer. Here, the task is to select all listed states where the 7-day cases incidence is increasing, which are a) North Rhine-Westphalia (the glyph above "Nordrhein-Westfalen" in Fig. 6.19) and d) Saarland (the glyph above "Strasbourg" in Fig. 6.19). The North Rhine-Westphalia-glyph (correctly identified by 86%) has a higher contour wave amplitude, and thus is more salient than
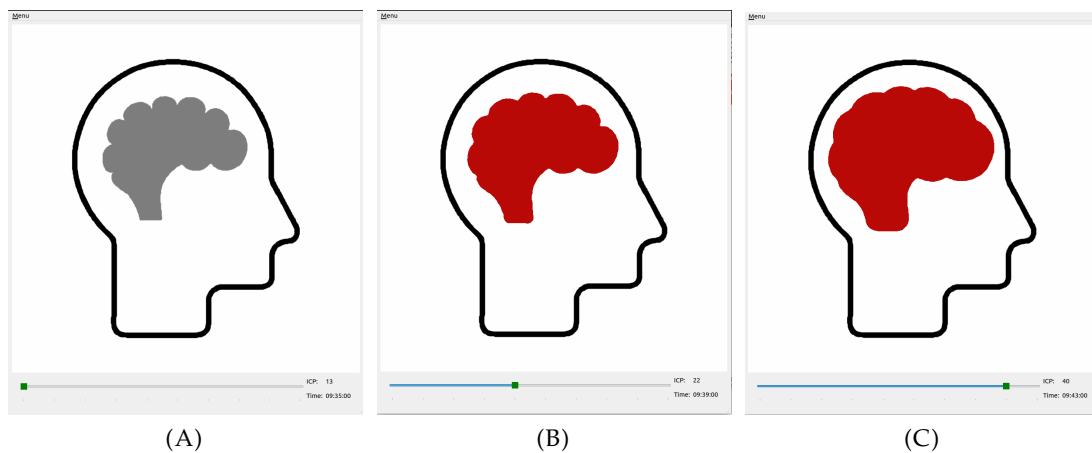
(A) (B) (C)

FIGURE 6.20: Prototype for visualization of intracranial pressure (ICP). (A): Normal pressure.
(B): A pressure value inside the pathological range. The increased value is visualized by a higher brain volume and shallower sulci; the discrete "pathological" category is represented by dark red colour. The head contour remains unchanged.
(C): A still higher pathological value. The category and, therefore, the brain colour remain same but the volume keeps growing with the continues pressure value.

the Saarland-glyph, which might have diverted the attention of the participants away from the latter.

6.2.6 Conclusion

In this section, a novel approach for automated contour-based icon modifications was introduced. This approach provides the *PArametric Contour-basEd MODification (PACEMOD)* concept for automated, parametric manipulations of the geometry and colour of a given base icon and, thus, contributes to the efficient creation of metaphoric glyphs. In particular, the automation potentially supports user-centred design, allowing a fast or even real-time feedback integration, and thus short iterations in the collaborative visualization developing, e.g., with domain experts. Furthermore, a specific implementation of the PACEMOD concept for periodic contour modifications was provided and the suitability of the PACEMOD functionalities for transferring existing abstract glyph-based approaches, e.g. in uncertainty visualization, to metaphoric glyphs as well as for developing of new glyph designs, such as gradient glyphs, was demonstrated. In both application examples, the generated glyphs allow for a visually integrated assessment of the multivariate data, which is commonly only achievable using several views.

Moreover, the proposed generic approach can be combined with further modification strategies such as segment-related variation. In particular, to demonstrate the feasibility of this concept application, besides the rain drop example, discussed above, a prototype for visualization of intracranial pressure was developed in collaboration with neurosurgeons and social scientists [Pre+23a]. It uses a base icon that schematically shows a human head with brain inside (see Fig. 6.20). The scalar pressure value is mapped to volume/folding of the brain, i.e. an icon's segment, using the "inflation under pressure" metaphor, that is the higher the value, the larger

the volume and the shallower the grooves. Additionally, the discrete pressure categories, i.e. normal, elevated, pathological, are mapped to the brain's inner colour, exploiting the traffic light metaphor.

Chapter 7

Conclusion

This thesis introduced a novel concept of Visually Integrated Clinical Cooperation (VICC), aimed to enhance the information transfer between medical personnel, especially the conveyance of the patient's current status. The concept was designed basing on a interdisciplinary collaboration in the context of the project A06 in CRC 1187 "Media of Cooperation" at the University of Siegen, i.e. exploiting the findings of a field study on the neurosurgical ward, conducted by social scientists, as well as ideas and feedback collected during the frequent discussions with physicians and social scientists.

The thesis addressed the problem of a visual support for information transfer in a hospital on different levels of abstraction. First, describing an overall visualization concept (see Chapter 3). Second, providing specific implementations of some of its components, comprising the creation of prototypical systems and their evaluation. In particular, these components are:

- *Anatomically integrated visualization*, which allows for a synoptical in-place representation of the patient's data with anatomical reference, such as clinical symptoms, on a 3D human body model (see Chapter 4).

- A method for robust *range camera pose estimation* in the context of mobile online scene reconstruction, using inertial data and an EKF-based sensor fusion algorithm (see Chapter 5). This is a preliminary step of the anatomical body model personalization, which can be optionally applied for the aforementioned anatomically integrated visualization.

- An approach for (semi-)automated generation of *iconic glyphs* with an augmented design space due to contour-based visual variables (see Chapter 6). Besides, a *perception and quantization model* for this type of visual variables, derived from the results of a user study, was proposed (see Sec. 6.1). The iconic glyphs can represent, inter alias, patients' multivariate data without anatomical reference as well as their temporal trends.

A prototype for the anatomically integrated in-place visualization of patient's symptoms by a spinal disc herniation was evaluated by a group of neurosurgeons, obtaining a positive feedback. Some gradient glyphs, created with the proposed iconic glyph approach, were subject of an online user study, whose results demonstrated their suitability for trend visualization.

At the same time, since the VICC concept handles a complex real-life problem, operating heterogeneous technical approaches, it should not be considered as a completed solution but rather as basis and inspiration for further research. Again, different directions for future work open at the concept level as well as at the level of its single components.

First, some concept components have not been implemented yet or, at least, are incomplete. In particular, in the personalization option for the anatomical body model, the 3D reconstruction step needs to be completed with an appropriate model fitting approach, as briefly discussed in Chapter 5.

Furthermore, a combined application of the anatomically integrated visualization and iconic glyphs appears to be a promising research direction. In this case, the automatically created glyphs can serve as quantifiable textures for respective anatomical structures, providing a more intuitive and precise appearance control than the hitherto used procedural textures. However, such an application also implies new challenges to face, for instance, the need of a suitable texture mapping method. Also, the use of glyphs for texturing of 3D objects would influence the perception of visual variables, e.g. due to a potentially limited glyph size, requiring appropriate compensatory mechanisms.

Second, the already implemented components still have a potential of improvement. Regarding the anatomically integrated visualization (Chapter 4), for example, the in-place symptom representation can be augmented with common spatial medical data such as CT imaging, or with 3D wound models for healing monitoring, which can be also created on a mobile device as described in [SPK19]. This augmentation requires a registration procedure between the anatomical body model and these external 3D data.

Several enhancement possibilities can be outlined for the proposed range camera pose estimation method (Chapter 5). The filter accuracy, e.g., can be potentially increased by more sophisticated, adaptive noise estimation methods (cf. Sec. 5.4.3). Moreover, the tracking robustness in scenarios with sparse geometric features can be improved by integrating a corresponding ICP failure detection logic.

The perception and quantization model for glyph contour modifications, presented in Sec. 6.1, is based on a simplifying assumption of no or negligible weak dependencies between single visual variables. Even if this work provides first insights, which verify this assumption, it still requires specific user studies for a more in-depth investigation. Besides, the proposed scale-dependent adaptation rules (see Sec. 6.1.5) also need an experimental verification.

The parametric iconic glyph generation (Sec. 6.2) is conceptualized as a generalizable approach, i.e. its application is not limited to the design examples proposed in this thesis. Conceivable further implementations are more sophisticated schemes for locally adaptive applications of shape and colour modifications, e.g. changing appearance of specific semantic segments of the base icon, which, however, implies an automated segmentation method or more user interaction.

Bibliography

[AS16]   F. Aghili and C.-Y. Su. "Robust Relative Navigation by Integration of ICP and Adaptive Kalman Filter Using Laser Scanner and IMU". In: *IEEE/ASME Transactions on Mechatronics* 21.4 (2016), pp. 2015–2026.

[AOB08]   R. Allendes Osorio and K. W. Brodlie. "Contouring with uncertainty". In: *Proc. EG UK Theory and Practice of Computer Graphics*. Eurographics Association. 2008, pp. 59–66.

[AFC05]   B. Amidan, T. Ferryman, and S. Cooley. "Data outlier detection using the Chebyshev theorem". In: *Proc. IEEE Aerospace Conference*. 2005, pp. 3814–3819.

[An+10]   J. An, Z. Wu, H. Chen, X. Lu, and H. Duan. "Level of Detail Navigation and Visualization of Electronic Health Records". In: *Proc. Int. Conf. Biomedical Engineering and Informatics*. Vol. 6. 2010, pp. 2516–2519.

[Att54]   F. Attneave. "Some informational aspects of visual perception." In: *Psychological review* 61.3 (1954), pp. 183–193.

[Bar+21]   K. Bartol, D. Bojanić, T. Petković, and T. Pribanić. "A review of body measurement using 3D scanning". In: *Ieee Access* 9 (2021), pp. 67281–67301.

[Bel+18]   J. L. Belden, P. Wegier, J. Patel, A. Hutson, C. Plaisant, J. L. Moore, N. J. Lowrance, S. A. Boren, and R. J. Koopman. "Designing a medication timeline for patients and physicians". In: *Journal of the American Medical Informatics Association* 26.2 (2018), pp. 95–105.

[Ber83]   J. Bertin. *Semiology of Graphics*. University of Wisconsin Press, 1983.

[BM92]   P. J. Besl and N. D. McKay. "A Method for Registration of 3-D Shapes". In: *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)* 14.2 (1992), pp. 239–256.

[Bez+10]   H. Bezerra, E. Eisemann, D. DeCarlo, and J. Thollot. "Diffusion Constraints for Vector Graphics". In: *Proc. ACM Int. Symp. Non-Photorealistic Animation and Rendering*. 2010, pp. 35–42.

[BL95]   G. Blais and M. D. Levine. "Registering Multiview Range Data to Create 3D Computer Objects". In: *IEE Transactions on Pattern Analysis and Machine Intelligence* 17.8 (1995), pp. 820–824. DOI: 10.1109/34.400574.

[Boe80]   W. Boehm. "Inserting new knots into B-spline curves". In: *Computer-Aided Design* 12.4 (1980), pp. 199–201.

[BFK84]   W. Böhm, G. Farin, and J. Kahmann. "A survey of curve and surface methods in CAGD". In: *Computer Aided Geometric Design* 1.1 (1984), pp. 1–60. DOI: https://doi.org/10.1016/0167-8396(84)90003-7.

[Bon+14]   G.-P. Bonneau, H.-C. Hege, C. R. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz. "Overview and state-of-the-art of uncertainty visualization". In: *Scientific Visualization*. Springer, 2014, pp. 3–27.

[Boo78]   C. de Boor. *A Practical Guide to Spline*. Vol. 27. Springer, 1978.

[Bor+13]   R. Borgo, J. Kehrer, D. H. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward, and M. Chen. "Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications." In: *Eurographics (STARs)*. 2013, pp. 39–63.

[BAK07]   A. A. T. Bui, D. R. Aberle, and H. Kangarloo. "TimeLine: Visualizing Integrated Patient Records". In: *IEEE Transactions on Information Technology in Biomedicine* 11.4 (2007), pp. 462–473. DOI: 10.1109/TITB.2006.884365.

[BH95]   G. Burel and H. Henoco. "Determination of the orientation of 3D objects using spherical harmonics". In: *Graphical Models and Image Processing* 57.5 (1995), pp. 400–408.

[Bur+78]   B. Burns, B. E. Shepp, D. McDonough, and W. K. Wiener-Ehrlich. "The Relation Between Stimulus Analyzability and Perceived Dimensional Structure." In: ed. by G. H. Bower. Vol. 12. Psychology of Learning and Motivation. Academic Press, 1978, pp. 77–115. DOI: https://doi.org/10.1016/S0079-7421(08)60008-0.

[BF08]   M. Burns and A. Finkelstein. "Adaptive Cutaways for Comprehensible Rendering of Polygonal Scenes". In: *ACM Trans. Graph.* 27.5 (2008). ISSN: 0730-0301. DOI: 10.1145/1409060.1409107. URL: https://doi.org/10.1145/1409060.1409107.

[Cai+15]   Z. Cai, Y.-N. Li, X. S. Zheng, and K. Zhang. "Applying feature integration theory to glyph-based information visualization". In: *Proc. IEEE Pacific Visualization Symp.* 2015, pp. 99–103.

[Cam+15]   M. Camurri, S. Bazeille, D. G. Caldwell, and C. Semini. "Real-time depth and inertial fusion for local SLAM on dynamic legged robots". In: *Proc. IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE. 2015, pp. 259–264.

[Car12]   S. Card. "Information Visualization". In: *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. Ed. by J. A. Jacko. 3rd. CRC press, 2012. Chap. 23, pp. 515–548.

[Car03]   M. S. T. Carpendale. *Considering Visual Variables as a Basis for Information Visualisation*. Tech. rep. Calgary, AB: University of Calgary, 2003.

[Car+99]   D. B. Carr, A. R. Olsen, S. M. Pierson, and J. Courbois. "Boxplot variations in a spatial context: An Omernik ecoregion and weather example". In: *Statistical Computing & Statistical Graphics Newsletter* 9 (1999), pp. 4–13.

[CR04]   G. Casciola and L. Romani. "A generalized conversion matrix between non-uniform B-spline and Bézier representations with applications in CAGD". In: *Multivariate Approximation: Theory and Applications*. Springer, 2004. DOI: 10.6092/unibo/amsacta/853.

[Cav+14]   A Cavallo, A Cirillo, P Cirillo, G De Maria, P Falco, C Natale, and S Pirozzi. "Experimental Comparison of Sensor Fusion Algorithms for Attitude Estimation". In: *IFAC Proceedings Volumes* 47.3 (2014), pp. 7585–7591.

[Cha06]   W. W.-Y. Chan. "A survey on multivariate data visualization". In: *Department of Computer Science and Engineering. Hong Kong University of Science and Technology* 8.6 (2006), pp. 1–29.

[CM92]   Y. Chen and G. Medioni. "Object modelling by registration of multiple range images". In: *Image and vision computing* 10.3 (1992), pp. 145–155.

[Cho+14]   J. C. Chow, D. D. Lichti, J. D. Hol, G. Bellusci, and H. Luinge. "Imu and multiple RGB-D camera fusion for assisting indoor stop-and-go 3D terrestrial laser scanning". In: *Robotics* 3.3 (2014), pp. 247–280.

[Chu+16]   D. H. S. Chung, D. Archambault, R. Borgo, D. J. Edwards, R. S. Laramee, and M. Chen. "How Ordered Is It? On the Perceptual Orderability of Visual Channels". In: *Computer Graphics Forum* 35.3 (2016), pp. 131–140. DOI: https://doi.org/10.1111/cgf.12889. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12889.

[Chu+13]   D. H. Chung, P. A. Legg, M. L. Parry, R. Bown, I. W. Griffiths, R. S. Laramee, and M. Chen. "Glyph sorting: Interactive visualization for multi-dimensional data". In: *Information Visualization* 14.1 (2013), pp. 76–90.

[Cra11]   D. Craig. "An EHR interface for viewing and accessing patient health events from collaborative sources". In: *Proc. Int. Conf. Collaboration Technologies and Systems (CTS)*. 2011, pp. 319–324.

[CMM18]   J. M. Cunha, P. Martins, and P. Machado. "Emojinating: Representing concepts using emoji". In: *Proc. Int. Conf. Case-Based Reasoning (ICCBR) Workshop*. Vol. 185. 2018.

[Cun+18]   J. M. Cunha, E. Polisciuc, P. Martins, and P. Machado. "The many-faced plot: strategy for automatic glyph generation". In: *Proc. Int. Conf. Information Visualisation*. IEEE. 2018, pp. 71–77.

[CL96]   B. Curless and M. Levoy. "A volumetric method for building complex models from range images". In: *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. 1996, pp. 303–312.

[DOL03]   M. F. De Oliveira and H. Levkowitz. "From visual data exploration to visual data mining: A survey". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 9.3 (2003), pp. 378–394.

[Dry+17]   I. Dryanovski, M. Klingensmith, S. S. Srinivasa, and J. Xiao. "Large-scale, real-time 3D scene reconstruction on a mobile device". In: *Autonomous Robots* (2017), pp. 1–23.

[ESHN08]   N. El-Sheimy, H. Hou, and X. Niu. "Analysis and Modeling of Inertial Sensors Using Allan Variance". In: *IEEE Trans. Instrumentation and Measurement* 57.1 (2008), pp. 140–149.

[Ema+17]   S. Emani, D. Y. Ting, M. Healey, S. R. Lipsitz, A. S. Karson, and D. W. Bates. "Physician beliefs about the meaningful use of the electronic health record: a follow-up study". In: *Applied Clinical Informatics* 8.04 (2017), pp. 1044–1053.

[Far02]   G. Farin. *Curves and surfaces for CAGD: a practical guide*. Morgan Kaufmann, 2002.

[FR81]   B. Flury and H. Riedwyl. "Graphical representation of multivariate data by means of asymmetrical faces". In: *Journal of the American Statistical Association* 76.376 (1981), pp. 757–765.

[FB93]   B. Fowler and R. Bartels. "Constraint-based curve manipulation". In: *IEEE Computer Graphics and Applications* 13.5 (1993), pp. 43–49.

[Fuc+16]   J. Fuchs, P. Isenberg, A. Bezerianos, and D. Keim. "A systematic review of experimental studies on data glyphs". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 23.7 (2016), pp. 1863–1879.

[Fuc+15]   J. Fuchs, D. Jäckle, N. Weiler, and T. Schreck. "Leaf Glyph : Visualizing Multi-Dimensional Data with Environmental Cues". In: *Proc. Int. Conf. Information Visualization Theory and Applications (IVAPP)*. 2015, pp. 195–208.

[FPRARM15]   J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha. "Visual simultaneous localization and mapping: a survey". In: *Artificial Intelligence Review* 43.1 (2015), pp. 55–81.

[GRT17]   T. Gerrits, C. Rössl, and H. Theisel. "Glyphs for space-time jacobians of time-dependent vector fields". In: *Journal of WSCG* 25 (Jan. 2017), pp. 31–38.

[GP16]   V. Gikas and H. Perakis. "Rigorous Performance Evaluation of Smartphone GNSS/IMU Sensors for ITS Applications". In: *Sensors* 16.8 (2016). Article 1240. DOI: 10.3390/s16081240.

[Gli+19]   B. S. Glicksberg et al. "PatientExploreR: an extensible application for dynamic visualization of patient clinical history from electronic health records in the OMOP common data model". In: *Bioinformatics* 35.21 (June 2019), pp. 4515–4518. DOI: 10.1093/bioinformatics/btz409.

[Gör+17]   J. Görtler, C. Schulz, D. Weiskopf, and O. Deussen. "Bubble treemaps for uncertainty visualization". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 24.1 (2017), pp. 719–728.

[GA14]   M. S. Grewal and A. P. Andrews. *Kalman filtering: Theory and Practice with MATLAB*. John Wiley & Sons, 2014.

[GWA07]   M. S. Grewal, L. R. Weill, and A. P. Andrews. *Global positioning systems, inertial navigation, and integration*. John Wiley & Sons, 2007.

[Hak+17]   A. Hakone, L. Harrison, A. Ottley, N. Winters, C. Gutheil, P. K. Han, and R. Chang. "PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication". In: *IEEE Transactions on Visualization and Computer Graphics* 23.1 (2017), pp. 601–610.

[HI72]   S. Handel and S. Imai. "The free classification of analyzable and unanalyzable stimuli." In: *Perception & Psychophysics* (1972).

[He+15]   S. He, D. Ou, C. Yan, and C.-H. Lee. "A chord error conforming tool path B-spline fitting method for NC machining based on energy minimization and LSPIA". In: *Journal of Computational Design and Engineering* 2.4 (2015), pp. 218–232.

[HBG12]   T. Hervier, S. Bonnabel, and F. Goulette. "Accurate 3D Maps from Depth Images and Motion Sensors via Nonlinear Kalman Filtering". In: *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*. IEEE, 2012, pp. 5291–5297.

[Hol+19]   N. S. Holliman, A. Coltekin, S. J. Fernstad, M. D. Simpson, K. J. Wilson, and A. J. Woods. "Visual entropy and the visualization of uncertainty". In: *arXiv preprint arXiv:1907.12879* (2019).

[HZY15]   J Huai, Y Zhang, and A Yilmaz. "Real-time large scale 3D reconstruction by fusing Kinect and IMU data". In: *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* II-3/W5 (2015), pp. 491–496.

[Inf15]   Infineon. *Real3 (TM) image sensor family - 3D depth sensing based on Time-of-Flight*. http://www.infineon.com/dgdl/Infineon-REAL3%20Image%20Sensor%20Family-PB-v01_00-EN.PDF?fileId=5546d462518ffd850151a0afc2302a58. Product Brief. 2015.

[Iza+11]   S. Izadi et al. "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera". In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 2011, pp. 559–568.

[Jaa+19]   E. A. A. Jaatun, M. Fallon, A. Kofod-Petersen, K. Halvorsen, and D. F. Haugen. "Users' perceptions on digital visualization of neuropathic cancer-related pain". In: *Health Informatics Journal* 25.3 (2019), pp. 683–700.

[Jek12]   C. Jekeli. "Inertial navigation systems with geodetic applications". In: *Inertial Navigation Systems with Geodetic Applications*. de Gruyter, 2012.

[Jes16]   S. Jeschke. "Generalized diffusion curves: An improved vector representation for smooth-shaded images". In: *Computer Graphics Forum* 35.2 (2016), pp. 71–79.

[JCW09]   S. Jeschke, D. Cline, and P. Wonka. "A GPU Laplacian solver for diffusion curves and Poisson image editing". In: *ACM Trans. Graphics (Proc. SIGGRAPH Asia)* (2009), pp. 1–8.

[JCW11]   S. Jeschke, D. Cline, and P. Wonka. "Estimating color and texture parameters for vector graphics". In: *Computer Graphics Forum* 30.2 (2011), pp. 523–532.

[Käh+15]   O. Kähler, V. A. Prisacariu, C. Y. Ren, X. Sun, P. Torr, and D. Murray. "Very High Frame Rate Volumetric Integration of Depth Images on Mobile Devices". In: *IEEE Transactions on Visualization and Computer Graphics* 21.11 (2015), pp. 1241–1250.

[KK96]   D. A. Keim and H.-P. Kriegel. "Visualization techniques for mining large databases: A comparison". In: *IEEE Trans. Knowledge and Data Engineering* 8.6 (1996), pp. 923–938.

[Kel+13]   M. Keller, D. Lefloch, M. Lambers, S. Izadi, T. Weyrich, and A. Kolb. "Real-Time 3D Reconstruction in Dynamic Scenes Using Point-Based Fusion". In: *Proc. Int. Conf. 3D Vision (3DV)*. IEEE, 2013, pp. 1–8.

[KS14]   A. Kerren and F. Schreiber. "Why Integrate InfoVis and SciVis?: An Example from Systems Biology". In: *IEEE Computer Graphics and Applications* 34.6 (2014), pp. 69–73. DOI: 10.1109/MCG.2014.122.

[KR96]   J. Kirby and A. L. Rector. "The PEN&PAD Data Entry System: From prototype to practical system". In: *Proc. AMIA Annual Fall Symposium*. 1996, pp. 709–713.

[Kli+15]   M. Klingensmith, I. Dryanovski, S. S. Srinivasa, and J. Xiao. "Chisel: Real Time Large Scale 3D Reconstruction Onboard a Mobile Device using Spatially-Hashed Signed Distance Fields". In: *Robotics: Science and Systems*. 2015.

[KP15]   A. Kolb and F. Pece. "Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality". In: AK Peters / CRC Press, 2015. Chap. Range Imaging, pp. 51–64.

[KT75]   D. H. Krantz and A. Tversky. "Similarity of rectangles: An analysis of subjective dimensions". In: *Journal of mathematical Psychology* 12.1 (1975), pp. 4–34.

[KS08]   R. G. Kuehni and A. Schwarz. *Color Ordered: A Survey of Color Systems from Antiquity to the Present*. Oxford University Press, 2008.

[KYW14]   T.-H. Kwok, K.-Y. Yeung, and C. C. Wang. "Volumetric template fitting for human body reconstruction from incomplete data". In: *Journal of Manufacturing Systems* 33.4 (2014), pp. 678–689.

[LCD12]   R. L. Larkins, M. J. Cree, and A. A. Dorrington. "Analysis of binning of normals for spherical harmonic cross-correlation". In: *Proc. SPIE-IS&T Electronic Imaging*. Vol. 8290. 2012.

[Lef+17]   D. Lefloch, M. Kluge, H. Sarbolandi, T. Weyrich, and A. Kolb. "Comprehensive use of curvature for robust and accurate online surface reconstruction". In: *IEEE transactions on pattern analysis and machine intelligence* 39.12 (2017), pp. 2349–2365.

[Leg+12]   P. A. Legg, D. H. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. "MatchPad: interactive glyph-based visualization for real-time sports performance analysis". In: *Computer Graphics Forum* 31.3pt4 (2012), pp. 1255–1264.

[LMW10]   J. Li, J.-B. Martens, and J. J. van Wijk. "A Model of Symbol Size Discrimination in Scatterplots". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 2553–2562. ISBN: 9781605589299. DOI: 10.1145/1753326.1753714. URL: https://doi.org/10.1145/1753326.1753714.

[LLZ15]   Y.-N. Li, D.-J. Li, and K. Zhang. "Metaphoric transfer effect in information visualization using glyphs". In: *Proc. Int. Symp. Visual Information Communication and Interaction*. 2015, pp. 121–130.

[Liu+16]   D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, and Y. Wu. "Smartadp: Visual analytics of large-scale taxi trajectories for selecting billboard locations". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 23.1 (2016), pp. 1–10.

[Lu+20]   S. Lu, X. Ding, F. Gao, and J. Chen. "Shape Manipulation of Diffusion Curves Images". In: *IEEE Access* 8 (2020), pp. 57158–57167.

[Lu+19]   S. Lu, W. Jiang, X. Ding, C. S. Kaplan, X. Jin, F. Gao, and J. Chen. "Depth-aware image vectorization and editing". In: *The Visual Computer* 35.6 (2019), pp. 1027–1039.

[Mag+12]   E. Maguire, P. Rocca-Serra, S.-A. Sansone, J. Davies, and M. Chen. "Taxonomy-Based Glyph Design—with a Case Study on Visualizing Workflows of Biological Experiments". In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2603–2612.

[MS13]   M. S. A. Malik and S. Sulaiman. "Doctor's Perspective For Use of EHR visualization systems in public hospitals". In: *Proc. Science and Information Conference*. 2013, pp. 86–92.

[Mar+20]   N. Martignene, T. Balcaen, G. Bouzille, M. Calafiore, J.-B. Beuscart, A. Lamer, B. Legrand, G. Ficheur, and E. Chazard. "Heimdall, a Computer Program for Electronic Health Records Data Visualization". In: *Studies in health technology and informatics* 270 (June 2020), pp. 247–251. DOI: 10.3233/SHTI200160.

[Mar03]   R. C. Martin. *Agile Software Development: Principles, Patterns, and Practices*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2003.

[Mas+19]   L. J. Masulo, M. L. S. Martins, D. R. Costa, and R. A. Nicolau. "Methods for qualitative and quantitative analysis of pain and quality of life validated in Brazil: systematic review". In: *RGO - Revista Gaúcha de Odontologia* 67 (2019).

[MQR03]   C. R. Maurer, R. Qi, and V. Raghavan. "A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions". In: *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)* 25.2 (2003), pp. 265–270.

[May+08]   S. May, D. Droeschel, D. Holz, C. Wiesen, and S. Fuchs. "3D pose estimation and mapping with time-of-flight cameras". In: *IROS Workshop on 3D Mapping*. 2008.

[May79]   P. S. Maybeck. *Stochastic models, estimation, and control*. Vol. 1. Academic Press, 1979.

[May82]   P. S. Maybeck. *Stochastic models, estimation, and control*. Vol. 2. Academic Press, 1982.

[Meh70]   R. K. Mehra. "On the identification of variances and adaptive Kalman filtering". In: *IEEE Transactions on Automatic Control* 15.2 (1970), pp. 175–184.

[Mur+16]   O. Muratov, Y. Slynko, V. Chernov, M. Lyubimtseva, A. Shamsuarov, and V. Bucha. "3DCapture: 3D Reconstruction for a Smartphone". In: *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Workshops*. IEEE, 2016, pp. 291–299.

[New+11]　R. A. Newcombe et al. "KinectFusion: Real-time dense surface mapping and tracking". In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2011, pp. 127–136. DOI: 10.1109/ISMAR.2011.6092378.

[NDF14]　M. Nießner, A. Dai, and M. Fisher. "Combining Inertial Navigation and ICP for Real-time 3D Surface Reconstruction". In: *Eurographics (Short Papers)*. EG, 2014, pp. 13–16.

[Nie+13]　M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger. "Real-time 3D reconstruction at scale using voxel hashing". In: *ACM Transactions on Graphics (ToG)* 32.6 (2013), pp. 1–11.

[NSS05]　T. Nocke, S. Schlechtweg, and H. Schumann. "Icon-based visualization using mosaic metaphors". In: *Proc. Int. Conf. Information Visualisation*. 2005, pp. 103–109.

[OKI15]　P. Ondrúška, P. Kohli, and S. Izadi. "MobileFusion: Real-Time Volumetric Surface Reconstruction and Dense Tracking on Mobile Phones". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 21.11 (2015), pp. 1251–1258.

[Orz+08]　A. Orzan, A. Bousseau, H. Winnemöller, P. Barla, J. Thollot, and D. Salesin. "Diffusion curves: a vector representation for smooth-shaded images". In: *ACM Trans. Graphics* 27.3 (2008). https://maverick.inria.fr/Publications/2008/OBWBTS08/, pp. 1–8.

[Pla+98]　C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman. "LifeLines: using visualization to enhance navigation and analysis of patient records". In: *Proc. AMIA Symp.* 1998, pp. 76–80.

[pmd15]　pmd technologies GmbH. *Reference Design Brief CamBoard pico flexx*. http://pmdtec.com/picoflexx/downloads/PMD_RD_Brief_CB_pico_flexx_V0201.pdf. Reference Design Brief. 2015.

[PRH10]　J.-S. Prassni, T. Ropinski, and K. Hinrichs. "Uncertainty-aware guided volume segmentation". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 16.6 (2010), pp. 1358–1365.

[PBP02]　H. Prautzsch, W. Boehm, and M. Paluszny. *Bézier and B-spline techniques*. Vol. 6. Springer, 2002.

[PBK23]　D. Presnov, M. Berels, and A. Kolb. "Pacemod: parametric contour-based modifications for glyph generation". In: *The Visual Computer* (2023). DOI: 10.1007/s00371-023-03040-4. URL: https://doi.org/10.1007/s00371-023-03040-4.

[PK22]　D. Presnov and A. Kolb. "Perception and Quantization Model for Periodic Contour Modifications". In: *Journal of Imaging* 8.11 (2022). Article 311. ISSN: 2313-433X. URL: https://www.mdpi.com/2313-433X/8/11/311.

[Pre+23a]　D. Presnov, J. Kurz, J. Dillmann, D. Alt, C. Schubert, V. Braun, M. Yavuz, and A. Kolb. *ICP monitoring visualizes short term ICP fluctuation. Approaches recognizing dependencies between the ICP magnitude and duration are not intended for live patient monitoring. Icon-based visualization of ICP is a promising alternative new concept with benefits in immediate trend assessment.* https://www.dgnc.de/fileadmin/media/jahrestagung/2023/Dokumente/DGNC2023_Abstractband.pdf, p.

571. Poster presented at 74th Annual Meeting of the German Society of Neurosurgery (DGNC). 2023.

[Pre+23b] D. Presnov, J. Kurz, J. Willkomm, J. Dillmann, D. Alt, R. Zilke, V. Braun, C. Schubert, and A. Kolb. "Anatomically Integrated In-Place Visualization of Patient Data for Cooperative Tasks with a Case Study on a Neurosurgical Ward". In: *Health Informatics Journal* 29.2 (2023). PMID: 37137867, p. 14604582231171878. DOI: 10 . 1177 / 14604582231171878. URL: https : / / doi . org / 10 . 1177 / 14604582231171878.

[PLK18] D. Presnov, M. Lambers, and A. Kolb. "Robust range camera pose estimation for mobile online scene reconstruction". In: *IEEE Sensors Journal* 18.7 (2018), pp. 2903–2915.

[Rin+13] A. Rind, T. D. Wang, W. Aigner, S. Miksch, K. Wongsuphasawat, C. Plaisant, and B. Shneiderman. "Interactive information visualization to explore and query electronic health records". In: *Foundations and Trends in Human–Computer Interaction* 5.3 (2013), pp. 207–298.

[Ris08] J. S. Risch. "On the role of metaphor in information visualization". In: *arXiv preprint arXiv:0809.0884* (2008).

[Rob22] Robert Koch Institut. *COVID-19-Dashboard*. https : / / experience . arcgis.com/experience/478220a4c454480e823b17327b2bf1d4. last accessed Nov. 21st, 2022. 2022.

[Rob92] Robert Koch Institute. *Coronavirus SARS-CoV-2*. https://www.rki . de / DE / Content / InfAZ / N / Neuartiges _ Coronavirus / Daten / Inzidenz-Tabellen.html. 7/09/2022.

[RHHL02] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. "Real-time 3D model acquisition". In: *ACM Transactions on Graphics (TOG)* 21.3 (2002), pp. 438–446.

[Sab11] A. M. Sabatini. "Kalman-Filter-Based Orientation Determination Using Inertial/Magnetic Sensors: Observability Analysis and Performance Evaluation". In: *Sensors* 11.10 (2011), pp. 9182–9206. DOI: 10 . 3390/s111009182.

[Sal+07] J. Salvi, C. Matabosch, D. Fofi, and J. Forest. "A review of recent range image registration methods with accuracy evaluation". In: *Image and Vision computing* 25.5 (2007), pp. 578–596.

[Sch+15] T. Schöps, T. Sattler, C. Häne, and M. Pollefeys. "3D Modeling on the Go: Interactive 3D Reconstruction of Large-Scale Scenes on Mobile Devices". In: *Proc. Int. Conf. 3D Vision*. IEEE, 2015.

[SK21] C. Schubert and A. Kolb. "Designing Technology, Developing Theory: Toward a Symmetrical Approach". In: *Science, Technology, & Human Values* 46.3 (2021), pp. 528–554. DOI: 10 . 1177 / 0162243920941581.

[SMM12] M. Sedlmair, M. Meyer, and T. Munzner. "Design study methodology: Reflections from the trenches and the stacks". In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2431–2440.

[Seg+92]   M. Segal, C. Korobkin, R. van Widenfelt, J. Foran, and P. Haeberli. "Fast Shadows and Lighting Effects Using Texture Mapping". In: *Proc. ACM SIGGRAPH Comput. Graph.* 26.2 (1992), pp. 249–252.

[Shi+18]   L. Shi, J. Sun, Y. Yang, T. Ling, M. Wang, Y. Gu, Z. Yang, Y. Hua, and J. Zhang. "Three-Dimensional Visual Patient Based on Electronic Medical Diagnostic Records". In: *IEEE Journal of Biomedical and Health Informatics* 22.1 (2018), pp. 161–172.

[SPK19]   T. Shirley, D. Presnov, and A. Kolb. "A Lightweight Approach to 3D Measurement of Chronic Wounds". In: *Journal of WSCG* 27.1 (2019), pp. 67–74.

[SSC14]   F. Steinbrücker, J. Sturm, and D. Cremers. "Volumetric 3D mapping in real-time on a CPU". In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2014, pp. 2021–2028.

[Ste57]   S. S. Stevens. "On the psychophysical law". In: *Psychological Review* 64.3 (1957), pp. 153–181.

[SG63]   S. S. Stevens and M. Guirao. "Subjective scaling of length and area and the matching of length to loudness and brightness." In: *Journal of Experimental Psychology* 66.2 (1963), p. 177.

[SSS14]   M. Stone, D. A. Szafir, and V. Setlur. "An engineering model for color difference as a function of size". In: *Proc. Color and Imaging Conference*. Vol. 22. 1. 2014, pp. 253–258.

[Str+85]   A. L. Strauss, S. Fagerhaugh, B. Suczek, and C. Wiener. *Social Organization of Medical Work*. ISBN: 9780226777078. University of Chicago Press, 1985.

[STZ14]   T. Sun, P. Thamjaroenporn, and C. Zheng. "Fast Multipole Representation of Diffusion Curves and Points". In: *ACM Trans. Graph.* 33.4 (2014). ISSN: 0730-0301. DOI: 10.1145/2601097.2601187. URL: https://doi.org/10.1145/2601097.2601187.

[Sun+07]   E. Sundvall, M. Nyström, M. Forss, R. Chen, H. Petersson, and H. Åhlfeldt. "Graphical Overview and Navigation of Electronic Health Records in a Prototyping Environment Using Google Earth and openEHR Archetypes". In: *Proc. World Congress on Health Medical Informatics (MEDINFO)*. Ed. by K. A. Kuhn et al. 2007, pp. 1043–1047.

[Sur05]   H Surtola. "The effect of data-relatedness in interactive glyphs". In: *Proc. Int. Conf. Information Visualisation*. IEEE. 2005, pp. 869–876.

[SSG14]   D. A. Szafir, M. Stone, and M. Gleicher. "Adapting Color Difference for Design". In: *Proc. Color and Imaging Conference*. Vol. 22. 1. 2014, pp. 228–233. DOI: 10.2352/CIC.2014.22.1.art00040. URL: https://library.imaging.org/cic/articles/22/1/art00040.

[Tan+13]   P. Tanskanen, K. Kolev, L. Meier, F. Camposeco, O. Saurer, and M. Pollefeys. "Live Metric 3D Reconstruction on Mobile Phones". In: *Proc. IEEE Int. Conf. Computer Vision (ICCV)*. IEEE, 2013, pp. 65–72.

[TW04]   D. H. Titterton and J. L. Weston. *Strapdown inertial navigation technology*. IET, 2004.

[TR05]   N. Trawny and S. I. Roumeliotis. "Indirect Kalman filter for 3D attitude estimation". In: *University of Minnesota, Dept. of Comp. Sci. & Eng., Tech. Rep* 2 (2005), p. 2005.

[VW95]   R. C. Veltkamp and W. Wesselink. "Modeling 3D curves of minimal energy". In: *Computer Graphics Forum*. Vol. 14. 3. Wiley Online Library. 1995, pp. 97–110.

[VKG04]   I. Viola, A. Kanitsar, and M. E. Groller. "Importance-Driven Volume Rendering". In: *Proc. IEEE Int. Conf. Visualization*. 2004, pp. 139–145.

[Wan99]   J. Wang. "Stochastic Modeling for Real-Time Kinematic GPS/-GLONASS Positioning". In: *Navigation* 46.4 (1999), pp. 297–305.

[War08]   M. O. Ward. "Multivariate Data Glyphs: Principles and Practice". In: *Handbook of Data Visualization*. Berlin, Heidelberg: Springer, 2008, pp. 179–198. DOI: 10.1007/978-3-540-33037-0_8.

[War09]   C. Ware. "Quantitative texton sequences for legible bivariate maps". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* 15.6 (2009), pp. 1523–1530.

[Wei+06]   D. Weiskopf, K.-L. Ma, J. J. van Wijk, R. Kosara, and H. Hauser. "Scivis, infovis-bridging the community divide". In: *Proceedings of the IEEE Visualization Conference. IEEE*. 2006.

[WBH14]   V. L. West, D. Borland, and W. E. Hammond. "Innovative Information Visualization of electronic health record data: a systematic review". In: *Journal of the American Medical Informatics Association* 22.2 (2014), pp. 330–339.

[Wic+12]   H. Wickham, H. Hofmann, C. Wickham, and D. Cook. "Glyph-maps for visually exploring temporal patterns in climate data and models". In: *Environmetrics* 23.5 (2012), pp. 382–393.

[Wij+08]   M. Wijffelaars, R. Vliegen, J. J. Van Wijk, and E.-J. Van Der Linden. "Generating Color Palettes using Intuitive Parameters". In: *Computer Graphics Forum* 27.3 (2008), pp. 743–750. DOI: https://doi.org/10.1111/j.1467-8659.2008.01203.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2008.01203.x.

[Wij05]   J. van Wijk. "The value of visualization". In: *VIS 05. IEEE Visualization, 2005*. 2005, pp. 79–86. DOI: 10.1109/VISUAL.2005.1532781.

[WWH98]   F. Wilkinson, H. R. Wilson, and C. Habak. "Detection and recognition of radial frequency patterns". In: *Vision Research* 38.22 (1998), pp. 3555–3568. ISSN: 0042-6989. DOI: https://doi.org/10.1016/S0042-6989(98)00039-X. URL: https://www.sciencedirect.com/science/article/pii/S004269899800039X.

[Xie+14]   G. Xie, X. Sun, X. Tong, and D. Nowrouzezahrai. "Hierarchical diffusion curves for accurate automatic image vectorization". In: *ACM Trans. Graphics* 33.6 (2014), pp. 1–11.

[Yin+21]   L. Ying, T. Tang, Y. Luo, L. Shen, X. Xie, L. Yu, and Y. Wu. "GlyphCreator: Towards Example-based Automatic Generation of Circular Glyphs". In: *IEEE Trans. Visualization and Computer Graphics (TVCG)* (2021), pp. 400–410. DOI: 10.1109/TVCG.2021.3114877.

[ZMG21]   G. M. H. Zahan, D. Mondal, and C. Gutwin. "Contour Line Stylization to Visualize Multivariate Information". In: *Graphics Interface 2021*. 2021.

[Zan+16]   P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo. *Time-of-Flight and Structured Light Depth Cameras: Technology and Applications*. Springer, 2016.

[Zha00]   Z. Zhang. "A flexible new technique for camera calibration". In: *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*. 2000, pp. 133–1334.

[ZK10]   C. Ziemkiewicz and R. Kosara. "Embedding Information Visualization Within Visual Representation". In: *Advances in Information and Intelligent Systems* (May 2010), pp. 307–326.

[Zwi+01]   M. Zwicker, H. Pfister, J. Van Baar, and M. Gross. "Surface splatting". In: *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. 2001, pp. 371–378.