Darmstadt University of Technology

Department of Computer Science
Telecooperation
Prof. Dr. Max Mühlhäuser

Diploma Thesis

MICON: A Music Stand for Interactive Conducting

Advisors: **Prof. Dr. Max Mühlhäuser**
Prof. Dr. Jan Borchers

Aristotelis Hadjakos

May 2006

Contents

Ehrenwörtliche Erklärung

Hiermit versichere ich, die vorliegende Diplomarbeit ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus den Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 13. Juli 2006

Abstract

The MICON is an electronic music stand extending Maestro, the latest in a series of interactive conducting exhibits that use real orchestral audio and video recordings. The MICON uses OpenGL-based rendering to display and animate score pages with a high degree of realism. It offers three different score display formats to match the user's level of expertise. An animated visual cueing system helps users with their conducting. The MICON has been evaluated with music students and has been presented at various events, including an one-day exhibition at the Hessischer Rundfunk, a radio and television broadcaster.

Zusammenfassung

Der MICON ist ein elektronischer Notenständer für Maestro, dem neuesten einer Reihe von Dirigierexponaten basierend auf echten Video- und Audioaufnahmen. Er nutzt OpenGL für die realistische Darstellung und Animation der Partitur und bietet verschiedene Darstellungsformen des musikalischen Materials für Benutzer mit verschiedener musikalischer Vorbildung an. Der Benutzer wird durch animierte visuelle Zeichen beim Dirigieren unterstützt. Der MICON wurde mit Musikstudenten getestet und bei verschiedenen Anlässen gezeigt, darunter auch eine Präsentation beim Hessischen Rundfunk anlässlich eines Tages der offenen Tür.

Acknowledgements

It has been a big pleasure to create this work and to see people enjoy the exhibit—I have to thank: I thank Prof. Dr. Max Mühlhäuser for making it possible and for his encouragement. I thank Prof. Dr. Jan Borchers for his support and his advice. Thank you Henning Kiel and Eric Lee for your friendliness, the interesting discussions and for your help getting MICON and Maestro to talk to each other. Thank you Gerhard Austaller for your friendly help and for the beautiful installation of the system in the media room. Special thanks go to my parents Birgit and Kyriakos not only for their support during my studies and to my brother Stephanos. Special thanks go to my wife Won-Kyoung.

Chapter 1

Introduction

Some people enjoy to conduct while hearing music on their CD-players. They are involved in the musical process but have no active influence.

This is where Personal Orchestra comes in.

With Personal Orchestra, the user can control some aspects of a real recording: The user can control tempo, volume and can put emphasis on an instrument group. Within boundaries, given by the recording and the limits of the system, the user can change the expression of the music and become a more active part of the performance. Personal Orchestra is an interactive exhibit. Different versions are currently being presented at the HOUSE OF MUSIC in Vienna, at the CHILDREN'S MUSEUM in Boston and at the BETTY BRINN CHILDREN'S MUSEUM in Milwaukee.

This thesis deals with the design of an electronic music stand for Personal Orchestra: the MICON (Music stand for Interactive CONducting). A paper about the MICON will be published in NIME 2006 (Borchers, Hadjakos and Mühlhäuser [2006]).

1.1 Motivation

The music stand for Personal Orchestra was created in order to enrich the experience with the notational aspect of music. Furthermore, it should give the user visual feedback about his actions and support the user while conducting. A real conductor, normally, has the score before him while conducting the orchestra. So a music stand for Personal Orchestra, being part of the interaction, would create a more realistic atmosphere.

1.2 Overview

This introduction concludes with the presentation of the proposed system. Chapter 2 presents already existing electronic music stands and gives an overview of recent Personal Orchestra systems and other interactive conducting systems. The requirements for the design of MICON are given in chapter 3. The design considerations and the resulting decisions can be found in chapter 4. Chapter 5 deals with the content creation process for MICON. Chapter 6 gives an overview of the inner workings of the MICON. Chapter 7 evaluates the MICON. Screenshots of the MICON can be found in appendix B.

1.3 Proposed System

The proposed system (figure 1.1) consists of two parts: Maestro and MICON. Maestro analyzes the conducting and modifies the tempo of the music with a time-stretching algorithm. The video of the playing orchestra is also time-modified to fit the tempo. The MICON renders and animates the musical score. Maestro informs MICON over a communication medium about the current position, and the MICON renders the appropriate section of the score.

Maestro is a modification of the already existing Personal Orchestra system. Maestro was developed by the Media Computing Group at RWTH Aachen.

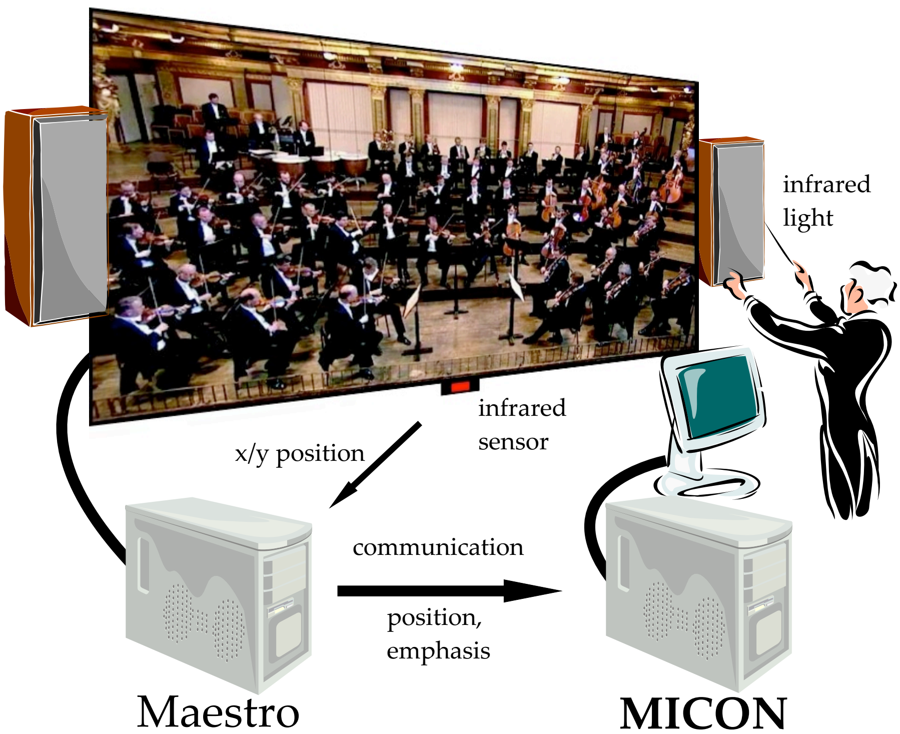


Figure 1.1: The proposed system

Chapter 2

Related Work

2.1 Electronic Music Stands

2.1.1 MOODS

MOODS (Music Object Oriented Distributed System) is a synchronous real-time cooperative editor for music scores, i.e. each performed change is immediately made visible to all users. MOODS was created by Bellini et al. [2002]. It is intended to be used by the musicians of an orchestra during rehearsal.

MOODS consists of three different types of lecterns: DLIOO (distributed lectern interactive object-oriented), MASE (main score editor) and MASAE (main score auxiliary editor). DLIOOs are single part lecterns for the instrumentalists and are used for editing and visualizing. The MASE is used by the conductor in order to view and modify the main score. MASAE is a general manager and music editor to be used by the archivist. The archivist can make modifications on the main score during rehearsal and score revision. The system allows not only multiple DLIOOs but also multiple MASEs and MASAEs.

Editing in MOODS is based on different permissions for the archivist, the conductor, the concertmaster and the instrumentalist. MOODS supports semi-automatic page-turning. The rate of music execution is initially set on the basis of a metronomic indication in terms of beats per minutes. This trend can be adjusted by the operator of the MASAE. The score on the DLIOO is separated horizontally. Below the separator is the score that is currently being played. Above the separator the next page is being shown. As the music advances, the separator moves downwards. The score on the MASE and the MASAE is separated vertically. The right half shows the current score, and the left half shows the beginning of the next page.

MOODS assumes a constant tempo for a piece; tempo variations have to be adjusted manually by a human operator during the performance making it unsuitable for our purposes.

2.1.2 muse

muse is a digital music stand for musicians in a symphony orchestra (Graefe et al. [1996]). It consists mainly of a portable display and a matching stand. It has an integrated metronome with auditive and visual feedback. It has a pitch-generating tuner. It allows on-screen annotation and has intrasymphonic communication capabilities. The pages are turned either automatically or manually.

muse is powered by battery. There is wireless communication between the music librarian's computer and the muses. This communication is encrypted to prevent copyright violation. Every muse has an attached microphone. The microphone is needed for the tuner and for automatic page turning; the sound of the instruments is compared against the score, and the pages are turned in the appropriate moment.

muse was presented to the musicians of the Pittsburgh Symphony Orchestra.

muse it does not highlight the current score position, and its page-turning requires per-instrument microphones making it unsuitable for our purposes.

2.1.3 eStand

The company eStand Inc. (`http://www.estandmusic.com`) produces electronic music stands (figure 2.1). The eStand is a hybrid hardware and software product. The hardware consists of a notebook with a 14" display and a footswitch. The display has a rotating hinge for tablet functionality. A special pen is provided with the notebook that lets the user draw directly on the screen. The hardware is very silent. The software displays the score,

Figure 2.1: The eStand

which has to be downloaded from the Internet or created with note-setting software. The musician can turn pages with the footswitch. The musician can annotate the score with the provided pen. The eStand has a built-in metronome and tuning. It also includes music library management. For ensembles and orchestras, eStand offers networking capabilities: Scores can be sent to each other, and annotations and comments can be shared.

eStand does not highlight the current position in the score and does not advance the score automatically making it unsuitable for our purposes.

2.2 Interactive Conducting Systems

2.2.1 The Radio-Baton

The Radio-Baton (Boie et al. [1989], Boulanger and Mathews [1997]) is a controller for live computer performances. It is a MIDI instrument designed to work with MIDI synthesizers and MIDI-based sequencing and programming software.

The Radio-Baton system can determine the 3D position of the batons, which are equipped with small radio antennas. A sensor board is equipped with five receiving antennas. The closer the baton comes to an antenna, the higher is the intensity of the signal. This effect is used to determine the positions of the batons.

The Radio-Baton can send trigger signals when the baton hits an imaginary plane. Besides sending trigger and position signals to a listening computer, the Radio-Baton processor can also run other programs. In Conductor Mode, the Radio-Baton plays a previously loaded MIDI score according to the performer's conducting gestures by sending MIDI messages to an attached synthesizer.

2.2.2 The MIDI Baton

Keane and Gross [1989] created the MIDI baton, a sensor that can be used to control an interactive computer music system with traditional conducting gestures. The MIDI baton can sense acceleration that occurs when the conductor places a beat.

The sensor consists of a metallic ball that is being held by a steel spring wire inside a brass tube (figure 2.2). The acceleration of the baton presses the ball against the brass tube and closes an electronic contact generating a signal. The MIDI baton was used to control synthesized music. While leaving the basic principle unchanged, the system was improved, and the successors MIDI baton II (Keane et al. [1990]) and MIDI baton III (Keane et al. [1990]) were created.

2.2.3 Computer Music System which Follows a Human Conductor

Morita et al. [1989] developed a system which follows a human conductor. The system uses computer vision to analyze the conducting. Special, feature extraction hardware searches for a white area in images that are

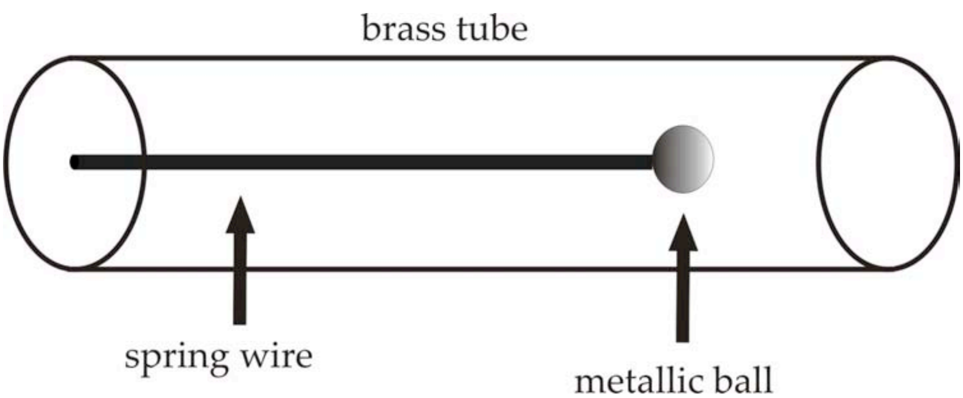


Figure 2.2: The MIDI baton

provided by a CCD camera. The tip of the conductor's baton has a white marking attached to it, or the conductor alternatively wears white gloves and conducts without baton. The gesture is a analyzed, and tempo and volume are derived.

2.2.4 Light Baton

Bertini and Carosi [1992] created the light baton. The system can derive beat time, beat order and amplitude from traditional conducting gestures. The system also recognizes end of movement and absence of the baton. The data is used to control a computer-generated musical performance. The baton they use has a light source attached on the tip. A CCD camera provides the computer with images of the conducting area, which are examined to find the position of the baton and to analyze the gesture.

2.2.5 Conductor Follower

The Conductor Follower (Lee et al. [1992], Brecht and Garnett [1995]) uses a Mattel Power Glove and a Buchla Lightning baton for input and creates synthesized MIDI music. The emphasis of the project lies in beat time prediction.

2.2.6 The Digital Baton

The Digital Baton was created by Marrin and Paradiso [1997]. It consists of an infrared LED on the baton's tip, piezo-resistive strips for finger and palm pressure and three orthogonally placed accelerometers. The position of the baton is determined by an infrared sensor. The sensor produces horizontal and vertical positioning information; the intensity of the infrared light is a rough measure for distance.

2.2.7 The Conductor's Jacket

Marrin and Picard [1998] created the Conductor's Jacket, an array of sensors that are built into normal clothing. The sensors are used to collect and analyze data from conductors. The Conductor's Jacket detects physiological change and physical motion of the wearer without encumbering his motions. The Conductor's Jacket has electromyography sensors on each forearm bicep and tricep and sensors for respiration, temperature, skin conductance and heart rate (figure 2.3). Additional 3D position and orientation sensors are attached to the jacket.

Nakra [2000] used a modified jacket for controlling a MIDI synthesized orchestra. The system recognized traditional conducting parameters like tempo, holds, pauses, volume and emphasis and, additionally, let the user control instrument panning, pitch transposition and more.

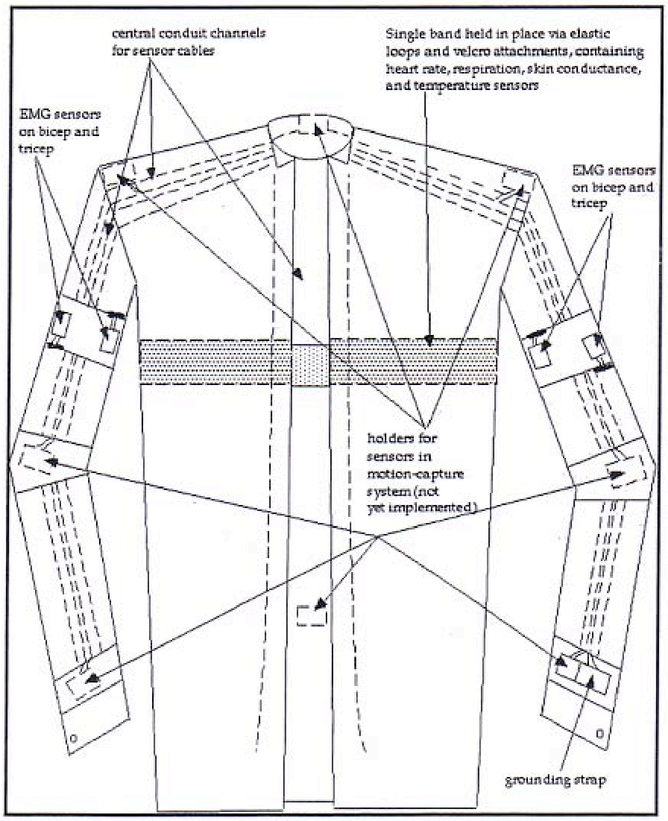


Figure 2.3: The Conductor's Jacket

2.2.8 Multi-Modal Conducting Simulator

Usa and Mochida [1998] created the Multi-Modal Conducting Simulator, which uses input from different sources to control a synthetic MIDI orchestra. The system recognizes beginning and end of a piece, cue for a player with the eyes, beat timing and beat number in the measure, some aspects of articulation, and breathing. Acceleration, breathing and eye-tracking sensors provide the system with user input.

2.2.9 Virtual Conducting Practice Environment

The Virtual Conducting Practice Environment by Garnett et al. [1999] is a system to support students of conducting to learn and practice conducting with the computer. The system offers auditive and visual feedback about the student's performance. Conducted tempo is represented visually with sliders, time versus tempo graphs and numerical values. The primary visual representation, however, is the Beat Window (figure 2.4), which shows the positions of the recognized beats. The student can improve his beat technique with this information. Auditory feedback about the conducting gesture is given by different qualities of beeps, which are played when a conducted beat is recognized. Loud beeps and soft beeps represent forte and piano conducting. Short and long beeps represent staccato and legato conducting. The student can activate a metronome and try to overlay the conducted and the metronomic beep.

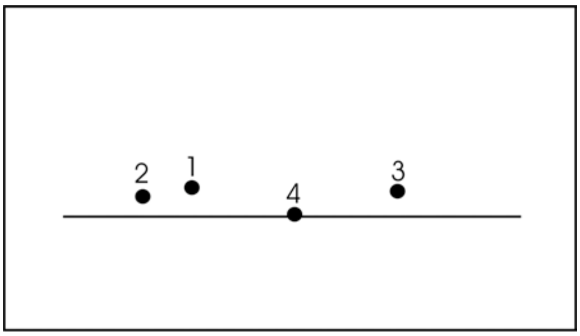


Figure 2.4: The beat window

2.2.10 Interactive Virtual Ensemble

Garnett et al. [2001] use a wireless MotionStar tracker by Ascension Technologies to recognize standard conducting gestures. This sensor gives 3D position and orientation, which is in particular important to recognize the orientation of the conductor's left hand to convey dynamics and for giving cues. The Virtual Ensemble creates MIDI output.

2.2.11 DIVA's Virtual Orchestra

The Virtual Orchestra of DIVA, the Digital Interactive Virtual Acoustics research group, is a band of four virtual musicians. The musicians are animated 3D characters (figure 2.5) and move their hands to play their instruments. Two persons are involved in the interaction: the conductor and the listener. The position of the listener is determined, and the sound at that space is calculated and sent to the listener's headphones (Ilmonen [2003]). The system uses magnetic trackers, which consist of a magnetic pulse generator and several small sensors, for motion tracking (Ilmonen [1998, 1999], Ilmonen and Takala [1999]). A sensor behind the conductor's neck serves as reference point. The system uses neural networks for beat recognition. The conductor can control tempo and volume of the synthesized MIDI band.

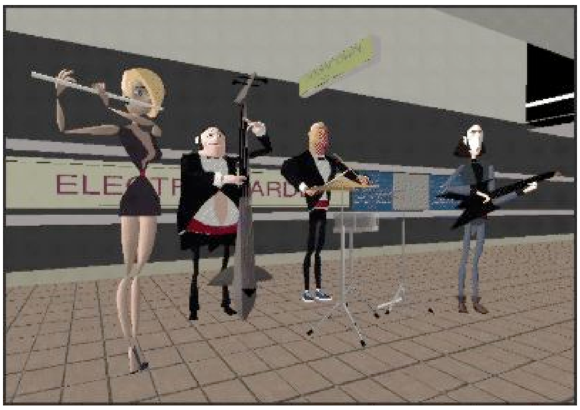


Figure 2.5: The Virtual Orchestra

2.2.12 Conducting Gesture Recognition, Analysis and Performance System

The Conducting Gesture Recognition, Analysis and Performance System by Kolesnik [2004] uses the input of two cameras giving a front and side view of the conductor, who wears a colored glove for automatic feature extraction. The system analyzes the movement of the right hand for beat placement and the left, for expressive information. The system features a real audio/video recording.

2.2.13 Personal Orchestra and WorldBeat

WorldBeat

Borchers [1997] created the WordBeat system. The system features different modules that lets even persons unexperienced in music and computers generate pleasing music. One of the modules lets the user conduct a synthesizer allowing to control tempo and volume. The system uses a modified version of the gesture recognition algorithm developed by Lee et al. [1992]. The system evaluates the vertical movement of a Buchla Lightning electronic baton to determine beat placement and evaluates the size of the gesture to determine volume.

Virtual Conductor

Samminger [2002], Borchers et al. [2002, 2004] created the first version of Personal Orchestra, which is being exhibited under the name "Virtual Conductor" at the HOUSE OF MUSIC in Vienna (figure 2.6). This is the first multimedia system to conduct a realistic electronic orchestra. The user can conduct an audio and video recording of the famous Vienna Philharmonic.

The input device for Virtual Conductor is an electronic baton. An infrared light shines inside the translucent baton. The baton's position light is detected by a sensor, and position data is sent to the computer where it is analyzed. Virtual Conductor interprets the user input as a simple single-hand up-down gesture. Volume is controlled by the size of the user's conducting gesture. A beat is marked by the change of the baton's motion from downwards to upwards. By placing the beats densely or sparsely the user

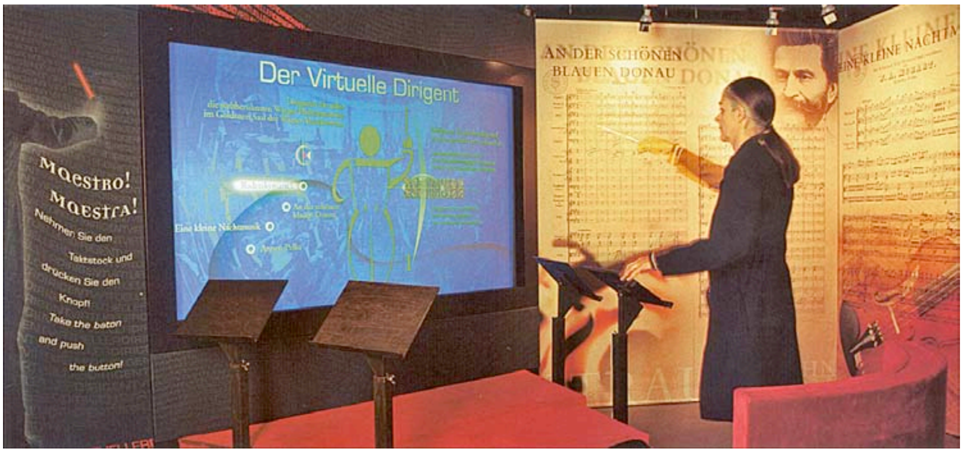


Figure 2.6: Personal Orchestra 1 at the HOUSE OF MUSIC in Vienna

can control the tempo of the music. The user can also put an emphasis on an instrument group. In order to do so, the user has to conduct toward the desired instrument group. The Virtual Conductor finds the center of the gesture and emphasizes the selected instrument group by increasing the volume of the corresponding audio-track. The recording with the Vienna Philharmonic was therefore done with additional microphones for the different instrument sections. Time-stretching depends on pre-computed tracks. The system shifts between these tracks to achieve tempo changes. The Virtual Conductor was designed aiming for best possible audio quality, as it would feature a recording with the Vienna Philharmonic.

You're the Conductor

The next version of Personal Orchestra, which is being presented under the name *You're the Conductor* at the CHILDREN'S MUSEUM in Boston (figure 2.7), was designed specifically for children. The exhibit features "Stars and Stripes Forever" played by the Boston Pops/Boston Symphony Orchestra. It was created by Lee et al. [2004]. In order to create a system that can be successfully used by small children, the conducting gesture was based on the velocity of the baton rather than beat placement: The velocity of the baton is directly mapped to the tempo of the music. The size of the gesture, which is determined by the points where the direction of the baton's motion changes, is mapped to volume. A physically more robust baton than the baton in use for the Virtual Conductor was developed and constructed for this exhibit.

Figure 2.7: Personal Orchestra 2 in the CHILDREN'S MUSEUM in Boston

You're the Conductor uses a new time-stretching algorithm. This algorithm, an improved phase vocoder algorithm, works in real-time and does not depend on pre-stretched tracks. With the new time-stretching algorithm *You're the Conductor* can follow the user more accurately, as it is not limited to the tempi of the pre-stretched tracks.

Maestro

The next version of Personal Orchestra, called Maestro (figure 2.8), was developed by Lee et al. [2006c]. It supports three different conducting gestures: the wiggle gesture from *You're the Conductor*, the up-down gesture from the Virtual Conductor and a professional conducting gesture. In order to be usable by a broad audience, Maestro analyzes the user's conducting and selects the appropriate conducting model. A framework for conducting gesture analysis (Gruell [2005], Lee et al. [2006a]) is used to recognize beat times and to derive volume and instrument emphasis. Maestro uses the PhaVoRit algorithm (Karrer [2005]) for time-stretching. The design of Maestro is based on a framework, which uses the concept of se-

Figure 2.8: Personal Orchestra 3 in the BETTY BRINN CHILDREN'S MUSEUM

mantic time versus real time for a more elegant way of handling audio and video streams with variable time-base (Lee et al. [2006b]). The MICON was designed for this system.

Chapter 3

Requirements

3.1 Design Goals

The functional goals of the MICON were:

- to display the musical score to Maestro visitors while they are conducting and
- to indicate the current position in the score and to automatically advance the pages following the music to help visitors with their conducting.

The key constraints in creating the MICON were:

Production quality. Since the Maestro system features professional orchestras, its look and feel has to fulfill high standards to be accepted by the museum and orchestra in question. The MICON has to look and behave as professional as the rest of the exhibit, which requires excellent visual quality of the score display and fluid, non-distracting interaction with it.

Visitor profile. Typical visitors using the system would be one-time users with a short dwelling time. The system has to provide for this through a particularly simple, self-explanatory and obvious interface that requires little or no interaction apart from the conducting itself.

Musical knowledge. Some visitors might be amateur or even professional conductors, but most would have no prior experience in conducting. The MICON has to provide these beginners with alternatives to the complexity of a full orchestral score document. Hopefully, by interacting with the system visitors would learn a little more about conducting and experience some of its challenges and rewards.

Design philosophy. Being a part of the interactive conducting exhibit Personal Orchestra, the MICON has to fit with the design philosophy of Personal Orchestra.

Listener architecture. The MICON has to extend the Maestro system. To minimize dependencies between these two co-evolving projects, their communication interface has to be kept as narrow as possible, with the MICON essentially listening to the timing information that was already being generated by the Maestro gesture recognition engine.

Screen geometry. The MICON has to be adaptable to different screen sizes and geometries so that it can be run on a small display for a presentation as well as on a big display in a full installation of the exhibit.

3.2 Design Philosophy

3.2.1 Design Patterns

Borchers [2001] presents a set of patterns for interactive exhibits. A selection of these patterns, which apply to the Personal Orchestra system, are described here.

ATTRACT-ENGAGE-DELIVER*: An interactive exhibit is used by visitors for their entertainment. They have no actual need to use the interactive exhibit. Usually, an interactive exhibit has a message for the user or wants to teach something. The duration of the interaction has to be limited so that the exhibit can be used by many persons over the day. When the message has been delivered, the interaction should be ended. Otherwise, users will stop the interaction when they get bored or because they cannot figure out how to properly end the interaction.

The interaction should take place in three phases: First, the users have to be attracted. Then within the interaction, the message of the exhibit has to be delivered. After the message has been delivered, the interaction should be ended in a positive way so that the length of the interaction corresponds to the desired throughput of the exhibit. References: ATTRACTION SPACE.

ATTRACTION SPACE*: In a museum many exhibits compete for the visitor's attention. In order to deliver a message to the user, an exhibit has to attract the user first. An exhibit should not disturb other near-by standing exhibits. When the exhibit is not been used, the first thing a passing-by visitor sees is its appearance in the idle-state. The exhibit in idle-state can attract the user with visual and auditory clues. Visual signals are only perceived when the user looks into the right direction. Auditory signals are omni-directional. They are also more intrusive and stronger. Strong visual clues (big displays, animation) can be perceived in the visual periphery of the user.

An attraction space has to be defined around the exhibit. This space should be as large as possible, yet it should not penetrate the other exhibits' attraction spaces. An exhibit should not frequently violate the border of its attraction space. To achieve this, static stimuli can be used: The appearance of the exhibit should be attractive by its own. Excessive animation and the frequent use of undirected stimuli should be avoided as they violate the other exhibits' attraction spaces. References: INNOVATIVE APPEARANCE, SIMPLE IMPRESSION, DOMAIN-APPROPRIATE-DEVICES.

COOPERATIVE EXPERIENCE**: Often people visit exhibitions in groups. However, most interactive exhibits can only be used by a single person a time. When used by more then one person, the interaction of the humans with each other becomes the main part of the experience. There are, of course, many situations where the control has to stay with a single person. Apart from cooperative use, an interactive exhibit should allow bystanders to watch what is going on. This allows to increase the number of persons that can at least passively experience the exhibit. While watching, a bystander learns about the system and can gain form this knowledge when using the exhibit himself. If a cooperative experience is not planned, the visitors will create it themselves.

Two people should be able to use an interactive exhibit at the same time. At least five bystanders should be able to watch or listen while another person uses the exhibit. References: EASY HANDOVER, IMMERSIVE DISPLAY.

EASY HANDOVER*: Most interactive exhibits assume that each user begins the interaction form the very beginning. However, it often happens that a visitor leaves the controls to another person during the interaction. The new user takes over the control in the middle of the interaction and does possibly not know the interaction history. When the new user takes over, he has already gained some knowledge by watching his predecessor. This can compensate for missed explanations.

Knowledge of the interaction history that a new user needs to perform successfully should be as little as possible. A simple means to return to the initial state should be given. Critical parameters, like language settings, should be accessible at every step of the interaction. References: LANGUAGE INDEPENDENCE.

LANGUAGE INDEPENDENCE: Museums and exhibition centers often have international visitors, who might not speak the local tongue. Most interactive exhibits show text messages to the user. It does not suffice to provide a language selection at the beginning of the interaction. The language selection has to be accessible at every step of the interaction because a new user might take over the control in the middle of a previous interaction.

At every step of interaction the language selection should be accessible. This can be done in a multilingual or pictorial way. References: none.

SIMPLE IMPRESSION*: Interactive exhibits often have complex messages and present new technology to the visitors of an exhibition. Typically the visitors are first-time and one-time users. Therefore, the users have to learn to use the system very quickly.

The user interface of an interactive exhibit should be simple. This simplicity should be represented in the overall appearance of the exhibit. References: INVISIBLE HARDWARE, INCREMENTAL REVEALING.

INCREMENTAL REVEALING**: Systems that look complex do not seem to be inviting. If an exhibit does not have enough interaction possibilities, it can quickly become boring.

Initially the exhibit should present a simple overview of the exhibits functions. When the user becomes active showing interest in a specific part, more information should be revealed about it, and the user is given new choices. References: FLAT AND NARROW TREE.

FLAT AND NARROW TREE*: Many interactive exhibits contain a set of different pages. The user navigates through these pages. Large hierarchies and unordered networks can confuse the user.

A tree-like hierarchy should be used for navigation. There should be no more than seven items to choose from at every step. The tree should be no more that five levels deep. References: none.

DOMAIN-APPROPRIATE DEVICES*: Although interactive exhibits deal with a wide range of different domains, they use mostly standard computer input and output devices. Most interactive exhibits use metaphors from the application domain, yet they use standard keyboard and mouse for input and a monitor for output. The metaphors remain virtual. Of course, building innovative input and output hardware is more expensive, but often it is not even attempted to estimate the extra cost and its payoff. Psychological research suggests that users are more successful when using I/O-devices from the application domain.

Input and output devices that resemble real-world objects should be preferred over standard periphery. Whenever the user receives information or has to input something, it should be considered if standard hardware should really be used, or if special, domain-appropriate, hardware can serve as periphery for the exhibit. References: INNOVATIVE APPEARANCE, INVISIBLE HARDWARE, ONE INPUT DEVICE.

INNOVATIVE APPEARANCE*: Standard hardware is usually not considered to be exciting.

Input and output hardware that looks differently form standard hardware should be preferred, as it might rage the curiosity of visitors. The usage of innovative I/O-devices offers new ways of interaction to the users. References: INVISIBLE HARDWARE, ONE INPUT DEVICE.

IMMERSIVE DISPLAY*: Exhibits are usually visited by groups of people. When people interact with the exhibit, they want to immerse themselves into the world of the exhibit. Large displays enable bystanders to watch what is going on, while another person uses the system.

A single exhibit with a large screen (minimum 1.5 m) should be preferred over several similar exhibits with smaller displays. The user should not be shielded from the other visitors: Head-mounted displays do not let bystanders take place in the interaction. The viewing distance should be roughly equal to the width of the screen. References: INVISIBLE HARDWARE.

INVISIBLE HARDWARE*: Modern interactive exhibits often use advanced and expensive technology. If the technology is visible, users might be driven away by it. The display of the technology can overshadow the real contents of the exhibit.

The computer hardware used in an exhibit should be hidden. Ideally, only relevant devices, devices the user interacts with, should be visible. These devices should create an image, which is adequate to the application domain. References: ONE INPUT DEVICE.

ONE INPUT DEVICE*: Many interactive exhibits use a variety of input devices. Some exhibits recognize gestures without any physical input device. But exhibits are used by people without prior knowledge about it. Many input devices might cause them to feel confused. No input device at all might make them feel uneasy.

All input should be mapped to one input device. References: DOMAIN-APPROPRIATE DEVICES.

3.2.2 Design Patterns in Personal Orchestra

The interaction with Personal Orchestra is divided into three phases (ATTRACT-ENGAGE-DELIVER): The user's attention is drawn to the exhibit by auditive and visual clues. The interaction takes place, and the musical message is delivered. When the user finishes the piece, applause marks the end of the interaction. The user can leave the exhibit or, alternatively, start over from the beginning. The pieces are roughly three minutes long. Depending on the tempo the user conducts, the interaction takes between 1.5 and 6 minutes.

The Personal Orchestra exhibit has an attractive appearance (figure 3.1). This attracts visitors without violating the borders the ATTRACTION SPACES of other exhibits. When Personal Orchestra is in idle-state, the user's attention is drawn by auditive and visual signals: Periodically, ca. every 30 seconds, the orchestra can be seen and heard tuning its instruments for a while. The sound is soft and unobtrusive enough to not disturb the atmosphere of nearby standing exhibits.

Personal Orchestra cannot be used by more than one person at a time. In spite of this, Personal Orchestra is designed for a COOPERATIVE EXPERIENCE: Bystanders can hear the music from the speakers and see the orchestra on a big screen, while someone else interacts with the system.

When a user gives the control, the baton, to a new user in the middle of the interaction, the new user does not need to know much about the interaction history (EASY HANDOVER): The interaction history determines only the music piece that is being conducted and the language setting. However, the language setting is solely used to provide the proper piece-selection screen. The new user might not have read the conducting advice that are presented

Figure 3.1: Personal Orchestra 1 exhibit

in the beginning. He will, however, have watched the predecessor and have learned how to interact from this source. If the new user does not conduct, the system returns to the idle-state after a while.

In order to ensure LANGUAGE INDEPENDENCE, the user chooses his language. This is done in the very first step of the interaction. A multilingual language selection is presented at the beginning of the interaction.

Personal Orchestra gives a SIMPLE IMPRESSION. The entities the user deals with are the baton, the music from the loudspeakers and the video of the orchestra on the screen.

The user selects the preferred language. Then he selects a music piece from a list of four items, and the conducting begins. The selection has the structure of a FLAT AND NARROW TREE.

The input device for Personal Orchestra is an electronic baton, resembling a real conductor's baton. The electronic baton is a DOMAIN APPROPRIATE DEVICE. The baton also provides Personal Orchestra with an INNOVATIVE APPEARANCE.

The exhibit uses a large screen for visual output. This allows the user to get immersed into the situation of standing before a large orchestra (IMMERSIVE DISPLAY).

In the Personal Orchestra exhibit much of the computer hardware is hidden. INVISIBLE HARDWARE lets the real contents of the exhibit—music shine. Users can be driven away if complicating looking and advanced hardware is shown.

In Personal Orchestra every interaction is mapped to ONE INPUT DEVICE: the electronic baton. The baton is not only used for conducting but also for selecting language and piece in the beginning of the interaction by pointing toward the desired item and pushing a button on the baton.

3.2.3 Design Principles

Norman [2002] provides design principles to aid designers to create good products.

Human action has to phases: execution and evaluation. After an action is executed, the results are evaluated and compared to the desired outcome: the goal. Action execution can be divided into smaller steps. The goal, a very high-level description (e.g. get more light), has to be transformed into a specific intention (put on the lamp or, alternatively, use a candle). The intention is then transformed into a series of planned actions, which are then performed and change the state of the world. Evaluation can also be divided into smaller steps. The physical world is perceived, and a mental model of the world state is built: the interpretation. The interpretation is evaluated, and the goals are modified (or may remain unchanged if the action was unsuccessful). The seven stages of human action are: perception, interpretation, evaluation, goals, intention, action sequence, execution. In many tasks problems arise when mapping mental intention to physical states and physical states to interpretations.

Two gulfs separate mental sates from physical states: The gulf of execution is a measure of the difficulty for the user to map his intention to possible actions on a device. The gulf of evaluation is a measure of the difficulty to interpret and evaluate the physical state of the device. Design should support the user at every stage of action: The design should make clear what can be achieved with the device (goal level) and show the possible actions (intention level). It should support the user to determine the mapping between intention and physical plan (action sequence level), and the device should be physically operated with ease (execution level). The design should enable the user to determine the state of the device (perception level), help the user to interpret the perception (interpretation level) and support the user to tell if the device is in the desired state (evaluation level). These requirements can be accomplished by using principles of good design: good conceptual models, good mappings visibility and feedback.

A good conceptual model allows the user to predict the effects of his actions. A conceptual model is a representation of a device in the mind of a person. It is a representation of the relationships of actions and outcomes. A wrong or inadequate conceptual model can lead to difficulties in using the device. Therefore, the designer is obliged to provide the user with a good conceptual model. First the designer has to create a good conceptualization of the device: an easy to learn conceptualization that describes the device sufficiently. This model, the design model, has to be communicated to the user. A very important communication medium is the device itself. Manuals are often ignored by users. The user develops his mental model, the user model, mainly by interacting with the device. Therefore, the system image of a device is critical: It has to lead the user to a good understanding of the device.

A mapping is the relationship between actions on the device and effects in the world. A natural mapping takes advantage of analogies and cultural conventions. In a spatial analogy, e.g., the layout of the control elements could resemble the physical or functional structure of the device. Other natural mappings take advantage of principles of human perception. A natural mapping leads the user immediately to understanding the device. Natural mappings can easily be learned by users. Arbitrary mappings, however, are difficult to learn and to use.

The relevant parts of a device should be visible. The user should be able to easily distinguish important parts from other parts like decorative items. Aesthetic considerations should not affect this visibility principle: An important part should not be left out for a better appearance of the device. Feedback gives the user information about the effects of his actions. It gives every action an immediate effect. Especially visible feedback, in form of a well-designed display, can improve the usability of a device significantly. Without feedback the user will wonder whether his actions have been executed. The user might then repeat his actions without need, possibly harming the result. The user might conclude that the device does not react and restart his interaction unnecessarily.

When a person uses a new device, a big number of possible actions might cause problems. The number of possible actions can be reduced by constraints. There are different types of constraints: physical, semantic, cultural and logical constraints. Physical constraints depend on properties of the physical world. A big peg, e.g., does not fit to a small hole. Physical constraints are recognized by users without special training. If a physical constraint can easily been seen, the number of possible actions is immediately reduced: The user will not make obviously impossible actions. Semantical constraints rely on the meaning of the situation and on the user's knowledge. Cultural constraints rely upon cultural conventions. Logical constraints rely on logics and analogy.

The affordance of an object gives the user a strong hint what actions can be done: A button affords ("is for") pushing; a knob affords turning. Even materials have affordances: Glass, e.g., affords looking through it (and also breaking it). With affordances, the user can know what to do without the need of labels, instructions or pictures. Simple things should be self-explanatory.

A task should be simple to perform by the user. If a task is unnecessary complex, the task should be changed. Technology can aid to simplify the task. In order to simplify a task, a designer has several alternatives: The task can be left unchanged, and the designer provides the user with mental aids to remember what actions to do. Alternatively, the designer could use technology to make a hidden part visible. This improves the feedback from the device and helps the user to gain better control on the device: A diagram on a computer screen could, e.g., help the user. To simplify a task some of it could be automated, or the task could be changed completely with the aid of technology.

Chapter 4

Design of the MICON

4.1 Functional Design

The MICON features four different musical representations:

- piano roll,
- pulse notation,
- full orchestral score and
- piano extract score.

The piano roll and the score representations provide visualizations of the musical material, enhancing the Maestro experience with the notational aspect of music. The pulse notation was designed to support the user with his conducting: It visualizes the, previously inaccessible, metrical structure of the music.

4.1.1 Piano Roll

The piano roll represents notes as boxes of uniform height. Box width defines the length of a note, the vertical position defines pitch, and the horizontal position defines onset time. Music software such as MIDI sequencers often uses piano roll (figure 4.1) so that some users of the MICON may already be familiar with this representation.

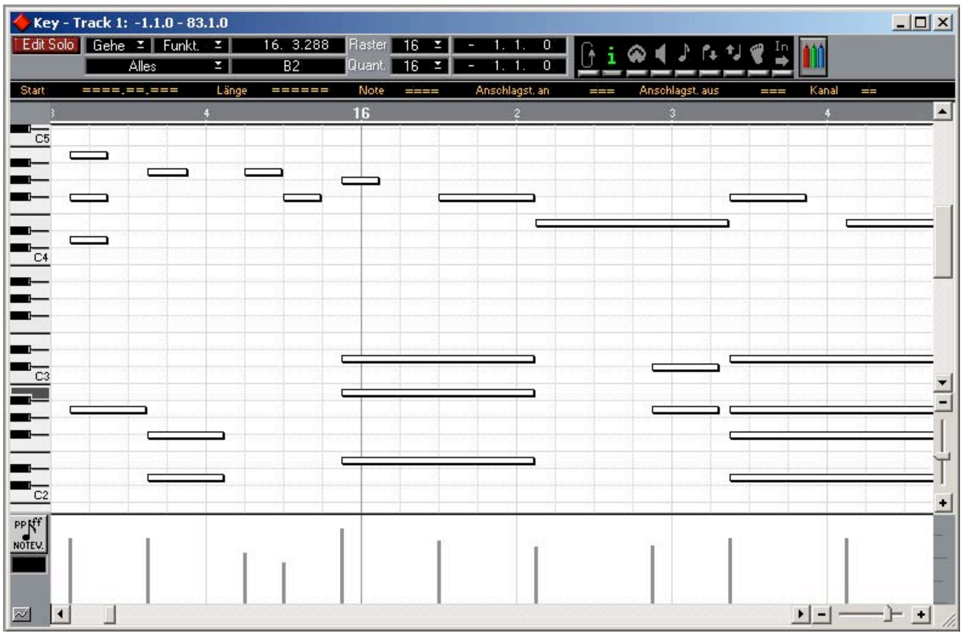


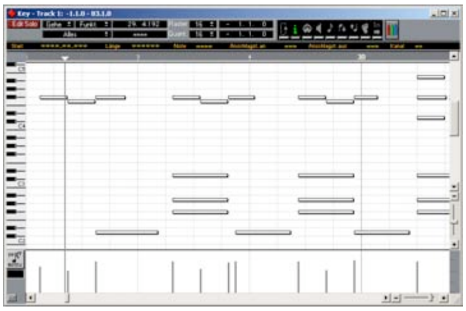
Figure 4.1: Piano roll in CubasisVST 5

For understanding the piano roll, the user has to recognize two correlations:

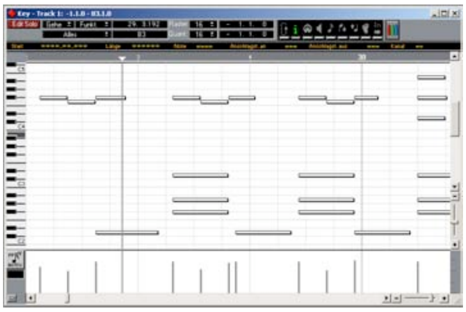
1. between pitch and vertical position and
2. between time and horizontal position.

The visible piano roll section has to be advanced during interaction. There are two possibilities: continuous and non-continuous advancement. Non-continuous advancement is characterized by a cursor moving from left to right while the visible section stays immobile (figure 4.2). When the cursor reaches the right end of the visible section, the next section is immediately displayed. For the MICON the other approach was taken: Instead of flipping between visible sections, the MICON continuously advances so that the current position is always in the middle of the screen. New notes appear on the right end of the screen, and old notes disappear on the left.

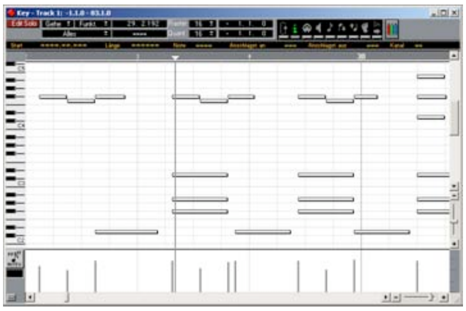
The continuous approach has no interruptions of the reading flow. The current position stands still in the middle of the screen, and the notes move along as time proceeds. The user does not have to search the current position even if he only occasionally looks at the piano roll.



(a) The cursor is at the end of the first note.



(b) The cursor is at the middle of the first bass note.



(c) The cursor is at the beginning of the first chord.

Figure 4.2: CubasisVST 5 piano roll: A cursor marks the current position.

The notes that are located in the center of the screen are a representation of the currently heard musical material. To additionally mark the position, the MICON highlights the currently played notes. As time proceeds the notes move to the left, and the highlights move to the following notes. The moving highlights create a visual rhythm that corresponds exactly to the rhythm of the music. When the user hears a note onset, he sees a note that gets highlighted in exactly the same moment. This can lead a user without prior experience about the piano roll notation to understand it.

Another alternative to mark the current position was considered but rejected: A cursor line would stand still in the center of the screen and mark the current position: This would lack the visual rhythm, and no natural connection between the visual representation and the heard music would be present. The piano roll would have to be explained to the typical first-time and one-time user.

Piano roll notation has some drawbacks. The absolute pitch of a note is hard to identify on a piano roll. Dynamics and instrumentation are not represented on a plain piano roll, and articulation can only be partly derived by examining the length of the notes. The most problematic property of the piano roll notation in context of an interactive conducting system, however, is the lack of representing high-level rhythmical structures like meter. The pulse notation supplements the piano roll with a representation of a high-level rhythmical structure.

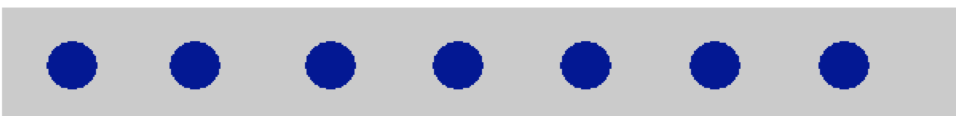
4.1.2 Pulse Notation

The pulse notation was created to support the Personal Orchestra user with his conducting. The deviation of the conducted beat from the beat in the music is essential for controlling the tempo of Maestro. Conducting before the beat increases, conducting after the beat decreases the tempo. However, in a piece with a lot of rubato, the timing of the beats is hard to predict increasing difficulty. The pulse notation visualizes the beat of the music and helps the user to coordinate his conducting gesture with the beat.

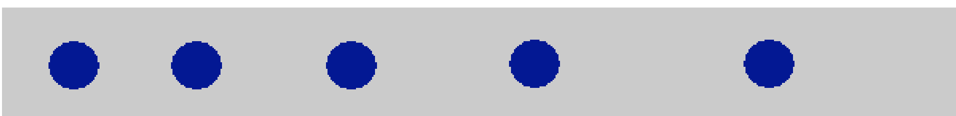
The beats are represented as circles, which lie on a horizontal line. The distance between two circles on that line depends on the temporal distance between the two represented beats.

The pulse notation outlines the metrical landscape of the music around the current beat, giving the user a representation of the recently passed and the upcoming tempo. Figure 4.3(a) shows a metric landscape for a steady tempo, and figure 4.3(b) shows a ritardando that has just begun.

The visible section moves continuously to the right so that the reading-flow is not interrupted. New circles appear on the right end of the screen, and old circles disappear on the left. To mark the current position clearly and to make the pulse notation self-explanatory, the currently heard beat is visu-



(a) Steady tempo.



(b) Ritardando

Figure 4.3: Metrical landscapes

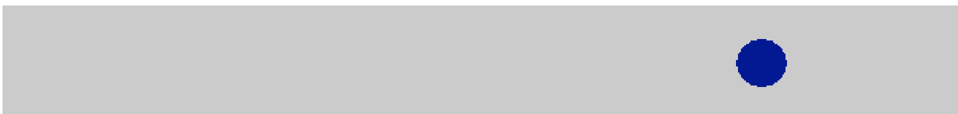
ally highlighted (figure 4.4). This leads to a visual rhythm that corresponds to the beats of the music and enables the user to understand the pulse notation by himself.

Before the user begins to conduct, a tempo indication should give the initial tempo of the piece. This gives an important cue on how to begin the piece, which was a frequent problem with users of prior Personal Orchestra exhibits and a Maestro-MICON prototype (section 7.1). A non-moving pulse can give this tempo indication.

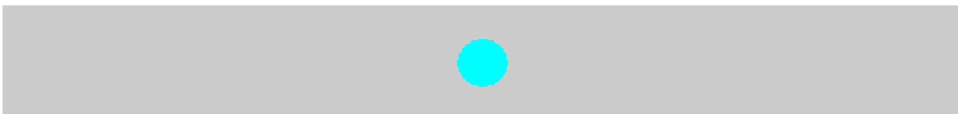
4.1.3 Combining Piano Roll and Pulse Notation

Piano roll and pulse notation supplement each other. The pulse notation provides a representation for a high-level rhythmical structure, and the piano roll provides a representation of the musical material including rhythm. Therefore, piano roll and pulse notation are simultaneously shown on the MICON.

The indication of a beat in the pulse notation very often coincides with new notes getting highlighted on the piano roll. These two events have to occur exactly simultaneously to ensure that the visual rhythms of the piano roll and the pulse notation are perfectly synchronized. Aside from aesthetic considerations, badly synchronized pulse and piano roll representations would harm the clarity of the visual rhythm, making it more difficult to



(a) The pulse approaches the center of the screen.



(b) The beat occurs.

(c) The pulse leaves the center of the screen.

Figure 4.4: The pulses are visually highlighted.

understand the representations. The content creation for the piano roll and the pulse notation is interwined and provides the desired synchronicity (section 5.4).

4.1.4 Score

The piano roll and the pulse notation do not require a background of formal musical training. However, as music exhibit Personal Orchestra attracts the attention of people with interest and knowledge about music. The MICON should offer visualizations that enable musically educated users to profit from their abilities. To this end the score representations were included.

Aside form offering more or easier accessible information about pitch, articulation, dynamics, instrumentation, etc., a score representation is easier to read than a piano roll for users with good music-reading skills. They recognize bigger structures like chords, scales, etc., which enables them to read the score fluently.

MICON advances the score automatically. Several strategies were considered:

- The score is presented like a book. When the cursor reaches the end of the right page, the next two pages become visible.
- When the music reaches the end of the left double-page, the right page moves to the left, and the new page is displayed on the right.
- The entire score is represented without line breaks, and the line continuously moves while music plays.

The last two choices allow the user to read the score in advance at every moment. In spite of this, it was decided to use the book metaphor in order to keep the representation as simple as possible and to not introduce another, musically unimportant, item (SIMPLE IMPRESSION Borchers [2001]).

A cursor marks the current position in the score (figure 4.5). The user, sometimes watching the video of the orchestra or exclusively concentrating on conducting, can always easily find the current position in the score.

The user can emphasize an instrument section on Maestro by conducting toward it. The resulting change in the timbre of the orchestra is hard to perceive. In order to give visual feedback, the emphasized instrument section

Figure 4.5: A cursor marks the current position.

is marked with a color on the full score representation (figure 4.6). This visual hint, hopefully, leads some users to observing the change of the timbre.

Even for users with good music reading skills, a full score can have too much information to comprehend in the rush of time. To give these users a score representation, which fits their abilities, the MICON also features a piano extract score. Figure 4.7 shows a comparison of the same musical idea expressed as full score and piano extract score. The piano extract score representation on the MICON does not highlight the emphasized instrument section but is otherwise an exact clone of the full score representation.

Figure 4.6: A highlight marks the emphasized instrument group.

(a) Full score

(b) Piano extract score

Figure 4.7: Full score and piano extract score of the first six measures of the Annenpolka

4.1.5 Combining Score and Pulse Notation

The beat-timing information offered by the pulse notation is a valuable aid for the user and is hence shown together with the score representations. To connect the score and the pulse notation, a jumping ball, which marks the beats, was added to the score representation (figure 4.8). The ball jumps from beat-position to beat-position in the score. Beats occur when the ball hits the "ground". The shapes of the jumping ball and the pulses are similar, and the rhythmical properties of their motions, i.e. hitting the ground respectively pulsating, coincide and chain the two representations together.

Figure 4.8: A jumping ball marks the beats and connects score and pulse notation.

4.2 Aesthetic Design

4.2.1 Score

The score pages have a curved form, which resembles physical paper more accurately and gives a more beautiful impression than a flat rendering would (figure 4.9).

Page-turning is animated. During page-turn animation the shape of the turned page continuously changes: At the beginning, the page turned has a double-curved, S-shaped form. Reaching 90 degrees it is perfectly round and has a C-shaped form. Finally, the MICON reestablishes the S-shaped form when the animation completes (figure 4.10). A backward page-turn animation can be seen in figure 4.11. The backward turning animation is slightly different as the page is bowed into the opposite direction.

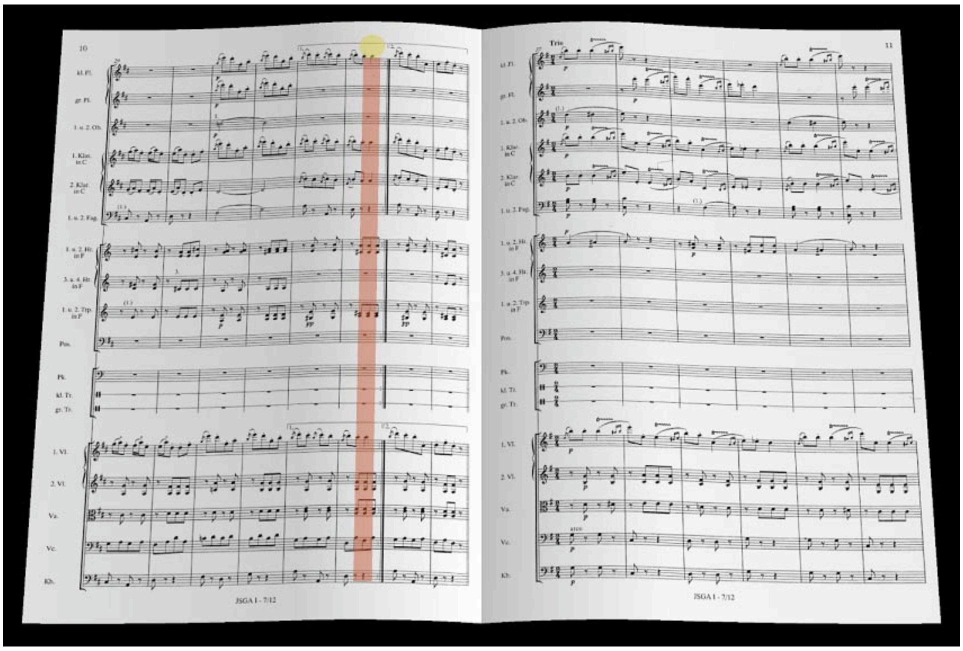
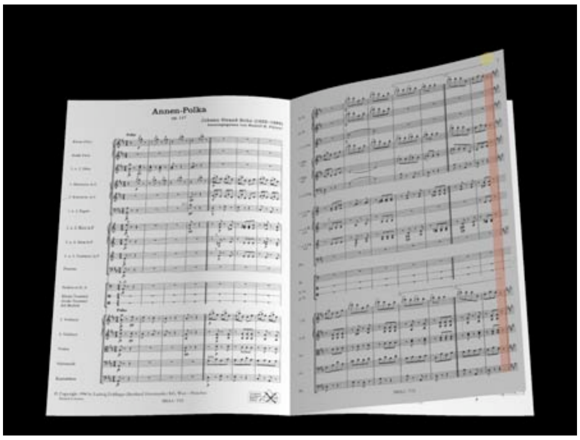
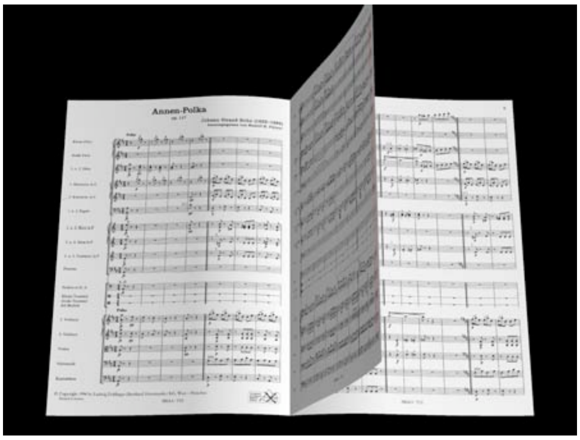


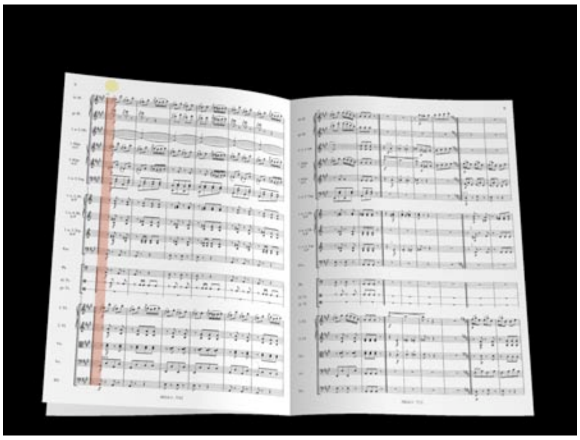
Figure 4.9: Full score on the MICON



(a) Page-turning begins.

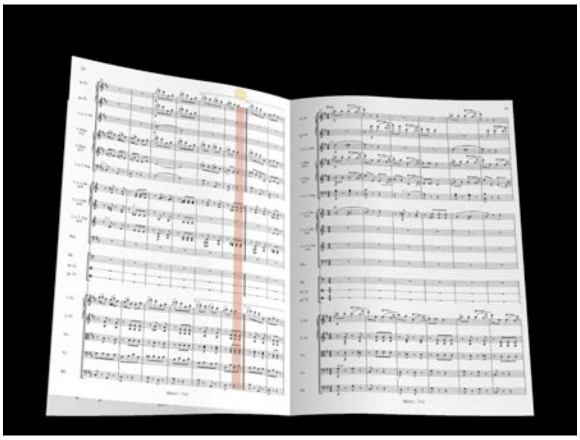


(b) The page is half-way turned.

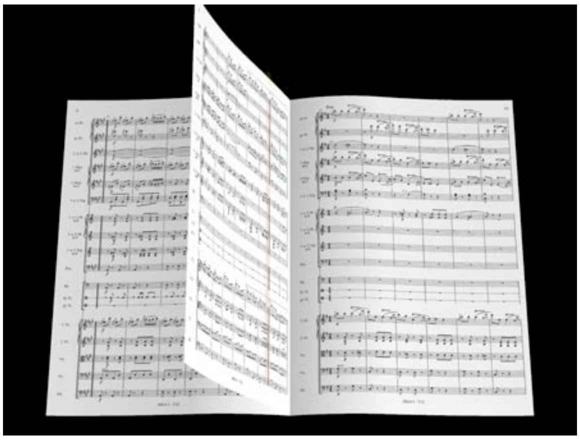


(c) Page-turning has nearly completed.

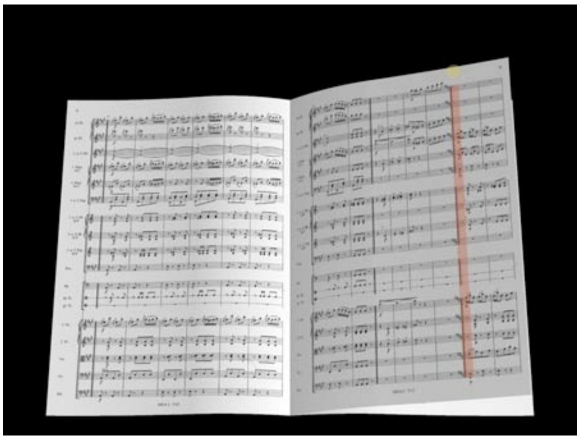
Figure 4.10: Page-turn animation (forward). See appendix B for more and bigger screenshots.



(a) Page-turning begins.



(b) The page is half-way turned.



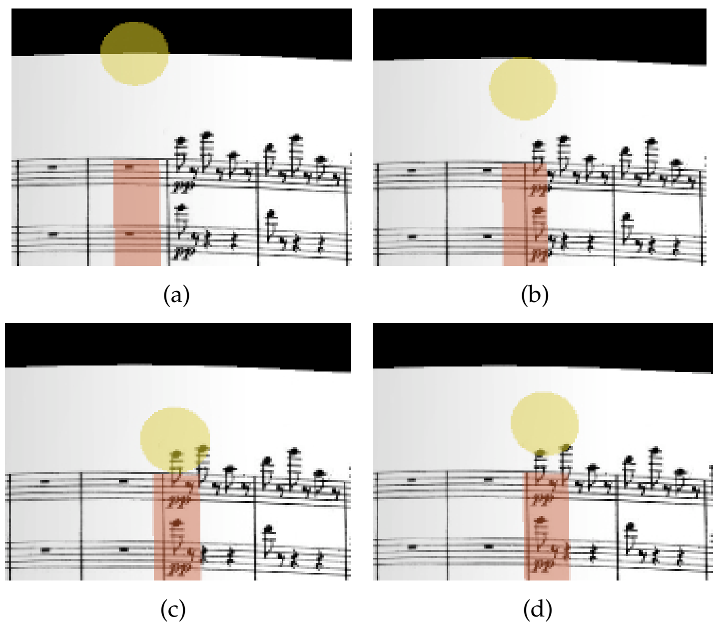
(c) Page-turning has nearly completed.

Figure 4.11: Page-turn animation (backward)

The cursor and the jumping ball move along the curvature of the pages. The jumping ball can be seen from close in figure 4.12.

The MICON gives feedback about the current instrument emphasis by coloring the appropriate lines in the full score. The intensity of the emphasis is reflected by the intensity of the highlight color (figure 4.13). Maestro changes instrument emphasis only every few seconds. In order to avoid a jumpy impression, the MICON continuously modifies the highlight, slowly reaching the new emphasis intensity.

When moving to another note system, the cursor fades out smoothly. After vanishing from the old note system, it smoothly fades in at the new. The cursor spends half-a-beat time in the old and half-a-beat time in the new note system. The instrument highlight and the jumping ball fade in and out accordingly (figure 4.14).

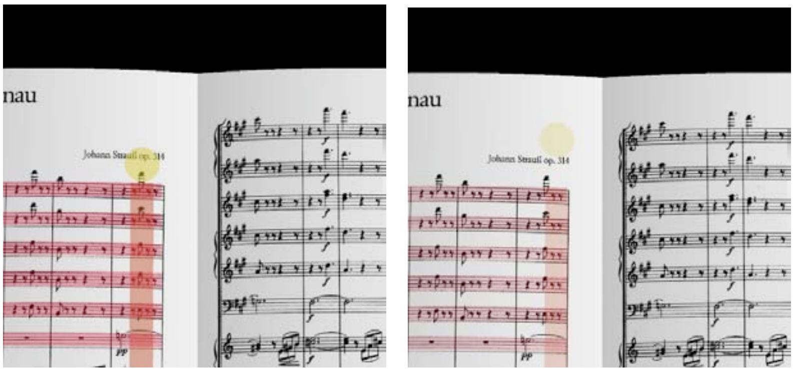


(a) (b) (c) (d)

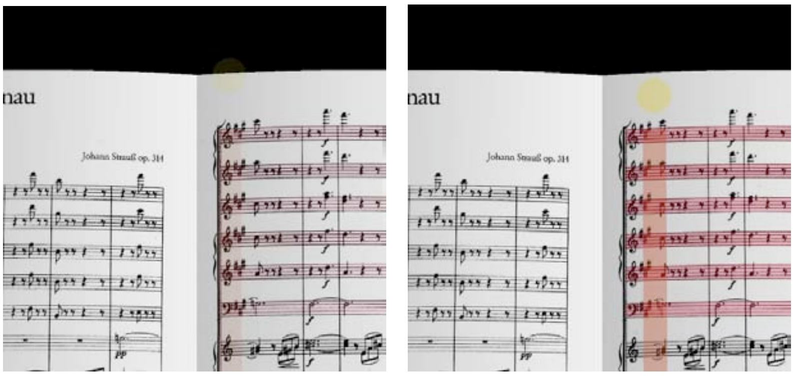
Figure 4.12: The jumping ball: (a) and (b): before the beat, (c) the beat occurs, (d) after the beat

(a) 100% emphasis (b) 30% emphasis

Figure 4.13: Emphasis intensities



(a) Fading out begins. (b) The ball reaches the highest point.



(c) The cursor reappears on the next page. (d) The fading process is nearly finished.

Figure 4.14: The cursor moves to another note system.

4.2.2 Piano Roll

Figure 4.15 shows some time-proximate screen shots of the MICON presenting the "Blue Danube" as piano roll. The notes that are currently played by the orchestra are in the center of the screen and are additionally highlighted with a brighter color. While the boxes continuously move, the highlights jump from note to note in the rhythm of the music.

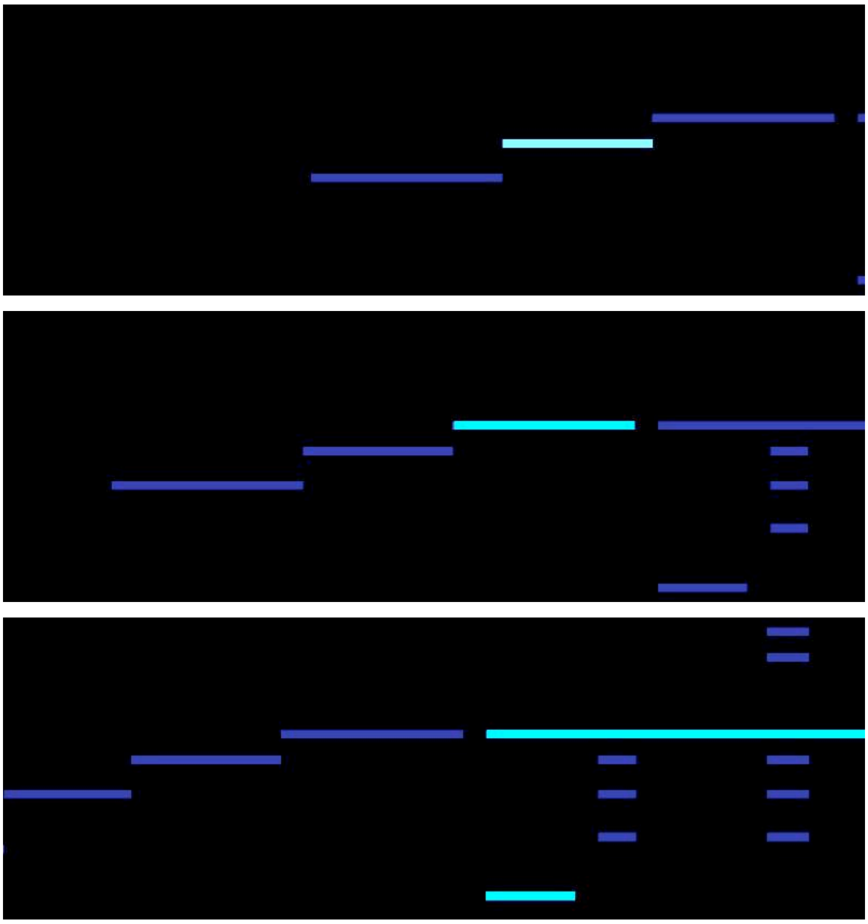


Figure 4.15: The highlights mark the current position. The highlights were artificially emphasized for better clarity in print. An unmodified piano roll screenshot can be found in appendix B.

4.2.3 Pulse

The beats are marked with a pulsating animation of the moving circles (figure 4.16). Coming from the right and approaching the center of the screen, the circles grow in size. Far from the center they grow only slowly, but as they approach the center their rate of growth increases. They reach their maximum size in the center of the screen, marking the beat. Afterward, they shrink again, first rapidly then slowly, to their original size.

Before the user begins to conduct, a pulsating circle in the middle of the screen indicates the initial tempo of the piece. The other circles stand still (figure 4.17). When the user begins conducting, the center circle slowly fades out while the other circles begin moving (figure 4.18).

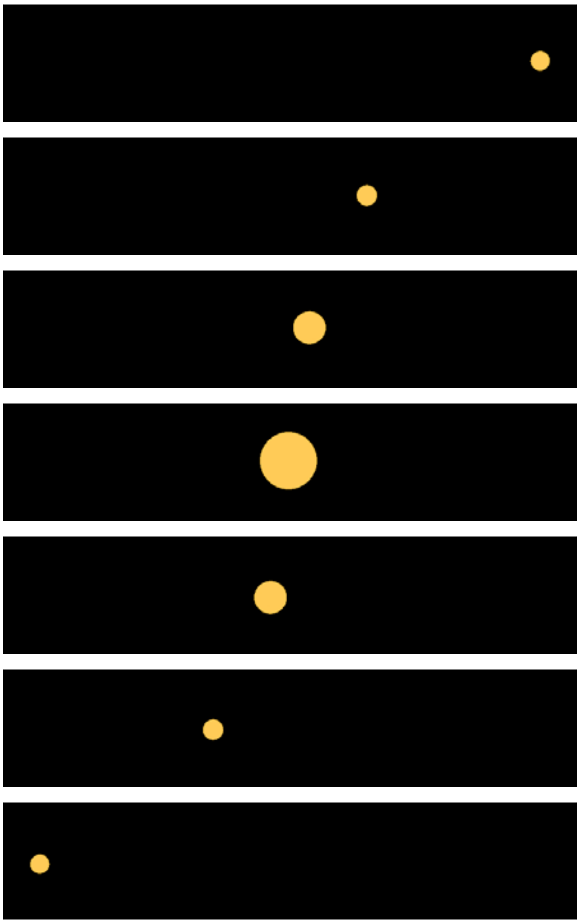


Figure 4.16: The pulses grow as they approach the center of the screen. The beat occurs, and the pulses shrink again to their original size.

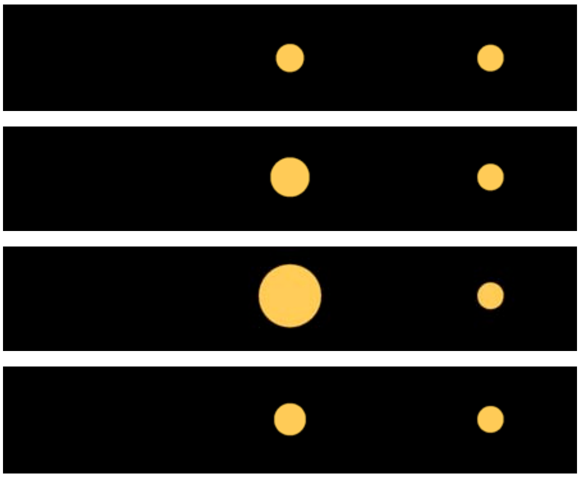


Figure 4.17: A circle pulsates in the middle.

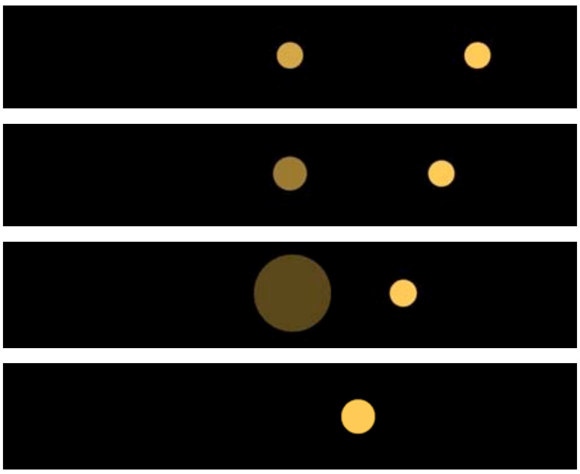


Figure 4.18: The circle in the middle fades out.

4.2.4 MICON Display Options

Several properties of the representations are configurable (figure 4.19) and can be altered to fit well with the design of the exhibit installation.

For the full score and the piano extract score,

- the color of the cursor, of the emphasis highlight and of the jumping ball and
- the thickness of the cursor and the size of the jumping ball can be set.

The options are handled for the full score and the piano part score separately. Thus it is possible, for example, to have different cursor and jumping ball sizes for these two representations.

For the pulse notation,

- the color of the circles,
- the size of the circles,
- the magnification factor and
- the point in time when magnification begins can be set.

For the piano roll, the color of the boxes can be set.

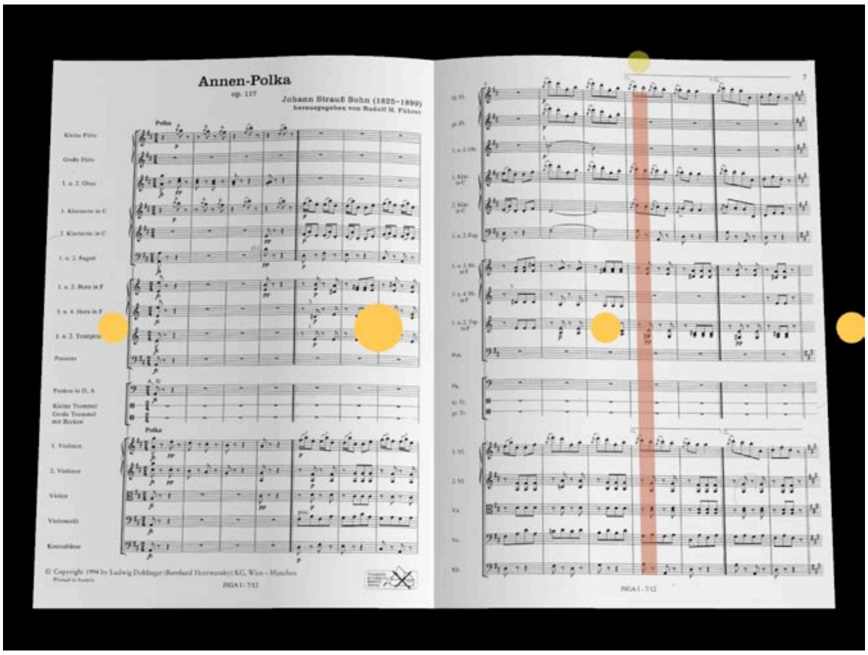
The MICON allows to place the representations in a virtual 3D space and change their orientation and scale their size freely. This flexibility allows to change the layout of the MICON screen, enabling to save space on a small display while placing the representation more orderly on a large display (figure 4.20).

(a) Score

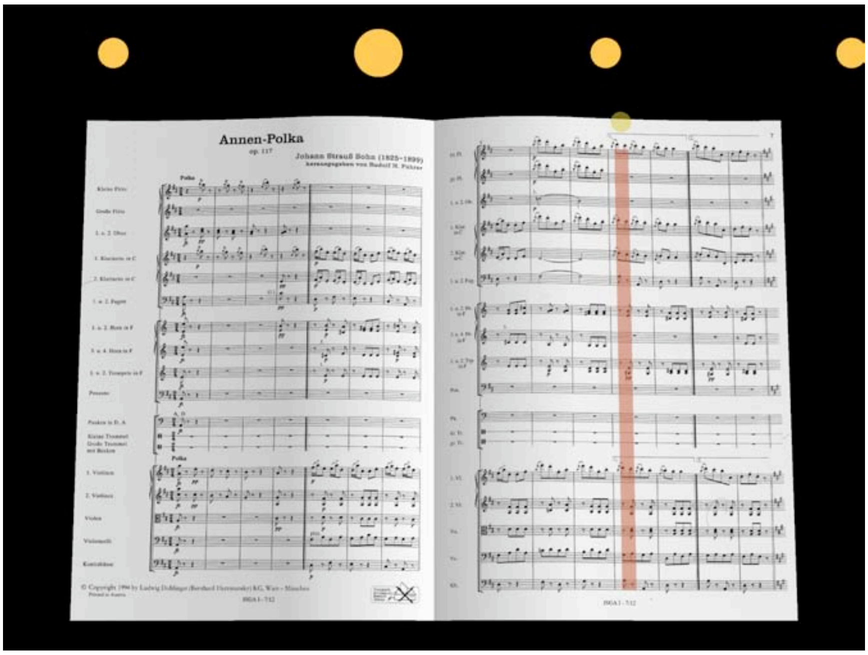
(b) Pulse

(c) Piano roll

Figure 4.19: Options for the score, piano roll and pulse notation



(a) Small display



(b) Large display

Figure 4.20: Suggested layouts

4.3 Integrating MICON and Maestro

Figure 4.21 shows the layout of a typical MICON-Maestro installation. The user stands in front of a large screen. Two loudspeakers are directed toward him. The computer hardware is hidden. In front of the user stands the music stand. A sensor tracks the baton and sends its position to the computer.

The user sees a list of music pieces on the large screen, points the baton toward the desired item and pushes the button on the baton. In the same way he selects the representation of the musical material on the music stand. He has three choices: full score, piano part score or piano roll. In either case, the MICON shows the pulse notation at the top of the selected representation (figure 4.22). After the selection, the orchestra appears, the user begins to conduct, and the orchestra starts playing with the score display advancing automatically.

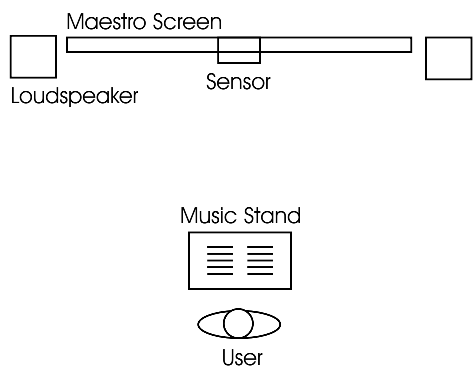
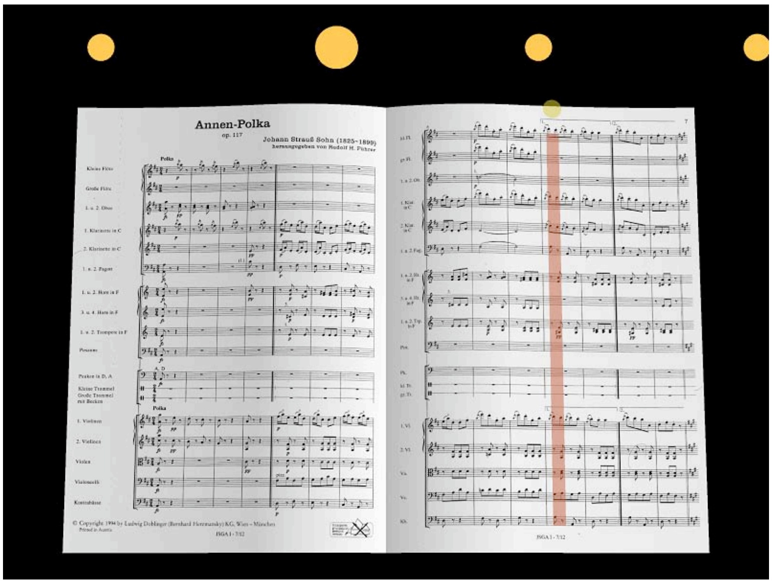
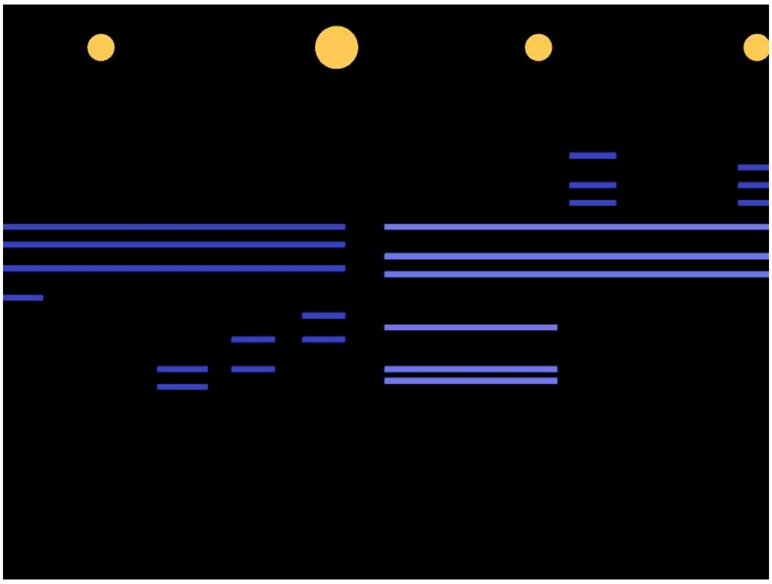


Figure 4.21: Installation of Personal Orchestra



(a) Score



(b) Piano roll

Figure 4.22: The pulse notation is shown in top of the score representations and the piano roll.

Chapter 5

Content Creation

For a new music piece, Maestro needs:

- a synchronized audio and video recording of the piece and
- timing information about the beats in the recording provided by a Beats file.

For a new music, piece MICON needs:

- scans of the full score and the piano extract score,
- information about the score provided by ScoreInfo files,
- timing information about the beats in the recording provided by a Beats file and
- a piano roll representation given by a MIDI file that is synchronized with the audio recording.

Adding a new music piece to Personal Orchestra 4 requires various manual processing steps. Figure 5.1 shows the suggested workflow for that task. The beat concept is of central importance. It determines the variables in the creation of the ScoreInfo file and the Beats file. The beat concept is the regular pulse of the piece, the base unit that the user conducts. The base unit given by the composer is a good choice (e.g., a quarter note in a 4/4 measure). Sometimes, however, a larger base unit is easier to conduct with the quite heavy electronic baton. The MICON and the Maestro have to

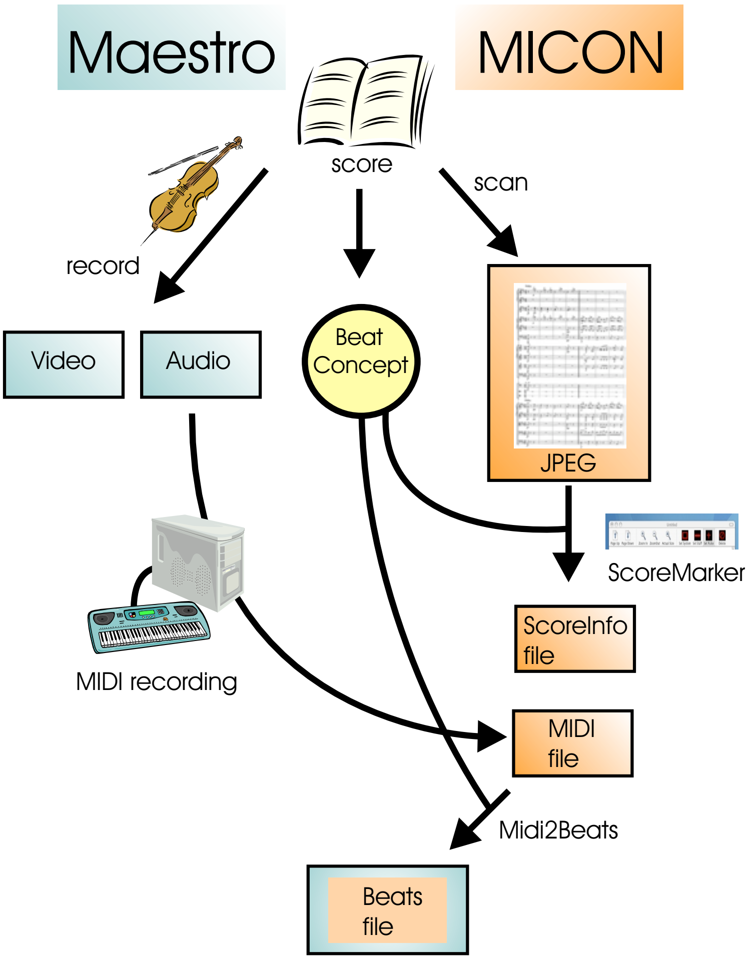


Figure 5.1: Workflow for adding a new piece to the Personal Orchestra system

be based on the same beat concept; otherwise, the pulse notation would be pointless. This is best achieved by using the same file for both parts. Because of the combination of the pulse notation with the piano roll, the MIDI file and the Beats file have to be perfectly synchronized. Therefore, the Beats file is derived from the MIDI file and is then given to Maestro.

5.1 Score Information Model

As far as the MICON is concerned, a score is made of three structures:

- note systems,
- beats and
- staffs.

A note system is a rectangular area in the score. It has an upper and a lower external leading (figure 5.2). The rectangular area defines the area where the cursor moves from left to right during score animation. The note systems are always axis-aligned and cannot be sloped.

A beat is a x/y-position in the score. Every beat is embraced by exactly one note system. A beat belongs to a note system if it lies within the rectangle spanned by the left and right end of the note system and the upper and lower external leading. Figure 5.3 shows a note system with adjunct beats.

A staff is a line in the note system. Each staff is embraced by exactly one note system. A staff belongs to a note system if it lies within the rectangle spanned by the left and right end of the note system and the upper and lower external leading. Figure 5.4 shows a note system with adjunct staffs. Each staff has an id-number that represents the instrument group that corresponds to that staff.

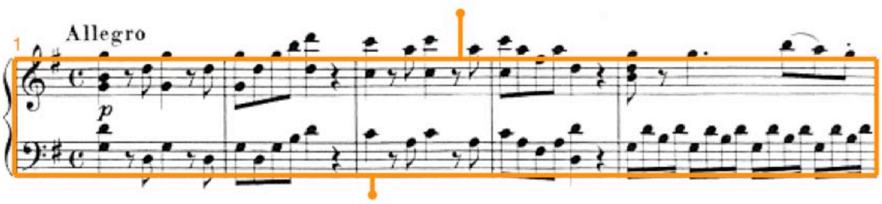


Figure 5.2: Note system

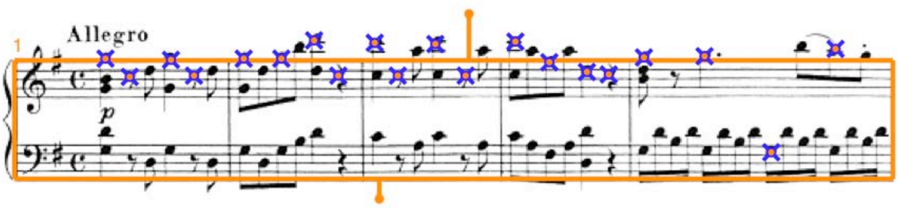


Figure 5.3: Note system with adjunct beats

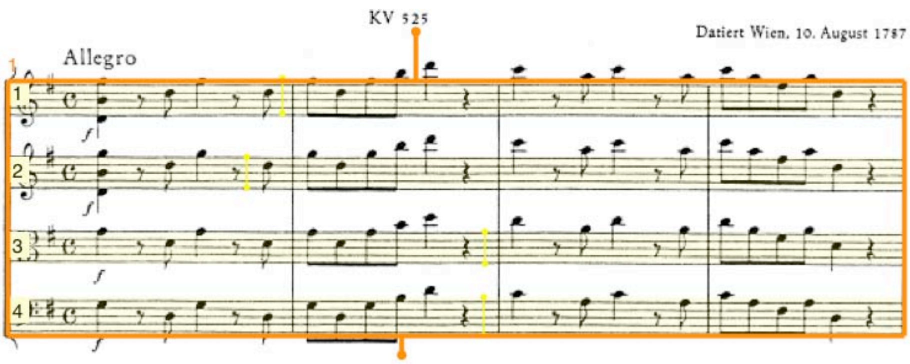


Figure 5.4: Note system with adjunct staffs

The note systems are the building blocks of the score animation. The single note systems can be arbitrarily connected, and the cursor can jump freely from one system to another. Within one note system, however, the cursor has to move the entire way from left to right before the it can move to another note system. The note systems are undividable (atomicity).

5.2 ScoreMarker

The MICON gets information about the positions of the note systems, the beats and the staffs from ScoreInfo files. These files can be created conveniently with the ScoreMarker program.

The requirements for the ScoreMarker were:

Efficient operation. The goal of ScoreMarker was to minimize the workload for providing the MICON with information about the score.

Usability. The ScoreMarker should be an easy-to-use program. Specifically, the order of user input should make no difference to the MICON.

5.2.1 Creating the ScoreInfo file

The ScoreMarker program (figure 5.5) offers a simple means to create ScoreInfo files. After loading the scans, the user defines note systems, beats

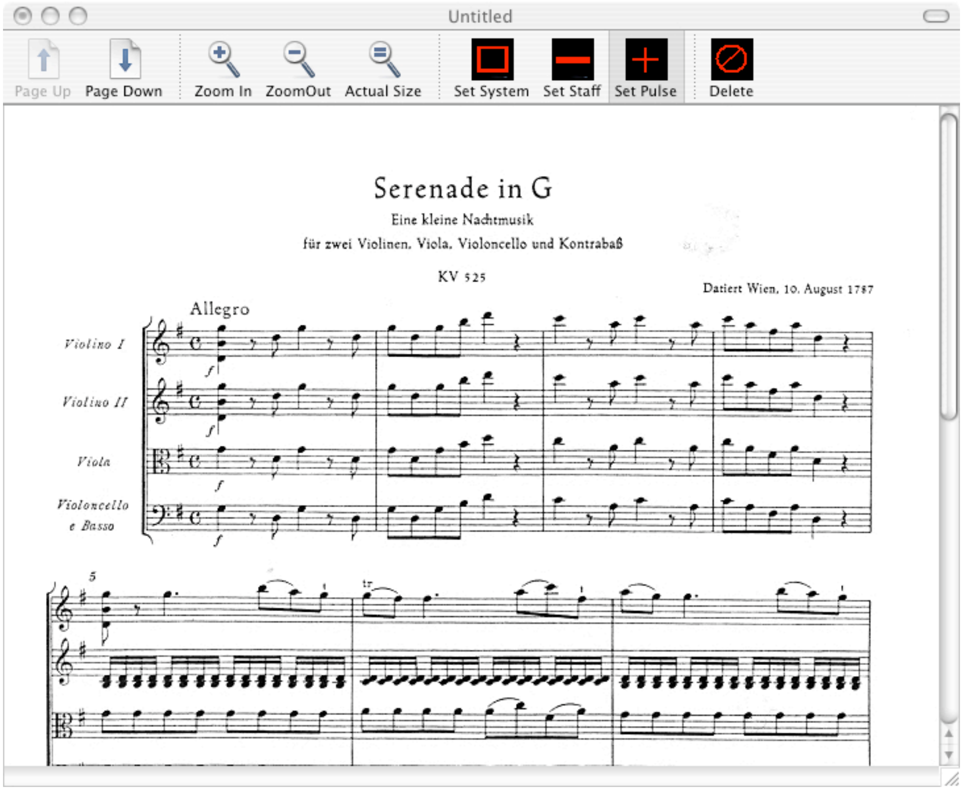


Figure 5.5: ScoreMarker after loading the scans

and staffs by clicking with the mouse. ScoreMarker provides immediate visual feedback. Facilities for zooming the score and for deleting previous input supports the user of ScoreMarker at his task. The ordering of the note systems, i.e. the sequence in that the cursor visits them, can be defined by the user. For pieces without repeats and jumps, ScoreMarker automatically finds the correct sequence of note systems.

A thorough description of ScoreMarker in form of a tutorial can be found in appendix A.

5.3 Creating the MIDI file

The piano rolls on the MICON are visualizations of MIDI files. To hear and see the same musical material at the same time, MIDI files have to be produced that exactly match the orchestral recordings. MIDI files have to be created that, besides matching the notes, closely resemble the tempo changes of the original recordings.

The central idea is to make MIDI recordings with an electronic piano attached to a computer: The player hears the orchestral recording and adjusts his tempo to closely fit the orchestra's play. The recording is manually post-processed in order to correct minor errors, to further improve synchronization and to achieve a better appearance of the piano roll. The creation of the MIDI files for the piano roll is a three step process:

1. Select and edit the score.
2. Practice and record.
3. Post-Process.

The piano extracts of the music pieces were used as the score for the MIDI recording. Some differences between the piano extract and the orchestral score, which were introduced in the transcription process, had to be reversed since the goal was to create a good match between the heard music and the piano roll; a piano extract, on the other hand, is required to provide a good sounding (and not a good looking or perfectly accurate) transcription of the piece. The changes include the deletion of newly introduced notes and the reversion of pianistic idioms like Alberti-basses and tremoli to their orchestral counterparts.

Having prepared and practiced the score, the MIDI recording was made with CubasisVST 5, a sequencer program allowing to import audio tracks. The orchestral recording was loaded into the application so that it was heard while the MIDI recording took place. For convenience, the entire piece was not recorded at once, but the piece was split both vertically and horizontally: It was not recorded all-at-a-time, but it was split into horizontal parts; it was not recorded all-voices-at-once, but it was split into horizontal parts: the different voices.

The final step is to manually edit the recording in order to get rid of errors, enhance the synchronicity with the orchestral recording and to refine the piano roll under aesthetic considerations. The recording was quantized, i.e. the beginning and the end of the notes were mapped from a seemingly continuous time to a multitude of a discrete time unit. Based on this fixed time raster, post-processing was done by moving around, cutting and elongating the notes with the mouse. Occasional errors were corrected. Wherever synchronicity issues arose, which were identified by attentive listening and comparing the audio and the MIDI recording, the notes were moved one-by-one to a better time position. Finally, the recording was refined under aesthetic considerations: Notes belonging to a logical compound should begin and end in the correct positions. When, for example, a chord is played, the notes that make up the chord should all start and end at the same horizontal position (figure 5.6). For legato passages, neighboring notes should end respectively start at the same horizontal position (figure 5.7).

5.4 Creating the Beats file

To ensure that the piano roll and the pulse notation have a common visual rhythm, the MIDI file and the Beats file of the music pieces have to be perfectly synchronized. This can be accomplished by deriving the Beats file from the MIDI file with the MIDI2Beats program.

Midi2Beats extracts the rhythm out of a MIDI file. I.e., every note-on in the MIDI file is stored as a beat event in the Beats file. As it is actually not wanted to extract the rhythm but only the beats, the MIDI file has to be edited beforehand. A MIDI file has to be created that has a note-on exactly then when a beat occurs. For achieving this, the majority of the notes have to be deleted. In the circumstance that there is no note-on when a beat occurs, a note has to be added there. After editing, the Midi2Beats program can be used to create the desired Beats file. MIDI2Beats is console application. It has two command-line arguments. The first argument is an integer number giving the time-scale in ticks per second. The second argument is the filename of the MIDI file to be processed. The beat times are written to the console and can then be redirected to a file.

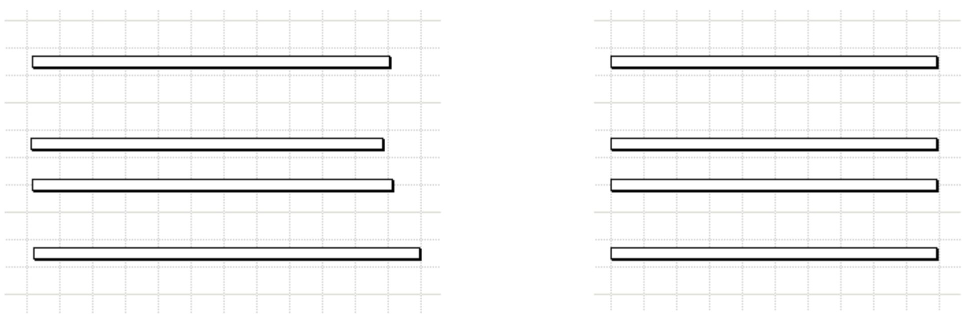


Figure 5.6: A chord: original and refined

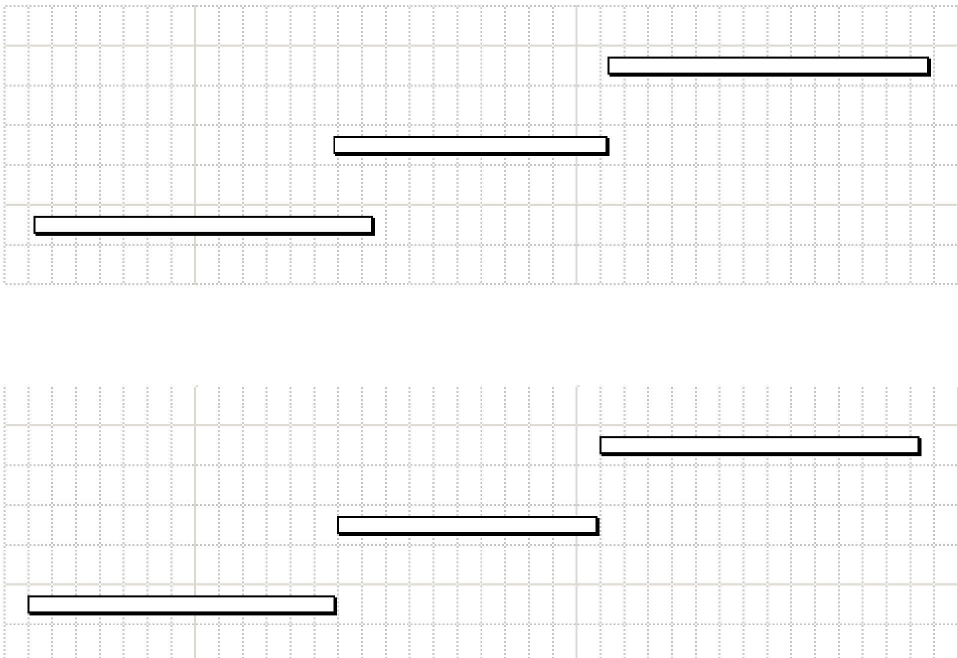


Figure 5.7: A legato passage: unrefined and refined

Chapter 6

Implementation

Three programs were developed:

- MICON,
- ScoreMarker and
- Midi2Beats.

ScoreMarker was the first program to be written. ScoreMarker consists of two parts: a back-end, implemented in C++, and a front-end, implemented in Obj-C using the Obj-C Cocoa framework. The back-end reads and writes ScoreInfo files while the front-end handles the user interface.

Second, the MICON was written. The MICON consists of four parts: file-I/O, communication, assembling and rendering. The file-I/O part consists of a ScoreInfo file facility, which was reused from the back-end of Score-Marker, and additional facilities to read Beats and MIDI files. The file-I/O part was written in C++. The rendering part of the MICON uses OpenGL. The rest was written in Obj-C using the Cocoa framework.

The last program to be written was the Midi2Beats converter. It was needed because the Beats files that were produced for the Maestro system were not sufficiently synchronous with the MIDI files used for the piano roll. The implementation of MIDI2Beats reused the MIDI-facility from the MICON.

6.1 MICON

During conducting Maestro continuously informs the MICON about the current position of the music. The position is expressed in seconds since beginning of the piece. Using this temporal information, the MICON assembles the representations and renders them.

This section follows the way of a time-packet. This section describes how the time-packet is communicated between the Maestro and the MICON, and how the MICON uses this information to assemble the representations. Then, the rendering process will be presented. However, at start-up MICON first loads and parses Beats files and ScoreInfo files. So, this exploration starts there.

6.1.1 File Formats

MICON uses JPEG and MIDI files for displaying the score and piano roll. MICON uses two non-standard file-formats, Beats files and ScoreInfo files, that are explained in this section.

Beats File

A Beats file contains a representation for the beats of a music piece. A Beats file is separated into a header and a body. The header defines the time-scale and gives the duration of the piece. The time-scale is defined in terms of ticks per second. The beat times are contained in the body of the Beats file. They are given according to the time scale. Here is an excerpt from a Beats file:

```
###########################
##  BeatTapper BeatFile  ##
###########################

# Timescale in ticks per second
[scale]
600

# Duration of the movie in ticks
[duration]
```

```
129875

# Beattimes measured in ticks
[beats]
755
1511
2266
3021
3808
4666
...
```

This beat-file defines the first beat on position 755. This is 755/600 seconds according to the time-scale. The next beat is on 1511/600 seconds and so on.

The Beats-file format has not been originally developed for the MICON, but it has been in use in prior version of Personal Orchestra to present beat times. Using the same file-format in the Maestro and the MICON increases their interoperability and diminishes workload in the set-up process of the Personal Orchestra 4 system.

ScoreInfo File

ScoreInfo files give the MICON information about the score, which is provided as a set of JPEG files. This information is mainly about geometrical properties of the score that are of importance to the MICON.

Here, an excerpt form a ScoreInfo file:

```
JPEG
10 files
/Users/telis/Noten/donau/DONAU1.JPG
/Users/telis/Noten/donau/DONAU2.JPG
[...]

10 pages

page 1
1 note systems
163 1127 895.333 107.123 534.29 1174.38 506.518 79.35
```

```
12  18 beats
13  232 339
14  247 338
15  ...
16
17  18 instruments
18  0 307 571 307 546
19  0 307 519 307 496
20  1 308 1127 308 1102
21  1 310 1077 310 1055
22  1 309 1028 309 1007
23  1 311 981 311 956
24  1 312 931 312 907
25  1 313 882 313 860
26  [...]
27
28  page 2
29  [...]
30
31  sequence
32  7 boxes
33  1 2 3 4 1 2 3
```

The preamble of a ScoreInfo file (lines 1–7 in the above example) gives the file-type (JPEG or PDF) and the filenames of the files making the score. The last line of the preamble gives the number of pages, which may be different from the number of files in PDF-mode. After the preamble come the score page descriptions (page 1: lines 9–26, page 2: lines 28–29). Each page description contains data about the note systems, beats and instrument groups of that page. The note systems are described by four 2D coordinates (8 numbers): upper left point, lower right point, upper external leading, lower external leading. The beats are described by one 2D coordinate. The instrument groups are described by an instrument group id-number followed by two 2D coordinates: upper boundary and lower boundary. The beats and the note systems are independently stored, and the fact that a beat is always contained in a note system is not explicitly reflected in the structure of the ScoreInfo file. The last part of the ScoreInfo file (lines 31–33) gives information about the sequence of the note systems. The cursor visits the note systems during score animation on the MICON in the given order. The note systems are numbered according to their natural order in the score (first page before last page, top before bottom, left before right). This is also the order in which they appear in the ScoreInfo file.

6.1.2 Communication

Maestro and the MICON communicate via an UDP-based protocol. Communication is one-way: Maestro sends messages to the MICON, which then updates its internal state. During the initial user selections, Maestro informs the MICON about the chosen piece and the score representation. During conducting, Maestro continuously sends the current position to the MICON. Every few seconds, Maestro sends the current instrument emphasis.

There are 5 different messages:

- select piece,
- select representation,
- set time (aka. time packet),
- set emphasis and
- end of piece.

The parameters of these messages are listed in table 6.1: Figure 6.1 sketches a communication cycle.

6.1.3 Assembling the Representations

Assembling a representation means to compute all time-variable parameters that affect its appearance. Assembly depends very much, but not solely, on the last received time-packet.

Message	Parameter Type	Description
select piece	integer	the id of the piece
select representation	integer	the id of the representation
set time	float	time passed in seconds
set emphasis	integer float	the id of the instrument group current emphasis between 0.0 and 1.0
end of piece	no parameter	

Table 6.1: The communication messages and their parameters

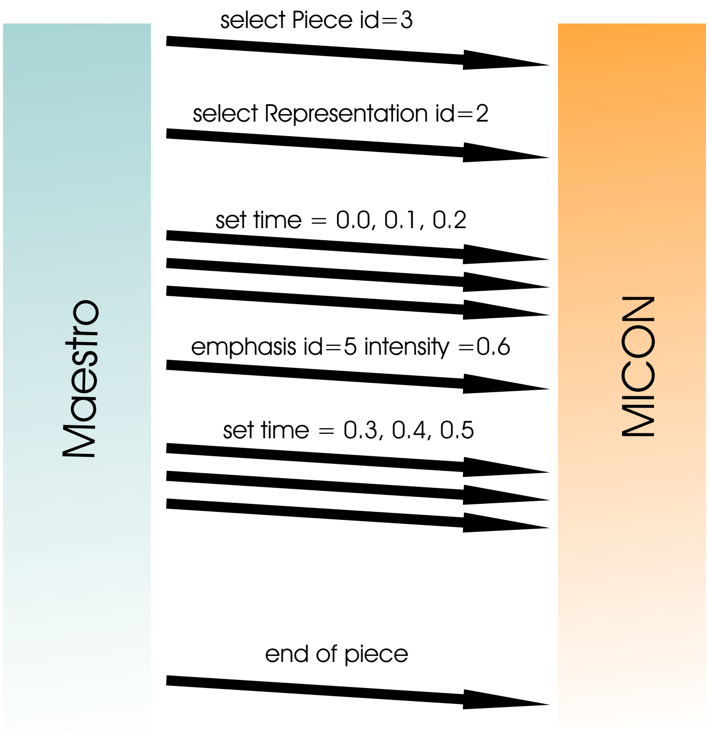


Figure 6.1: Communication between Maestro and MICON

Score

The time-variable parameters of the score representation and hence the targets of the assembly process are:

- the page-numbers of the left and right page,
- the position of the cursor,
- the height of the jumping ball above the cursor,
- the positions of the highlighted staffs,
- the translucency (alpha value) of the cursor, the jumping ball and the emphasis highlight,
- whether a page turn currently occurs and if so
- the page-numbers of the front and back page of the turning page and
- how far the page is turned ($0°$ to $180°$).

For determining the position of the cursor, the timestamps, which originate from Maestro, are first converted into a format that is more meaningful in the context of a musical score: time measured in beats. A value of 9.5, for example, describes the point in time halfway between beat 9 and 10. To convert form seconds-based to beats-based time-format, MICON uses information provided by the Beats files.

The MICON, now provided with the converted time format, has to find the positions of the preceding and the succeeding beat in the score. The MICON looks up the positions in an internal data structure: an array that is built upon the ScoreInfo file. The array contains the positions, i.e. x/y-coordinates and page numbers, of all beats. Beginning with the first note system given in the sequence-section of the ScoreInfo file, the contained beats are added to the positions array starting with the leftmost beat. Repeating this for the note systems in the order of their appearence in the sequence-section of the ScoreInfo file, makes the positions array contain all beats in the correct sequence. A beat can occur in the positions array more than once, if the embracing note system is contained more than once in the sequence-section of the ScoreInfo file, in a piece with repeats. Having looked up the positions of the beats, MICON linearly interpolates the position of the cursor. Let x_1 and x_2 be the horizontal positions of the first and second of the two beat (figure 6.2) and t_1, t_2 be the points in time when the two beats occur. then

$$\text{cursorPosition}(t) = x_1 + \frac{t - t_1}{t_2 - t_1} \cdot (x_2 - x_1)$$

is the horizontal position of the cursor. The vertical position and extent of the cursor is taken form the note system that embraces the two beats. In order to calculate the height of the ball above the cursor, with vertical cursor position x, a quadratic formula is used (see figure 6.3 for a plot); x_1

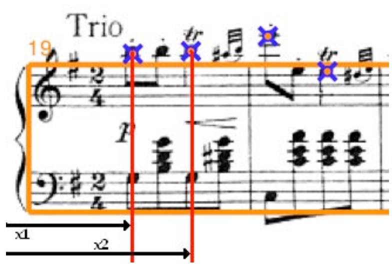


Figure 6.2: The positions of the beats in the score are provided by the ScoreInfo file.

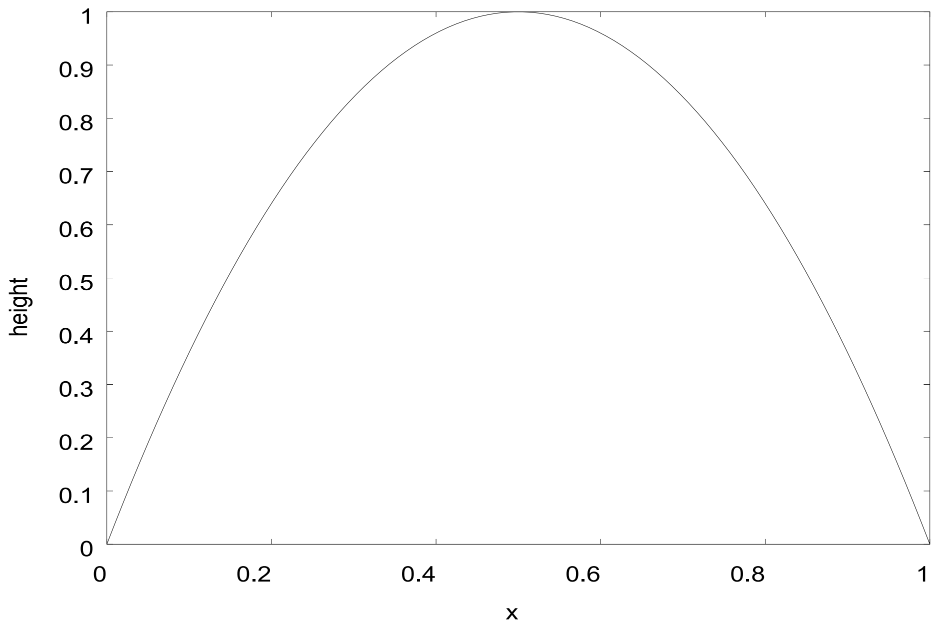


Figure 6.3: A Plot of the ball jump function, jumping from $x_1 = 0$ to $x_2 = 1$

and x_2 are same as above:

$$\text{ballHeight}(x) = 1 - (2 \cdot \frac{x - x_1}{x_2 - x_1} - 1)^2$$

In the case that first and second beat are contained in different note systems, the formula can nevertheless be applied with modified parameters t_1, x_1 and t_2, x_2, ensuring that the ball has maximal height when leaving the first and entering the second note system. In case the two beats are contained in different note systems, the cursor smoothly fades out at the end of the first note system and then fades in at the second note system. The translucency (alpha value) for the ball and the cursor are linearly interpolated between normal stroke and invisibility. But if the two note systems are very close to each other, translucency is unaffected: the ball and the cursor are drawn with normal stroke. This allows having more than one note system on a single line, which is often needed if a piece has repeats.

MICON considers two different types of pages: static, one-sided pages and animated, two-sided pages. The left and the right page are static and one-sided. The two sides of the middle page become successively visible during page-turn animation. In the beginning MICON shows the first two pages of the score. Whenever the cursor reaches the last beat of the right page, page-turning animation is started. MICON assigns new page numbers for the static right page and the two sides of the animated middle page. Let $t_{\text{real start}}$ be the point in real-world time when the page-turn starts, and let t_{real} be the current point in time (seconds); let $t_{\text{beat before}}$ and $t_{\text{beat after}}$ be the beat times as given by the Beats file and t_{Maestro} the value of the last time-packet the MICON has received from Maestro in converted time-format. According to this formula the middle page is turned between 0° and 180°:

$$\text{pageTurn}(t_{\text{real}}, t_{\text{Maestro}}) = 180 \cdot \left(\frac{t_{\text{real}} - t_{\text{real start}}}{1.5sec} + \frac{t_{\text{Maestro}} - t_{\text{beat before}}}{t_{\text{beat after}}} \right)$$

If the music does not proceed, page turning is nevertheless completed in 1.5 seconds (1.5-seconds-limit). If the user conducts and music proceeds, the new page is reached slightly ahead of time. When page-turning is completed, the MICON assigns new page numbers to the static left and right pages and the middle page vanishes.

Pulse

The time-variable parameters of the score representation and hence the targets of the assembly process are:

- the vertical positions of the pulses and
- the size of the pulses.

The positions of the pulses are determined by extracting the beat times from the beats file and translating these absolute positions along the x-axis to reflect the point in time given by the last time-packet.

The pulses move along form right to left with constant speed and gain size as they approach the center of the screen. The desired maximal magnification m is configurable; so is the distance d_{center} between the center and the point when the magnification begins. Let $distance$ be the current distance between the pulse and the center of the screen. The size of the pulse is then given by:

$$\text{magnification}(distance) = \frac{a}{b + distance}$$

where

$$a = \frac{d_{\text{center}} \cdot m}{m - 1}$$

$$b = \frac{d_{\text{center}}}{m - 1}$$

See figure 6.4 and 6.5 for plots of the pulse function. For values between 0 and d_{center} ($D(\text{magnification}) = [0, d_{\text{center}}]$), the function returns values between 1 and m ($W(magnification) = [1, m]$). If the distance between the pulse and the center of the screen is less than d_{center}, the MICON uses the pulse function to compute the magnification of the circle; otherwise magnification is defined 1. As $\text{magnification}(d_{\text{center}}) = 1$, the pulses seamlessly transition from unaltered to magnified state. At zero distance from the center of the screen the maximal magnification is reached ($\text{magnification}(0) = m$).

Piano Roll

The time-variable parameters of the score representation and hence the targets of the assembly process are:

- the positions of the notes and
- whether a note is highlighted or not.

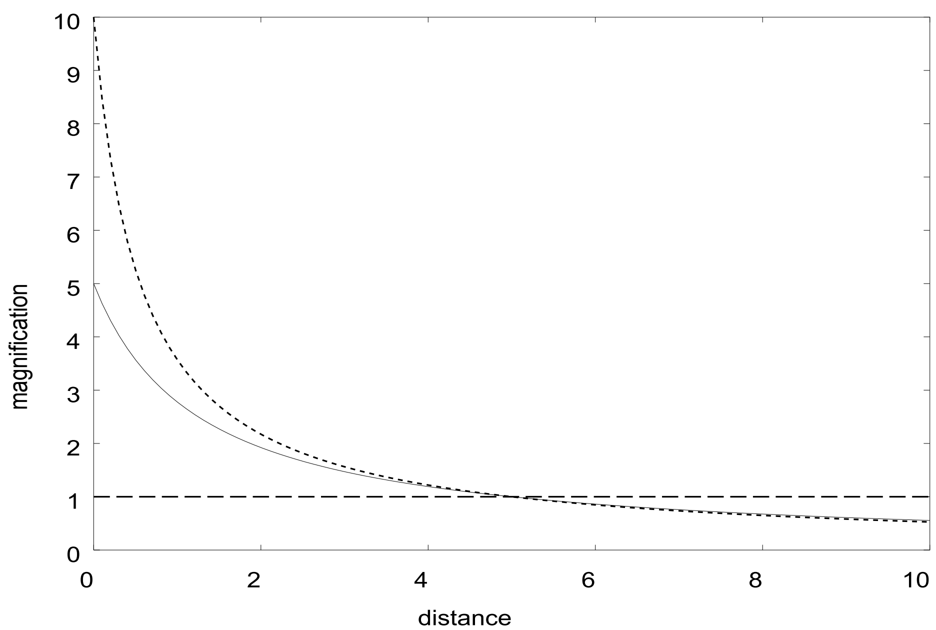


Figure 6.4: A plot of the pulse function with $m = 5$ and $m = 10$

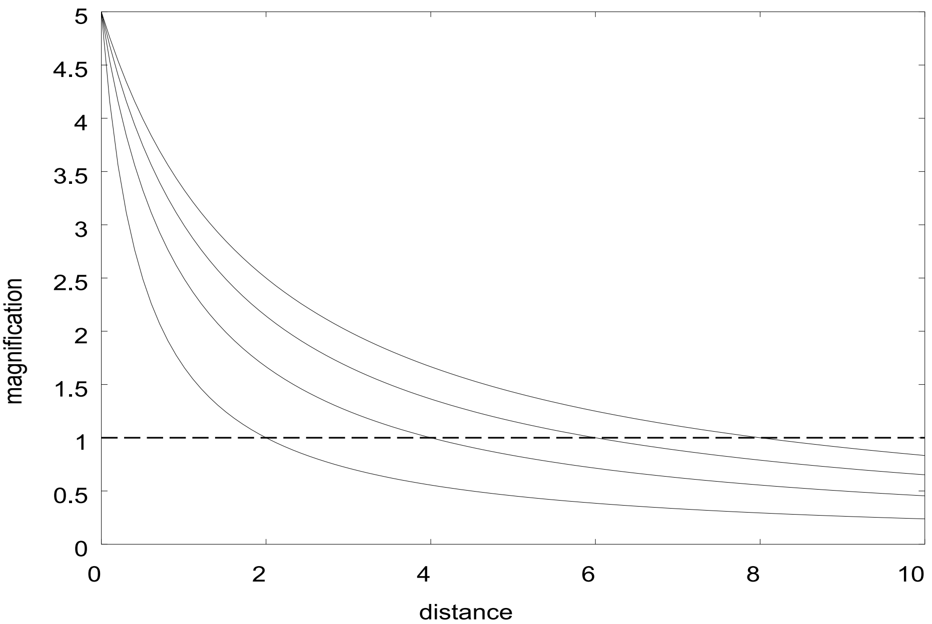


Figure 6.5: A plot of the pulse function with $d_{\text{center}} = 2, 4, 6$ and 8. Note that the functions cross $y = 1$ at $x = 2, 4, 6$ and 8.

For each note the absolute position is first derived from the timing and pitch information that is contained in the MIDI file. The absolute position is then translated along the x-axis to reflect the passed time given by the last received time-packet. A note is highlighted if it is currently played. MICON, therefore, compares the current time with the onset- and offset-times of the note and sets the highlight accordingly.

6.1.4 Score Rendering

MICON uses OpenGL to create realistic looking pages that resemble physical paper more accurately than a normal flat 2D rendering would (figure 6.6).

To create these pages, each page is sliced vertically into 100 equidistant pieces, which are then reconnected, each slice with its two neighbors, to form the curved page. The angles between these neighboring slices were defined beforehand for several key frames of the page-turning animation. Rendering the slices for a given key frame is then straightforward: Starting at the origin, a slice is drawn along the x-axis (the coordinate system is

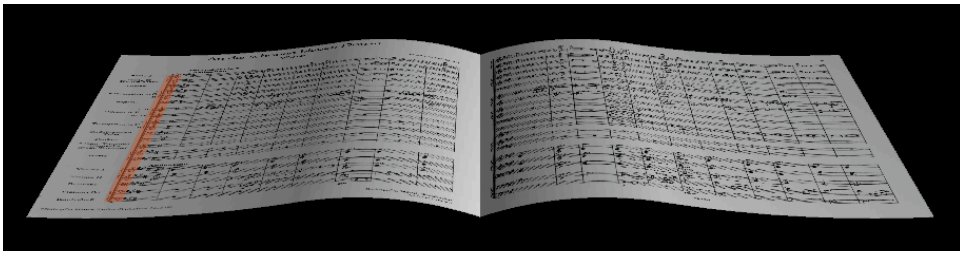


Figure 6.6: Score representation from the side

shown in figure 6.7). The origin is then translated along the x-axis and points to the end of the drawn slice after translation. The origin is now rotated around the y-axis with the predefined angle, and drawing of the next slice begins.

As the page turns, it changes shape: At the beginning (0°) and end (180°) of the page turn, the page has the normal shape. Additional configurations of the 99 angles were specified for a key frame with the page turned 90°. For all other frames during page-turn animation, the 99 inter-slice angles smoothly transition between these key frame configurations using linear interpolation. When turning the page backwards, a modified key frame at 90° has to be taken as the page is bowed into the opposite direction.

The cursor, the jumping ball and the instrument emphasis highlight have to follow the shape of the page. The cursor and the jumping ball touch the page like a tangent. This behavior is accomplished by mimicking the

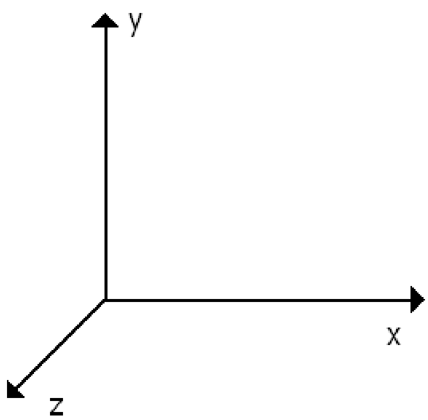


Figure 6.7: Coordinate system

page rendering process. Suppose that the cursor has to be drawn at a horizontal position of 2.5 slice widths. The origin of the coordinate system is translated one slice width along the x-axis. The coordinate system is rotated around the y-axis with the first predefined inter-slice angle. There are still 1.5 slice widths to go, which leads to: translation of 1 slice width, rotation and translation of 0.5 slice width. The coordinate system is now located in the exact middle of the 3rd slice; the x-y-plane of the coordinate system and the 3rd slice are now parallel. The coordinate system is now attached to the 3rd slice. The cursor and the jumping ball are drawn on the two-dimensional x/y-plane.

6.1.5 System Architecture

The functional structure of the MICON is mirrored by the structure of the program. The most classes making up the program can be classified into communication, file-I/O, assembly and rendering. Other classes deal with aspects of minor complexity like setting the preferences and managing the user interface.

The communication subsystem receives and evaluates the UDP-packets sent by Maestro. It manages a State-object, which represents every aspect of the Maestro system the MICON is concerned about: the selected piece, the selected representation, the current time, the instrument emphasis, etc. The state object is than distributed to other parts of the MICON system.

The rendering subsystem consists of a single object: the RenderEngine. The RenderEngine offers high-level drawing primitives for drawing boxes used for the piano roll notation, circles used for the pulse notation and entire score-pages with cursor, instrument emphasis and jumping ball for the score notation. Special classes, the Assemblers, use these drawing-primitives to assemble the scene.

The RenderEngine functions as central point of the entire application: A timer regularly invokes the RenderEngine to draw a frame. RenderEngine then distributes this message to previously registered Assemblers, who in their turn use the drawing functions of the RenderEngine to create the actual rendering (figure 6.8).

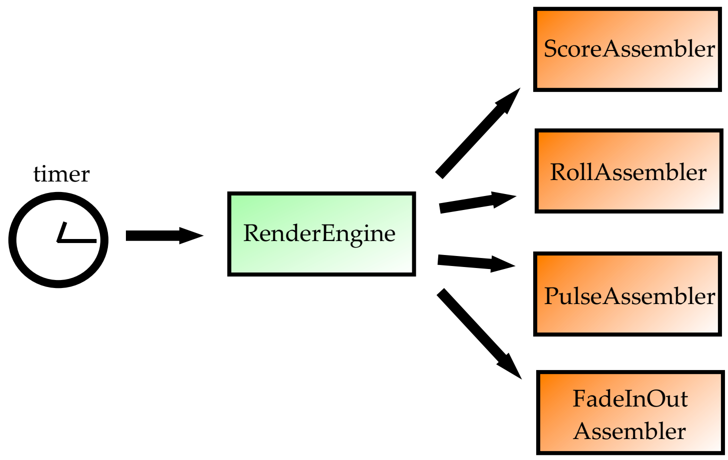


Figure 6.8: MICON draws a frame.

The RenderEngine has also a central position in the distribution of the State-object. The RenderEngine is the mediator between the communication subsystem and the Assemblers. When the state of Maestro changes, be it that the music advances, MICON is informed by an UDP message. The communication subsystem modifies the State-object to reflect that change. The communication subsystem sends the State-object to the RenderEngine, which then distributes the State-object to previously registered Assemblers (figure 6.9).

The two program-flow paths, the path for drawing a frame and the path for distributing the State-object, are independent from each other. The path for distributing the State-object is taken whenever a new data-packet is received by the communication subsystem. A timer triggers the frame-drawing path at a rate of 30fps.

The Assemblers use the drawing-primitives of the RenderEngine to create the scene. There are four different Assemblers:

- the ScoreAssembler,
- the RollAssembler,
- the PulseAssembler and
- the FadeAssembler.

For every music piece, two ScoreAssembler-objects (one for the full score and one for the piano extract), one RollAssembler and one PulseAssembler are created and registered to the RenderEngine.

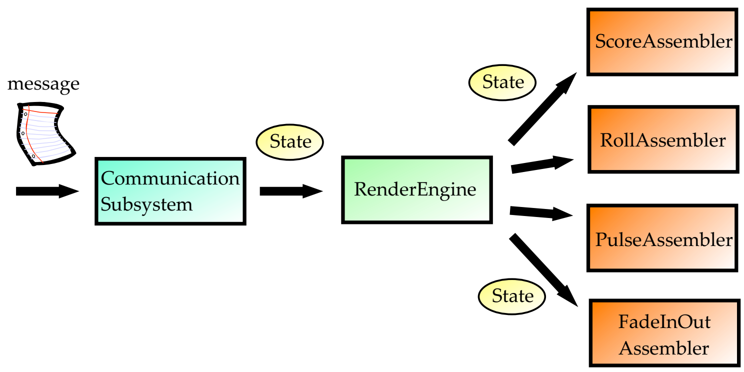


Figure 6.9: Distribution of the State-object

The FadeAssembler is responsible for smoothly fading in or out the representations when the piece begins or ends. The FadeAssembler does not use drawing functions provided by the RenderEngine but issues OpenGL-commands directly. FadeAssembler generates the fading effect by drawing a black rectangle with changing transparency.

Chapter 7

Evaluation

7.1 Testing the MICON with Experts

The Personal Orchestra system was presented in the "Hochschule für Musik und Darstellende Kunst" in Frankfurt to music students. Ten students participated in the test. They were:

- four pianists,
- one composer,
- four other instrumentalists (violinist, cellist, flutist, guitarist).

The students had different experience in conducting: Six of them had learned conducting in a course at the Hochschule or other institutes. Two of them had practical conducting experience (brass orchestra, church choir). Two students had no experience, neither theoretical nor practical, of conducting.

For each student the test took 30 minutes. The student conducted the system at least two times: once with the combined full score and pulse notation, once with the combined piano-roll and pulse notation. If there was still time, the student would conduct a third time with the piano-part score on the note stand. The "Blue Danube" was at that time this the only piece available on the MICON prototype. After each run, a discussion with the student took place. The discussion had three focus points: checking whether the representation were understood, getting feedback about the

MICON and getting feedback about the entire system. The interaction was filmed with a video camera. The experiment set-up can be seen in figure 7.1. The student sees the score animation on a 19″ monitor. The video of the conducted orchestra was not shown. At the time of this user-test the instrument emphasis feature of Maestro did not function properly. Therefore, this aspect of the Maestro-MICON system was left out in the user test.

In the discussion, the students were asked what parameters of the performance they could control. All students recognized that tempo and six, that volume can be controlled.

The students were asked to explain the representations that they had just seen. None of the music students had previously known the piano roll notation, but nine of them were able to figure out meaning of the representation themselves. The student that failed to explain the piano roll had problems to conduct the Maestro system as she did not take the initiative

Figure 7.1: Experiment setup

to control the tempo but waited and reacted to the heard beat. The tempo became extremely slow as both the student and Maestro waited for each other. This "spiral of death" has been a known usability problem of the Personal Orchestra series (Lee et al. [2005]). The piano roll was showed to her a second time with another person conducting, and she understood the notation herself. The pulse notation was understood in general by all students: They all attributed the pulsating circles to the beat. Some students, however, thought that they had created the pulses on the MICON themselves with their conducting gesture.

Several students criticized that the score was too small and that the notes could not be easily read. Two students mentioned that the page occasionally turns unnaturally slow. They both referred to the bars 45–46 of the "Blue Danube" (figure 7.2), and they both conducted very slowly there. The students all liked the appearance of the MICON. Especially, the curved score pages and the page-turn animation were appreciated.

7.1.1 Effects on the Final MICON Version

Three problems became evident during user testing:

(a) First page: bars 42–45

(b) Second page: bars 46–49

Figure 7.2: Bars 42–49 of the "Blue Danube"

- Pulse notation was slightly misunderstood by some students.
- The score was too small and, therefore, difficult to read.
- Page-turning was performed too slow if the user conducted very slowly.

The analysis of the misinterpretation of the pulse notation in terms of Norman's [2002] design principles suggest that the user model is inadequate. In order to communicate the design model to the user the device itself, here the pulse notation, has to be modified. Some students falsely interpreted the pulse notation as a visualization of their input. The MICON-prototype showed only a very narrow section of the pulse notation so that most of the time only two pulses were visible: One that had passed and one that was just being conducted to. Therefore, the wrong hypothesis was consistent with the MICON-prototype's output. Increasing the size of the visible section thus seeing more pulses at once, makes the wrong hypothesis falsifiable as the MICON has no indication of the user's conducting in more than a beat future. This can be easily done by modifying the scale of the pulse notation (section 4.2.4) and no modification of the MICON was necessary.

To make page-turning more natural, the animation should not be tightly coupled to tempo. This inspired the 1.5-seconds-limit (section 6.1.3): Page-turning is completed in at most 1.5 seconds in the final version of the MICON.

For an installation of the MICON in an exhibition, a bigger display should be used. (The 19″ monitor was the biggest available monitor for this test.)

7.2 Presentation at the Hessische Rundfunk

Personal Orchestra 4 was presented during an one-day exhibition at the Hessische Rundfunk (HR), a radio and television broadcaster. A 21″ flat-screen monitor, which would be used for the MICON, was mounted to a music stand and positioned in front of the podium. The system did run the entire day without any technical problems and became a major attraction of this event (figure 7.3). Personal Orchestra 4 was covered in a news reportage of the HR television channel.

Figure 7.3: Personal Orchestra 4 at the HR

Chapter 8

Summary

The MICON is a music stand for the interactive conducting exhibit Maestro. The MICON displays musical material in various formats: full score, piano extract score, piano roll and pulse notation. This enables a broad public to use the MICON, while still allowing expert music score readers to benefit from their abilities.

The pulse notation is a visualization of the metrical structure of the music. An animated pulsation marks the beat and creates a visual rhythm that corresponds to the heard beat, enabling an unexperienced user to understand the notation without further explanation. The pulse notation aids the user in his conducting by providing information about the beat times in advance.

The piano roll, a common representation in MIDI editing software, does not require music reading skills. The currently played notes, which are represented by moving boxes, are highlighted. This creates a visual rhythm that corresponds to the heard rhythm, enabling an unexperienced user to understand the notation without further explanation.

The full score and the piano extract score are intended for user with music reading abilities. The current position is marked in the score. In the full score representation the current instrument emphasis is visualized by coloring the relevant sections of the score. The MICON renders its output via OpenGL and features natural-looking curved pages and realistic-looking page-turning.

The MICON was presented to music students and their feedback was incorporated into the final MICON version.

Appendix A

ScoreMarker—A Tutorial

ScoreMarker is a program that offers a GUI to conveniently create ScoreInfo files. You will learn how to operate the ScoreMarker program in this tutorial.

First, load the score into the ScoreMarker. Select "ScoreMarker→Load JPEG…" from the menu. A dialog pops up letting you provide the filenames of the score pages (figure A.1). First, specify the number of pages you want to load. Enter the number into the pages-field and press "change". The list element at the bottom of the dialog, now, has as many entries as you have specified in the pages-field. One after another select each list item and press the open-button. A file-dialog for loading a JPEG file becomes visible. After selecting the JPEG, the filename shown in the selected list item. Do this for all pages. When you are finished click "OK".

The score is now being displayed on the ScoreMarker (figure A.2). You can magnify and minimize the score with the plus- and minus-buttons on the toolbar. The actual-size-button takes you back to the initial magnification. By pressing the page-up and page-down you can browse the score.

Next to loading the files, you have to define the positions of the note systems. For a piece without repeats or jumps, place one note system on each line. Select "Set Note System" from the toolbar. The note system is then marked with four mouse-clicks in the score (figure A.3). Place the first **click** on one corner of the note system. Now move the mouse to the opposite corner. While moving you see a rectangle spanned by the position of the first click and the current mouse position. **Click** on the opposite corner. You may have, meanwhile, used the scrollbar or changed the magnifica-

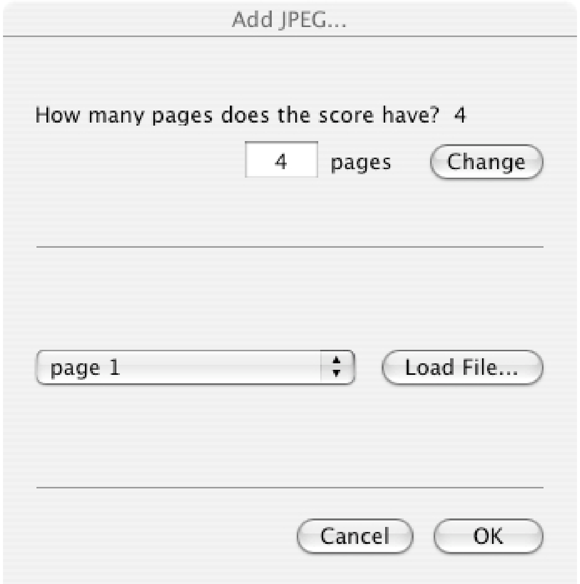


Figure A.1: ScoreMarker: Load JPEG. . .

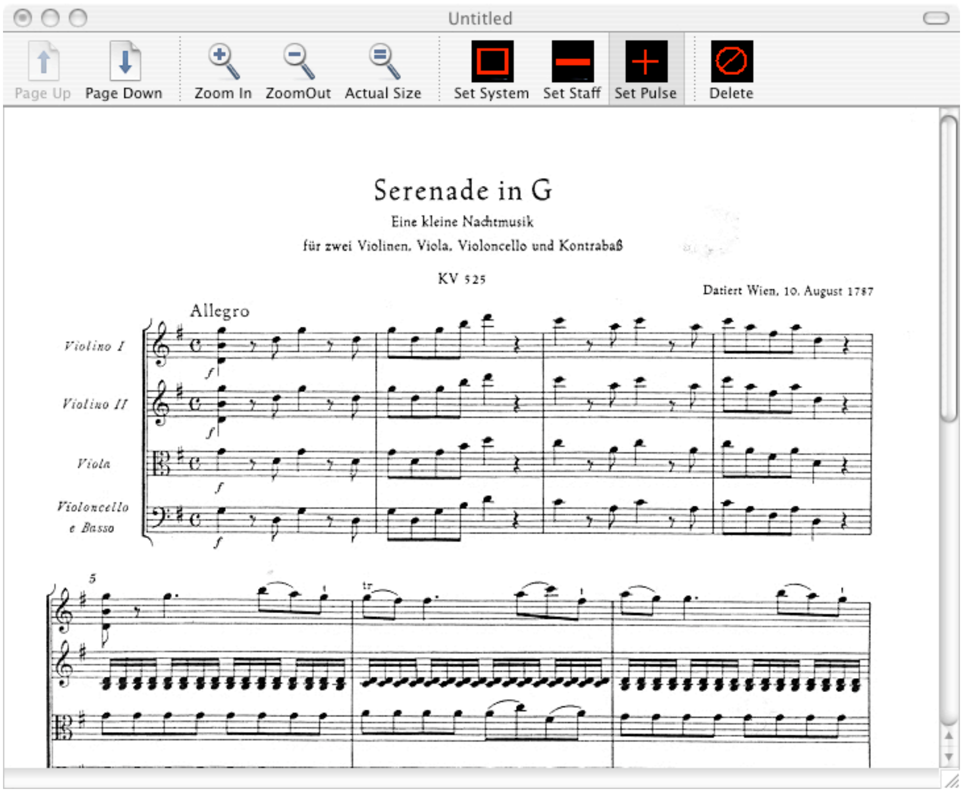


Figure A.2: After loading, the scans are visible on the ScoreMarker.

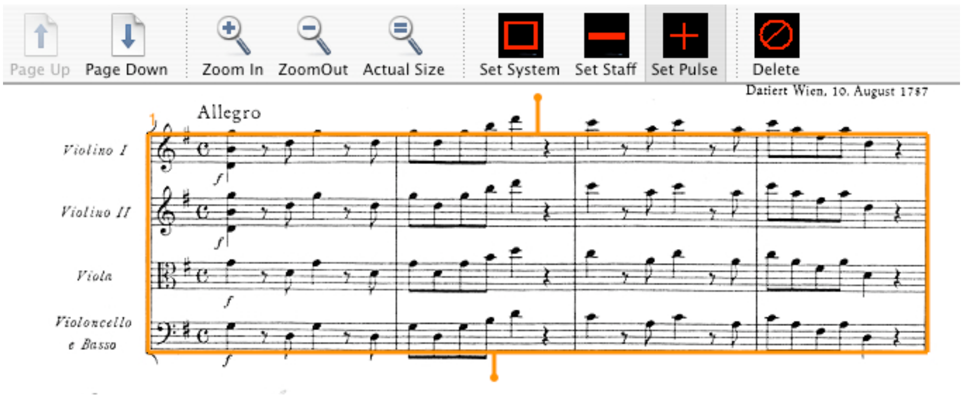


Figure A.3: Define note system

tion. To complete the task, you have to specify an upper and lower leading to the note system. Place a **click** above and a **click** below the rectangle (or below and then above if you wish). You have now completely described the note system to ScoreMarker. On the top left corner of the note system there is a small number: the sequence number of the note system (figure A.4). During score animation the cursor visits the note systems in the order of the sequence numbers. For pieces without repeats or other jumps, ScoreMarker automatically assigns the correct sequence numbers. We will see how to manually assign sequence numbers at the end of this tutorial.

Once you have decided on your beat concept, the process of placing the beats in the score is fairly simple. Select "Set Pulse" from the toolbar and click in the score. A blue cross now marks the beat (figure A.5). Be aware that every beat has to be contained in a note system; so that if you click outside the area of any note system, i.e. outside the rectangle left/right side upper/lower leading, your input will be ignored, and you will hear a sound informing you of that.

Now that the note systems and the beats are specified, it is time to provide

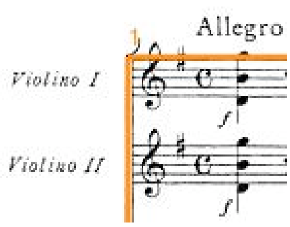


Figure A.4: Sequence number

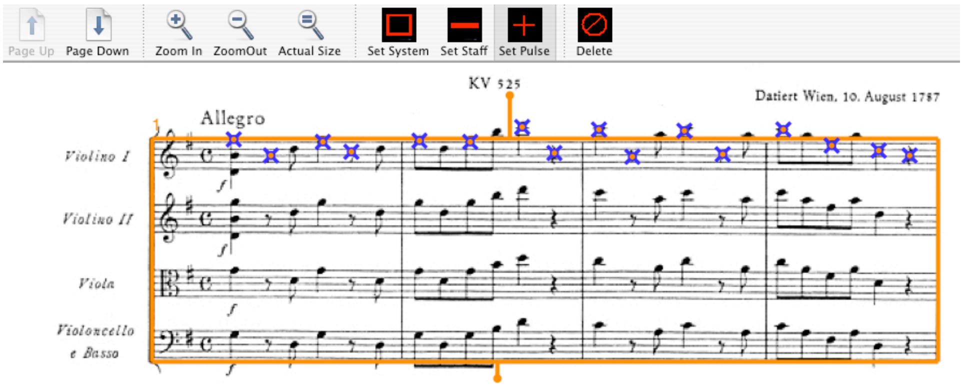


Figure A.5: ScoreMarker: Note system with beats

the ScoreMarker with information about the positions of the instrument groups in the score. This is done on a per-line basis. You inform ScoreMarker about the position of each line and to which instrument group this line belongs. You determine the vertical position and extent of the line with two mouse clicks. The horizontal position and extent are derived from the embracing note system. Select "Set Staff" from the toolbar. Click at the upper or lower boundary of a the staff belonging to the instrument group you want to mark. Move the mouse toward the opposite boundary of the line. While moving the mouse the selection is marked with a translucent yellow color. Reaching the opposite boundary, click again. The marking of the line becomes permanent (figure A.6). A dialog pops up and asks you for the instrument-id of the line (figure A.7). Enter the appropriate integer number in the id-field and click "OK" to complete the definition of the instrument group. The instrument-ids have to be agreed upon by Maestro

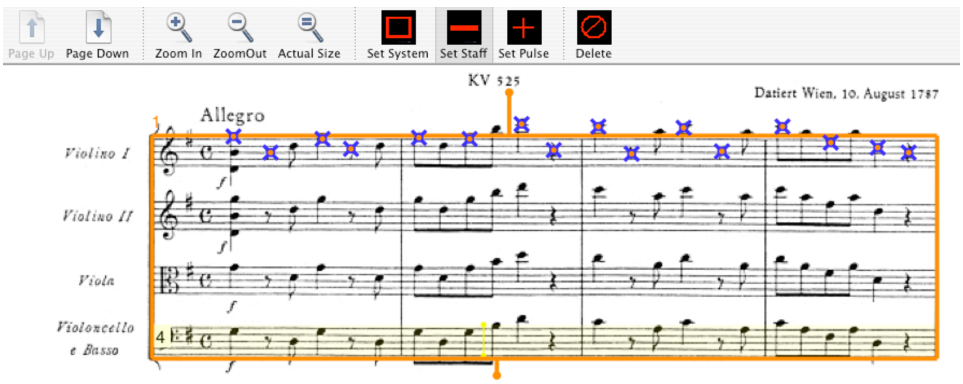


Figure A.6: ScoreMarker with instrument group

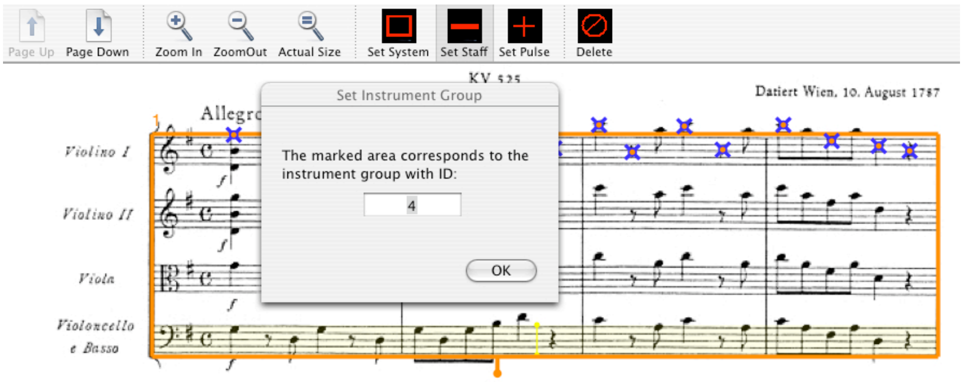


Figure A.7: Instrument-id dialog

and MICON.

If you want to delete a note system, a beat or an instrument group choose "Delete" from the toolbar. The mouse cursor changes shape informing you that ScoreMarker is in delete-mode. Move the mouse to the item you want to delete. When the mouse is sufficiently near, the item is highlighted with red color (figure A.8). Click to remove the highlighted item. When you remove a note system, the contained beats and instrument groups are not automatically deleted. If you place a new note system around these items, they are automatically assigned to it so that you can easily change the shape of a note system (a typical last-minute change) without having to redefine beats and instrument groups. If you, however, want to entirely erase a note system you have to erase every contained beat and instrument line by hand or you will have an inconsistent ScoreInfo file.

You have now learned how to annotate a basic piece with ScoreMarker. "ScoreMarker→Save" brings up a file dialog and lets you save your work.

As promised, we will now, at the end of this tutorial, see how to annotate a music piece with repeats or jumps. The order in which the cursor visits the note system during score animation can be manually defined. Within a note system the cursor moves along from left to right entirely. Only after completing a note system the cursor may move on to another note system or start over with the same note one. The note systems are undividable (atomicity). So: Wherever the cursor jumps out during score animation, there is a right end of a note system in ScoreMarker. Wherever the cursor jumps in during score animation, there is a left end of a note system in ScoreMarker. So you might want to define several note systems on a single line

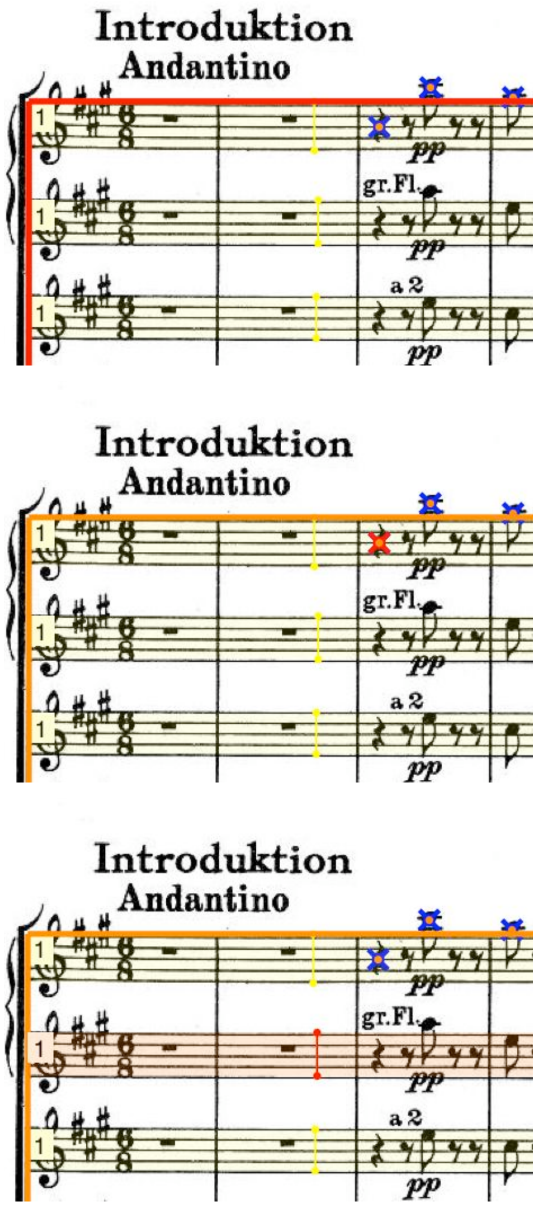


Figure A.8: Higlighted for deletion: the note system, the beat and the instrument group (figure A.9). If you define a note system with the left side closely aligned to the right side of another note system, ScoreMarker recognizes that these two note systems belong to the same line and automatically aligns them.

Having placed all note systems, we will define the order in which the note systems will be visited during score animation. Select “Set Sequence” form the menu. Click the note systems in the sequence you want to see them been visited during score animation. The same note system can be visited several times (figure A.10). When you erase a note system or add a new one, ScoreMarker considers the manually assigned sequence numbers to be invalid and automatically assigns new sequence numbers. This will break the manually assigned values. You can, of course, redo the ordering at any time by selecting “Set Sequence” from the menu.

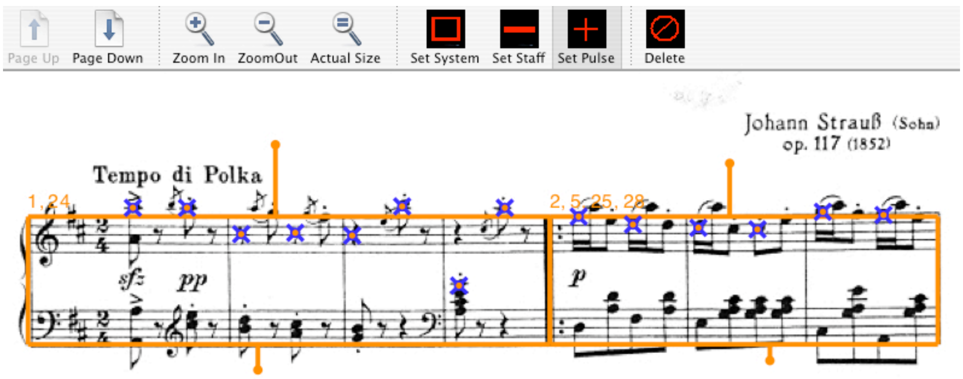


Figure A.9: Note systems of the beginning of “Annenpolka”

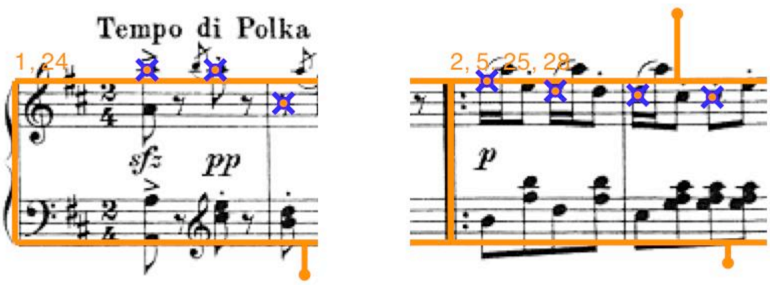
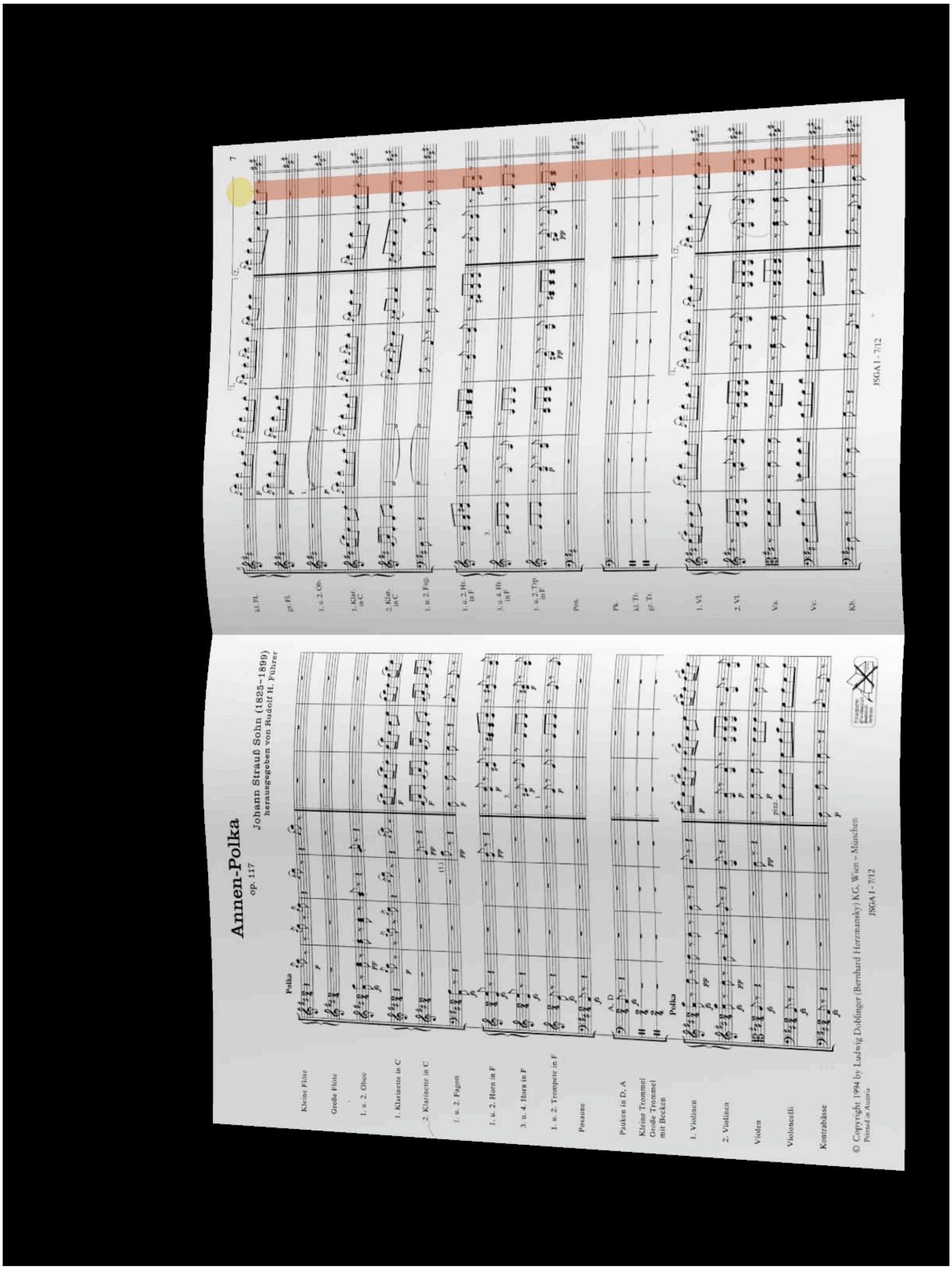
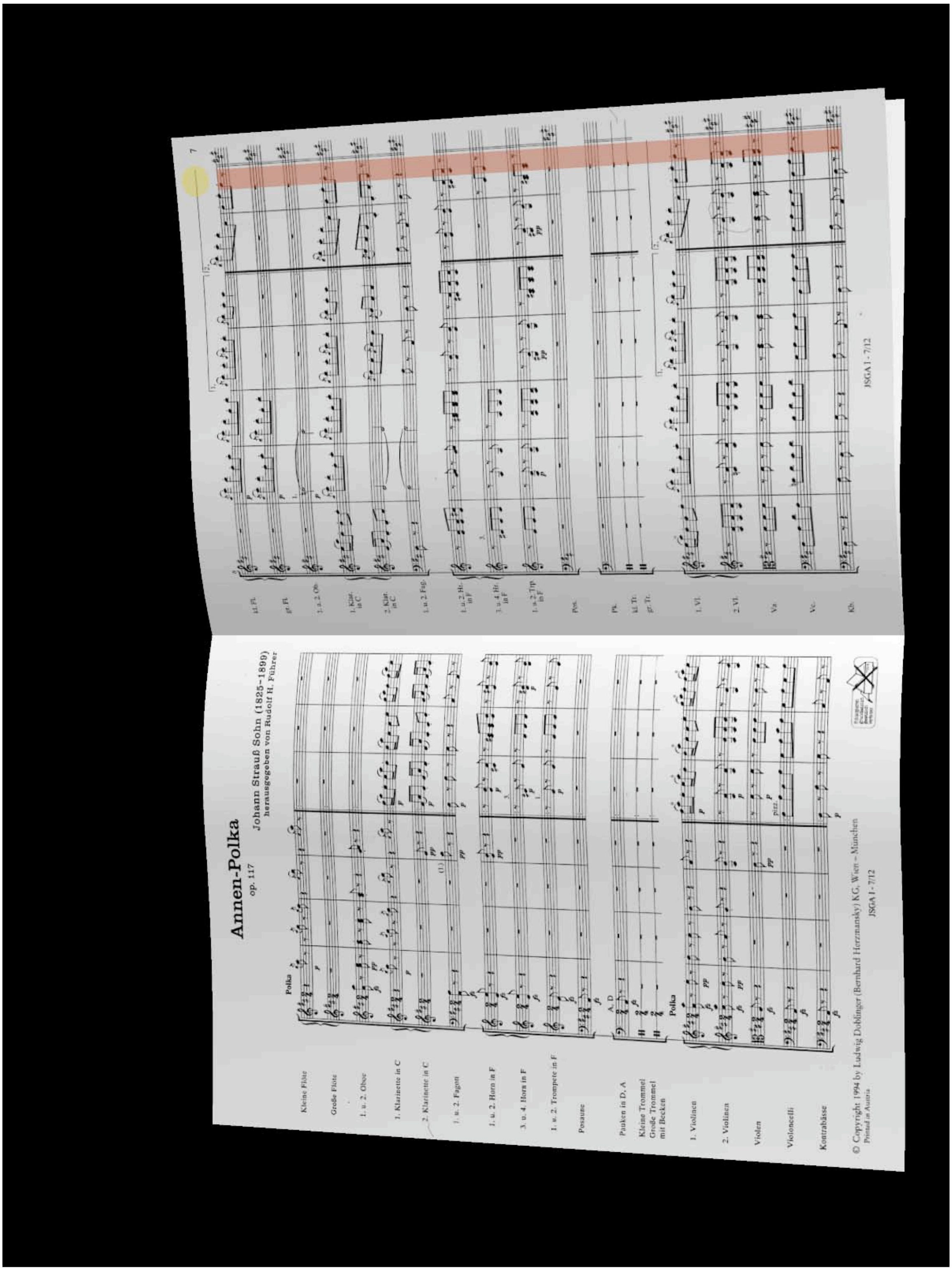


Figure A.10: Sequence numbers of the first and second of Annenpolka’s note systems: The first note system is to be visited in the 1st and 24th place, the second note system in the 2nd, 5th, 25th and 28th place during score animation.

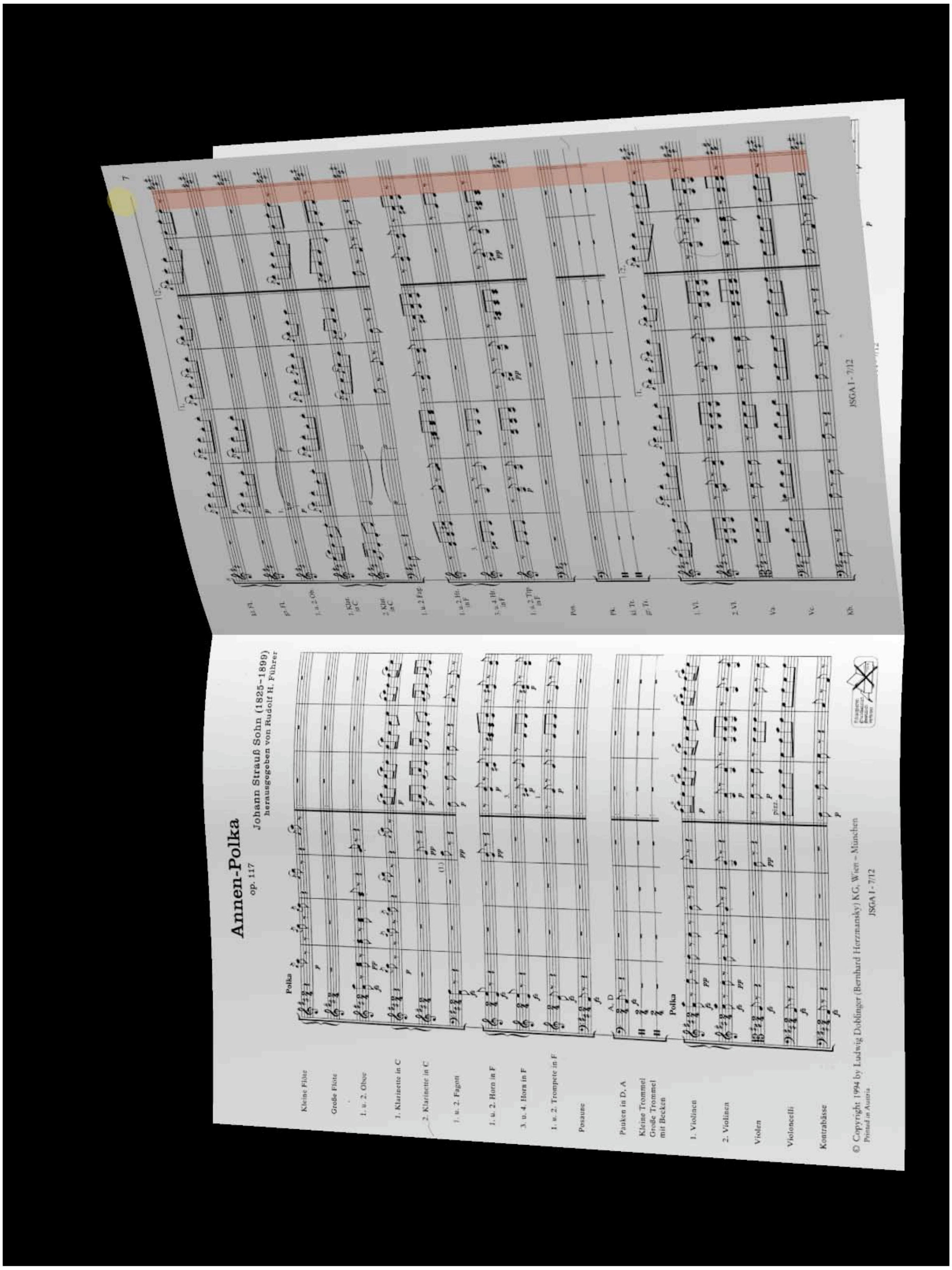
Appendix B

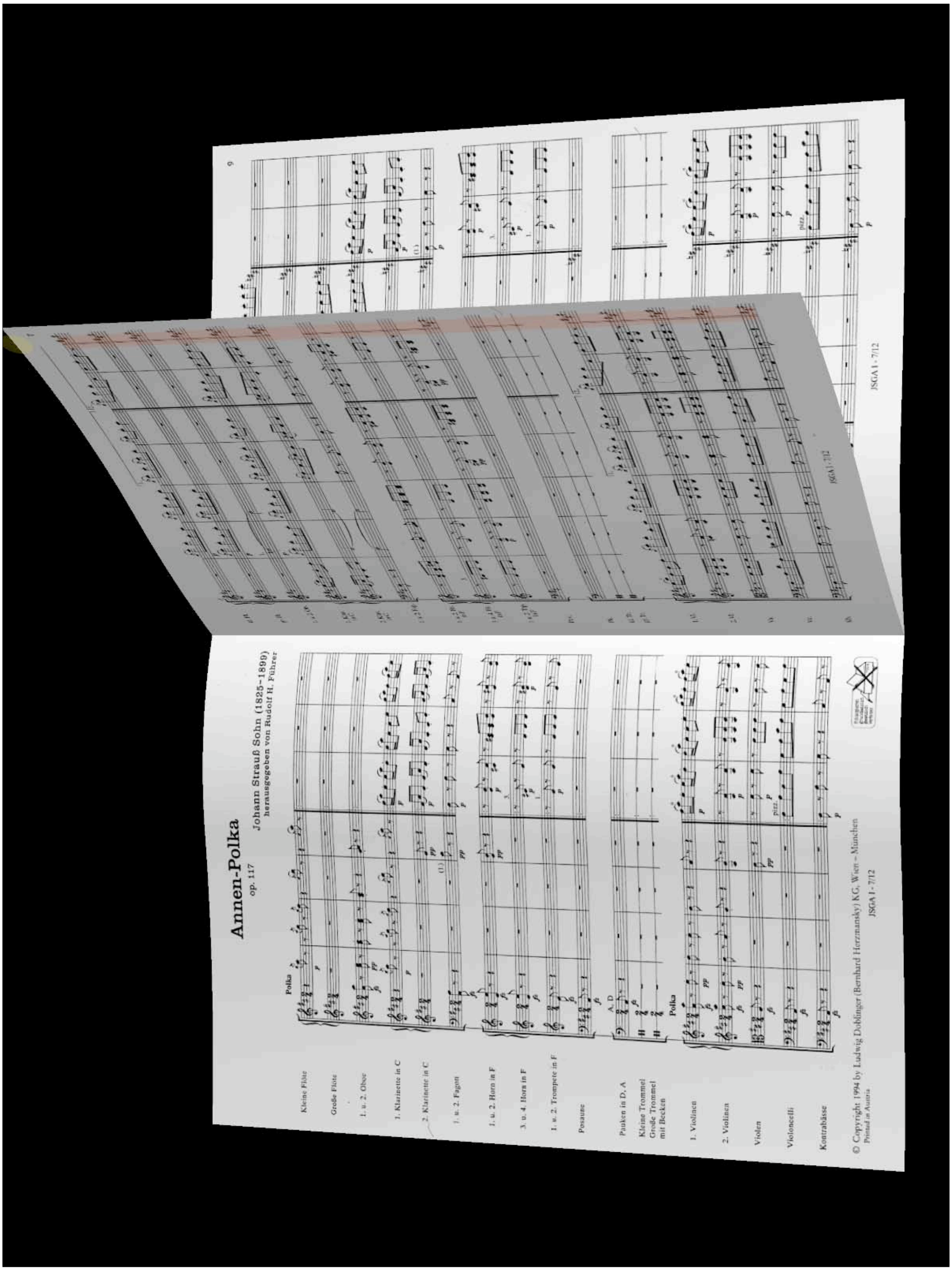
Screenshots

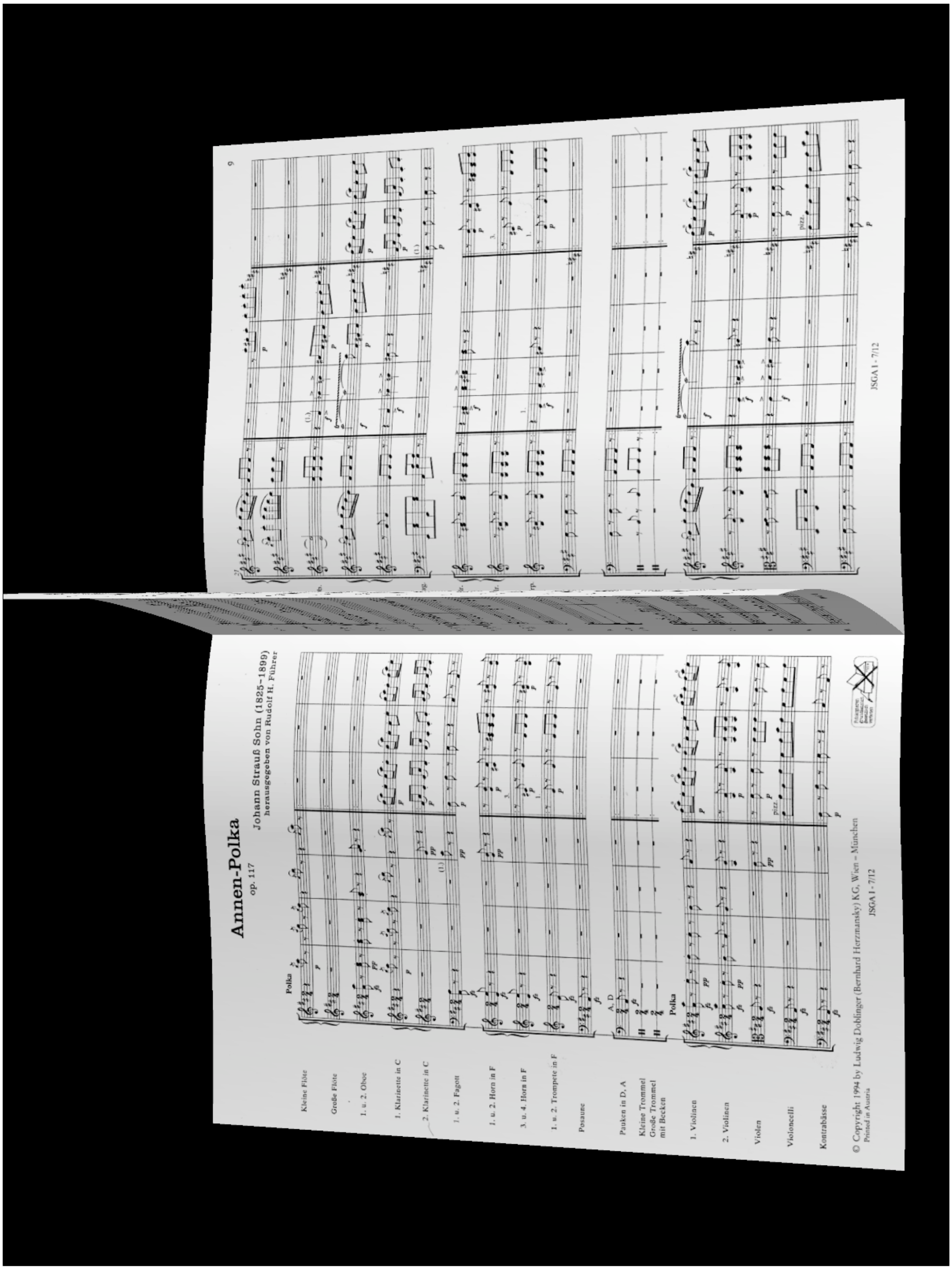

Annen-Polka
op. 117
Johann Strauß Sohn (1825–1899)
herausgegeben von Rudolf H. Führer
Polka
Kleine Flöte
Große Flöte
1. u. 2. Oboe
1. Klarinette in C
2. Klarinette in C
1. u. 2. Fagott
1. u. 2. Horn in F
3. u. 4. Horn in F
1. u. 2. Trompete in F
Posaune
Pauken in D, A
Kleine Trommel
Große Trommel mit Becken
1. Violinen
2. Violinen
Violen
Violoncelli
Kontrabässe
© Copyright 1994 by Ludwig Doblinger (Bernhard Herzmansky) KG, Wien – München
JSGA I - 7/12

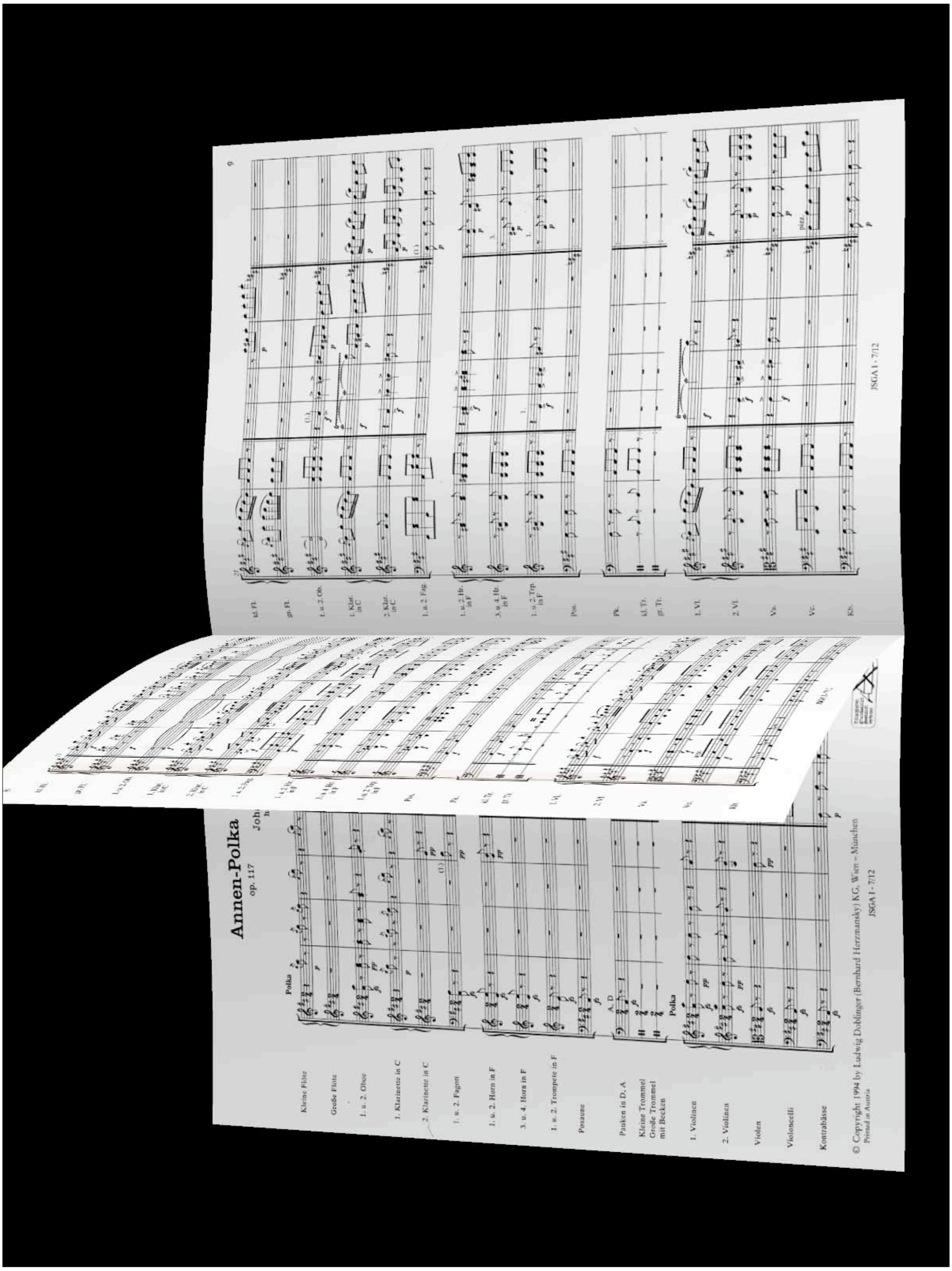


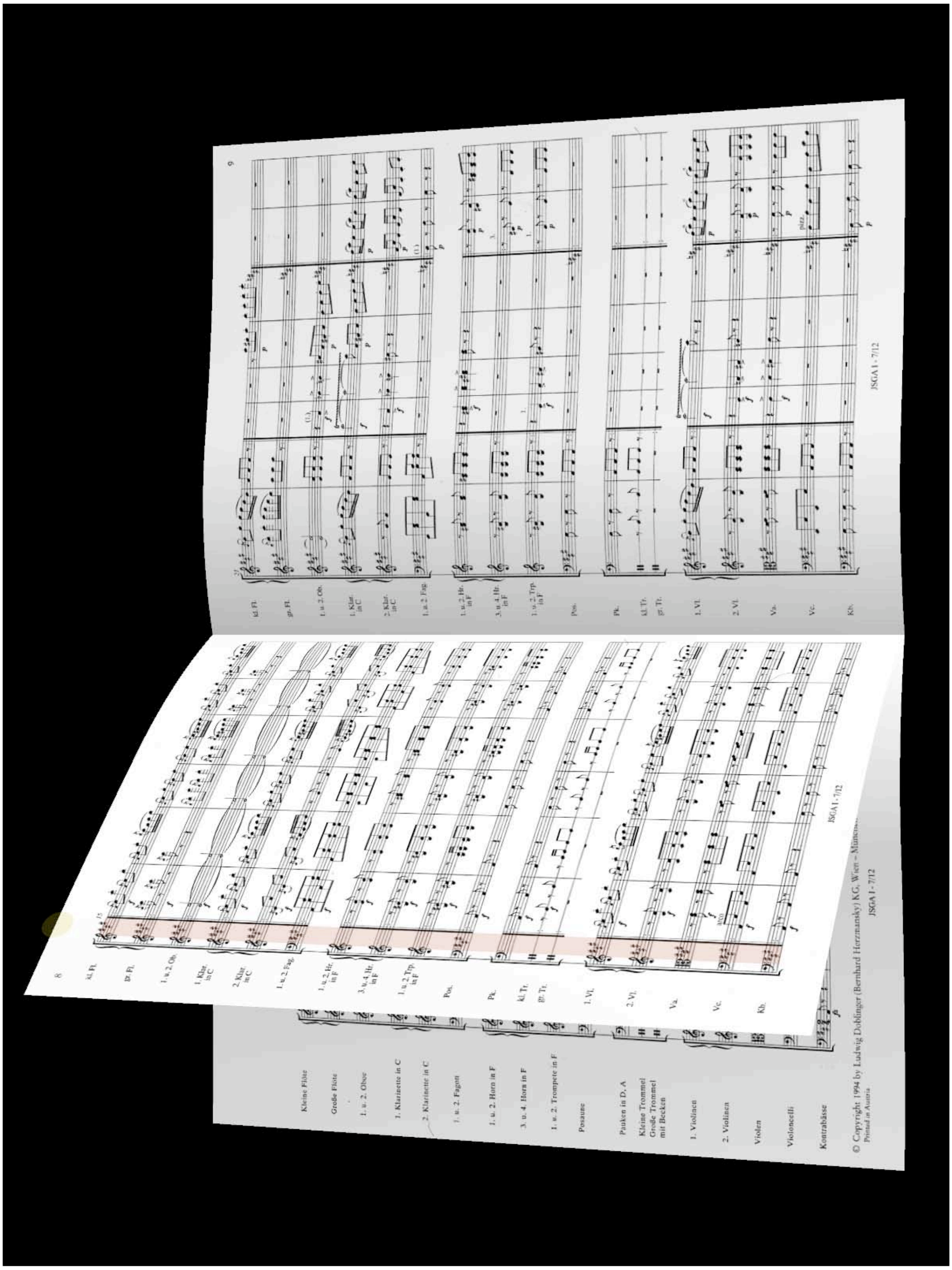
Annen-Polka
op. 117
Johann Strauß Sohn (1825–1899)
herausgegeben von Rudolf H. Führer

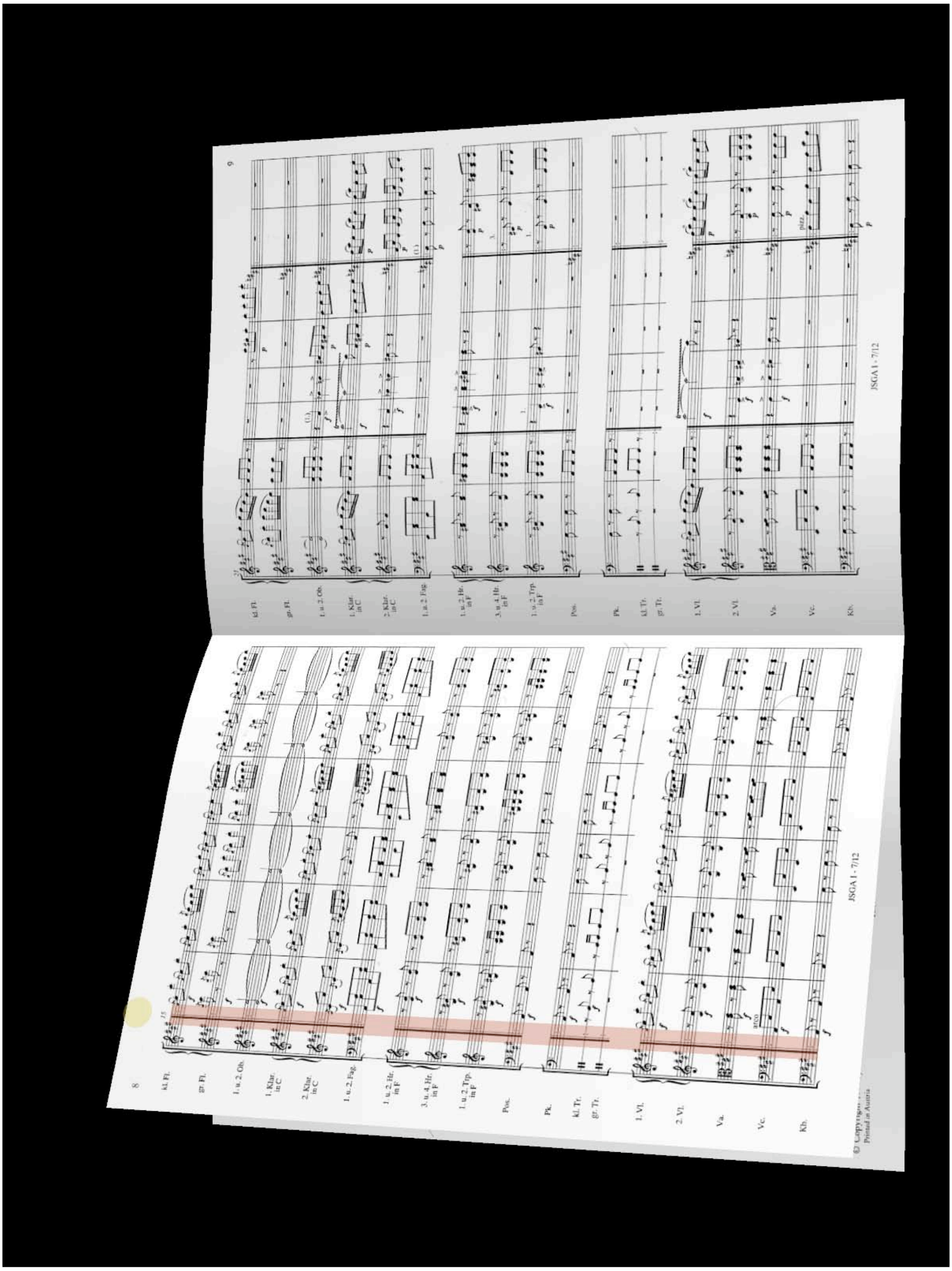

Annen-Polka
op. 117
Johann Strauß Sohn (1825–1899)
herausgegeben von Rudolf H. Führer

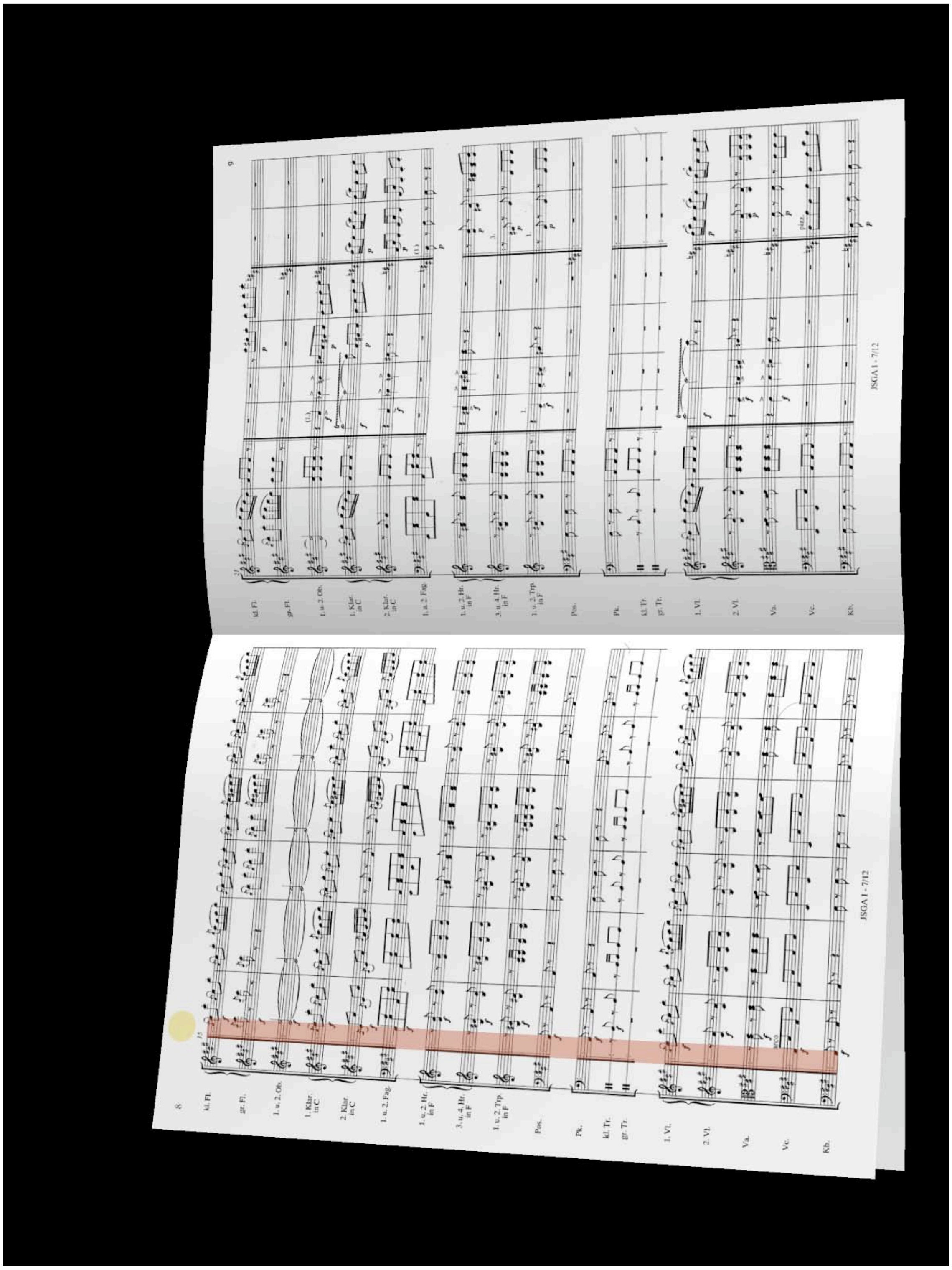

Annen-Polka
op. 117
Johann Strauß Sohn (1825–1899)
herausgegeben von Rudolf H. Führer

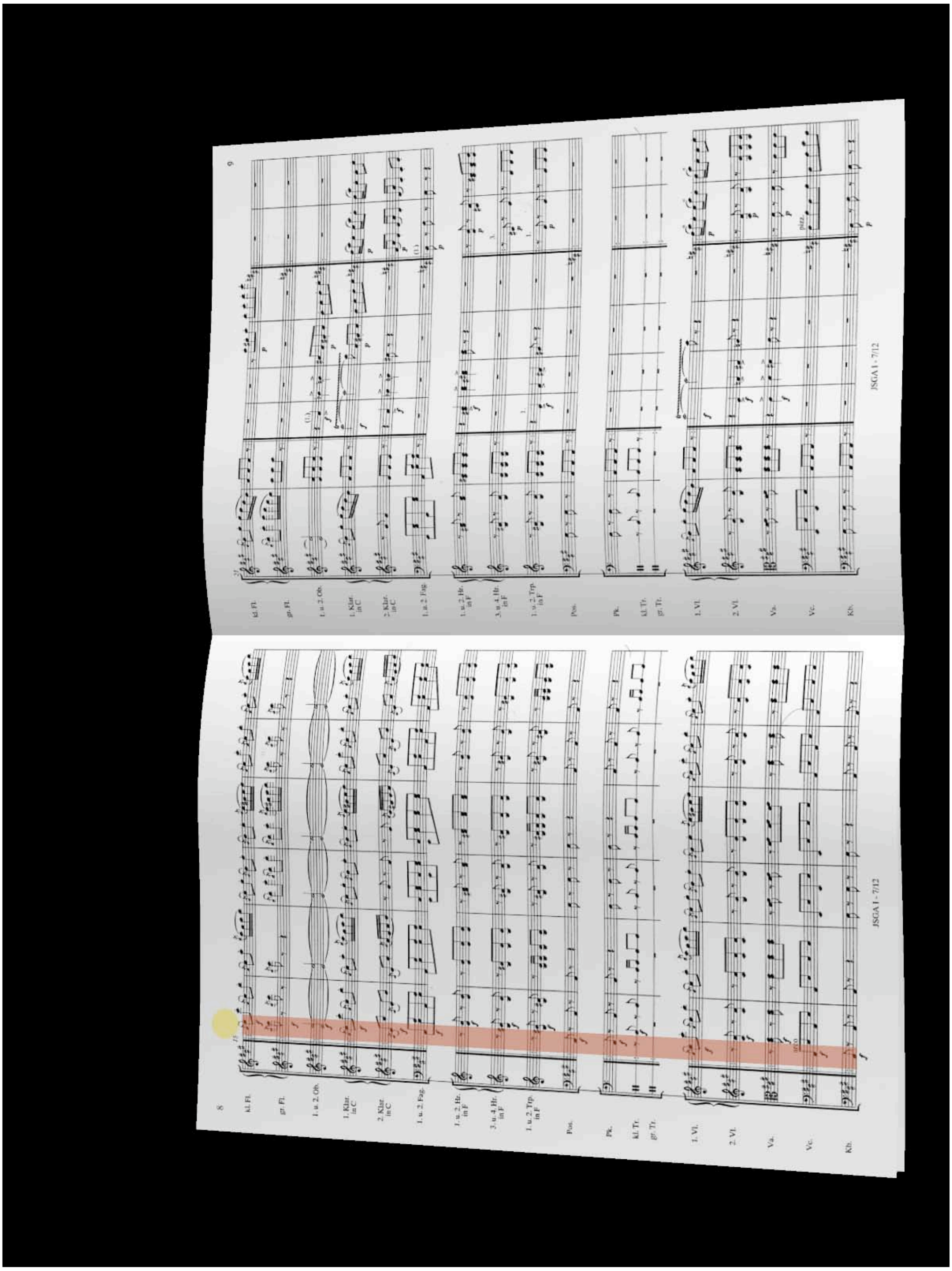

Annen-Polka
op. 117
Johann Strauß Sohn (1825–1899)
herausgegeben von Rudolf H. Führer

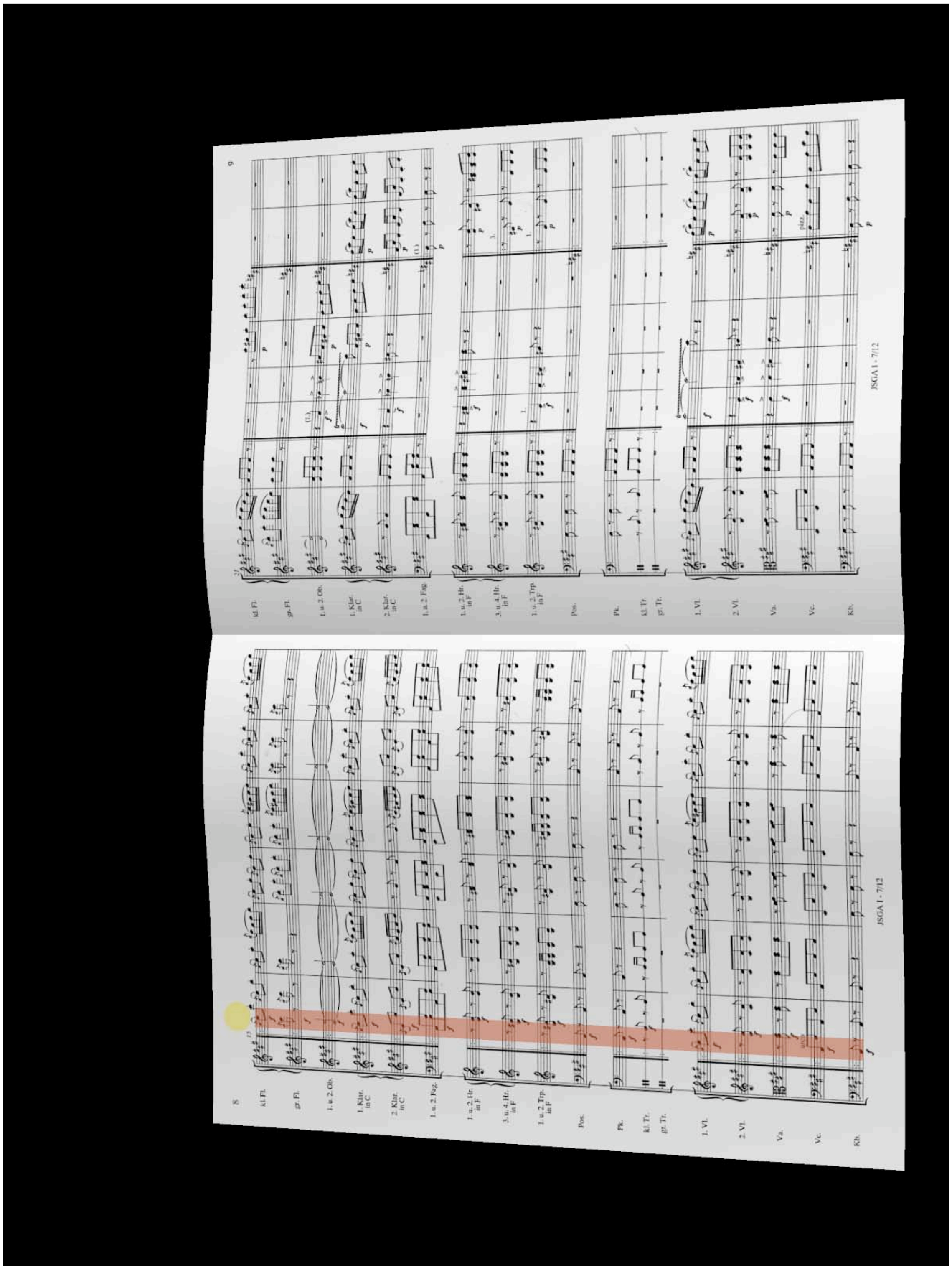

Annen-Polka
op. 117
Polka
Kleine Flöte
Große Flöte
1. u. 2. Oboe
1. Klarinette in C
2. Klarinette in C
1. u. 2. Fagott
1. u. 2. Horn in F
3. u. 4. Horn in F
1. u. 2. Trompete in F
Posaune
Pauken in D, A
Kleine Trommel
Große Trommel
mit Becken
1. Violinen
2. Violinen
Violen
Violoncelli
Kontrabässe
© Copyright 1994 by Ludwig Doblinger (Bernhard Herzmansky) KG, Wien – München
JSGA I - 7/12

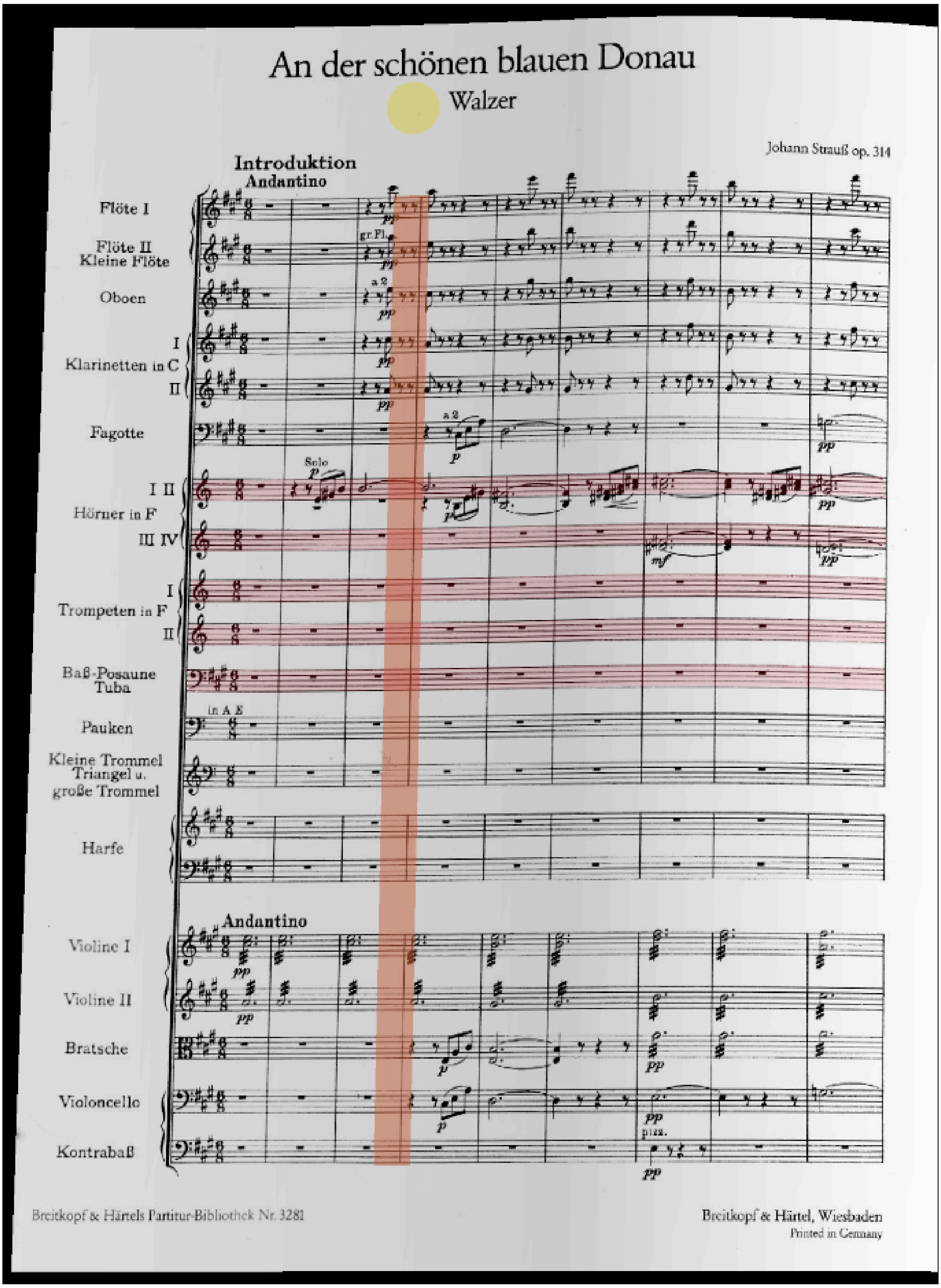


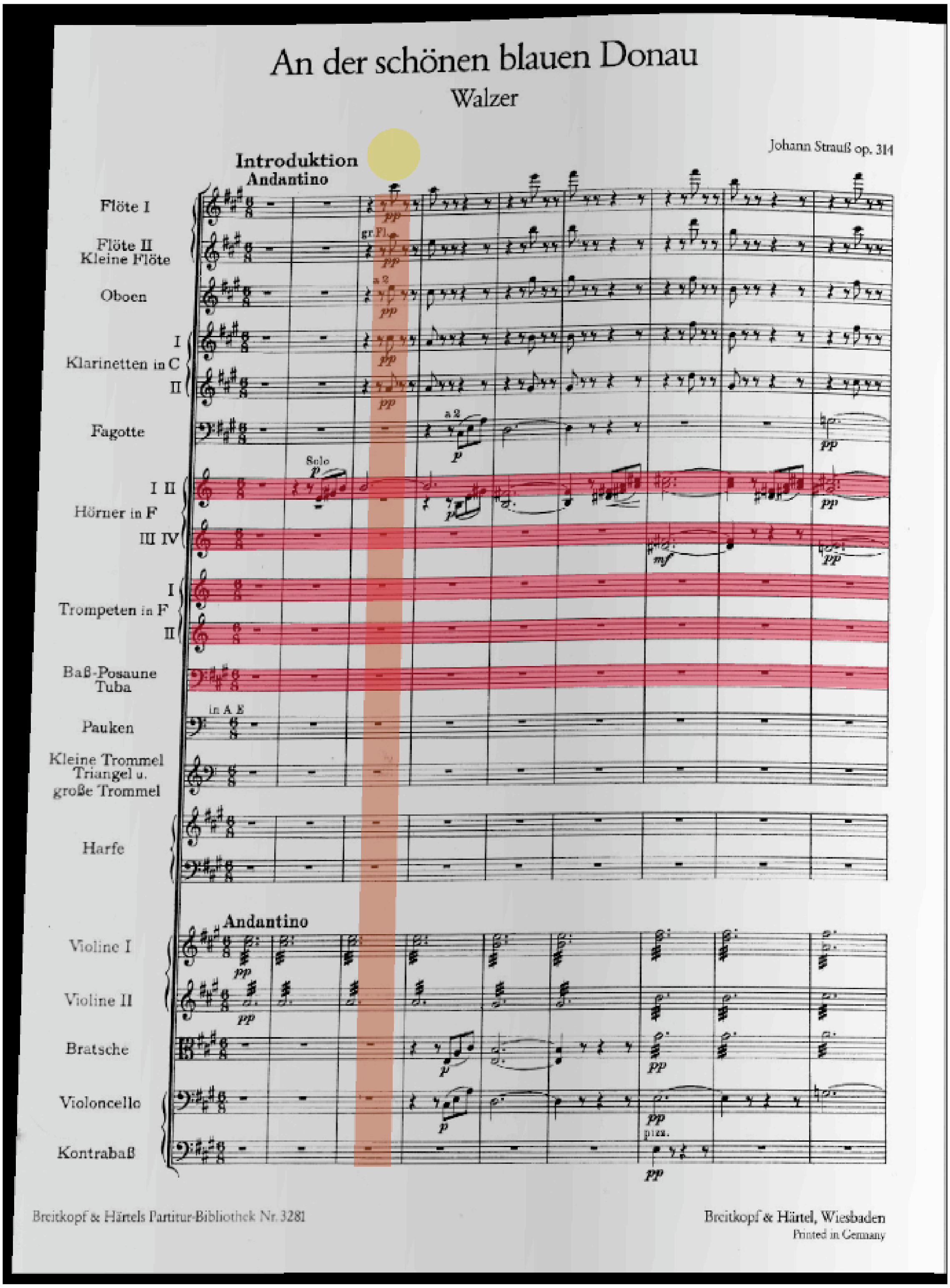




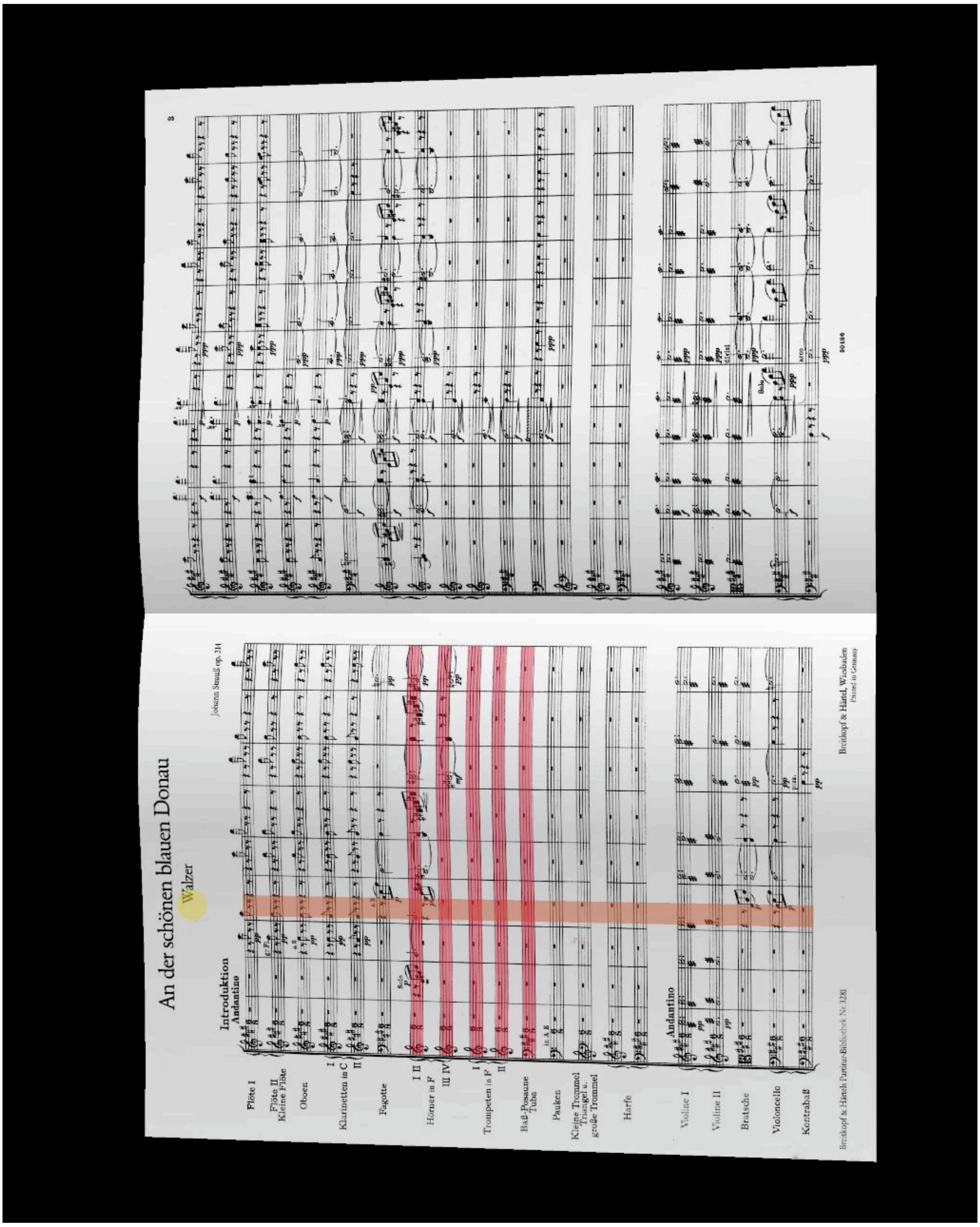

kl. Fl.
gr. Fl.
1. u. 2. Ob.
1. Klar. in C
2. Klar. in C
1. u. 2. Fag.
1. u. 2. Hr. in F
3. u. 4. Hr. in F
1. u. 2. Trp. in F
Pos.
Pk.
kl. Tr.
gr. Tr.
1. Vl.
2. Vl.
Va.
Vc.
Kb.
JSGA I - 7/12
JSGA I - 7/12

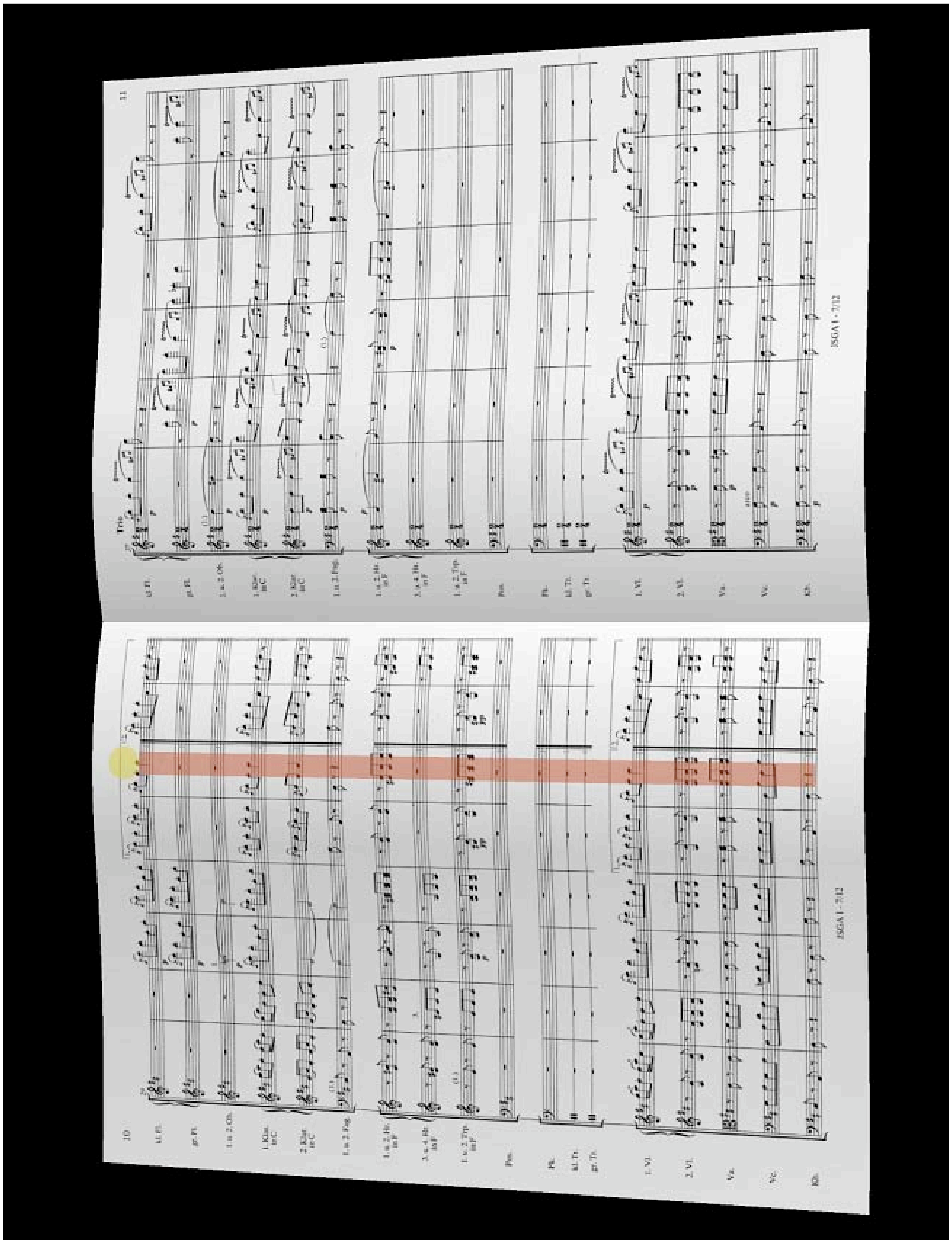


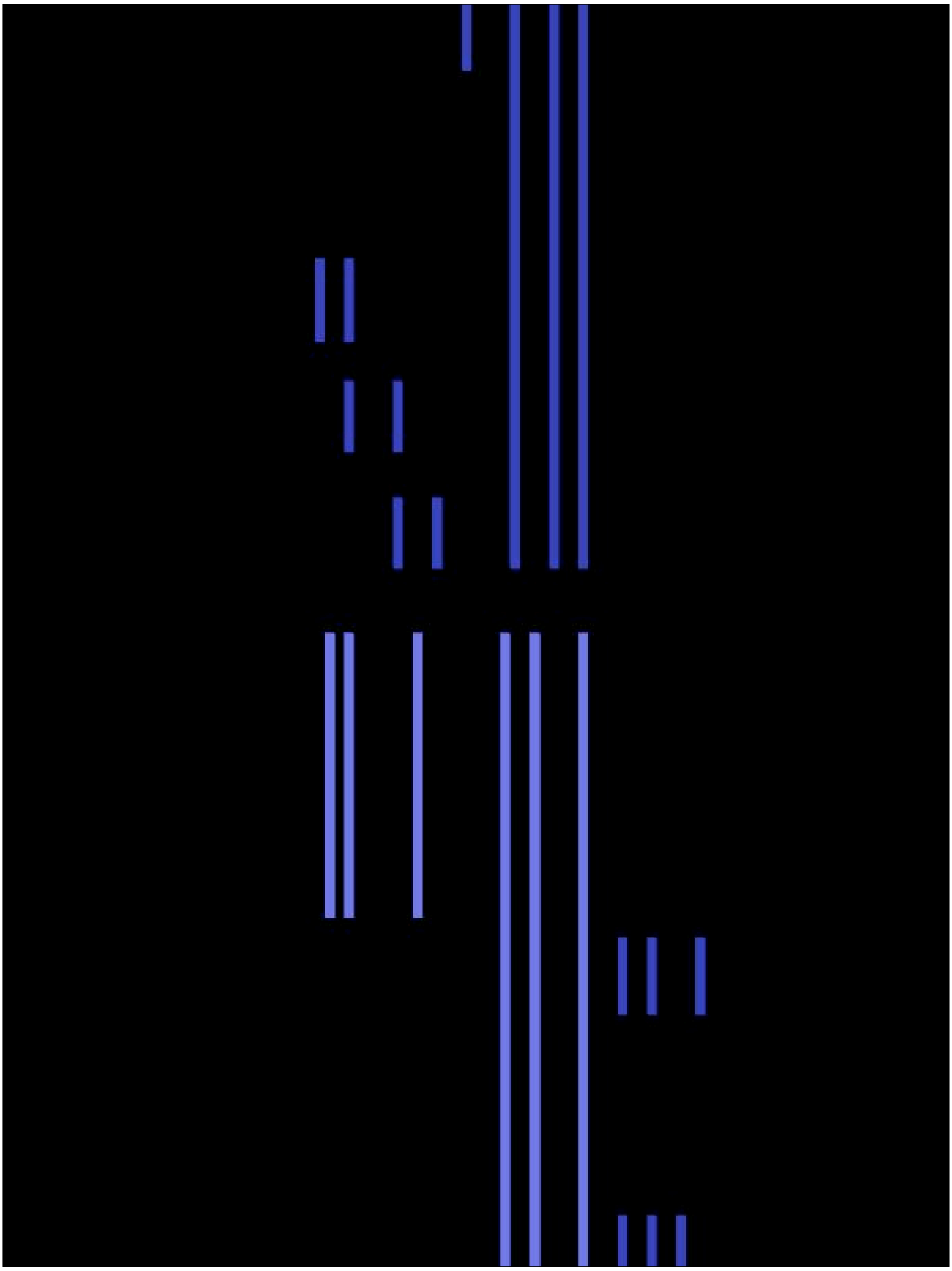

An der schönen blauen Donau
Walzer
Johann Strauß op. 314
Introduktion
Andantino
Flöte I
Flöte II
Kleine Flöte
Oboen
Klarinetten in C
Fagotte
Hörner in F
Trompeten in F
Baß-Posaune
Tuba
Pauken
in A E
Kleine Trommel
Triangel u.
große Trommel
Harfe
Andantino
Violine I
Violine II
Bratsche
Violoncello
Kontrabaß
Solo
gr. Fl.
a 2
pizz.
Breitkopf & Härtels Partitur-Bibliothek Nr. 3281
Breitkopf & Härtel, Wiesbaden
Printed in Germany



An der schönen blauen Donau
Walzer
Johann Strauß op. 314
Introduktion
Andantino
Flöte I
Flöte II
Kleine Flöte
Oboen
Klarinetten in C
Fagotte
Hörner in F
Trompeten in F
Baß-Posaune
Tuba
Pauken
Kleine Trommel
Triangel u.
große Trommel
Harfe
Andantino
Violine I
Violine II
Bratsche
Violoncello
Kontrabaß
Breitkopf & Härtels Partitur-Bibliothek Nr. 3281
Breitkopf & Härtel, Wiesbaden
Printed in Germany


An der schönen blauen Donau
Walzer
Introduktion
Andantino
Johann Strauß op. 314
Flöte I
Flöte II
Kleine Flöte
Oboen
Klarinetten in C
Fagotte
Hörner in F
Trompeten in F
Baß-Posaune
Tuba
Pauken
Kleine Trommel
Triangel u.
große Trommel
Harfe
Violine I
Violine II
Bratsche
Violoncello
Kontrabaß
Breitkopf & Härtel, Wiesbaden





Bibliography

P. Bellini, P. Nesi, and M. B. Spinu. Cooperative visual manipulation of music notation. *ACM Transactions on Computer-Human Interaction*, 9(3): 194–237, 2002.

Graziano Bertini and Paolo Carosi. Light baton: A system for conducting computer music performance. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, 1992.

Bob Boie, Max Mathews, and Andy Schloss. The radio drum as a synthesizer controller. In *Proceedings of the International Computer Music Conference*, pages 42–45. International Computer Music Association, 1989.

Jan Borchers. Worldbeat: Designing a baton-based interface for an interactive music exhibit. In *Proceedings of the ACM CHI International Conference on Human Factors in Computing Systems*, pages 131–138. ACM, March 1997.

Jan Borchers. *A Pattern Approach to Interaction Design*. John Wiley & Sons, Baffins Lane, Chichester, West Sussex, PO19 1UD, England, March 2001.

Jan Borchers, Eric Lee, Wolfgang Samminger, and Max Mühlhäuser. Personal orchestra: A real-time audio/video system for interactive conducting. *ACM Multimedia Systems Journal Special Issue on Multimedia Software Engineering*, 9(5):458–465, March 2004.

Jan Borchers, Aristotelis Hadjakos, and Max Mühlhäuser. Micon: A music stand for interactive conducting. In *NIME 2006 International Conference on New Interfaces for Musical Expression*, Paris, France, June 2006.

Jan O. Borchers, Wolfgang Samminger, and Max Mühlhäuser. Engineering a realistic real-time conducting system for the audio/video rendering of a real orchestra. In *IEEE MSE 2002 Fourth International Symposium on Multimedia Software Engineering*, December 2002.

Richard Boulanger and Max Mathews. The 1997 mathews radio-baton and improvisation modes. In *Proceedings of the International Computer Music Conference*, pages 395–398. International Computer Music Association, 1997.

Bennett Brecht and Guy E. Garnett. Conductor follower. In *Proceedings of the International Computer Music Conference*, pages 185–186. International Computer Music Association, 1995.

Guy E. Garnett, Fernando Malvar-Ruiz, and Fred Stoltzfus. Virtual conducting practice environment. In *Proceedings of the International Computer Music Conference*, pages 371–374. International Computer Music Association, 1999.

Guy E. Garnett, Mangesh Jonnalagadda, Ivan Elezovic, Timothy Johnson, and Kevin Small. Technological advances for conducting a virtual ensemble. In *Proceedings of the International Computer Music Conference*, pages 167–169. International Computer Music Association, 2001.

Christopher Graefe, Derek Wahila, Justin Maguire, and Orya Dasna. muse: a digital music stand for symphony musicians. *ACM interactions*, 3(3): 26–35, 1996.

Ingo Gruell. conga: A conducting gesture analysis framework. Master's thesis, University of Ulm, Ulm, Germany, April 2005.

Tommi Ilmonen. Tracking conductor of an orchestra using artificial neural networks. Master's thesis, Helsinki University of Technology, Esboo, Finland, April 1999.

Tommi Ilmonen. Tracking conductor of an orchestra using artificial neural networks. In *STeP'98, Human and Artificial Information Processing*. Finnish Conference on Artificial Intelligence, 1998.

Tommi Ilmonen. Diva - digital interactive virtual acoustics. `http://www.tml.tkk.fi/Research/DIVA/`, March 2003.

Tommi Ilmonen and Tapio Takala. Conductor following with artificial neural networks. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, 1999.

Thorsten Karrer. Phavorit: A phase vocoder for real-time interactive time-stretching. Master's thesis, RWTH Aachen University, Aachen, Germany, November 2005.

David Keane and Peter Gross. The midi baton. In *Proceedings of the International Computer Music Conference*, pages 151–154. International Computer Music Association, 1989.

David Keane, G. Smecca, and Kevin Wood. The midi baton ii. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, 1990.

Paul Kolesnik. Conducting gesture recognition, analysis and performance system. Master's thesis, Mc Gill University, Montreal, June 2004.

Eric Lee, Teresa Marrin Nakra, and Jan Borchers. You're the conductor: A realistic interactive conducting system for children. In *NIME 2004 International Conference on New Interfaces for Musical Expression*, pages 68–73, Hamamatsu, Japan, June 2004.

Eric Lee, Marius Wolf, and Jan Borchers. Improving orchestral conducting systems in public spaces: Examining the temporal characteristics and conceptual models of conducting gestures. In *Proceedings of the CHI 2005 Conference on Human Factors in Computing Systems*, pages 731–740, Portland, Oregon, USA, April 2005. ACM.

Eric Lee, Ingo Grüll, Henning Kiel, and Jan Borchers. conga: A framework for adaptive conducting gesture analysis. In *Proceedings of the New Interfaces for Musical Expression*, 2006a.

Eric Lee, Thorsten Karrer, and Jan Borchers. Toward a framework for interactive systems to conduct digital audio and video streams. *Computer Music Journal*, 30(1):21–36, Spring 2006b. The video for this article appears in the Computer Music Journal Sound and Video Anthology 29(4), 2005.

Eric Lee, Henning Kiel, Saskia Dedenbach, Ingo Gruell, Thorsten Karrer, Marius Wolf, and Jan Borchers. isymphony: An adaptive interactive orchestral conducting system for conducting digital audio and video streams. In *Extended Abstracts of CHI 2006 Conference on Human Factors in Computing Systems*, Montreal, Canada, April 2006c. ACM Press.

Michael Lee, Guy Garnett, and David Wessel. An adaptive conductor follower. In *Proceedings of the International Computer Music Conference*, pages 454–455. International Computer Music Association, 1992.

Teresa Marrin and Joseph Paradiso. The digital baton: a versatile performance instrument. In *Proceedings of the International Computer Music Conference*, pages 313–316. International Computer Music Association, 1997.

Teresa Marrin and Rosalind Picard. The conductor's jacket: A device for recording expressive musical gestures. In *Proceedings of the International Computer Music Conference*, pages 215–219. International Computer Music Association, 1998.

H. Morita, S. Ohteru, and S. Hashimoto. Computer music system which follows a human conductor. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, 1989.

Teresa Marrin Nakra. *Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture*. PhD thesis, Massachusetts Institute of Technology, February 2000.

Donald A. Norman. *The Design of Everyday Things*. Basic Books, 2002.

Wolfgang F. Samminger. Personal orchestra: Interaktive steuerung synchroner audio- und videoströme. Master's thesis, Johannes Kepler Universität Linz, Linz, Germany, September 2002.

Satoshi Usa and Yasunori Mochida. A multi-modal conducting simulator. In *Proceedings of the International Computer Music Conference*, pages 25–32. International Computer Music Association, 1998.