ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 7, July 2017

Dynamic time warping music pattern detection

[1] Ms. Snehal Kshirsagar, [2] Prof. M.N. Annadate
[1][2] Department of ENTC Engg,MITCOE, KOTHRUD, PUNE
Savitribai Phule Pune University, India

***Abstract* - Music conduct by a group requires a conductor to play the music. These conductors perform various music patterns in front of band and depend on this pattern band play the music. The pattern or gesture perform by the conductor are require to be more precise and accurate to play the correct musical node. Because of this a conductor has to practice more and more to draw correct pattern. There is always requiring a teacher or a senior music moderator to judge wrong and right pattern drawn by beginner conductor. It is not always possible to observe each and every student whether they drawn correct or wrong pattern. To overcome this problem we proposed a new technique called Dynamic Time Warping for musical pattern detection. We have implemented this concept on basic three pattern detection. In this method we have used a camera to detect the hand movement and then apply DTW algorithm to detect the correct pattern drawn by the beginner conductor.**

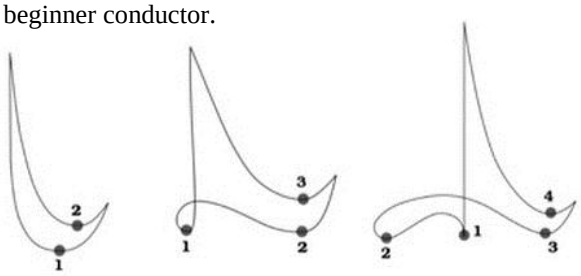
Index Terms: Distance learning, dynamic time warping (DTW), meter-mimicking, musical conducting, pattern recognition.

INTRODUCTION

Music performed by the western musical groups often requires a conductor to provide synchronization between the various instruments during the play that is a person whose gesture play the role to provide synchronizes. In music, conductor's body movement is used as resource to make decision and to play the correct musical node. Conductor can make their own set of pattern to perform but make sure it is understandable to the other to play the music. There are certain rules for how to draw the musical pattern to provide a good conduction amongst the different instruments while playing music.

At the initial stage of learning conducting technique is the study of basic patterns related to music time signature. The conduction of music measure is realized by the flow of the body generally the flow of hand in an organized pattern. These movement patterns are periodic and may be grouped into different sets of beats. The most common and usual type of music pattern are the basic three paterrn i.e. Duple, Triple and Quadruple as shown in fig1. Each of these patterns has a characteristic path drawn by the conductor hand and represent two, three or four bit of music.

The study of conductor at the beginning requires a music teacher to correct the student when they draw wrong patterns. It is not always possible to observe each and every student whether they drawn correct or wrong pattern. Some student may learn fast or some student may require some precise amount of time to learn each pattern to draw correctly. Teacher's job becomes tedious if a student requires considerable time to learn the basic pattern. There is a need to have a system who can predict whether the drawn pattern is right or wrong. To overcome this problem we proposed a new technique called Dynamic Time Warping for musical pattern detection. We have implemented this concept on basic three pattern detection. In this method we have used a camera to detect the hand movement and then apply DTW algorithm to detect the correct pattern drawn by the beginner conductor.



II. BACKDROP

Different type of methods and systems are implemented to detect the musical pattern. in early days the music pattern detection is done with the help of sensors like accelerometer and some infrared batons etc. some of the method used RGB-D camera to detect the body movement in each frame and from that they

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, March 2017

detect the pattern whether it is duple, triple or quadruple.

An accelerometer based digital pen to recognize handwritten pattern and gesture or digits is proposed by [2]. In this method they used an accelerometer connected to microcontroller to get the points of hand movement and send all this accelerometer data via transmitter to the computer to detect the correct pattern. In this method they used Trajectory recognition algorithm for the matching process.

An another accelerometer based gesture recognition method is given in [3] in which they comes with a problem in gesture recognition using the theory of random projection and proposed a solution over their by using the combination of two algorithm that is DTW (Dynamic Time Warping) and Affinity Propagation algorithm. They also shows that one-nearest DTW algorithm is not always sufficient to show the correct match for the recognition therefore they formulate the recognition problem as l1-minimization problem after projecting all gesture trace onto the same lower dimensional subspace.

[4] Present a generalized time warping algorithm which is the extension of the DTW algorithm for temporal aligning multi-mode sequence for multiple subject performing similar activity. They highlighted the three major drawback of DTW algorithm related multidimensional and large sequence of dataset. the proposed three points of GTW to overcome the problem of DTW are as 1) GTW gives a feature weighting layer to acquire different modalities. 2) GTW extends DTW by allowing more flexible time warping as collection of monotonic function. 3) Unlike DTW that typically incurs in quadratic cost.

A gesture recognition system that uses 3-axis accelerometer to sense the hand gesture. The main idea of this implementation is that it is based on the DTW algorithm for the pattern recognition. A pattern trace by the hardware is given to system which compares the trace pattern with the dataset and produce difference using DTW algorithm. If the difference is minimum then that pattern is match with that pattern [5].

Human action recognition is the most difficult task as it consider number of things while tracing and recognizing the movement. A new developed approach using depth sensor gains lots of popularity as it senses the human joints. A gesture trace by the depth camera may not produce correct trace as it often contain noise data. It produces wrong trace if serious occlusions appear which increases the intra class variations in the actions.[6] present an action model to represent each action to capture intracless variation. In addition to this they also proposed new approach for easy depth data.

A good recognition system for words from the combination of three handwriting recognizer is given in [7]. The main approach of this method is the HMM based recognizes which consist of dynamic and contextual information for better representation of handwriting unit. They also introduce a state-tying process based on contextual modeling.

[8] Gives a continuous real-time gesture following and recognition system. The proposed system is uses HMM algorithm for the recognition. The proposed system produces output by comparing the trace pattern with the store paten in dataset. The method based on a complete modeling of multidimensional temporal curves.

A UBS Virtual Maestro: interactive conducting system is given in [9]. This system is produce to fill the common people like they conducting the live music. In this they developed a WII remote which is operated on the accelerometer. When a person's moves his hand the data taken by the accelerometer is send to controller and system changes its audio and video tempo according to that hand movement speed.

[11] a study where participants perform "air drumming" gestures in time to rhythmic sounds. These movements are recorded, and the timing of various movement features with respect to the onset of audio events is analyzed. A novel algorithm for detecting sudden changes in direction is used to find the end of the strike gesture. We find that these occur on average after the audio onset and that this timing varies with the tempo of the movement.

III. PROJECT APPROCH

Here in this project we proposed a simple method for the music pattern detection with the help of simple camera and DTW algorithm.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, March 2017

Many methods are proposed related to gesture recognition in which most of the method or technique are based on sensors. Some of the new technique used RGB-D camera for the detection of body movements all the previous method which are sensor based or by using RGB-D camera are expensive as it involves highly accurate sensors or camera sensors. So we proposed or implemented a less expensive music pattern detection technique. This method is dome with the help of simple webcam of system. A continuous video is taken from the system cam. From the input video a single frame is captured one by one and various image processing operations is dome on the frame to detect the red color present in the frame. If a red color is present in the frame then it assigns a box to the frame and finds its centroid and gives X and Y coordinates values according to the frame scale in which we have open the camera frame. For the tracing of movement of hand we used this XY coordinates and store this XY points in txt file for a 100 frames according to hand movement. From this XY points we plot this points and find the movement of hand. This XY coordinates is then compare with the reference XY coordinates to predict whether the trace pattern is right or wrong.

For the detection of the correct pattern a test pattern is compare with the reference patterns which are stored in a folder. For the detection we used DTW algorithm which gives output in difference form. If two patterns that is test or pattern preset in train folder are same or near to each other then it gives less difference and if the two patterns are not same then it produce large difference. From this DTW output we classify which pattern is detected or whether the detect pattern is right or wrong

System Architecture

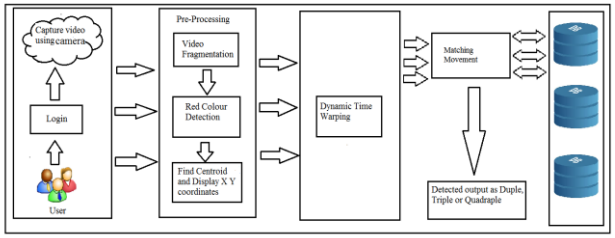


Fig. 2. System Architecture

Fig 2 shows the sytem architecture of our implemented system. As shown in fig it consist of vedio caprute part, a image pre-processing part and last is the matching part. In the first process a continous video is capture from the system cam or through the web cam. This continous video is consideras input to our system. In the next part or in process one by one a single image is capture from the input video and single frames are generated from them. On each single frame red color is detected after detecting a red color a box is generated surrounding to that box and its centroid is find. After getting its centroid its XY coordinates are displayed in video. This XY coordinates is then staored in text file for the comarision. In the matching section, stored XY coordinates is comapre with the dataset which also contain XY coordinates. The matching is done with the help of DTW algorithm and at the last it display whether it detected as Duple, Triple or Quadraple or wrong pattern.

A. MATHEMATICAL EVOLUTION

S={s, e, X, Y, }

Where,

s = Start of the program.

1. Login with System.

2. Create musical movement.

e = End of the program.

Find musical sign in ranking format.

X = Input of the program.

Input of this system is data of different musical sign movement images.

Y = Output of the program.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 4, March 2017

First we are going to extract the different type of images combine into video to find duple, triple and quadruple movement and apply for DTW and HMM algorithm for matching gesture images in database images.

X, Y U

Let U be the Set of System.

U= {Hm}

Where Hm are the elements of the set.

Hm, = Hand movement.

B. DATAFILE ASSESMENT

Dataset used in this system is the collection of different XY coordinates stored in text file. A dataset contain total 41 data samples which contains 12 Duple samples, 12 Triple samples and 17 Quadruple samples which is made by system admin on the basis of trial and error method.

OUTCOME ASSESMENT

Below fig shows the output of our implemented system on which an input is a continuous video stream from the system cam or through the web cam. From this video input a frame is generated and detects the red color. If red object id find then it finds its centroid and place the XY coordinate to the red object. By the help of this red object we detect the movement of hand and trace whether the pattern is Duple, Triple or Quadruple. Fig (a), (b) and (c) shows detected output as Duple, Triple and Quadruple respectively.

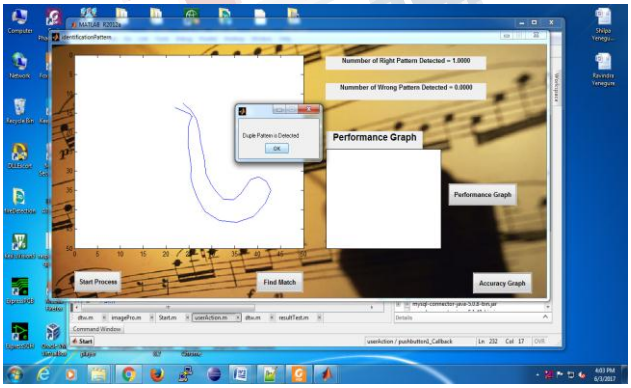


Fig (a)

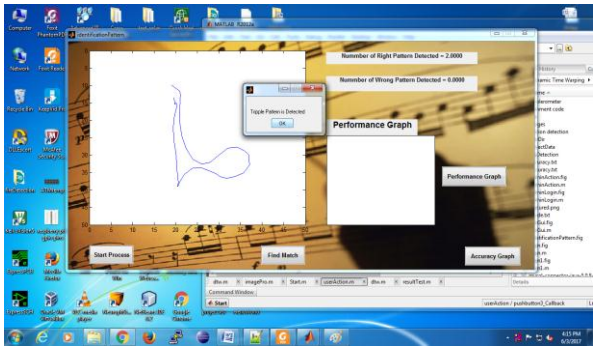


Fig (b)

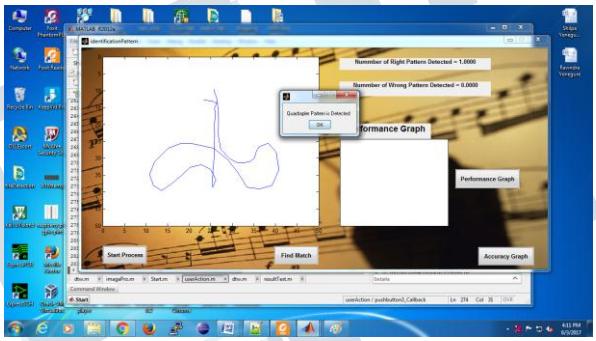


Fig (c)

I. CONCLUSION:

In this technique we have developed and implemented new music gesture detection with the help of simple system camera and DTW algorithm. In this system a movement of hand is detected by tracing a red color in frame and finding its coordinates and from this detected continuous XY points stored in txt file. This text file coordinates is then further compare with the dataset vales and final Duple, Triple or Quadruple detected pattern is displayed on the screen. This system is main implemented for the practice for the music learning student

.

REFERENCES:

[1] M. Mandanici and S. Sapir, "Disembodied voices: A Kinect virtual choir conductor," in Proc. 9th Sound Music Comput. Conf., 2012, pp. 271–276.

[2] J.-S. Wang and F.-C. Chuang, "An accelerometer-based digital pen with a trajectory recognition algorithm for handwritten digit and gesture recognition," IEEE Trans. Ind. Electron., vol. 59, no. 7, pp. 2998–3007, Jul. 2012.

[3] A Akl, C. Feng, and S. Valaee, "A novel accelerometer-based gesture recognition system," IEEE Trans. Signal Process., vol. 59, no. 12, pp. 6197–6205, Dec. 2011.

[4] F. Zhou, F. D. la, and T. Frade, "Generalized time warping for multimodal alignment of human motion," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2012, pp. 1282 –1289 .

[5] S. Hussain and A. Rashid, "User independent hand gesture recognition by accelerated DTW," in Proc. Int. Conf. Informat., Electron. Vis. 2012, pp. 1033–1037.

[6] J. Wang, Z. Liu, Y. Wu, and J. Yuan, "Mining actionlet ensemble for action recognition with depth cameras," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2012, pp. 1290–1297.

[7] A.-L. Bianne-Bernard, F. Menasri, R.-H. Mohamad, C. Mokbel, C. Kermorvant, and L. Likforman-Sulem, "Dynamic and contextual information in HMM modeling for handwritten word recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 10, pp. 2066–2080, Oct.2011.

[8] F. Bevilacqua, B. Zamborlin, A. Sypniewski, N.Schnell, F. Guédy, and N. Rasamimanana, "Continuous realtime gesture following and recognition," in Proc. 8th Int. Conf. Gesture Embodied Commun. HumanComput. Interaction, 2010, pp. 73–84.

[9] T. M. Nakra, Y. Ivanov, P. Smaragdis, and C. Ault, "The UBS virtual maestro: An interactive conducting system," in Proc. Int. Conf. New Interfaces Musical Expression, 2009, pp. 250–255.

[10] L. Dahl, "Triggering sounds from discrete air gestures: What movement feature has the best timing?," in Proc. Int. Conf. New Interfaces Musical Expression, London, U.K., Jun. 2014, pp. 201–206.

[11] L. Dahl, "Triggering sounds from discrete air gestures: What movement feature has the best timing?," in Proc. Int. Conf. New Interfaces Musical Expression, London, U.K., Jun. 2014, pp. 201–206.

[12] aggelos pikrakis, sergios theodoridis, dimitris kamarotos," recognition of isolated musical patterns using context dependent dynamic time warping", ieee transactions on speech and audio processing, vol. 11, no. 3, may 2003