

A Vision-based Machine Learning Approach to Guitar Tabs Transcription

Arnav Kalekar, Vishaal Iyer, Om Muthyala, Sooji Lee

Existing Work



[3]

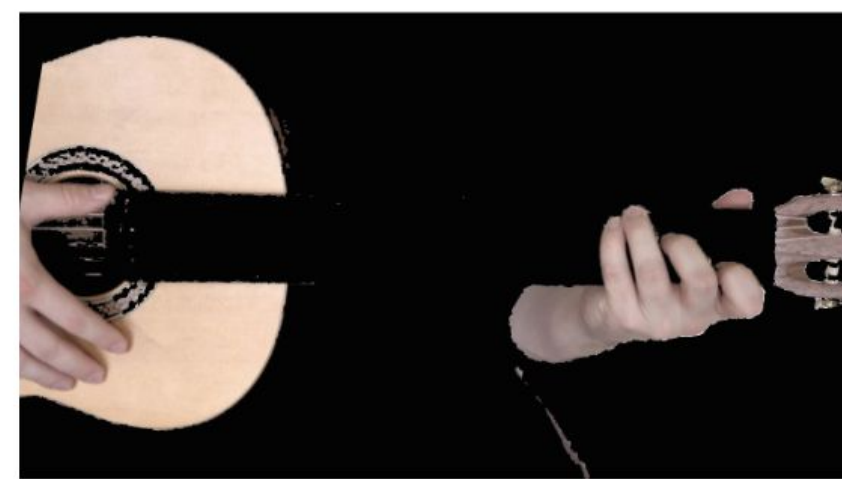


Fig. 5. Skin detection and segmentation.

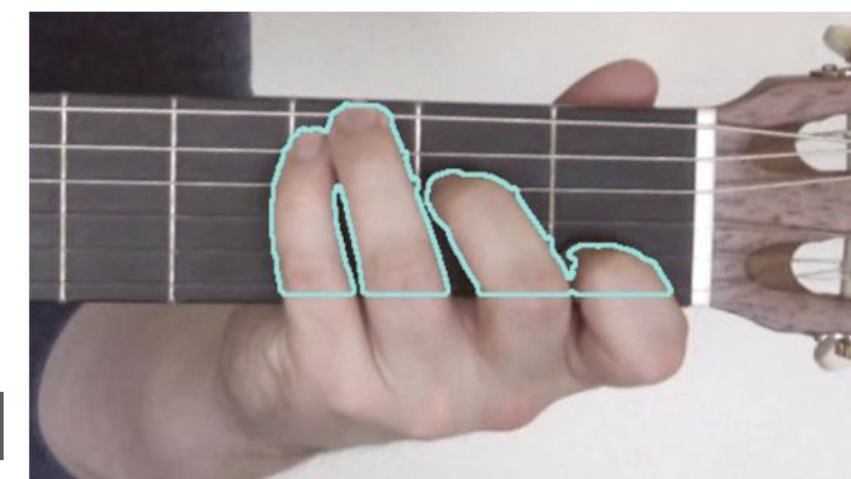
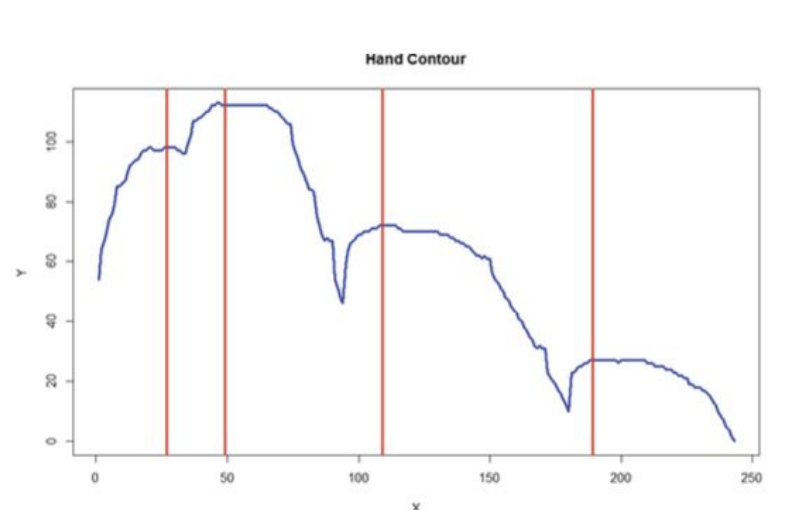
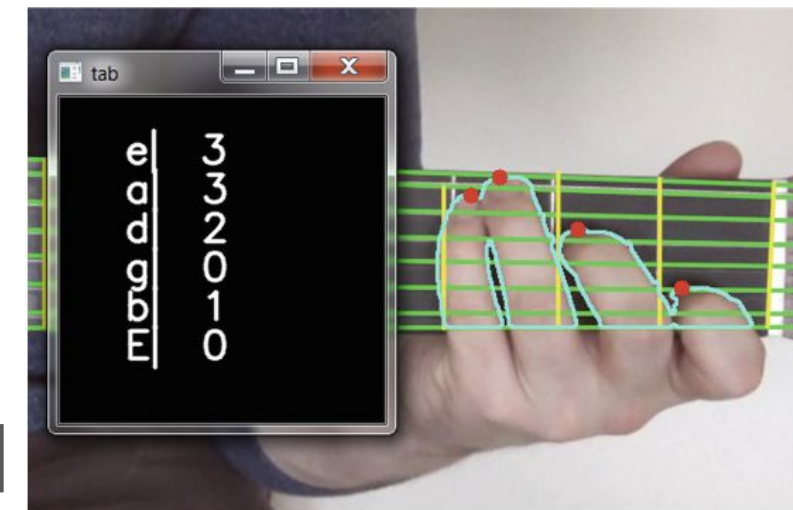


Fig. 6. Finger isolation and contour detection.



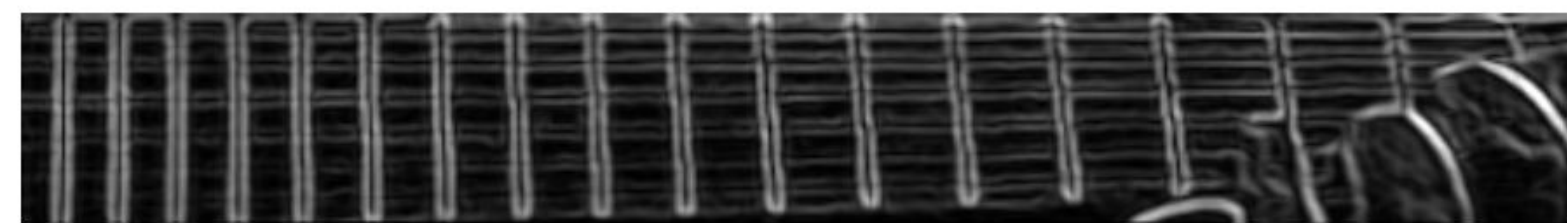
[6]



[6]

Fig. 7. Fingers contour with fingertips located at local maxima.

Fig. 8. System output. Tablature notation was flipped to correspond with strings in the video.



(a) The obtained image after applying both horizontal and vertical Sobel filters and adding their results.



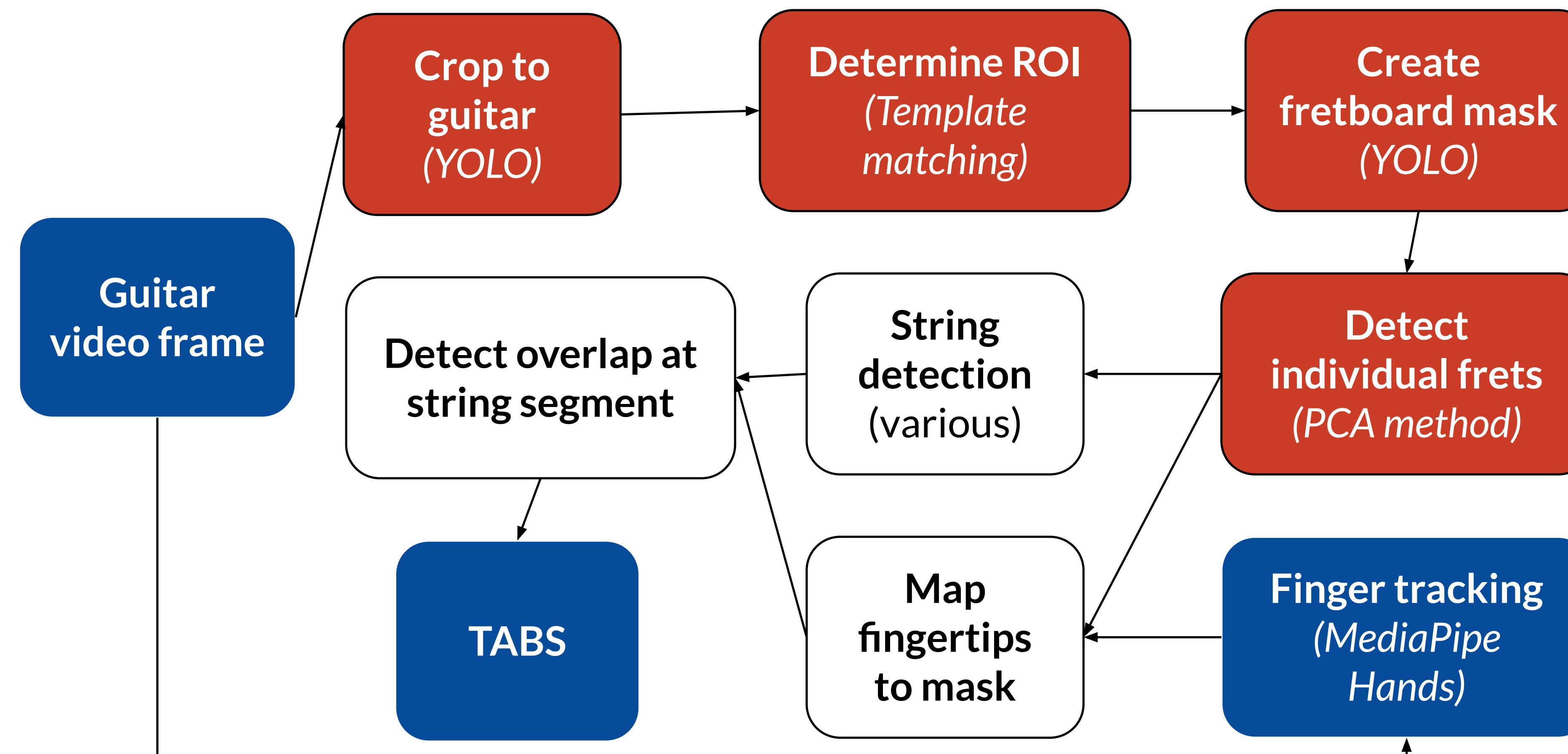
(b) The resulting image after using adaptive thresholding.

Acknowledgements

Thank you to the NSF for funding the AI in Music group and their projects, including the Automatic Music Transcription project. Additionally, thank you to the advisors of this project, Professor Yung-hsiang Lu and Professor Yeon Ji Yun.



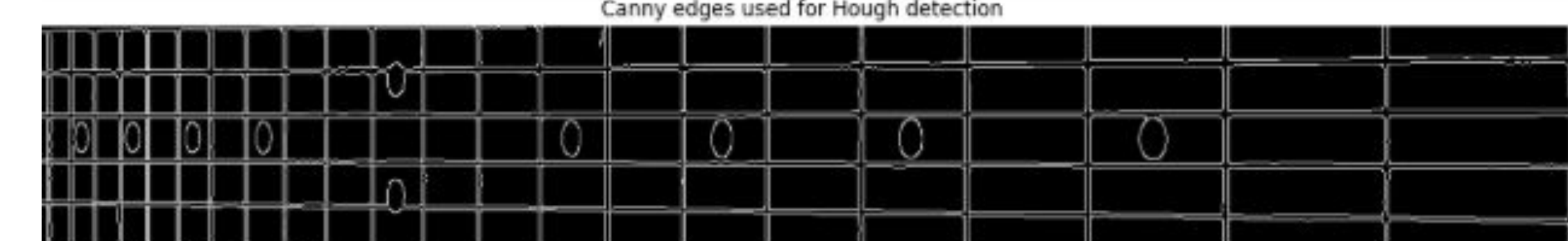
Main Work



Cropped fretboard ROI



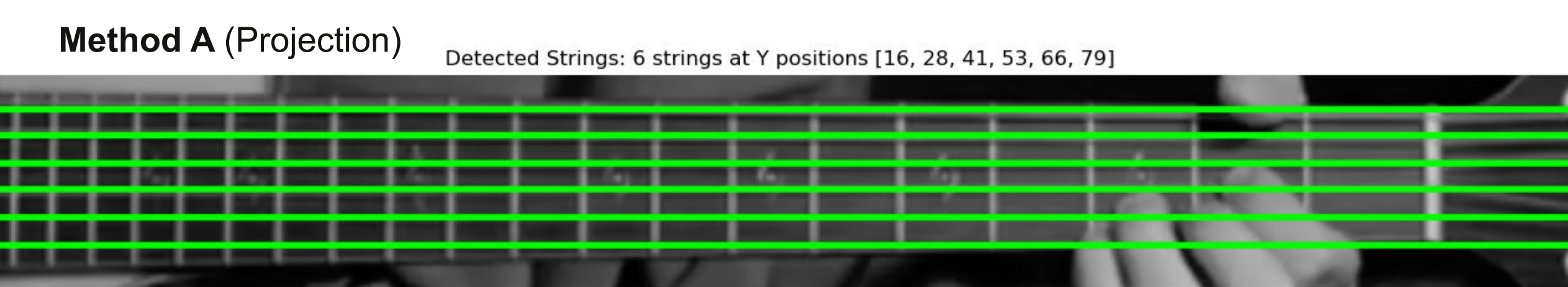
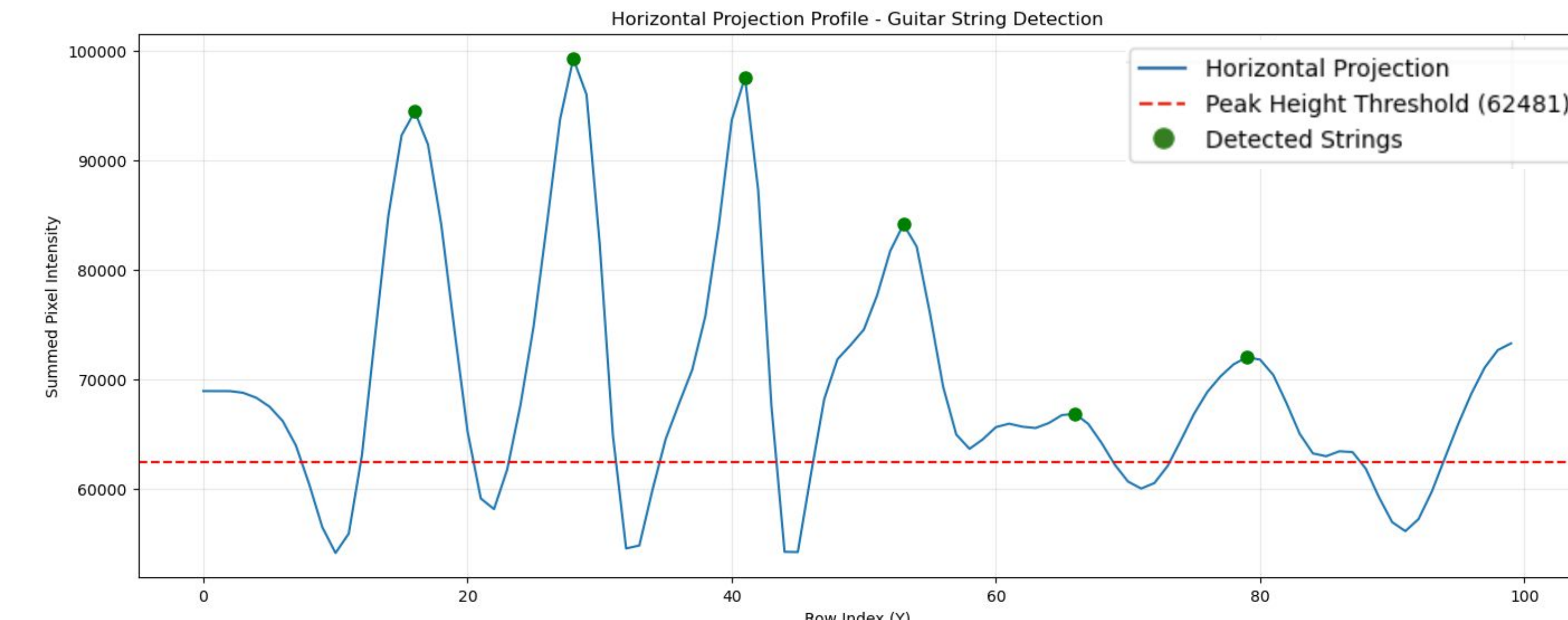
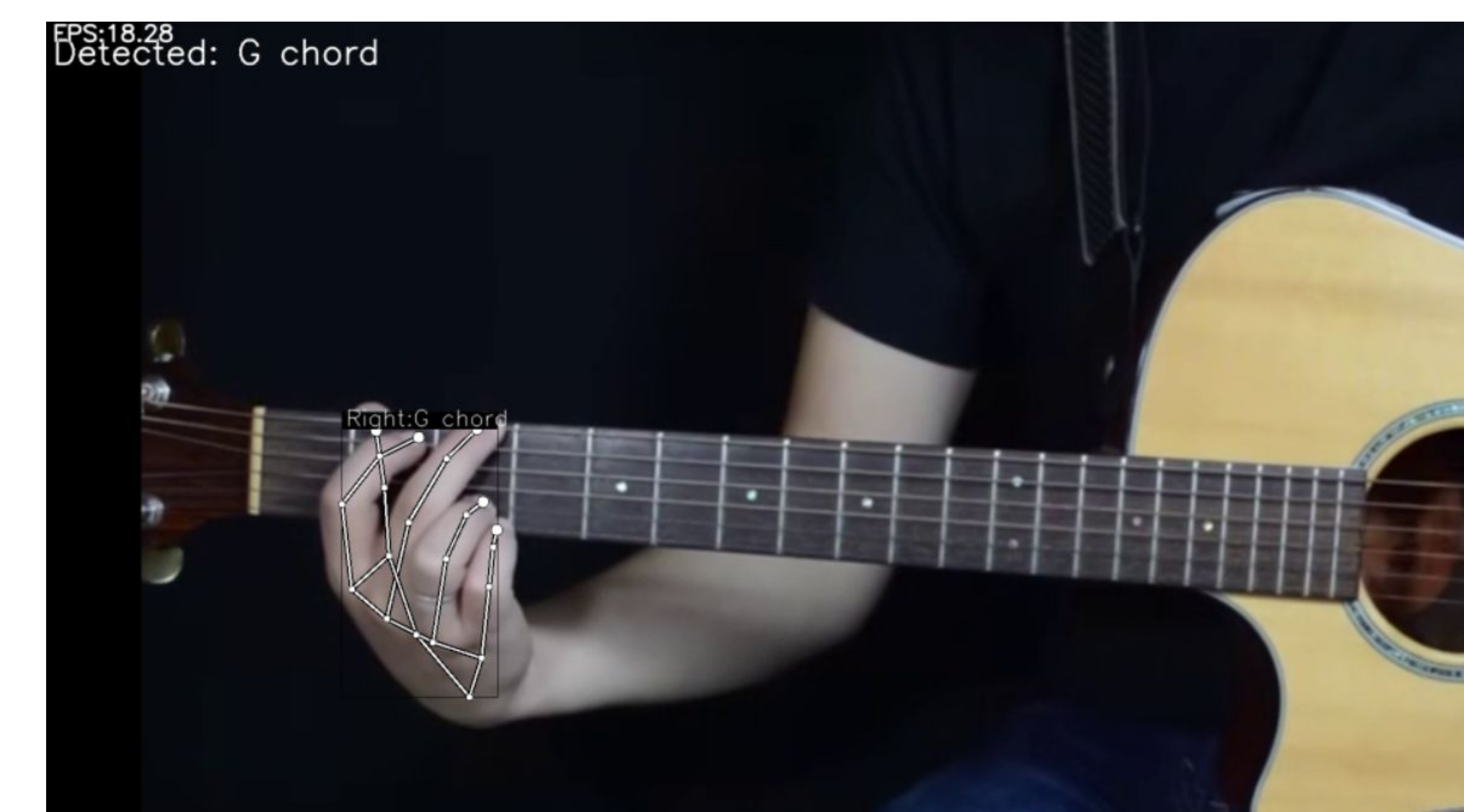
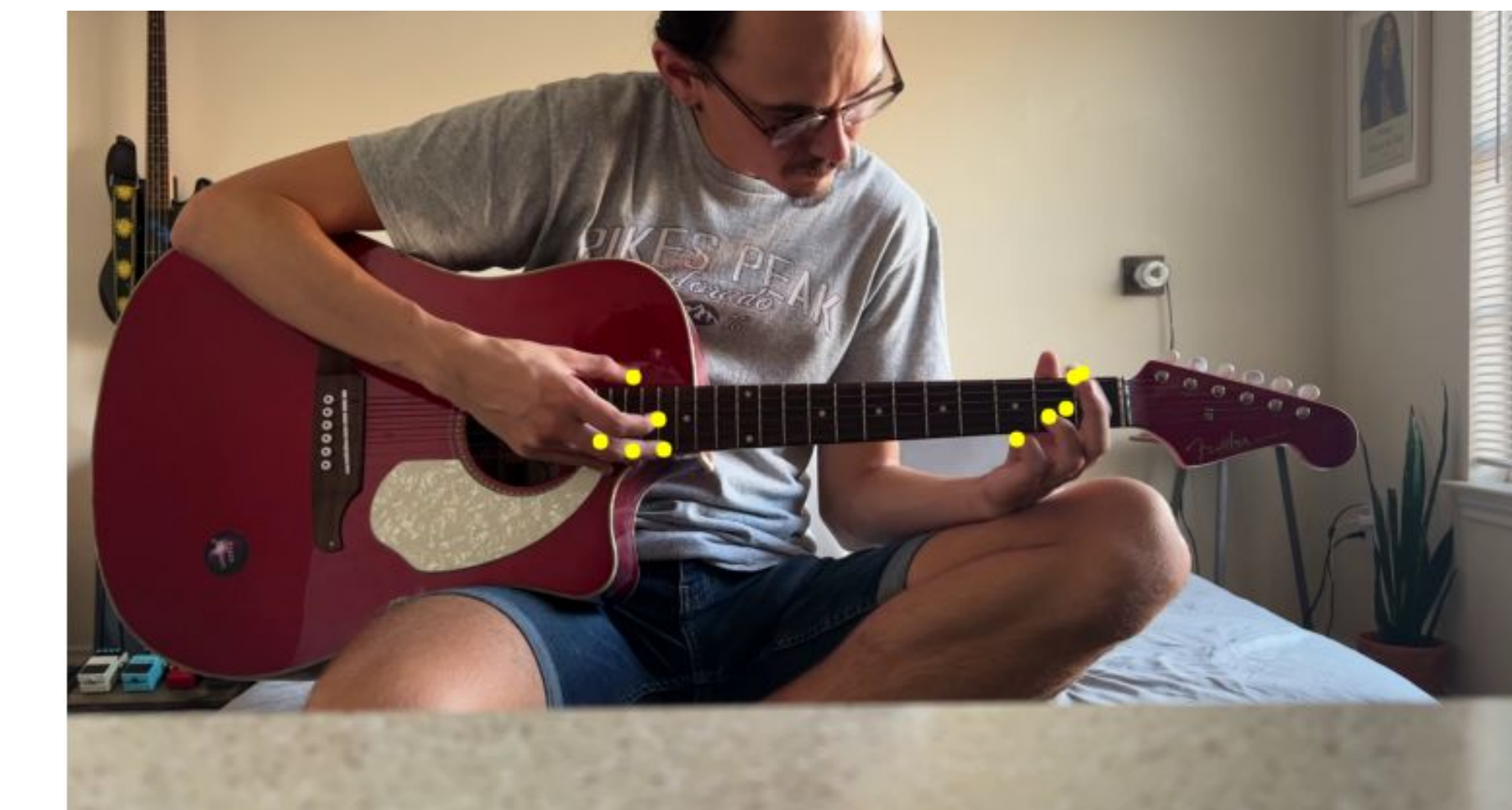
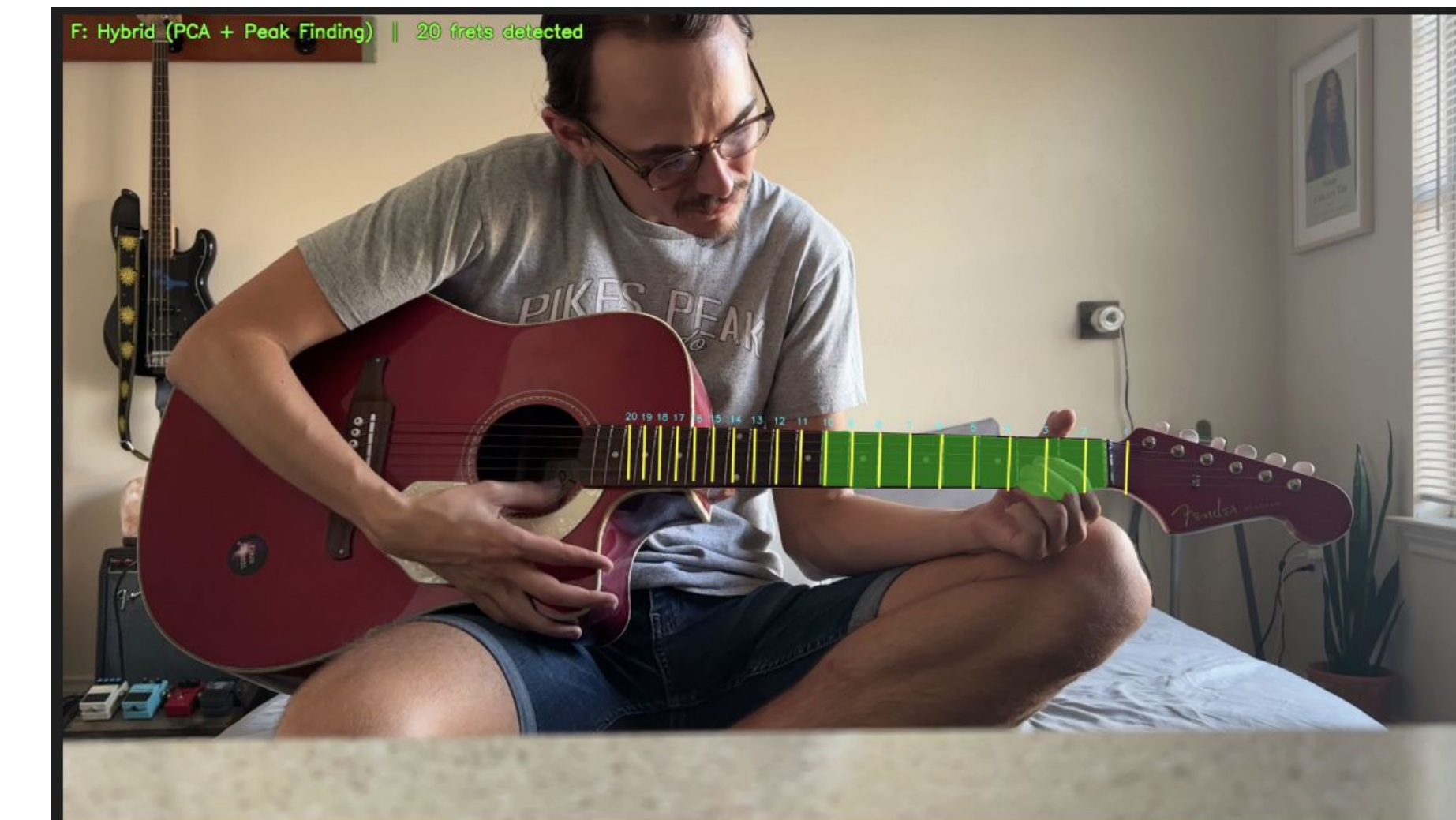
ROI fretboard mask



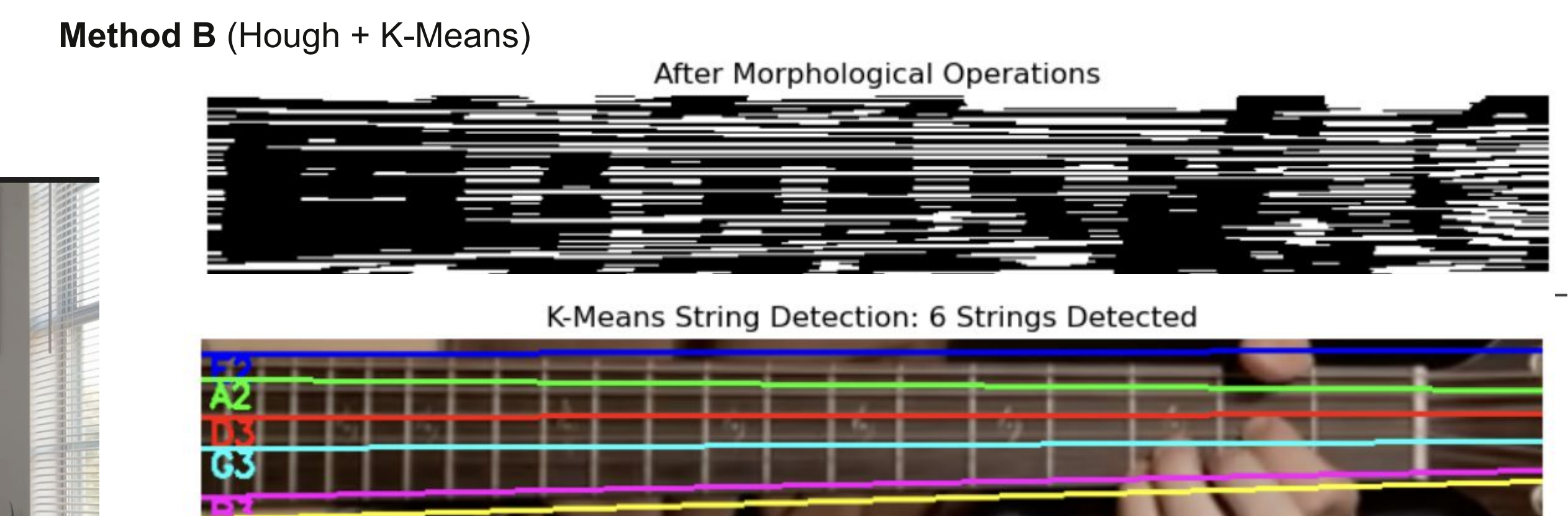
Canny edges used for Hough detection



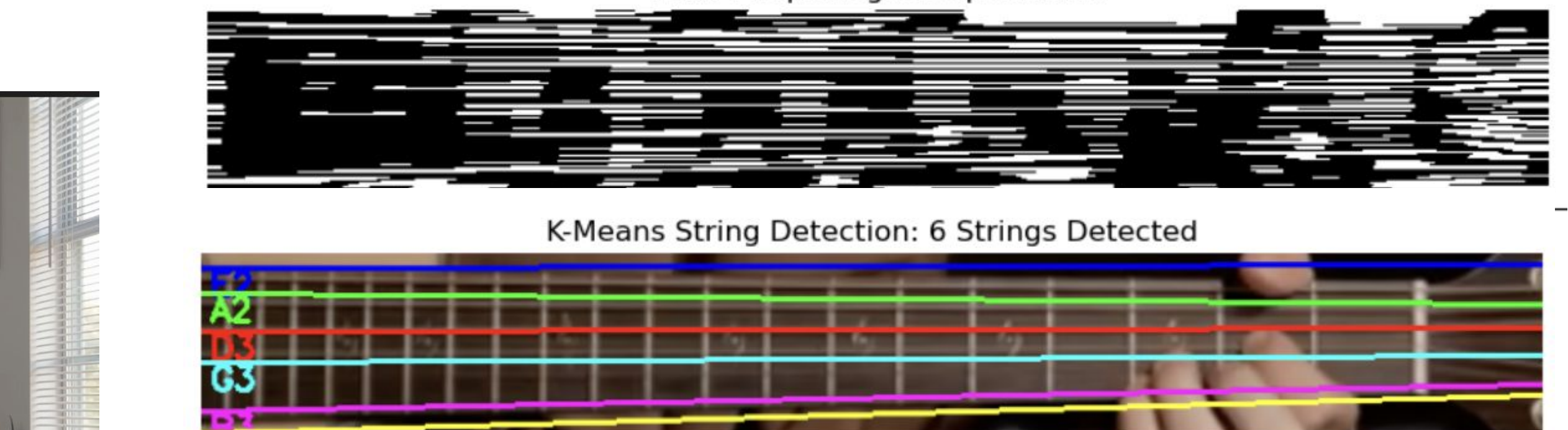
Rectified fretboard with Hough fret detections



Method A (Projection) Detected Strings: 6 strings at Y positions [16, 28, 41, 53, 66, 79]



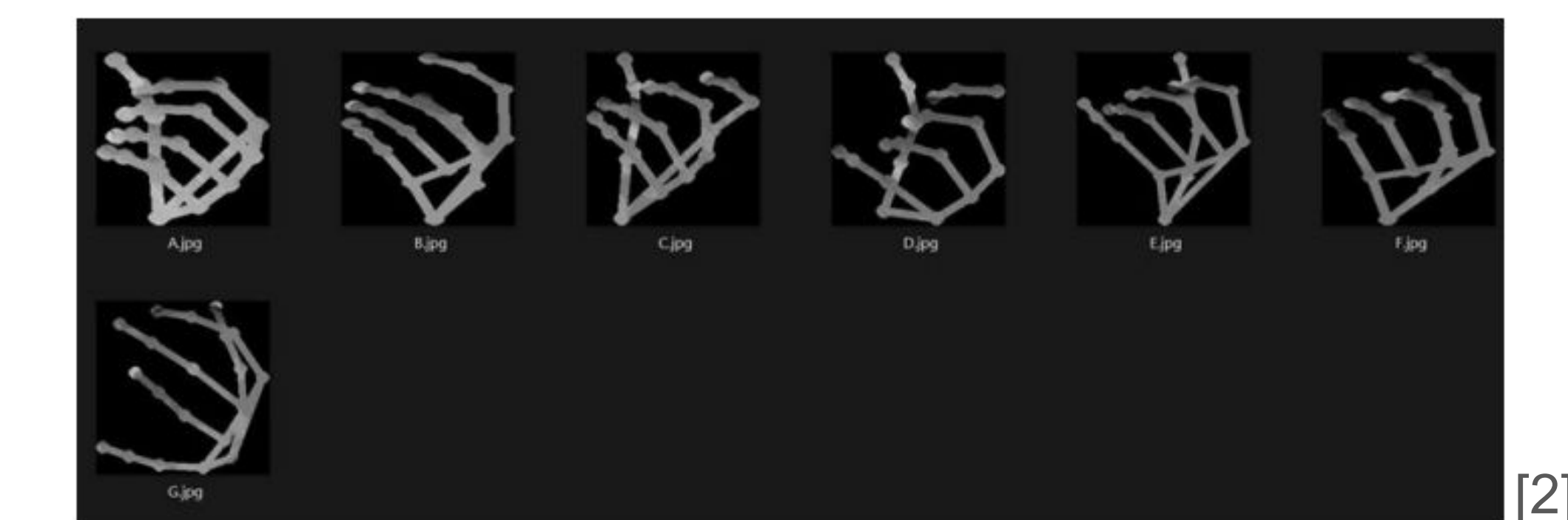
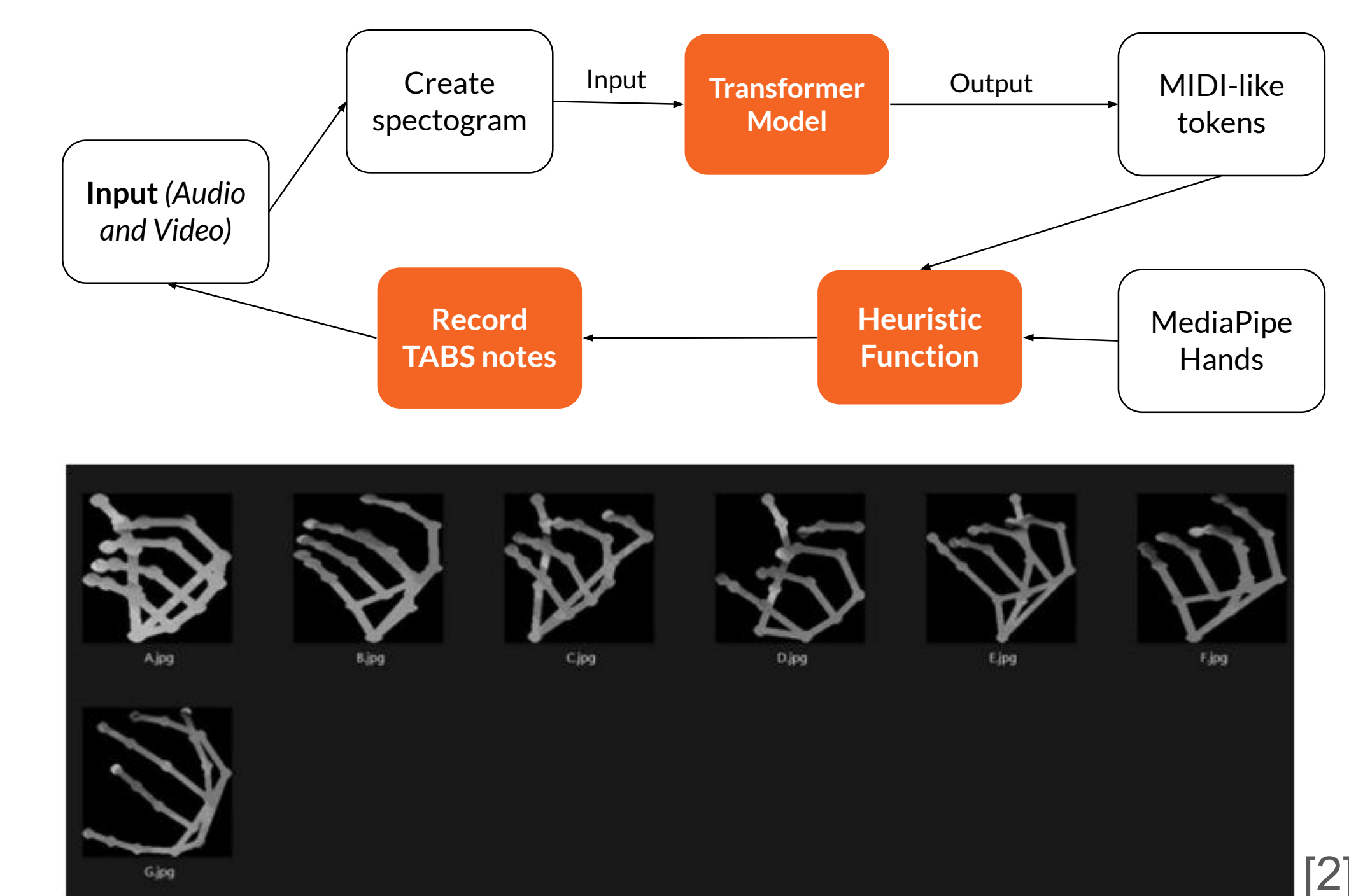
Method B (Hough + K-Means) After Morphological Operations



K-Means String Detection: 6 Strings Detected

Method	F1 @ 1%	F1 @ 2%	MPE
PCA + Warp	0.6333	0.8067	0.788
PCA (no warp)	0.6365	0.7963	0.762
Paleari & Huet	0.4567	0.5597	0.843
Template Matching	0.3748	0.4055	0.716

Future Work



[2]