Screen view compensation for eased reading in moving vehicle using sensor fusion

Dr. Satyanarayan Panigrahi
Associate Prof., School of Mechanical Sciences
Indian Institute of Technology (IIT) Bhubaneswar

Harleen Hanspal
School of Electrical Sciences
IIT Bhubaneswar

Rahul Kumar Kumawat
School of Mechanical Sciences
IIT Bhubaneswar

***Abstract*—Since the last decade laptops, smartphones and devices like kindle, etc. have become a very important part of our lives, being pervasive and extremely helpful in almost every activity of ours, among other things mainly due to their portability and mobility. In one of their applications i.e. reading while travelling in moving vehicles or while walking, these devices are subject to a lot of shaking and this makes reading very difficult, annoying and strenuous for the eyes. To address this issue, in this paper, we propose a method of adjusting the display on the laptops (or smartphones) so as to compensate for the device shaking and allow eased reading. The working prototype based on this method uses a 3-axis accelerometer + 3-axis gyroscope to decide the extent and direction of screen shift such that the eye of reader finds the content on the screen at roughly the same line of sight as it was and where the eye would expect it to be if the unexpected jerk or shaking was not encountered.**

I. Introduction

Reading in moving vehicles or while walking, where the screen is constantly shaking and moving is very strenuous, time consuming and frustrating for many. Studies have also cited it as one of the prominent reasons of motion sickness in vehicles and research has been undertaken to prevent motion sickness by providing comfortable and smooth viewing of the vehicle video displays. In our research, we tried to develop a screen view compensation system for the portable devices, which being prone to very unpredictable disturbances, jerks and vibrations make our task more challenging than for the fixed vehicle displays, necessitating extremely fast but smooth system response. On analysis, we find that more than compensation in terms of translation on screen, projective transforms in the screen view in response to the devices change in orientation can provide better reading and hence that is incorporated. A different approach of decomposing the acceleration values into linear and angular acceleration using the gyroscope and accelerometer data is tried as it is necessary to be able to measure and understand the type of disturbance being faced by the device at that moment and hence using 2 sensors help here to get better results. However, since this approach is difficult and requires experimentation, as in initial calibration, it cannot be easily employed by most people, therefore an alternate method using disturbance nature identification and classification using fuzzy logic is tried and the results discussed. The organization of the paper is as follows: In section I, the prior work done in this field is discussed and the need for the current study is established. In section II, the mathematically modeling of the system is explained. Section III elaborates our methodology and the experiments done on the problem and section IV substantiates the entire work with results.

II. Related Work done before

Till date, only a few papers and prototype softwares have been written/developed to achieve this goal of eased readability in mobile devices. One can refer to [1], where they propose and implement on an iPhone their compensation system named NoShake. The system is based on a physically-inspired mass damper model and the acceleration of the device is estimated using the tuning of the parameters of the impulse response of this model. The system is verified by user study and shows satisfactory improvement in readability in certain conditions. It does not give good results when a person is reading while holding the iPhone in his hands. They attribute it to the user expecting the bumps and automatically compensating their eyes for the motion. In [2], both accelerometer and gyroscope data is used as in our case but they are considering only the case when the person is reading while walking and not the various other situation as when in a moving vehicle, etc. because of which they can use Hidden Markov Model and position estimation to stabilise the screen, but this is not feasible in absolutely uncertain and uncorrelated motions like those in vehicles. Other than HMM, they have used PID control, which we have verified to be a good and necessary control algorithm for comfortable viewing, otherwise the compensated viewing becomes more jittery and unstable. Most other papers like [3],[4],[5] study the stabilizing of the video or camera feed which are mostly based on Image processing using optical flow, etc. which is slower than accelerometer-gyroscope sensor method for it to be done in real time and focus there is on the previously captured images and not on the displayed images.

III. Background on the problem

For a person sitting in a moving bus with a laptop in his lap, let us chose the moving bus as the frame of reference for specifying the persons and the devices movement. During a sudden jerk, both the person and the device displace from their mean position in a random manner. As per our empirical examination and analysis, our eyes are able to adjust themselves instantaneously to the change the person encounters in the jerk in an effort to try to keep his view focused on the same point of view as was before the jerk was experienced, but since the movement of the device is not known and is often much more dynamical and difficult to be predicted, the eye often loses its position on the screen and takes considerable time to relocate itself. However, given the former self adjustment of the eye to bodys movement, the screen compensation now clearly seems possible requiring just the negation of the effect of the devices movement in the frame of reference i.e. the bus. The focus can be put only on the device independently, not the device and persons relative movement with respect to each other. As has been mentioned in [7], human beings have a vestibulo-ocular reflex whereby the eyeballs turn reflexively in the opposite direction when the head is turned. That reflexive action readily enables vehicle passengers to maintain their gaze in a fixed direction even if their entire body pitches due to the vehicle's pitching motion or if the head pitches or turns relative to the upper body owing to the action of inertial force. That may explain the empirical results in [1], where NoShake proved to be the least successful in reading while holding the iPhone and walking, because people did not just control their own movement and hence their eye could compensate for that but they also held and therefore controlled the device and reflexively compensated for that too while that was being negated by the software hence making the compensation uncalled for.

IV. Mathematical modelling

For modeling the system, visualize the device screen as a perfectly frictionless body on a board attached to the device.

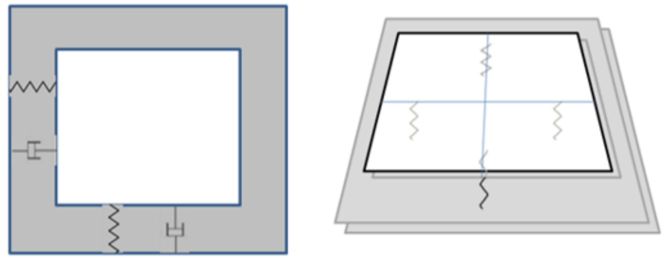
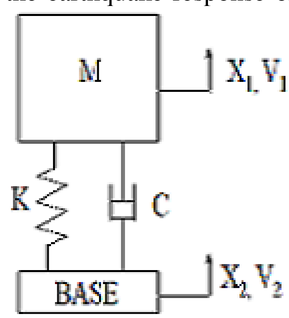


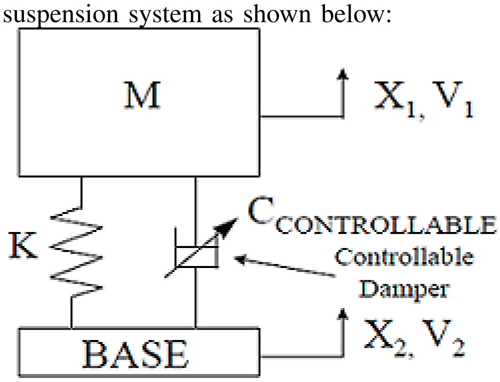
Fig.1- mass-spring-damper system with 6 degrees of freedom

Considering the linear accelerations and forces on such a body one can use the Base Excitation model to get the behavior of a vibration isolation system. The base of the spring (attached to the device) is given force, causing the mass to vibrate. This system is generally used to model vehicle suspension systems or the earthquake response of a structure.



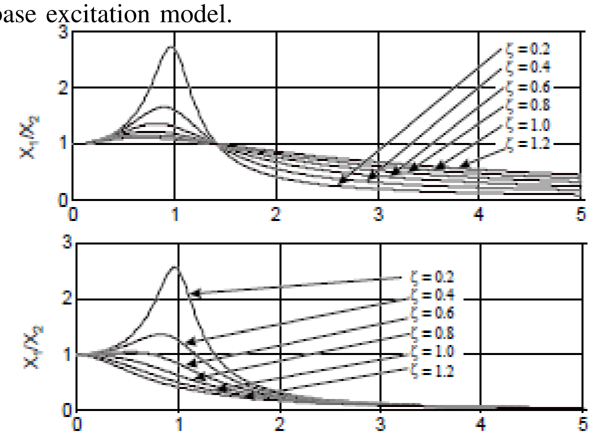
$$\frac{X_1}{X_2} = \frac{1 + j2\zeta_P(\frac{\omega}{\omega_n})}{1 - (\frac{\omega}{\omega_n})^2 + j2\zeta_P(\frac{\omega}{\omega_n})}$$

On plotting the transmissibility i.e. (X_1/X_2) as a function of the quantity (ω/ω_n) for various damping ratios for the suspension system shown above, we notice that at low passive damping ratios, the resonant transmissibility (around $w = w_n$) is relatively large while the transmissibility at frequencies above the resonant peak is quite low. The opposite is true for relatively high damping ratios, hence demonstrating the inherent trade-off that one has to deal with when using passive base excitation model, a model similar to the one which is being used in [1]. Therefore, to eliminate this trade-off between resonance control and high frequency isolation, we reconsider the configuration of the suspension system inspired by the Skyhook suspension system as shown below:



$$\frac{X_1}{X_2} = \frac{1}{1-(\frac{\omega}{\omega_n})^2 + j2\zeta_{SKY}(\frac{\omega}{\omega_n})}$$

where, in this case, ζ_{SKY} is the ideal skyhook damping ratio. In this model, as in the passive case, as the skyhook damping ratio increases, the resonant transmissibility decreases but increasing the skyhook damping ratio does not increase the transmissibility above the resonant frequency. The graphs below show transmissibility (X_1/X_2) as a function of ω/ω_n for the passive base excitation model and the Skyhook control based active base excitation model.



Many studies in literature back the effectiveness of the skyhook control policy and indicate that skyhook control is the optimal control in terms of its ability to isolate the suspended mass from the base excitations. For our application, we replace the passive damper in our model with a semiactive damper i.e. vary the damping coefficient and therefore the damping force linearly between high and low levels of damping depending upon the nature of disturbance, decided by the magnitude of the velocity of the device, the reading view and the velocity relative to each other.

With the gyroscope using the library by Jeff Ross the data input is directly in the form of yaw, pitch and roll, i.e. angles eliminating the need for integration.

Projective transform is used to compute this

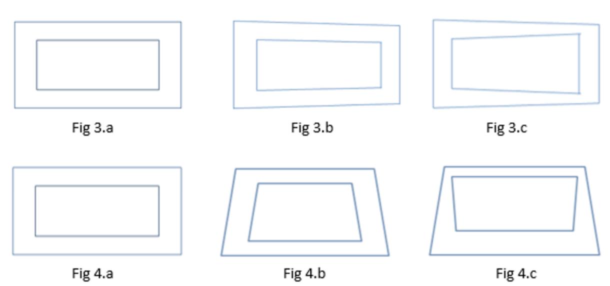


Fig. 1: Fig 3.a and 4.a- the initial or base laptop screen view; Fig 3.b- view when laptop is rotated about the y-axis; Fig 4.b- view when laptop flap is rotated about the x-axis away from the user; Fig 3.c and 4.c- intelligent compensation in the reading screen space to compensate and allow eased viewing

complimentary transformation of the view by using the standard 4X4 transformation matrix.

$$\begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

V. Proposed Method

In the modeling above, integration of the accelerometer data to get $x(t)$ alone does not yield very good results practically and excessive jitters are experienced in the content rendered on screen due to excessive noise which is characteristic of accelerometer data. This is removed by setting a threshold on the acceleration values and by using a third order moving average filter. Using a filter of higher order gives smoother response but makes the system slower which is not acceptable in this application. The most important and different aspect of our approach is that it treats the linear displacements and the angular pose changes of the device differently since the changes in the view when the device is rotated about an axis are different from when it is translated about the other perpendicular axes. This which would have been the same if the compensation was to be purely acceleration based. Therefore, we incorporate data from 2 sensors

(IMU) and use the gyroscope data to find the acceleration component recorded by the accelerometer due to that angular displacement. The accelerometer reading then minus this angular displacement produced acceleration gives the linear acceleration. Still this linear acceleration is given very less weight for compensation as a safeguard against unwanted jitters.
In case of rotation about any axis, say z:

$$a(t) = a_r(t) + a_l(t) \tag{1}$$

$$a(t) = \dot{w}(t)r + \omega^2(t)r + a_l(t)$$

Here ω is calculated by differentiating yaw angle found from the gyroscope measurement. In the last equation this 'r' depends on the distance of the instantaneous axis of rotation of the device from the center of the IMU. If the linear component of the acceleration a_l is small then we can find this 'r' by using two IMUs at some distance from each other. Then both the IMUs are on the same rigid body, they both will record the same angular velocity ω and hence 'r' can be found from the readings of the two accelerometers. But that is not of use here, since our aim behind this activity is to separate linear and angular components of acceleration from the accelerometer data. Also, the assumption of a_l being small cannot be taken in the moving vehicles. Therefore, at best, the static acceleration component arising due to a change in the pose i.e. θ, ϕ, ψ is subtracted from the acceleration readings recorded and the difference is taken as the linear acceleration in our system. The system response feels less noisy as a result. It may also be because the amplitude of the linear acceleration taken is lesser than the whole acceleration.

Using these corrected acceleration data then X_2(t) is found as per the model. The parameter like a and b where $pixelshift = a * X_2(t) + b * \dot{X}_2(t)$ is decided by hit and trial for all the various cases and kinds of disturbances and this is the major difficulty with making this system an easily usable one for the user, say through an app which can then use the inbuilt accelerometer of the phone or laptop. The following cases and their corresponding parameters are considered and computed:

- Displacement along x-axis shifting of pixels up and down (α).
- Displacement along y-axis shifting of pixels along left and right (α).
- Displacement along z-axis no change
- Angular displacement about x-axis reprojection of the views image about the x-axis and after the perspective correction, enlarging of the reprojected image (θ).
- Angular displacement about y-axis reprojection of the views image about the y-axis and after the perspective correction, enlarging of the reprojected image (ϕ).
- Angular displacement about z-axis rotation of the view in the plane of the device (ψ).

Reprojection is done taking a pinhole camera view model (involving a 4X4 perspective transform matrix) is sufficient and physically relatable for the transformations involved in the project since the object being read is planar and more complex warping is not required.

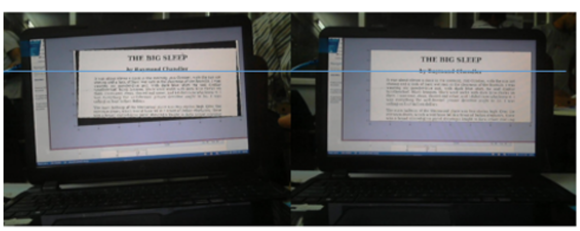
The compensation controller parameters can be set by either hit and trial and proper gain scheduling or by Fuzzy logic as we have tried.
Both accelerometer and gyroscope data are prone to systematic errors. The accelerometer provides accurate data over the long term, but is noisy in the short term. The gyroscope provides accurate data about changing orientation in the short term, but the necessary integration causes the results to drift over longer time scales. Generally, Kalman filter (or EKF) or a Complimentary filter is used to fuse the accelerometer and gyroscope values to obtain refine and validate the sensor data against the noise of the accelerometer, the gradual drift of the gyroscope measurements and also spurious noise. However, these are useful when the linear acceleration is less in magnitude and negligible as compared to changes in pose, which is not the case in moving vehicles where jerks and vibrations correspond to high linear accelerations and hence without separating the accelerometer readings into linear acceleration component and angular acceleration component gives erroneous results. After segregating the acceleration values, complimentary filter is used to give the angular acceleration values. Complimentary filter is used instead of Kalman as the former is less intensive computationally. The obtained values are then multiplied with empirically determined weights to obtain the pixel shift required for compensation. However, the filter used here should not be very broad since the program updates its equilibrium/or at rest reading position at every instant and a broad filter will make the system very heavily dependent on the past sensor data and hence less responsive.

The pixel shifts obtained for each of the 6 displacements mentioned above produce satisfactorily results individually but when employed together in the system produce unexpected errors and unnecessary shifting. This is because the overall pixel shift required at a time is not necessarily equal to the sum of the pixel shifts computed for the individual 6 displacements. Therefore, at this stage, we find the most prominent linear displacement of the three displacements and compute the overall shift as the gaussian weighted summation of the individual pixel shifts, with the most prominent linear acceleration contributing more significantly than the others in the shift. The outcome thus is more smooth and hence comforting to the eye.

This method of decomposing the acceleration values into linear and angular is novel as it has not been described and used much before because this is highly erroneous in small angular accelerations and large linear accelerations.

Therefore, another method is implemented and studied in which the acceleration values are not decomposed, using Fuzzy Logic. The same set of 6 basic disturbances are taken and applying fuzzy logic model, with triangular characteristic functions on the 6 readings i.e. x,y and z accelerations and yaw, pitch and roll, any disturbance is classified into the basic disturbances. This method is easier to implement and can be made very basic for the user to customize, tune and use in their personal devices. However, in both the methods, a considerably small value of acceleration is considered as a voluntarily change in either the pose or position of the device and hence is not compensated for in the algorithm. This is reasonable to do since for such small and slow changes, our eye gets sufficiently enough time to adjust. A key-input (or a touch based input in mobiles) is then given to reset the base values of the gyroscope and accelerometer after such a voluntary change.

VI. Results



VII. Conclusion and Future work

The screen view compensation nullifying the device movement is possible and highly helpful for easing the reading in vehicles and amidst disturbances elsewhere, as substantiated by experiments. We employed a setup for this by fusing accelerometer and gyroscope, since

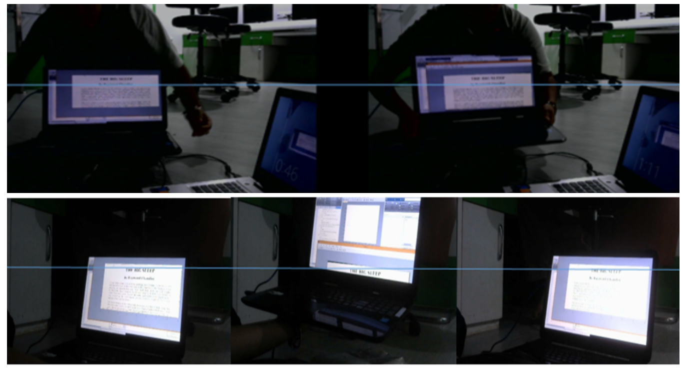


Fig. 2: Vertical movement(simulating jerks in the vehicle) compensation

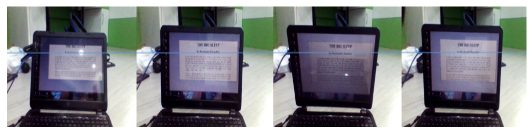


Fig. 3: Flap shaking/fluttering compensation

an accelerometer as used in all the earlier proposed algorithms does not provide compensation possibility for all types of disturbances and response is either inadequate or jittery. The paper discusses the image transforms we have used and why they are more suitable than the vertical and horizontal movements alone. The novel concept of decomposing accelerometer values into linear and angular acceleration values is discussed and results substantiate the methodologies employed. However, the system is highly heuristic based and a proper tuning strategy needs to be laid out for common people to use it. The prototype on which tests were done is Matlab based and very slow than what is desired for this application, hence it needs to be implemented as a dedicated program or an application. Using a DSP processor or a GPU to parallelize the algorithm will definitely make the system faster, hence that remains the work for implementation in our future versions.

References

[1] NoShake: Content Stabilization for Shaking Screens of Mobile Devices-Ahmad Rahmati, Clayton Shepard, Lin Zhong

[2] Walking with your Smartphone: Stabilizing Screen Content - Kevin Jeisy

[3] A Study for Understanding Carsickness Based on the Sensory Conflict Theory

[4] Video stabilization with moving object detecting and tracking for aerial video surveillance

[5] Simultaneous Video Stabilization and Moving Object Detection in Turbulence

[6] A Comparison of Complementary and Kalman Filtering

[7] Improvement of Ease of Viewing Images on an In-vehicle Display and Reduction of Carsickness