Visual Methods for the Retrieval of Guitarist Fingering

Anne-Marie Burns
Input Devices and Music Interaction Lab
Schulich School of Music, McGill University
555 Sherbrooke Street West, Montréal, Québec,
Canada, H3A 1E3
anne-marie.burns@mail.mcgill.ca

Marcelo M. Wanderley
Input Devices and Music Interaction Lab
Schulich School of Music, McGill University
555 Sherbrooke Street West, Montréal, Québec,
Canada, H3A 1E3
marcelo.wanderley@mcgill.ca

ABSTRACT
This article presents a method to visually detect and recognize finger gestures of the left hand of a guitarist. This method has been developed following preliminary manual and automated analysis of video recordings. These first analyses led to some important findings about the design methodology of a vision system for guitarist finger gestures, namely the focus on the effective gesture, the consideration of the action of each individual finger, and a recognition system not relying on comparison against a knowledge base of previously learned finger gestures. Motivated by these results, studies on three aspects of a complete finger gesture system were conducted: the first on finger tracking, the second on strings and frets detection, and the last one on movement segmentation. Finally, these concepts were integrated into a prototype and a system for left hand finger gesture detection was developed.

Keywords
gesture, guitar fingering, finger-tracking, Hough transform, line detection, gesture segmentation

1. INTRODUCTION
Fingering is an especially important aspect of guitar playing, as it is a fretted instrument where many combinations of string, fret, and finger can produce the same pitch. Finger retrieval is an important topic in music theory, music education, automatic music generation and physical modeling. Unfortunately, as Gilardino noted [6] [7], specific finger information is rarely indicated in scores.

Fingering information can be deduced at several points in the music production process. Three main strategies are:

- Pre-processing using score analysis;
- Real-time using MIDI guitars;
- Post-processing using sound analysis;

Radicino, Anselma, and Lombardo [10] retrieve finger information through score analysis. The score is fragmented in phrases, and the optimum fingering for each phrase is determined by finding the shortest path in an acyclic graph of all possible fingering positions. Weights are assigned to each position based on a set of rules. The problem with this approach is that it cannot account for all the factors influencing the choice of a specific fingering, namely phylological analysis (interpretation of a sequence of notes), physical constraints due to the musical instrument, and biomechanical constraints in the musician-instrument interaction. Outputs of these systems are similar to human solutions in many cases, but hardly deal with situations where the musical intention is more important than the biomechanical optimum fingering.

Other systems retrieve the fingering during or after a human plays the piece. One of these approaches uses a MIDI guitar. Theoretically, using a MIDI guitar with a separate MIDI channel assigned to each string, it is possible to know in real-time what pitch is played on which string, thus determining the position. In practice, however, MIDI guitar users report several problems, including a variation in the recognition time from one string to another and the necessity to adapt their playing technique to avoid glitches or false note triggers [13].

An approach using the third strategy is the study of the guitar timbre. Traube [11] suggested a method relying on the recording of a guitarist. The method consists of analyzing the sound to identify the pitch, find the plucking point and then determine the string length to evaluate the fingering point. Shortcomings of this method are that it cannot be applied in real time, it works only when one note is played at the time, and error of the string length evaluation can be as high as eight centimeters in the case of fretted strings [12].

This paper presents an alternative method for real-time retrieval of the finger information from a guitarist playing a musical excerpt. It relies on computer analysis of a video recording of the left hand of the guitarist. The first part of this article discusses the preliminary manual and automated analyses of multiple-view video recordings of a guitarist playing a variety of musical excerpts. The subsequent sections present studies on three aspects of visual analysis of a guitarist fingering: finger tracking, string and fret detection, and movement segmentation. Finally a system integrating these three components is presented.

2. PRELIMINARY ANALYSIS
During the preliminary analysis, different camera views were evaluated (global view, front view, and top view). The
Fro of the musician: (a) Global view with zoom on dif-
captured from a camera on a tripod placed in front
Figure 1: Two different views of a guitarist playing
and front view of right and left hands. (b) Top view
allowed to identify three
m stored in the for
racy and precision.

3.1 Circular Hough Transform
As illustrated in figure 3, the circular Hough transform
[3] is applied on the binary silhouette image of the hand. The edge-image is obtained by applying the Canny edge de-
tection algorithm [4] on the silhouette images. The circular Hough transform algorithm makes use of the fact that fin-
ger ends have a quasi-circular shape while the rest of the
hand is more linearly shaped. In this algorithm, circles of a
given radius are traced on the edge-images and regions with the highest match (many circles intersecting) are assumed
to correspond to the center of fingertips.
4. STRING AND FRET DETECTION
By tracking the fingertips it is possible to know where each finger is in space. In the case of guitar fingering, this space
can be defined in terms of string and fret coordinates. Prior
to the detection stage, the region of interest (in that case the
guitar neck) must be located in the image. Once the neck
has been located, the strings and frets are segmented from the
grayscale neck image by applying a threshold. A verti-
cal and a horizontal Sobel filter are applied on the threshold
image in order to accentuate the vertical and horizontal gra-
dients. A Linear Hough Transform [3] is then computed on the
two Sobel images. The linear Hough transform allows
detection of linearity in a group of pixels, creating lines.
These lines are then grouped by proximity in order to deter-
mine the position of the six strings and of the frets. Once
this is done, it is possible to create a grid of coordinates to
which fingertip positions will be matched.
5. MOVEMENT SEGMENTATION
Movement segmentation is essential in order to detect fin-
gering positions and not simply track fingertips during the
playing sequence. Furthermore, in order to save computer resources, this segmentation must be done early in the global process so that subsequent analysis steps are not performed unnecessarily. Movement segmentation is used to separate the nucleus phase of the gesture from the preparation and retraction phase [9].

In the preliminary analysis, movement segmentation was done by applying a threshold on the motion curve (figure 4 a) generated by the computation of the pixel difference between each frame. The characteristic lower velocity phase of the nucleus was easily detected between each chord. However, in other playing situations, such as when playing a series of notes, the separation between the movement transitory phases and the nucleus is not that clear (figure 4 b). This is due to a phenomenon called *anticipatory placements of action-fingers* that has been studied in violin [1] and piano [5]. In these cases, the preparation phase of other fingers occurs during the nucleus of the action-finger. Thus the motion is not serial and consequently, the global motion curve does not exhibit clear global minima like it is the case for chords. However, local minima can still be observed and detected as they can be assumed to correspond to the moment the note is triggered by the right hand. Local minima are found by computing the second derivative of the motion curve. As the prototypes work in real-time, this is done by subtracting the signal with its delayed version twice.

6. PROTOTYPE

The prototype was designed to fulfill the requirements for a fingering recognition system highlighted by the preliminary analysis. The focus on effective gestures is partially realized at the hardware level by affixing the camera to the guitar neck, thereby eliminating the motion of the neck caused by the ancillary gesture. Elimination of background elements is done by selecting a strict ROI (Region of Interest) around the neck and by applying a background subtraction algorithm on the image. Movement segmentation is performed by finding minima in the motion curve, obtained by computing the difference of pixel between each frame. The action of each individual finger is considered using the finger-tracking algorithm described above. The details of the algorithm are shown in figure 5.

During preliminary tests, the prototype was able to correctly recognize all fret positions. Due to the chosen camera view, the space between the strings is smaller for the high strings (E, B, G) than for the low strings (D, A, E), therefore affecting the accuracy of the recognition system. As demonstrated in [2], the circular Hough transform has an accuracy of 5 +/- 2 pixels with respect to the color marker references. The resolution of the camera used in this prototype is 640x480 pixels, therefore giving a 610x170 pixels neck region. The distance in pixels between the first and second string is of 12 pixels at the first fret and 17 at the fifth fret. Between the fifth and sixth strings, the distance in pixels is 16 and 20 pixels for the first and fifth fret, respectively.

Since the chosen algorithm attributes the string position to the finger by proximity, in the worst case the finger-tracking algorithm error exceeds half the space between the higher strings, therefore confusion happens. However, since this problem does not happen with lower strings were the distance between two strings is greater, the problem could be solved with an higher resolution camera. Another limitation is that in the current system only the first 5 frets are evaluated, but this could be solved with a wide angle camera. One problem that cannot be easily solved by changing the hardware is finger self occlusion. This problem only rarely happens, but exists in the case of fingerings were two fingers play at the same fret, for example in the case of C7 and Dm7. In future developments, this problem could potentially be solved by estimating the fingertip position using the finger angle.

7. CONCLUSIONS

This article discussed new strategies to capture fingerings of guitarists in real-time using low-cost video cameras. A prototype was developed to identify chords and series of notes based on finger-tracking and fret and string detection. It recognizes fingerings by matching fingertip positions to the strings and frets’ grid of coordinates, therefore not relying on any knowledge base. Results of the prototype are encouraging and open possibilities of studies on many aspects of a guitarist instrumental gesture, namely gesture segmentation, anticipatory movements, and bimanual synchronization. Applications of this research include automatic chord transcription, music education, automatic music generation.
and physical modeling.

8. ACKNOWLEDGMENTS

The study on finger-tracking was realized at InfoMus laboratory, D.I.S.T., Università degli studi di Genova and has been partially supported by funding from the Québec Government (PBCSE), the Italian Ministry of Foreign Affairs, and by the EU 6 FP IST ENACTIVE Network of Excellence. The first author would like to thank all students and employees of InfoMus lab who "lent a hand" for the realization of the tests. Special thanks go to Barbara Mazzarino, Ginevra Castellano for her help in compiling the results, Gualtiero Volpe for his contribution to the development of the EyesWeb blocks. Anne-Marie Burns would also like to thank Antonio Camurri for welcoming her as an internship student researcher.

The authors would like to thank the guitarists that participated in the tests: Jean-Marc Juneau, Riccardo Casazza, and Bertrand Scherrer. Thanks also to Mr. Donald Pavlasek, Electrical and Computer Engineering, McGill University, for the conception and implementation of the camera guitar mount.

9. REFERENCES