PERFORMANCE ISSUES OF GESTURE BASED COMMUNICATION SYSTEMS

[1]JAY KIRUTHIKA, [2]SOUEHIL KHADDDAJ, [3]RAUL HERRARA

[1,2,3]Kingston University
E-mail: jay.kiruthika@gmail.com, s.khadddaj@kingston.ac.uk, R.Herrera@kingston.ac.uk

Abstract- Gesture based communication is relatively adopted in areas like medicine, sports, inventory etc. Human factors are involved in the algorithms to determine the nature of the gesture and the computer tool translating it should be an autodidactic tool. Traditional systems use touch screens of any touch points to communicate with the system either via a control panel or following a user journey. Gesture communication varies with the traditional in the sense that human communication is transferred to technical form of algorithms for a system to decipher them. There are few issues relating to the performance such gesture communication systems namely accuracy, transfer of human knowledge to computer and various other observations needs to be addressed. A case study in the form of a simple experiment relating to a basic two handed gesture and human interaction in a 3D database is analyzed in this paper.

Keywords- Gesture based communication; Human computer interaction; Performance of Gesture based communications systems.

I. INTRODUCTION

The increasing ability of systems to decipher human communication is a complex process. The algorithm to translate the gesture to technical form has an element of psychology and should be able to differentiate between an agitated human behavior and a gesture which can be identifiable. The types of gesture communication can be classified into vision based, movement based, hand based or a combination of them. The technical obstacles or challenges are reliability, speed, interaction, intuitiveness and gesture spotting. This is especially true in gesture controlled games, medical assisted technologies for a faster response. The algorithms are designed to be robust and efficient where recognition performance is given preference and are focused on tracking and image processing. It mainly depends on user configuration and feedback which is especially vital for tools which are autodidactic. The system challenges like accuracy, performance, learnability are discussed next.

II. CHALLENGES

A. Accuracy/Performance

The main challenges for the performance measures of such gesture based systems are the detection recognition and tracking. It depends on the rapid motions and problems arise due to hand shape, variable lighting conditions, skin color, hand size, gender and the rate of motion of the gestures. Detection is a complex process as it has to identify or isolate the gestures from the surrounding environment and the motion detector has to be in the line of vision of the system not necessarily the user to record the movements and decipher them. A good accuracy can be attained by trail by error method or user feedback and depends on the application involved. There may be just few gestures necessary in applications like in a smartphone when the user has to use a hand wave to navigate the screens but complexity arises when recognizing a hand and finger movement. The learning-rate of the system varies from individual users and the algorithm designed should take this into account when calculating accuracy of the movements which are more or less 3D. Technologies like soft Kinect, Sony PSP, Nintendo (DS and Wii) include gesture recognition in their consoles which are based on waving and hitting the target and they are not dependent on the user's expressiveness. For example in Nintendo Wii fit games, a user can just sit on a sofa and shake the controller to fool the system in thinking that the user is doing a running rep whereas that might not be the case. The level of accuracy depends on the complexity of the system and cost/budget assigned to building the system.

B. Learnability

Learnability is one of the most challenging tasks any developer face when building an autodidactic system. The research when learning about human gesture identification should be wide and not set for individual users. Especially as users the gesture patterns differ according to the size, color and type. The control applications should be able to understand user cognitive skills to a certain extent in order to isolate gesture from moving. The recent video games learn from various users' ability to perform a task and stores in big data so as to learn from them and improve its accuracy.

C. Intuitiveness

The developers who develop the system should have a clear cognitive association with the users to develop the system and should clearly define the boundaries of functionality. For example stop may be an open palm; waving the finger to the object may be the direction where the said object is supposed to move and a finger pointer may be used to select the object in question etc. This is especially true when using multi hand or two handed gestures. This particular entity in discussion varies from individual to individual depending on cultural background, social interaction and experience. A simple gesture of nodding the head from left to right might be interpreted in some places as yes rather than a no.

There should be an element of clarity when planning such types of gesture recognition. One such option is to let user be given a selection of gestures for each type of activity so as to improve the 'ease of use' or usability function in the said system. Comfort of the user should also be taken into account when designing such systems. Repetitive use or strain of muscles or tissues will result in RSI (repetitive strain injury). This will be caused by repetitive tasks, forceful vibrations or exertions of the muscle or tissue concerned which are predominantly prevalent in hard gamers and children.

D. Re-Configurability

As mentioned earlier, users come in all sizes and the gesture based system should be able to cater to all types of users irrespective of sized to reach a wider audience. Most of the system has a limited vision range or sensor/ movement range to detect any motion which may be called as the interaction space. The designers should be able to adjust their systems to reach an average user.

E. Ubiquity and wear- ability

For a mobile based gesture recognition system, every aspect of daily activity irrespective of the location and context should be added into the functionality when developing the system. This is especially true with the current "wearable" technology. It should be spatially versatile and requires self-calibration.

F. Gesture Spotting

This can be named as a cue selection and is the ability of the system to differentiate between a gesture and moving. It is unreliable to define a start and end of a gesture. A defined time space between start and end of a gesture is also not possible to define as users differ based on gender and size. One option of recognizing the end of a gesture is to recognize just a portion of hand gesture movement and accept them by signaling via a ping in the controller to notify the user that the task has been completed or the gesture motion has been completed. The latest gesture recognition systems developed for Scotland Yard incorporates streamlining live videos of the underground traffic to isolate a passenger behaving abnormally. This requires highly tuned and specialized algorithm to define what is normal and abnormal. The algorithms should be able to have knowledge of typical body movement and characteristics in order to isolate non-behavior. This includes facial recognition, gesture recognition like gesture position, posture etc. There are challenges while developing such systems to differentiate between un-intended movement, resting and suspicious activities.

G. User Interaction and feedback

User interaction and feedback is one of the crucial components when building such gesture systems. User interaction may be limited to a finger or hand. These 'indicators' should be configured before-hand and the user feedback results should be fed into the system to improve the accuracy of the system. It depends on the depth of the gesture, color, appearance of the users said indicators. Popular methods to identify user interaction will be statistical methods, rule based methods, active shape models and shape context. Markov model can be used for classifying hand trajectory into dynamic gestures. Other gesture segmentation, spotting methods are dynamic time warping, Hough transform, mean-shift & Camshift and baynesian approaches.

III. 3D INTERACTION

3D interaction with computers is incorporated in gaming applications, printing, modelling etc. The system which supports it should be robust and reliable to launch such applications. Supporting live interaction with sensors and camera functionality to translate operational space of a user to 3D entity in computer is a complex process. Recent 'wearable' interactive devices like Google glass reflect back the live feed and are able to identify the user's geo location and are able to supply information. The performance of such systems needs to be constantly upgraded and modified to be able to cope with the complex processes.

IV. CASE STUDY

In this paper a case study based on a 3D interaction using Microsoft Kinect for hand gesture is examined. A 3D database programming environment based on 3D gesture interactions utilizing Kinect as an interface is presented. It is designed for 3D database to improve the development speed and the level of perception. A SQL query is designed for a 3D database to select data associated with gestures. The research is on time taken and the gesture deviation between the genders. This experiment was done only to a selected number of users and is not reflective of a real live environment.

H. 3D databases

Multidimensional databases are used to store complex structure. Recently, there is research done in 3D databases to search/retrieve data from the databases. Frati proposed an approach to perform the detection of the fingers where each end point represents a finger. But there is a limitation of degrees of freedom and lack of gesture detection. This is included in this experiment with the range or line of vision specified to be 60 to 90 centimeters between the user and the Kinect device for optimal hand and finger detection. The user's movements should be parallel to the device.

I. User Interaction

The experiment involved in this case study is to measure performance of pure hand gestures and no other device is involved to perform tasks. The actions are based on the number and position of fingers and palm. These actions are interpreted to equivalent tasks which are different and produce equivalent output from the system. This a two hand gesture model with finger selection included. The main interaction is divided into three major groups namely movements, selections and executions.

Movements: It translates to changes in position or orientation of the 3D graphic elements in the experiment

Selections: The selections are related to choosing or highlighting specific 3D elements or components and parts of them.

Executions: Interaction related with triggering an action could be result of a combination or a specified individual action.

Figure 1. 3D database model for the experiment

The faces of the 3D database have been designed to display details of the user in one and the next one contains age, weight etc. and the top are obviously to close the interface. It is kept simple to avoid confusion and complex functionalities.

There are 3 types of interactions:

- Rotation: The cube can be rotated from left to right and vice versa. To make the interface more intuitive only two faces of the cube are accessible, the front face (with the patients' personal information) and the right-side (with the weight/month information). This action is performed by keeping the left hand open and the movement action is performed by moving one finger of the right hand from left to right or vice versa, depending on the face that the user wants to see. There is no rotation on the horizontal axis.
- Selection: Users must display two fingers of their left hand and placing the moving finger over an element to select it.
- Clicking: The user must remain in selection mode (two fingers of the left hand) to access this mode. In order to perform the clicking, the indicator finger must be placed over a selectable element and "push" (move forward, towards the screen).

In this implementation the selectable elements of the cube are the button "close", the IDCode, Name and gender which are in the first face but to select the column area a finger selection to the column header will suffice. This experiment is limited to selection of all the rows and is not restricted to all columns and all rows. The interaction screen communicates with the user with the selected option so the user is aware of what is happening with the system. An example of the screen shot is in Figure 3, Figure 4 and the series of events for the selection and interaction is shown in Figure 5. This includes palm movement, finger gestures and pointed finger selection. For each task the user is timed.

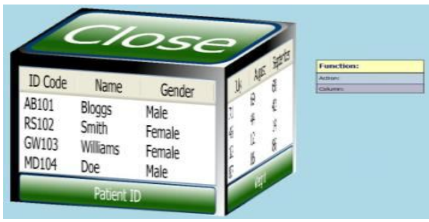
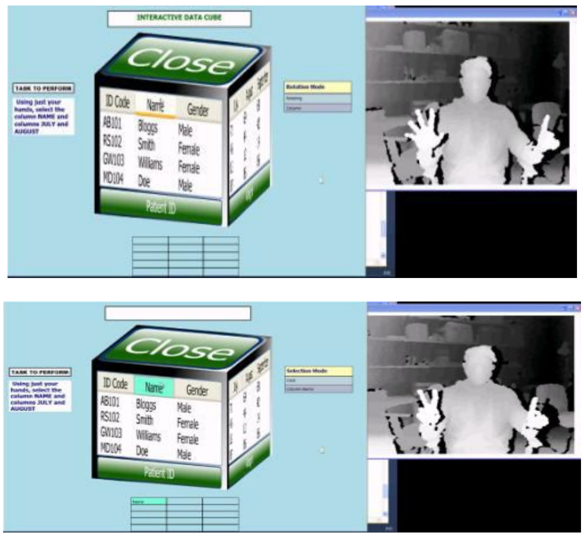


Figure 3 – Functionality selection

Figure 4. Interface after a task is performed.



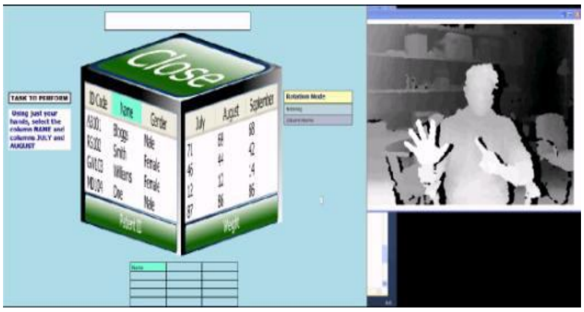


Figure 5: Series of user interaction and gestures.

The experiment was conducted on both genders and restricted to selected individuals with knowledge of computers but not necessarily knowledge of SQL or other database functions. At the end of the experiment each user was asked to fill in a questionnaire for feedback and knowledge of SQL etc. and a usability metrics is used to calculate the 'ease of use' functionality of the system. A scaling system as proposed by Lewis is used to measure the usability ranging from 'ease to use' (scoring 1) to 'difficult to use' (scoring 5).

V. USABILITY QUESTIONAIRE

1. Section 1

Was the interaction easy to understand?
Was it easy to manipulate?
Is the navigation system intuitive?

Section 2 - How would you rate:

The Interface?
The Performance?
The functionality?
The objective achieved?
The user experience?
The hand gestures selected?

Section 3- SQL interface compared with the proposed Visual approach

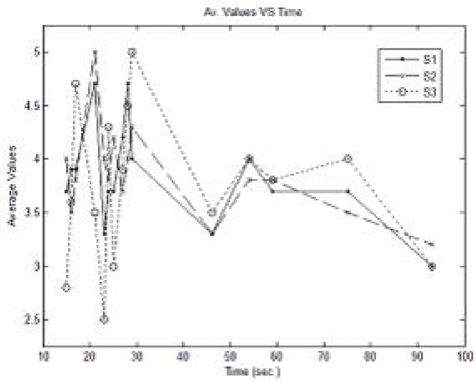
The selection is easier than SQL?
The task is more intuitive than SQL sentences?
Is it easier to learn the proposed visual approach than SQL?
Is the task faster to perform than with SQL?

SECTION 1	Q1	Q2	Q3
Average	4.5	3.2	3.9
Deviation	0.69	0.62	0.81

Table 1: Average value estimated from the questionnaire.

	Male	Female	SQL	NoSQL
Av.Time	33.47	30.2	33.57	30.5
Av.Age	36.07	31	33.57	36.33

Table 2 : Average time and age of users involved in the experiment.



This experiment clearly shows that there is a difference in the users' response time depending on the age and gender. The user's difficulty level is also measured using their knowledge of SQL. This experiment is a perfect example to gauge user's performance or response rate when developing such systems. The experiment results will vary for users who use applications where a timed response is needed. This does not measure such entities.

CONCLUSION

There are a number of variable factors when looking at the performance aspects of a gesture based systems, as human interaction should be included as a measurement entity. It should also consider entities like the performance time of the system, the time taken by a user to perform a single or set of tasks and the performance of the application. A summation and average deviation over the time will give a clear picture of the performance of such systems. Future work will include optimizing SQL query to retrieve/ display data or sets of data and to measure performance based on user interaction and SQL query complexity.

REFERENCES

[1] Herrera-Acuña, Raul A., Argyriou, Vasileios and Velastin, Sergio A. 2013. Graphical interfaces for development exploiting the third dimension using Kinect. In: 9th International Conference on Intelligent Environments - IE'13; 18-19 Jul 2013, Athens, Greece

[2] Belongie, S., Malik, J., and Puzicha, J. ,2002.Shape matching and object recognition using shape contexts. Transactions on Pattern Analysis and Machine Intelligence 24, v24 ,p509–522.

[3] Cootes, T.F. and Taylor, C.J. Active shape models: 'smart snakes.' In Proceedings of the British Machine-Vision Conference (Leeds, Sept. 22–24). Springer, Berlin, 1992, p266–275.

[4] Freeman, W. and Roth, M.,1995. Orientation histograms for hand-gesture recognition. In Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. (Zurich,.

[5] Nickel, K. and Stiefelhagen, R. 2007. Visual recognition of pointing gestures for human-robot interaction. Image and Vision Computing 25, v12, p1875–1884.

[6] Patel, R. and Roy, D.,1998. Teachable interfaces for individuals with dysarthric speech and severe physical disabilities. In Proceedings of the AAAI Workshop on Integrating Artificial Intelligence and Assistive Technology (Madison, WI, July 26–30), p40–47.

[7] Rauschert, I., Agrawal, P., Sharma, R., Fuhrmann, S., Brewer, I., and MacEachren, 2002.A.M. Designing a human-centered, multimodal GIS interface to support emergency management. In Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems (McLean, VA, Nov.8–9). ACM Press, New York, p119–124.

[8] Schlömer, T., Poppinga, B., Henze, N., and Boll, S. Gesture recognition with a Wii controller. In Proceedings of the Second International Conference on Tangible and Embedded Interaction (Bonn, Germany, Feb. 18–20). ACM Press, New York, 2008, 11–14

[9] K. Stefanidis, G. Koutrika and E. Pitoura, "A survey on representation, composition and application of preferences in database systems", ACM Trans.Database Syst, 2001. vol. 36, no. 4.

[10] Gelman, Andrew; Carlin, John B.; Stern, Hal S.; Rubin, Donald B. (2003). Bayesian Data Analysis, Second Edition. Boca Raton, FL: Chapman and Hall/CRC. ISBN 1-58488-388-X.

[11] Allen, J. G., Xu, Y. D., and Jin, J. S. (2006). “Object tracking using CAMShift algorithm and multiple quantized feature spaces.” In: Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing, Sydney, Vol. 36, pp. 3–7.

[12] J.R. Lewis, "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use", International Journal of Human - Computer Interaction 7-1(1993), 57-78.

[13] V. Frati and D. Prattichizzo, “Using Kinect for hand tracking and rendering in wearable haptics", IEEE World Haptics Conference (2011), 317-321.

[14] Junker, H., Amft, O., Lukowicz, P., & Tröster, G. (2008). Gesture spotting with body-worn inertial sensors to detect user activities. Pattern Recognition, 41(6), 2010-2024.

[15] R.Agrawal, A.Gupta and S.Sarawagi, "Modeling multi-dimensional databases", Proceedings of 13th. Int. Conf. On Data Engineering (ICDE)(1997),. 232-243.

[16] Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. Communications of the ACM, 54(2), 60-71.

★★★