Middle-East Journal of Scientific Research 11 (11): 1575-1583, 2012
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2012.11.11.1631

Adjustable Pedals by Gesture Recognition: A Novel Approach for User Interface in Automotive

Muhammad Haroon Yousaf and Hafiz Adnan Habib

Video and Image Processing Laboratory, University of Engineering and Technology Taxila, Pakistan

Abstract: This paper proposes the enhancement in the design of adjustable pedals in automotive by incorporating gesture recognition methodologies. The design addresses the three primary controls: Accelerator, Brake and Clutch. Mechanical input of driver's feet is deduced at location variant adjustable pedals in terms of feet gestures. Research work aims vision based technology for the feet gesture recognition. Machine vision algorithms are developed to recognize feet gestures. Low level features based upon background subtraction, connected component analysis, optical flow and motion trajectories are computed to infer spatial, symbolic, pathic and emotional information of feet gestures. Classification of motion trajectories is done by logic based analysis. Accurate classification of feet gestures proves a workable adjustable pedals framework based upon feet gestures. Research work recommends the substitution of conventional fixed pedals with user-defined gesture controlled adjustable pedals facilitating the drivers in terms of ease, flexibility, ergonomics and cost-effectiveness.

Key words: Gesture recognition · Adjustable pedals · Human machine interaction · In-car user interface

INTRODUCTION

In the recent era as the technology steps forward researchers and manufacturers are showing a wide inclination towards the use of Human-Machine Interaction techniques in their solutions and products. Human-Machine Interaction is the communication between humans and machines, whereas user-interface is the space where the interaction between human and the machine takes place. User interface provides a platform for the effective operation and control of the machine. It also provides feedback from the machine to assist in the decision making of operations. For any state of the art user-interface, proper ergonomics design must be considered for the ease of the users and to prevent the repetitive strain injuries.

Gesture recognition based techniques have been employed for majority of HMI applications e.g. sign language recognition [1], human machine interfacing using hand gestures [2], camera mouse control [3], controlling the secondary tasks in automotives [4], robotic control [5]. Since the gestures play an ample role in person to person communication, vision-based perceptual interfaces have made it possible to translate these actions in a machine understandable form - giving rise to Human-Machine interaction paradigm. The scale at which human gestures are recognized by machine is confined because vision algorithms are currently able to recognize only a set of gestures while humans can understand a wide variety of gestures with more accuracy. Vision-based perceptual devices are based on algorithms for real-time tracking, modeling and recognition of human gestures from video input. Such interfaces have been developed for a variety of applications, ranging from health care systems (i.e. non-contact interfaces for the operating room [6]) to video games. In majority of the applications [7-9] hand, head and body gesture recognition is performed to attain the goal. Feet gesture recognition is not too much explored in user-interface or application development, so it motivates the researchers to propose, design and develop new user-interfaces using feet gestures.

Research work is focused on user-interface of the pedals in the automotives. Currently almost all the automotives are provided with a fixed cluster of pedals to operate the clutch, brake and accelerator. This user interface lacks in ease and flexibility due to various reasons. Fixed location of these pedals raises problems for

Corresponding Author: Muhammad Haroon Yousaf, Video and Image Processing Laboratory, University of Engineering and Technology Taxila, Pakistan.

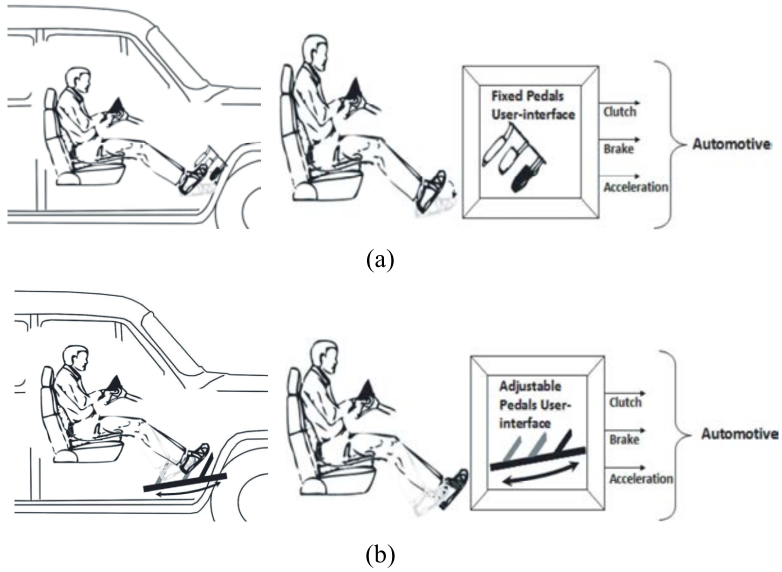


Fig. 1: Functional Comparison between Fixed and Adjustable Pedals

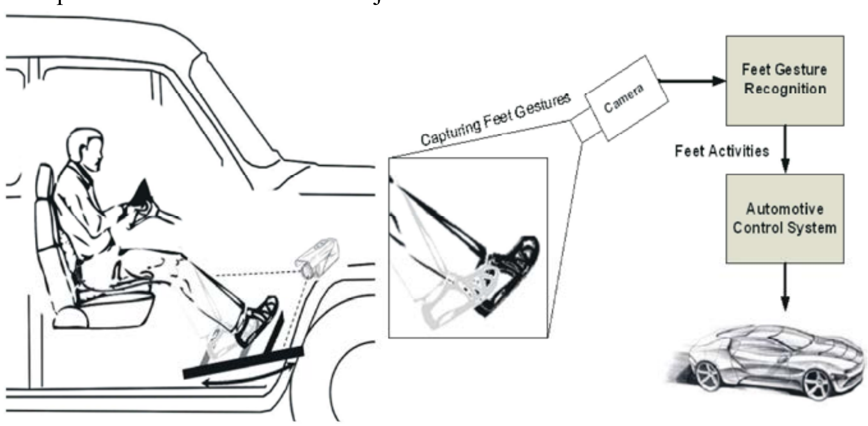


Fig. 2: Gesture Recognition Framework for Adjustable Pedals in Automotive

people with different heights to adjust themselves according to the position of the pedals. For example, even after the seat adjustment people with low height often complaint about their difficulties in accessing the pedals. Also people who have to drive a lot on routine basis like people whose primary job is driving, truck drivers driving on long routes etc. complaint about the fixed position of the pedals, due to which travelling turn out to be a real hectic job for them. Some other complaints have also been noticed by the people suffering from arthritis or some type of knee injury, due to which they feel distress to move their knees frequently and suffer with pain over a long period of driving time.

To deal with above mentioned issues, this research work is presenting a novel idea of gesture recognition based adjustable pedals in automotives. Adjustable pedals facilitate drivers to adjust the pedals according to their ease, comfort and requirement. Figure 1 illustrates the functional comparison between traditional fixed and adjustable pedals in automotives. Scheme suggests the floor-mounted pedals on a sliding panel having adjustable tendency according to the driver's desire. Pedals don't have any direct connectivity with the control system of automotive. Here pedals are only provided as a foot rest to give support to the driver from ergonomics point of view.

Feet gesture recognition is used as the foundation for the location variant adjustable pedals. Figure 2 explains the proposed framework for vision based gesture controlled adjustable pedals. To capture feet gestures, vision based technology is proposed. A camera is installed in the automotive under the dashboard.

Table 1: Possible Feet Activities

Foot ID	Possible Activities
Left Foot	Clutch Pressed
	Clutch Released
Right Foot	Accelerator Pressed
	Accelerator Released
	Brake Pressed
	Brake Released

Proposed scheme categorizes the activities performed by the driver e.g. pressing/releasing of acceleration, brake and clutch pedals. Also categorizes the emotional content made by gestures like, how much accelerator or brake has been pressed. Possible feet activities are categorized in Table 1. It recognizes the feet gestures and transmits the output signal to the control system of automotive for further necessary actions regarding acceleration, brake and clutch. Detailed architecture of the proposed system is discussed in section 3.

Related Work: Gesture recognition based user interfaces have already been proposed for the secondary tasks control in automotives. The automotive driving tasks are categorized in [10] into two types: Primary driving tasks and secondary driving tasks. Primary driving tasks are mainly consisting of controlling steering, gear and pedals cluster (accelerator, brake and clutch). In primary driving tasks, hands are utilized for steering control, gear change and other in-car controls on the instrument panel, whereas feet provides the output to the longitudinal controls i.e. accelerator, brake and clutch pedals. Whereas secondary controls may include the radio/CD player controls, HVAC switching and navigation system. Gesture recognition techniques have been implemented to operate the different in-car devices in an automotive. In [4], authors explored the use of gestures to control secondary tasks while driver is focused on driving. A detailed survey of hand gesture recognition based technology for secondary controls of automotives was discussed in [11]. To operate secondary tasks, vision and sensor based hand gesture recognition techniques have been employed. Major aim of these existing gesture based systems is to reduce the driver's distraction that helps in reducing the road accidents. Regarding the risk and driver's behavior towards the adjustable pedals, in [12], authors concludes the prediction of low accident risk using adjustable pedals. Talking about the gesture recognition based user-interfaces, head, hand and body gestures recognition based application have been proposed in [8]. A detailed survey of hand, head, face and body gesture recognition techniques and technologies is presented in [13]. But feet gesture recognition based applications still to explore. A visual keyboard using feet gesture recognition based on feet tracking to control the music-meta instruments is proposed in [14]. Foot tracking based system is developed to control a toy car in [15]. System efficiency was accurate and motivated us to propose a novel framework of adjustable pedals using feet gesture recognition for automotives.

Architecture of Proposed System: A gesture recognition based system demands proper gesture interface type, adequate gesture data collection techniques and technologies and most importantly good user acceptance rate. To achieve a better user acceptance rate for any gesture based system, natural gesture interface is preferred rather than symbolic or sign language gesture interfaces. Non-contact and dynamic gesture techniques proved for better system performance. As far as technologies are concerned, sensor and vision based approaches can be used, but the preferred technology has to be non-intrusive since this is the only one that is likely to gain wide spread user acceptance.

A generic vision based gesture recognition system can be viewed as proceeding from a data set to a high-level interpretation in a series of steps. The major steps involved are the following:

- Input video data
- Pre-processing of data
- Extraction of concise low-level features
- Gestures classification from low-level features

A gesture contains spatial, pathic, symbolic and emotional information. In gesture controlled adjustable pedals, spatial information tells about the location of feet gestures, pathic information meant for the path followed by the feet, symbolic information depicts what feet gestures are performing (pressing/releasing of accelerator, brake and clutch) and emotional information portrays how much feet have pressed/released any of the pedal. Methodology has been divided into following two sections.

Low-Level Feature Extraction and Trajectory Estimation: Prior to developing and deploying a product in the automotive, an experimental setup was established in the laboratory. Video sequence of driver's feet without any pedals support is captured using simple CCD camera. Because pedals only provide support to the driver, so they are not necessary to be the part of experimental setup. Figure 3 describes the algorithmic architecture for the feet trajectories estimation.

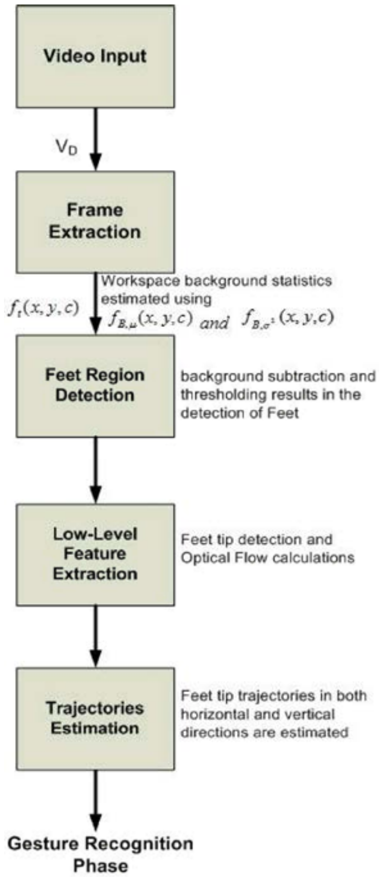


Fig. 3: Algorithmic Architecture for Feet Trajectories Estimation

Video sequence consists of 25fps; frame extraction has to be carried out continuously for analysis of every frame. Figure 4 shows the different poses of driver's feet representing pressing/releasing of pedals captured in laboratory setup. As a pre-processing step, background statistics were computed as in (1) and (2) when background is static means when feet are not present in the driver's workspace.

$$f_{B,\mu}(x,y,c) = \frac{1}{N}\sum_{t=1}^{N}[f_t(x,y,c)] \tag{1}$$

$$f_{B,\sigma^2}(x,y,c) = \frac{1}{N-1}\sum_{t=1}^{N}[f_t(x,y,c) - f_{B,\mu}(x,y,c)]^2 \tag{2}$$

Then quadratic difference is computed to generate the foreground image in the field of view. Foreground image is computed as:

$$f_{I,t}(x,y) = \begin{cases} 1 & \text{if} \quad f_{\Delta,t}(x,y) \ge \tau_b \\ 0 & \text{Otherwise} \end{cases} \tag{3}$$

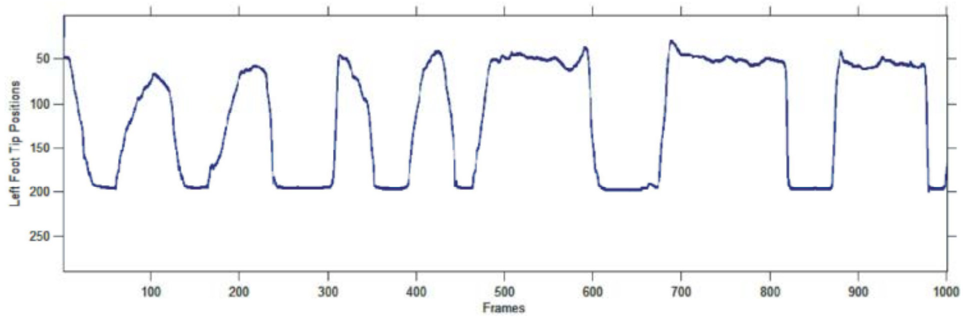
$$f_{\Delta,t}(x,y) = \frac{\sum_{c=1}^{3}[f_t(x,y,c) - f_{B,\mu}(x,y,c)]^2}{\sum_{c=1}^{3} f_{B,\sigma^2}(x,y,c)} \tag{4}$$

For the extraction of concise low-level features, background subtraction, optical flow, motion trajectories based techniques have been proposed in [16].

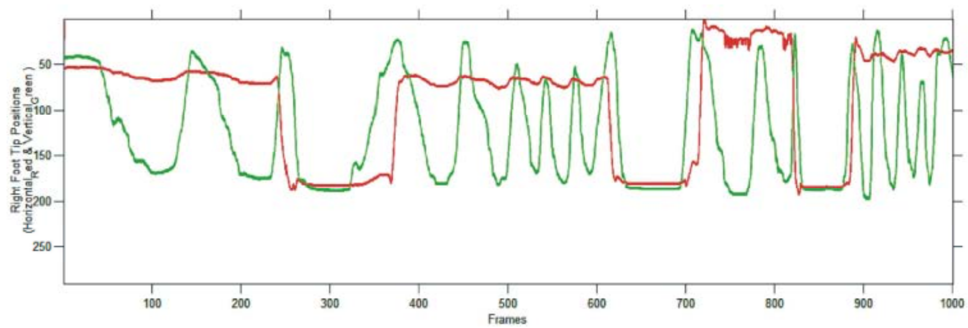
In gesture controlled adjustable pedals; motion trajectories are used as low-level features. Motion trajectories are estimated after a series of steps starting from background subtraction to feet tip tracking. Feet region detection is performed using the background subtraction technique. Connected component analysis and region labeling algorithms [17] are implemented on the subtracted image to find objects present in the driver's workspace. Feet regions are detected by assuming the largest regions in the workspace should be the feet. As feet tips plays major role in the pressing/releasing of pedals, so feet tips localization is mandatory. Feet tips are localized using region labeling and contour analysis of feet regions. Figure 5 shows the feet tips detected for above mentioned sequence of frames.

Fig. 4: Different Feet Poses for Pressing/Releasing of Pedals in Laboratory Setup

Fig. 5: Feet Tips Detection in Different Feet Poses



(a)



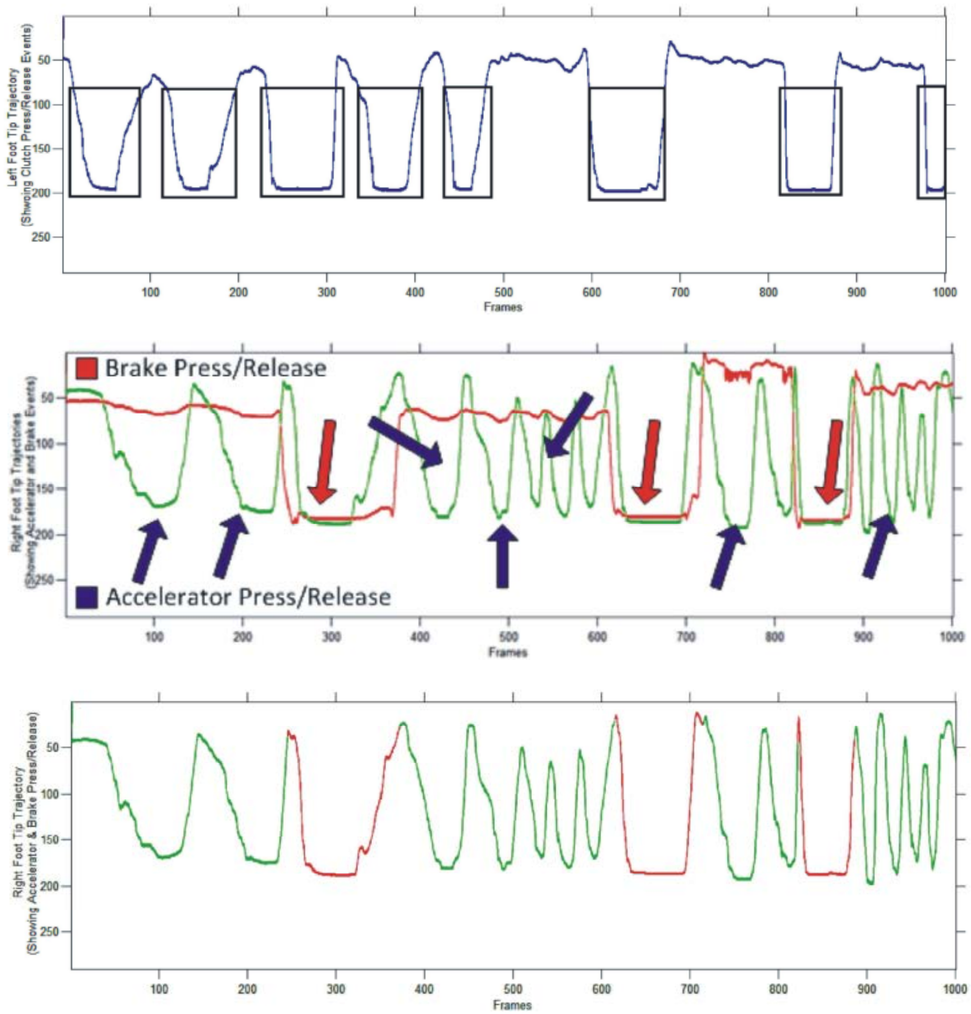
(b)

Fig. 6: Trajectories Estimated for Feet Tips (a) Plot of Left Foot Tip Positions (b) Plot of Right Foot Tip Position in both Horizontal and Vertical Direction

To estimate the feet tips trajectories, optical flow technique [18] is implemented for feet tip positions. As right foot is responsible for pressing/releasing of accelerator and brake, foot tip trajectories in both horizontal and vertical directions must be estimated. Whereas left foot is only meant to press/release clutch, so requires only vertical trajectory. Figure 6 presents the trajectories of both feet for a video sequence of 2000 frames. Trajectories are further scaled down here to 1000 frames just for the display purpose. Horizontal trajectory of right foot is represented by red color and vertical trajectory using green color. Both set of trajectories are further passed to the gesture recognition unit for recognition of actions performed by the feet.

Gesture Recognition: Major objective of gesture recognition phase is to accurately recognize the feet gestures in terms of spatial, pathic, symbolic and emotional context. Motion trajectories of feet tips are treated as the input to the gesture recognition phase. Human perception analysis of trajectories concludes.

- Left foot tip trajectory clearly indicates the pressing/releasing of clutch as described by the black rectangles in Figure 7a.
- Two trajectories generated for right foot deduced by red and blue color arrows in Figure 7b. Green trajectory means the press/release of pedal. Red trajectory decides the pedal whether it is accelerator



(a)

(b)

(c)

Fig. 7: a) Left Foot Trajectory Showing Possible Clutch Press/Release Activity b) Right Foot Trajectories Showing Accelerator and Brake Press/Release Activities c) Fusion of Right Foot Trajectories

or brake. Insignificance of red trajectory means right foot is pressing/releasing the accelerator showed by blue color arrows. Whereas significance of red signal shifts the control from accelerator to brake showed by red color arrows.

- Figure 7c shows the fusion of right foot trajectories. Green and red color of trajectory depicts the accelerator and brake activities respectively, performed by the driver's right foot.

For the machine vision perspective of these trajectories, supervised learning approach is adopted for the feet gesture recognition. Algorithm designed for the training and classification of trajectories is described in Figure 8.

Initially feature vectors comprising of the key spatial, pathic and probabilistic information are generated from the trajectories. If T_r represents the trajectory data then two components of feature vectors for K number of samples are calculated for both feet as eq. (5) and (6).

$$V^1 = \sum_{t=1}^{k} (T_r(t) - T_r(t-1)) \quad (5)$$

$$V^2 = \sum_{t=1}^{k} (\delta_d(t) - \delta_u(t-1)) \quad (6)$$

$$\text{Where } \delta_d(t) = \begin{cases} 1 & \text{if } (T_r(t) - T_r(t-1)) > 0 \\ 0 & \text{Otherwise} \end{cases}$$

$$\text{Where } \delta_u(t) = \begin{cases} 1 & \text{if } (T_r(t) - T_r(t-1)) > 0 \\ 0 & \text{Otherwise} \end{cases}$$

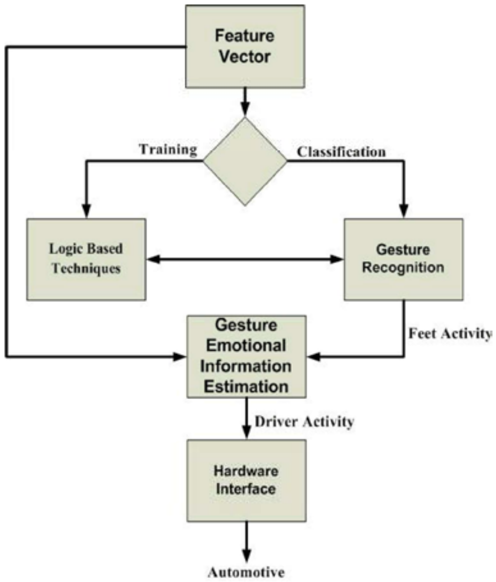


Fig. 8: Feet Gesture Recognition Classifierc

As the decision making in driving activity is dependent on the temporal data. Because gesture exists in both space and time - it's a spatiotemporal pattern. It is reasonable to assume that the temporal length of a gesture will vary amongst drivers. Thus we use a logic base analysis [16] to classify a gesture in the recognition system. Logical analysis takes input from the feature vectors. It classifies the feet activity on the basis of up-down sequences in feature vectors for respective foot. Logic based analysis decides the current activity of the feet i.e. pressing the clutch, brake or accelerator. Once an activity is identified, now it is the time to measure the emotional content of the activity. To estimate the emotional information, logic based classification is adopted using the feature vectors. Fusion of symbolic and emotional information of gestures decides about the feet activities, thus recognizing the feet gestures. Outcome of the classifier is transmitted to the hardware interface which will interact with the control system of automotives.

RESULTS AND DISCUSSION

To develop a video database for this project in laboratory setup; total 20 videos for 20 different drivers were recorded. For 15 videos, classifier is trained and tested across various videos just like section 3 by ways of three-fold cross validation. Experimentation results for feet activity recognition based adjustable pedals are evaluated using two key parameters. Parameters include gestures information (spatial, pathic, symbolic and emotional) retrieval and the feet activity recognition ratio. Activity recognition completely depends on the feet gesture information retrieval. Feet gesture recognition system for adjustable pedals is tested for 10 different drivers in the laboratory setup and results are collected. Figure 9 describes the experimentation results in terms of above said parameters. Gestures information retrieval for left foot is almost 100 % in all contexts resulting in 100 % accuracy for clutch press/release activity. Whereas, for accelerator and brake activities, right foot sharing influenced the emotional content of the gestures due to spatial disorder. This lack of efficiency in spatial and emotional information retrieval for right foot results in 94%, 96% accuracy results for accelerator and brake respectively. Automotives with automatic control functionality where clutch pedal do not exists; it becomes easy to apply this methodology. Because both feet can be reserved for each of the activity accelerator and brake resulting in the improvement of efficiency in feet gesture based adjustable pedals. Experimental results may differ in the real time scenario due to illumination and environment constraints. But this can be handled using state of the art vision technologies and specialized equipment.

Conclusions and Future Work: Research work presented in this paper proposed a novel user-interface of adjustable pedals in the automotives using feet gesture recognition. It recommends the replacement of traditional mechanical system of fixed pedals with gesture driven location variant adjustable pedals. Vision based feet gesture recognition is implemented and proved to be cost-effective and easy to deploy in automotives. User-interface was employed using recognition of natural, dynamic, non-intrusive feet gestures which results in a good user acceptance rate. Trajectory estimation and gesture classification was carried at their best to accomplish the goal. Outcome of feet gesture recognition can further be passed to drive-by-wire and brake-by-wire technologies which have replaced the traditional mechanical control systems with electronic control system in the automotive industry. Gesture recognition based adjustable pedals not only addresses the driver's complains about stress, uneasiness, less-accessibility and fatigue, but also the connectivity issues of the pedals with the mechanical system of vehicles by providing transmission of mechanical input in terms of electric signals achieved through gestures recognition. Feet gesture recognition can further been used in the context of automotive

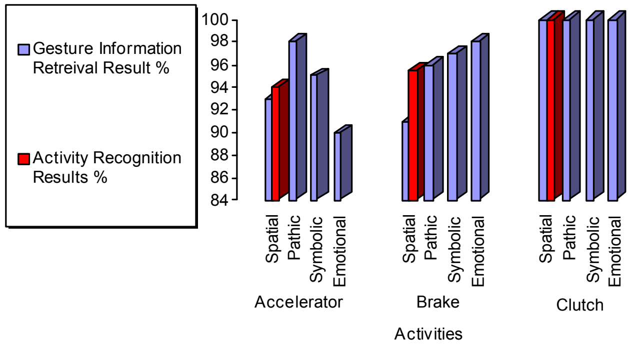


Fig. 9: Experimentation Results in Laboratory Setup

user-interfaces for the development of driver training systems. In future, a prototype system is planned to develop and deploy in the automotive to have the practical feedback from the driver as well as to judge the efficiency of the system. In a nut shell research work motivates the top automotive industries to explore the role of vision based gesture recognition systems for secondary as well as primary tasks in automotives.

ACKNOWLEDGMENTS

This research work is supported and funded by Directorate of Advance Studies, Research and Technological Development, University of Engineering and Technology Taxila, Pakistan. The authors are also thankful to postgraduate research fellows affiliated with the *video and image processing laboratory* for their suggestions and support in this research work.

REFERENCES

1. Bowden, R., A. Zisserman, T. Kadir and M. Brady, 2003. "Vision based interpretation of natural sign languages" in Proc. 3rd Int. Conf. Comput. Vis. Syst., Graz, Austria, Apr., pp: 391-401.
2. Yang Liu and Yunde Lia, 0000. "A robust hand tracking and gesture recognition method for wearable virtual interfaces and its applications" 3rd International conference on Image and Graphics (ICIG'04).
3. Luca Lumbardi and Marco Porta, 2007. "Adding Gestures to Ordinary Mouse Use: a New Input Modality for Improved Human- Computer Interaction" 14th International Conference on Image Analysis and Processing (ICIAP 2007).
4. Micah, Alpern and Katie Minardo, 2003. "Developing a Car Gesture Interface For Use as a Secondary Task" CHI', pp: 32-33.
5. Xiaoming Yin and Ming Xie, 2007. "Finger identification and hand posture recognition for human-robot interaction" Image and Vision Computing, 25: 1291-1300.
6. Graetzel, C., T. Fong, S. Grange and C. Baur, 2004. "A non-contact mouse for surgeon-computer interaction", Technol. Health Care, 12(3): 245-257.
7. Althoff, F., G. McGlaun, M. Lang and G. Rigoll, 2003. "Comparing an Innovative 3D and a Standard 2D User Interface for Automotive Infotainment Applications", Paper. Munich: Technical University of Munich.
8. Kai Nickel, Edgar Seemann and Rainer Stiefelhagen, 0000. "3D-Tracking of Head and Hands for Pointing Gesture Recognition in a Human-Robot Interaction Scenario" Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04).
9. Tabassam Nawaz, Muhammad Saleem Mian, Hafiz Adnan Habib, 2008. "Infotainment Devices Control by Eye Gaze and Gesture Recognition Fusion", IEEE Transactions on Consumer Electronics, 54(2): 277-282.
10. "Interacting with the Cars", 2005. IEE Computing and Control Engineering, pp: 31-33.
11. Pickering, Carl A., Burnham, Keith J. Richardson and J. Michael, 2007. "A Research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction" 3rd IET conference on Automotive Electronics" pp: 1-15.
12. Young, D.E., R.A. Schmidt, T.J. Ayres and D. Trachtman, 2001. Risk and driver behavior with adjustable pedals. In Proceedings of the Human Factors and Ergonomics Society 45th annual meeting. Santa Monica, CA: Human Factors and Ergonomics Society, pp: 1656-1660.

13. Sushmita Mitra and Tinku Acharaya, 2007. “Gesture Recognition: A Survey” IEEE Transactions on Systems, Man and Cybernetics -Part C: Applications and Reviews, 37(3): 311-324.
14. Fre´de´ric Jean and Alexandra Branzan Albu, 2008. “The visual keyboard: Real-time feet tracking for the control of musical meta-instruments”, Signal Processing: Image Communication, 23: 505-515.
15. Sameen Shaukat, Muhammad Haroon Yousaf and Hafiz Adnan Habib, 2010. “3rd International conference on Advance Computer Theory and Engineering” pp: v5-v6.
16. Pavan Turaga, Rama Chellappa, V.S. Subrahmanian and Octavian Udrea, 2008. “Machine Recognition of Human Activities: A survey” IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(11): 1473-1488.
17. Fu Chang, Chun-Jen Chen and Chi-Jen Lu, 2004. “A Linear-Time Component-Labeling Algorithm Using Contour Tracing Technique“ Computer Vision and Image Understanding, 93(2).
18. Horn, B.K.P. and B.G. Schunck, 1981. “Determining optical flow,” Artif. Intell., 17: 185-203.
19. Rabiner, L.R., 1989. “A tutorial on hidden Markov models and selected applications in speech recognition,” Proc. IEEE, 77(2): 257-285.
20. Richard Duda, Peter Hart and David Stork, 0000. “Pattern Classification” Book, 2nd Edition.