

A Method for Tracking Guitar Neck and Fingertips: Necking Tracking Robust against Occlusions Based on Geometry Analysis and Fingertips Tracking Based on Temporal Probability Map

Wang Zhao Jun Ohya

1. Introduction

One of the most important techniques in guitar learning is called fingering, and the precise fingering is a prerequisite for guitar performance. This paper proposes a computer vision-based method for tracking guitar neck and guitarist's fingertips. In guitar neck tracking part, we develop a feature points-based tracking method by analyzing geometric relationship between each two feature points in every frame. In fingertips tracking part, we present an approach by continuously monitoring the appearance and disappearance of the regions of fingertips candidates. Compared with related works, our method can relax some problems and inconvenience such as color mark attached at fingers or AR Tags fixed on the guitar neck. The experimental result demonstrates that the system is robust enough for tracking the guitar neck and fingertips under the complex contexts, such as complicated background, different illumination conditions, volatile hand shapes and so on.

2. Outline

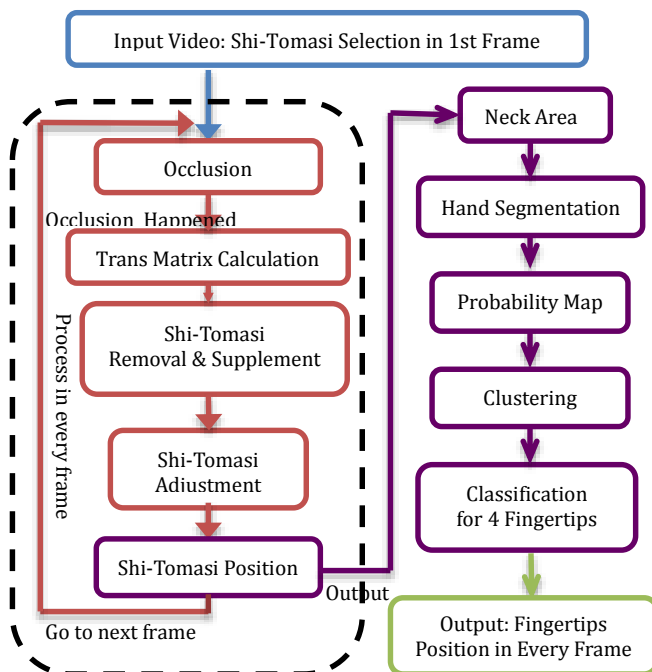


Figure 1 Outline of Proposed Method

Graduate School of Global Information and Telecommunication Studies, Waseda University

As shown in Fig.1, our proposed method consists of two steps: Step 1 (red textboxes in Fig.1) is the part of guitar neck tracking. The output of Step 1 is the tracking result of guitar neck, and it is also the input of Step 2, which is the part of fingertips tracking (purple textboxes in Fig.1). The output of our whole system is the position of four fingertips in every frame. Each step is detailed in Section 3 and Section 4 of this paper.

3. Guitar Neck Tracking

3.1 Shi-Tomasi Detector and Lucas-Kanade Optical Flow

In this guitar neck tracking part, we use Shi-Tomasi Detector as our tracking target, and Lucas-Kanade-based Optical Flow as our tracking method. In the first frame of input video, we select five Shi-Tomasi Detector within the area of guitar neck, two of them are at the left-most side of the guitar neck, other two points are at the right-most side, the other one is at middle side. Then, we calculate the distances between each two Detectors, save those distance into the memory in order to do the solution in Section 3.2. By tracking these Shi-Tomasi Detector precisely from second frame using Lucas-Kanade Optical Flow, the whole area of guitar is tracked accurately.

3.2 A Solution against Occlusion

The Optical Flow-based algorithm can track the whole guitar neck area by tracking the detector selected in Section 3.1. However, it cannot cope with the occlusion, which is an inevitable issue, as guitarist moves the left hand during playing. In order to deal with this problem, a solution is developed:

In any given frame, whether the detectors selected in Section 3.1 are occluded or not needs to be judged. The distances between each two detectors on this frame are calculated, and compared with the distances calculated in Section 3.1.

(1) Calculation of transformation matrix

After finding which detector is occluded, the location where this occluded point should be at is calculated as once the detector is occluded by guitarist, the tracking result of that detector is not accurate. For example, if lower-right detector is occluded, the transform matrix of the guitar neck is calculated based on other four unoccluded detectors in this frame and their correspondent detector selected in the first frame.

(2) Removal and supplement of new detector

Once the transform matrix of the guitar neck area between the first frame and the certain frame is calculated, it means that every pixel of the guitar neck area in the first frame can be projected to this frame. With the transform matrix calculated, the position that the occluded detector should be can be calculated, and at the position of that detector should be, we remove the occluded detector and supply a new one.

By calculating these geometric relationships between these detectors in every frame, the neck can be tracked accurately.

4. Fingertip Tracking

4.1 Probability Map

After segmenting the hand area (Fig.2.a) from the guitar neck tracking result, an image named probability map is generated. The probability map (Fig.2.b) has the same size with the segmentation result of the hand.



Figure 2 Conceptual Image of Probability Map

4.2 Clustering Algorithm for Fingertips^[2]

Compared with real fingertips, the noises' life spans are shorter. Figure 3 shows a visualized understanding of the clustering algorithm. The blue circles in Fig.3.a indicate all the pixels that have higher probability to be fingertips. The same color circles in Fig.3.b are grouped into a cluster, because in time sequence, the Euclid distances between the three same color circles are very short. On the other hand, when circles are contaminated from noise, as red circle in Fig.3.b, the Euclid distances between such a circle and other circles are not short enough to be clustered in the consecutive three probability maps. This is the intuitively understanding of the clustering algorithm. By monitoring the result of clustering algorithm in current frame and last frame of image sequence, fingertips of the guitarist can be tracked.^{[1][2]}

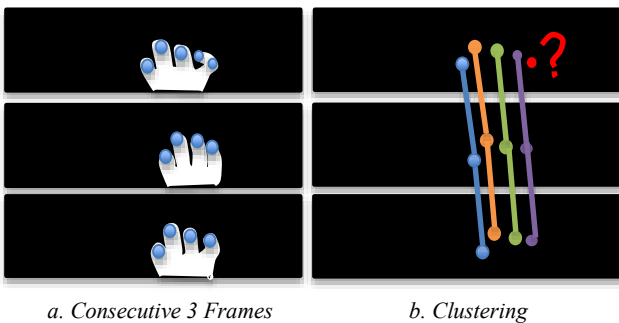


Figure 3 Conceptual Image of Clustering

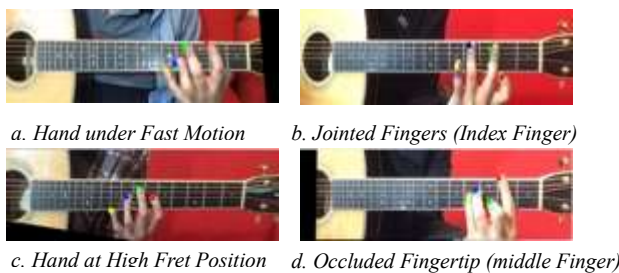


Figure 4 The Examples of Fingertips Tracking Result
Yellow: Small Finger Blue: Ring Finger
Green: Mid Finger Red: Index Finger

5. Evaluation

In this experiment, 24 videos (3704 frames) were used to test the validity of the Fingertips Tracking System. The examples of tracking result are shown in Fig.4.

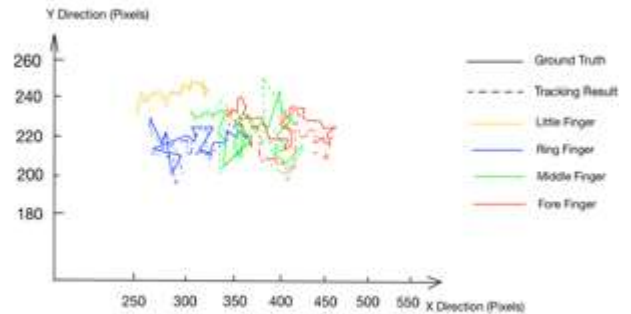


Figure 5 Trajectories of 4 Fingertips for an Image Sequence

Table 1 Mean Error for 4 Fingertips

	Little Finger	Ring Finger	Middle Finger	Fore Finger
Mean Error (Pixel)	6.15	5.65	7.4	10.2

Figure 5 shows the trajectories of four fingertips for an image sequence with 214 consecutive frames. The ground truth is manually observed by human eye, while the tracking result is obtained by our system. Table 1 shows the mean error of each finger in all 24 videos (total 3704 frames). With the mean error 6.16, 5.65, 7.4, 10.2 pixels for little finger, ring finger, middle finger and fore finger respectively, it is obvious that the system is robust enough for tracking the fingertip of guitarist during guitar playing.

6. Conclusion

Fingertips Tracking is the most challenging part in this research, because during guitar playing, sometimes the shape of hand changes so fast, sometimes two fingers jointed together, sometimes due to occlusion fingertips cannot be seen and etc. Another difficult problem is the system has to track four fingers at the same time while four fingers share same feature: they have almost same shape, almost same color.

In the future, machine learning-based fingering analysis is our first task. The features that can comprehensively and thoroughly express the trajectories of guitar and hand need to be selected. Then, in order to recognize good fingering, the algorithm of machine learning need to be developed.

Reference

- [1] Chutisant Kerdvibulvech and Hideo Saito, "Guitarist Fingertip Tracking by Integrating a Bayesian Classifier into Particle Filters". International Journal of Advances in Human-Computer Interaction (AHCII), Hindawi Publishing Corporation, ISSN 1687-5893, 10 pages, 2008.
- [2] Ng, William, et al. "A review of recent results in multiple target tracking." Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on. IEEE, 2005.