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 fingering 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 fingering, 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 fingering positions. Motivated by these results, studies on three aspects of a complete fingering 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 fingering 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. Fingering retrieval is an important topic in music theory, music education, automatic music generation and physical modeling. Unfortunately, as Gilardino noted [6] [7], specific fingering 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;

Radicioni, Anselma, and Lombardo [10] retrieve fingering 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 philological 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 then the biomechanical optimum fingering.

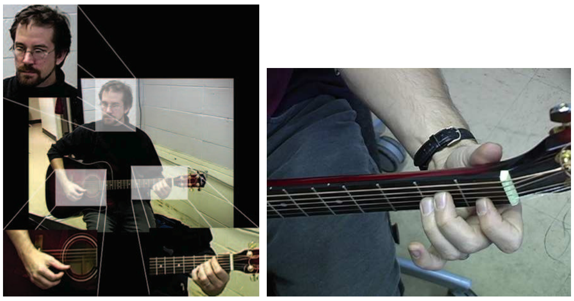
Other systems retrieve the fingering during or after a human plays the piece. One of these approaches uses a Midi guitar. Theoretically, using a Midi guitar with a separate Midi channel assigned to each string, it is possible to know in real-time what pitch is played on which string, thus determining fret 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 fingering 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.

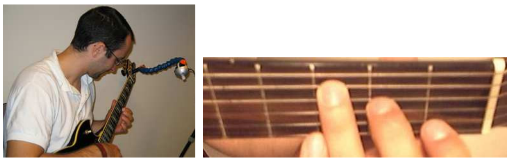
2. PRELIMINARY ANALYSIS

During the preliminary analysis, different camera views were evaluated (global view, front view, and top view). The



(a) Global view with zooms (b) Top view of the left hand

Figure 1: Two different views of a guitarist playing captured from a camera on a tripod placed in front of the musician: (a) Global view with zoom on different zones for gesture analysis: facial expression and front view of right and left hands. (b) Top view of the left hand.



(a) Camera mount (b) Camera view

Figure 2: Depiction of the guitar camera mount that was used to eliminate the ancillary gesture problem: (a) The camera mount installed on an electric guitar. (b) The camera on a classical guitar. In this example, the camera is placed to capture the first five frets.

aim was to find a viewpoint that allows the retrieval of the most information possible with the desired degree of accuracy and precision.

The top view (figure 1(b)) was retained for its interesting characteristics with respect to the problem, namely a detailed view of the fingers, the possibility for string and fret detection, and the ability to observe finger-string proximity. However, slow motion observations of the video recording showed that the neck is subject to many ancillary movements. Preliminary automated tests have shown that this type of movement can influence the computer's capacity to correctly identify fingering. Consequently, the tripod was replaced by a camera mount on the guitar neck (figure 2). The preliminary automated fingering recognition tests were performed by comparing two top view recordings of a musician playing musical excerpts against top view images of previously recorded chords played by the same performer stored in the form of Hu moments vectors [8]. These tests allowed to identify three main issues:

1. Using an appearance base method limits the system to previously learned material.
2. Using the global shape of the hand limits the system to the recognition of chords.
3. Using a knowledge base makes the recognition time grow with the knowledge base size.

From the above issues, the main specifications for a fingering recognition system are:

1. Focus on effective gestures by further reducing the presence of ancillary movements and background elements.
2. Use of a representation that considers the action of individual fingers.
3. Use of a recognition mechanism that eliminates the burden of a knowledge base and that is therefore not limited to previously learned material.

The first specification can be achieved using the guitar mount as presented in figure 2. In order to fulfill the other specifications, three studies were conducted. In a first study, the circular Hough transform was chosen to perform finger-tracking. The second study examined the use of the linear Hough transform for string and fret detection, and a third one explored movement segmentation.

3. FINGER-TRACKING

The circular Hough transform algorithm used in this paper was developed and implemented in EyesWeb [2]. It presents the following interesting characteristics:

1. It demonstrated to have a high degree of precision and accuracy;
2. It can be applied in complex environments and with partial view of the hand;
3. It can work on edge versions of the images.

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 detection algorithm [4] on the silhouette images. The circular Hough transform algorithm makes use of the fact that finger 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 vertical and a horizontal Sobel filter are applied on the threshold image in order to accentuate the vertical and horizontal gradients. 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 determine 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 fingering positions and not simply track fingertips during the

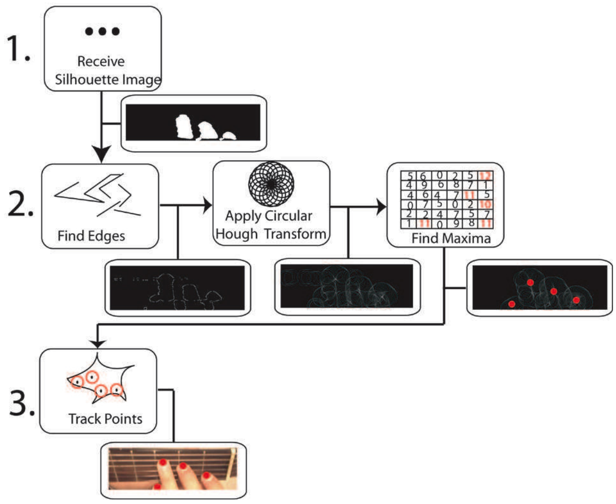


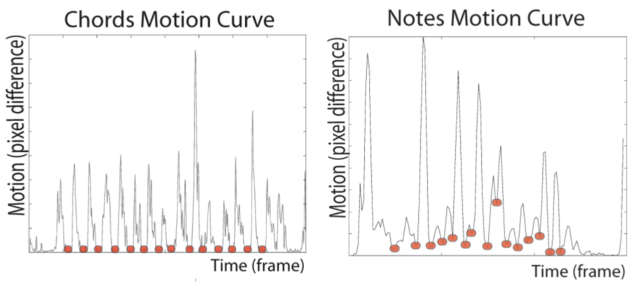
Figure 3: Fingertip detection using the circular Hough transform algorithm

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 occur 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 trigged 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



(a) Chords motion curve (b) Notes motion curve

Figure 4: (a) Motion curve of a guitarist playing chords (b) Motion curve of a guitarist playing a series of notes

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 fingering 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

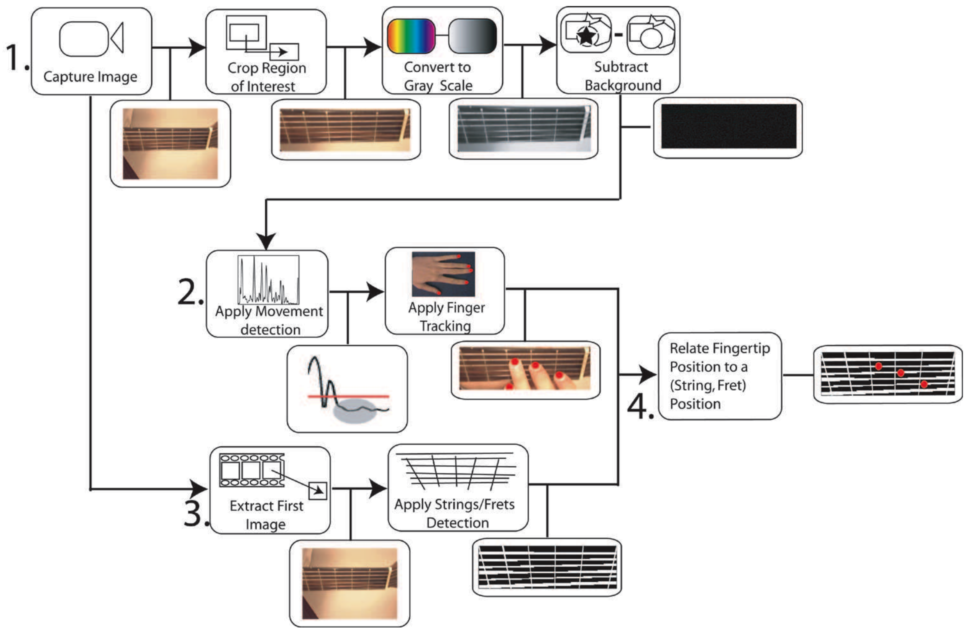


Figure 5: Prototype - algorithm

and physical modeling.

8. ACKNOWLEDGMENTS

The study on finger-tracking was realized at InfoMus laboratory, D.I.S.T., Universitá 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

[1] A. P. Baader, O. Kazennikov, and M. Wiesendanger. Coordination of bowing and fingering in violin playing. *Cognitive Brain Research*, 23:436–443, 2005.

[2] A.-M. Burns and B. Mazzarino. Finger tracking methods using eyesweb. In *S. Gibet, N. Courty and J.-F. Kamp (Eds)., Gesture Workshop 2005 Proceedings, LNAI 3881*, pages 156–167, 2006.

[3] A.-M. Burns and M. M. Wanderley. Computer vision method for guitarist fingering retrieval. In *Proceedings of the Sound and Music Computing Conference 2006*. Centre National de Création Musicale, Marseille, France, May 2006.

[4] J. A. Canny. Computational approach to edge detection. *, IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[5] K. C. Engel, M. Flanders, and J. F. Soechting. Anitcipatory and sequential motor control in piano playing. *experimental Brain Research*, 113:189–199, 1997.

[6] A. Gilardino. Il problema della diteggiatura nelle musiche per chitarra. *Il "Fronimo"*, 10:5–12, 1975.

[7] A. Gilardino. Il problema della diteggiatura nelle musiche per chitarra. *Il "Fronimo"*, 13:11–14, 1975.

[8] S. Paschalakis and P. Lee. Pattern recognition in grey level images using moment based invariant features image processing and its applications. *IEE Conference Publication*, 465:245–249, 1999.

[9] V. I. Pavlovic, R. Sharma, and T. S. Huang. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):677–695, 1997.

[10] D. Radicioni, L. Anselma, and V. Lombardo. A segmentation-based prototype to compute string instruments fingering. In *Proceedings of the Conference on Interdisciplinary Musicology*, Graz, 2004.

[11] C. Traube. *An Interdisciplinary Study of the Timbre of the Classical Guitar*. PhD thesis, McGill University, 2004.

[12] C. Traube and J. O. Smith III. Estimating the plucking point on a guitar string. In *Proceedings of the COST G-6 Conference on Digital Audio Effects*, Verona, Italy, 2000.

[13] J. A. Verner. Midi guitar synthesis yesterday, today and tomorrow. an overview of the whole fingerpicking thing. *Recording Magazine*, 8(9):52–57, 1995.