University of North Georgia

Mike Cottrel College of Business

Department of Computer Science & Information Systems

CSCI 4950 – Senior Project

Bryan Hutto

**Assignment 3**

**(Literature Review)**

1. Related Work

If a program is to scroll in time with the music on a page, it must keep pace with the reader’s focus in regards to the tempo and time signature of the piece, a speed which can be calculated algorithmically from those two inputs. However, the sight reader will be looking a distance ahead of the notes being currently played. This distance is designated the eye-hand span (EHS) by Rosemann, Altenmüller, and Fahle [9], who attempt to quantify this value inductively given different variables, such as piece complexity and the subject’s familiarity with it. They also give a comprehensive review of other researchers’ findings in regards to the EHS. A scrolling function would need initially wait an amount of time consistent with this value before matching the pace of the piece. A sight reader’s focus will frequently travel back from the eye-hand span’s locale to the notes currently being played, as described by Lechmann and Kopiez [8], as a kind of error checking, as such it is important that the previous measures, going back one EHS, remain in view.

Wallace, Savage, and Cockburn [7] experimented with what scrolling speeds, in terms of most comfortable and maximum tolerable, still allowed subjects to subjectively identify important features in a document, and what effect different magnifications had on these speeds, a relationship which was inversely proportional. While this experiment involved mixed text and image documents, the negative magnification effect is especially important to us, as it is one theoretical advantage of scrolling versus page turning. Quinn, Cockburn, Casiez, Roussel, and Gutwin [10] attempted to reverse engineer complex scrolling transfer functions, which transform human input into on-screen scrolling, with various drivers and devices. An understanding of how scrolling is implemented, especially with respect to user settings, is useful for our purposes, but especially helpful are the references to the Windows messages that control scrolling for focused applications.

Maestoso, Taele, Barreto, and Hammond [6], is an educational tool which, among many other features, can recognize hand-written measure breaks with a high degree of accuracy, using line tests. It is the goal of this project to design a similar line test and use it to scroll appropriately through a sheet music document, with a view towards facilitating comfortable sight reading for the user.

1. ***Audio Analysis as a Control Knob for Social Sensing***

Abstract: “While humans can act as effective sensors, human input is subject to a high degree of error and highly dependent on the context. Furthermore, extracting the signal from the noise for social sensing is a difficult challenge. One approach to improving the accuracy of social sensing is to use physical sensors as a control knob for social sensing algorithms. In this paper, we present an architecture for using audio sensors as a way to control an algorithm used for social sensing of interesting events. We present various use cases where the architecture is applicable, and go into the details of one specific use case, namely using crowd behavior in a golf-course to identify and control social media feeds related to the course.”

**ACM Ref:**

Dinesh Verma, Bong Jun Ko, Shiqiang Wang, Xiping Wang, and Graham Bent. 2017. Audio Analysis as a Control Knob for Social Sensing. In Proceedings of the 2nd International Workshop on Social Sensing (SocialSens'17). ACM, New York, NY, USA, 75-80. DOI: <https://doi.org/10.1145/3055601.3055616>

1. ***Advanced Auditory Menus: Design And Evaluation***

***of Auditory Scroll Bars***

Abstract: “Auditory menus have the potential to make devices that use visual menus accessible to a wide range of users. Visually impaired users could especially benefit from the auditory feedback received during menu navigation. However, auditory menus are a relatively new concept, and there are very few guidelines that describe how to design them. This paper details how visual menu concepts may be applied to auditory menus in order to help develop design guidelines. Specifically, this set of studies examined possible ways of designing an auditory scrollbar for an auditory menu. The following different auditory scrollbar designs were evaluated: single-tone, double-tone, alphabetical grouping, and proportional grouping. Three different evaluations were conducted to determine the best design. The first two evaluations were conducted with sighted users, and the last evaluation was conducted with visually impaired users. The results suggest that pitch polarity does not matter, and proportional grouping is the best of the auditory scrollbar designs evaluated here.”

**ACM Ref:**

Pavani Yalla and Bruce N. Walker. 2008. Advanced auditory menus: design and evaluation of auditory scroll bars. In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility (Assets '08). ACM, New York, NY, USA, 105-112. DOI=http://dx.doi.org/10.1145/1414471.1414492

1. ***Compressed Image File Formats JPEG, PNG, GIF, XBM, BMP***

First Paragraph of Preface: “The purpose of this book is to instruct the reader on how to write software that can read and write files using various 2-D image formats. I wanted to write a

book that explains the most frequently used file formats with enough depth for

the reader to implement them, as opposed to one that covered many different formats

at a high level or one that avoided the more difficult image formats. As a

result, I chose to cover the image file formats that are associated with Web

browsers. Those covered in this book (BMP, XBM, JPEG, GIF, and PNG) represent

the vast majority of image files that can be found on the Internet. They

employ a wide range of encoding techniques and range in implementation difficulty

from simple to very complex.

**ACM Ref:**

John Miano. 1999. Compressed Image File Formats: Jpeg, Png, Gif, Xbm, BMP. ACM Press/Addison-Wesley Publ. Co., New York, NY, USA.

1. ***Lightweight measures for timbral similarity of musical audio***

Abstract: “Timbral similarity measures basedon Mel-Frequency Cepstral Coefficients have been widely reported as the basis for a possible general music similarity function, which would have wide application to searching, browsing and recommendation. Many of the reported methods, however, have computational requirements that make them impractical for searching realistic collections using current hardware. We compare lightweight measures that appear to perform equally well, and introduce a simplification that reduces memory requirements and execution time by a further order of magnitude. This yields a similarity measure that will scale easily to large commercial collections. We give comparative results over two contrasting music collections, one of which has been widely studied, allowing direct comparison with previous work.”

**ACM Ref:**

Mark Levy and Mark Sandler. 2006. Lightweight measures for timbral similarity of musical audio. In Proceedings of the 1st ACM workshop on Audio and music computing multimedia (AMCMM '06). ACM, New York, NY, USA, 27-36. DOI=http://dx.doi.org/10.1145/1178723.1178728

1. ***Local Attributes for OAG-Based Attribute Evaluators***

Abstract: “Attribute grammars permit us to give meaning to a string in a context-free language. This meaning is expressed using attributes attached to symbols (terminals and non-terminals) of the underlying context-free grammar. In this paper we extend the concept of attributes, and allow to attach attributes to each production rule too. These rule-attached attributes are often called local attributes because of their scoping restriction. They can be defined and referenced only by attribution rules of the production rule to which they are attached. These attributes provide us the best place to store temporary values which can be used later to define other attributes in the same rule. First part of the paper describes the theoretical background to incorporate local attributes with OAG0based attribute evaluator generates such as [2,3]. Then it introduces the formal definition of local attributes along with a typical example where they are useful. “

**ACM Ref:**

Masayoshi Ishikawa. 1990. Local Attributes for OAG-Based Attribute Evaluators. In *Computer Science Technical Reports*. Retrieved June 17, 2017 from <http://scholar.colorado.edu/csci_techreports/454>

1. ***Maestoso: An Intelligent Educational Sketching Tool for Learning Music Theory***

Abstract: “Learning music theory not only has practical benefits for musicians to write, perform, understand, and express music better, but also for both non-musicians to improve critical thinking, math analytical skills, and music appreciation. However, current external tools applicable for learning music theory through writing when human instruction is unavailable are either limited in feedback, lacking a written modality, or assuming already strong familiarity of music theory concepts. In this paper, we describe Maestoso, an educational tool for novice learners to learn music theory through sketching practice of quizzed music structures. Maestoso first automatically recognizes students’ sketched input of quizzed concepts, then relies on existing sketch and gesture recognition techniques to automatically recognize the input, and finally generates instructor-emulated feedback. From our evaluations, we demonstrate that Maestoso performs reasonably well on recognizing music structure elements and that novice students can comfortably grasp introductory music theory in a single session.”

**ACM Ref:**

Paul Taele, Laura Barreto, and Tracy Hammond. 2015. Maestoso: An Intelligent Educational Sketching Tool for Learning Music Theory. *In Proceedings of the Twenty-Seventh Conference on Innovative Applications of Artificial Intelligence*. Association for the Advancement of Artificial In Conference on Innovative the Network ([www.aaai.org](http://www.aaai.org)). Retrieved June 17, 2017 from <http://faculty.cs.tamu.edu/hammond/publications/pdf/2015TaeleIAAIMaestoso.pdf>

1. ***Rapid Visual Flow: How Fast Is Too Fast?***

Abstract: “It is becoming increasingly common for user interfaces to use zooming visual effects that automatically adapt to user actions. The MacOs X 'dock' icon panel, for instance, uses a fisheye distortion to assist users in targeting items. Another example is 'speed-dependent automatic zooming', which has been shown to improve scrolling by automatically varying zoom level with scroll speed-when scrolling fast the document is zoomed out, but when scrolling slowly the document is fully zoomed in. When implementing automatic zooming interfaces, designers must calibrate the behaviour of their zooming systems so that the visual effects allow rapid navigation without stressing the human visual system. At present, these calibrations are derived from trial and error. This paper describes an attempt to determine metrics of visual flow to answer the question "how fast is too fast"? Our main focus is on automatic zooming in document scrolling tasks. We performed an experiment to measure participants' preferred and maximum-tolerable scrolling speeds at two different magnifications. We found that magnification affected the length of time that data needed to remain on screen. We also used the data to provide estimations regarding the appropriate calibration of threshold values in speed-dependent automatic zooming systems.”

**ACM Ref:**

Andrew Wallace, Joshua Savage, and Andy Cockburn. 2004. Rapid visual flow: how fast is too fast?. In Proceedings of the fifth conference on Australasian user interface - Volume 28 (AUIC '04), A. Cockburn (Ed.), Vol. 28. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 117-122.

1. ***Sight-reading***First Paragraph of Introduction: “Anyone who wants to perform the works of traditional Western music, and approach other musical styles using a similar performance practice, will most likely have to master music notation. Many musical cultures rely on a system of symbols to store and teach complex musical styles that are not, or only partly, grounded in improvisation. While the beginning music reader has to overcome the same problems as all readers do when learning to read other texts, namely going from a tedious matching of symbols to sounds to meaning, the expert reader has automatized the process of encoding and transforming the

signs into embodied action.

**ACM Ref:**

Andreas C. Lechmann and Reinhard Kopiez. 2009. Sight-reading. In *The Oxford Handbook of Music Psychology*, Oxford University Press, Oxford, United Kingdom, 344-351.

1. ***The art of sight-reading: Influence of practice, playing tempo, complexity and cognitive skills on the eye–hand span in pianists***Abstract: “Sight-reading is a skill required by musicians when they perform sheet music unknown to them.

It demands sequential anticipatory eye fixation of notes immediately followed by motor execution. The distance between eye (fixation of a note) and hand position (tapping the corresponding key) is called eye-hand span (EHS). The aim of our study was to investigate the influence of practice, playing tempo and complexity of the music on the size of the EHS, as well as its relation to performance and cognitive skills (shape recognition, working memory, and mental speed). We used a sight-reading paradigm where nine pianists accompanied a pre-recorded flute voice, which also served as a timekeeper. After a practice phase, a second measurement of the EHS with same tempo and a third and fourth measurement with a different playing tempo followed. We found that the practice phase only slightly affected the EHS but that the EHS significantly changed according to playing tempo and complexity of the music. Furthermore the EHS correlated with quality of performance after practice and mental speed skills. Hence we conclude that the EHS seems to be characteristic for each musician, is developed over years of practice and is relatively independent of a short practice phase.”

**ACM Ref:**

Stephanie Rosemann, Eckart Altenmüller, and Manfred Fahle. 2016. The art of sight-reading: Influence of practice, playing tempo, complexity and cognitive skills on the eye–hand span in pianists. Psychology of Music 44, 4 (2016), 658–673. DOI=http://dx.doi.org/10.1177/0305735615585398

1. ***Exposing and understanding scrolling transfer functions***

Abstract: “Scrolling is controlled through many forms of input devices, such as mouse wheels, trackpad gestures, arrow keys, and joysticks. Performance with these devices can be adjusted by introducing variable transfer functions to alter the range of expressible speed, precision, and sensitivity. However, existing transfer functions are typically "black boxes" bundled into proprietary operating systems and drivers. This presents three problems for researchers: (1) a lack of knowledge about the current state of the field; (2) a difficulty in replicating research that uses scrolling devices; and (3) a potential experimental confound when evaluating scrolling devices and techniques. These three problems are caused by gaps in researchers' knowledge about what device and movement factors are important for scrolling transfer functions, and about how existing devices and drivers use these factors. We fill these knowledge gaps with a framework of transfer function factors for scrolling, and a method for analysing proprietary transfer functions---demonstrating how state of the art commercial devices accommodate some of the human control phenomena observed in prior studies.”

**ACM Ref:**

Philip Quinn, Andy Cockburn, Géry Casiez, Nicolas Roussel, and Carl Gutwin. 2012. Exposing and understanding scrolling transfer functions. In Proceedings of the 25th annual ACM symposium on User interface software and technology (UIST '12). ACM, New York, NY, USA, 341-350. DOI: https://doi.org/10.1145/2380116.2380161

1. ***Aibiki: supporting shamisen practice with adaptive automatic score scroll***
2. Abstract: “We present a system called Aibiki, which can support users in practicing the shamisen, a three-stringed Japanese musical instrument, via an automatic and adaptive score scroll. We chose Nagauta, as an example of a type of shamisen music. Each piece typically lasts 10-40 min; furthermore, both hands are required to play the shamisen, and it is not desirable to turn pages manually during a performance. In addition, there are some characteristic issues that are particular to the shamisen, including the variable tempo of the music and the unique timbre of the instrument, which makes pitch detection difficult using standard techniques. In this work, we describe an application that automatically scrolls through a musical score, initially at a predefined tempo. Because there is often a difference between the predefined tempo and tempo with which the musician plays the piece, the application adjusts speed of the score scroll based on the input from a microphone. We evaluated the performance of the application via a user study. We find that the system was able to scroll the score in time to the actual performance, and that the system was useful for practicing and playing the shamisen.”

**ACM Ref:**

Takahito Hamanaka, Daisuke Sakamoto, and Takeo Igarashi. 2014. Aibiki: supporting shamisen practice with adaptive automatic score scroll. In Proceedings of the 11th Conference on Advances in Computer Entertainment Technology (ACE '14). ACM, New York, NY, USA, , Article 13 , 10 pages. DOI=http://dx.doi.org/10.1145/2663806.2663839