An Auto-Calibrated Laser-Pointing Interface for Collaborative Environments

Laberge, D.; Lapointe, Jean-François

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

An Auto-Calibrated Laser-Pointing Interface for Collaborative Environments *

Laberge, D., and Lapointe, J.-F.
October 2003

* published in Proceedings of the Ninth International Conference on Virtual Systems and Multimedia (VSMM 2003). October 15-17, 2003. Montreal, Québec, Canada. pp. 501-508. NRC 46524.

Canada

An Auto-Calibrated Laser-Pointing Interface for Collaborative Environments

Dominic LABERGE Jean-François LAPOINTE
Institute for Information Technology
National Research Council of Canada
1200 Montreal Road, Bldg. M-50
Ottawa, Canada, K1A 0R6
[Dominic.Laberge, Jean-Francois.Lapointe]@nrc-cnrc.gc.ca

Abstract. Most of the current laser-pointing interfaces use a vision-based approach, which requires camera calibration. This paper presents the use of a planar homography-based auto-calibrated technique to eliminate the camera calibration step, thus simplifying the setup process. The system performance of a user interface implemented with this technique is then analysed in details.

1. Introduction

One of the most used device during public presentations is the laser pointer, mainly for pointing and highlighting specific parts of the displayed image on a projection screen, particularly when it is out of reach of the user's arm. Unfortunately, used in this way, laser pointers[1] are just passive devices that offer no possibility of interaction with the content displayed on the screen. In the recent years, several publications have been made describing computer vision based systems that use a laser pointer as an input device [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. This new way of using laser pointers has great potential for inclusion in advanced collaborative environments such as CVEs.

All but one of the techniques described so far to track a laser spot on a screen require the previous calibration of the camera. We present here a technique similar but faster than a previous one presented but not detailed in [10]. This technique is based on a projective geometry tool that eliminates the calibration step, thus simplifying the setup process. This tool, called planar-based homographies, allows the mapping of camera coordinates to display screen coordinates [13].

We then present a detailed performance analysis of an implementation of the system in order to check if there is a match with the known basic requirements for effective human-computer interaction. The analysis is performed under several conditions, including the use of both a visible and of an invisible (infrared) laser pointer.

2. Previous work

One of the first papers on laser pointers is from Kirstein and Müller and describes the use of laser pointer as an input device and gives some advice on its implementation in [5].
Olsen and Nielsen who compared user performance between the mouse, a physical device and a laser pointer for a complex task involving pointing at widgets and manipulating their functions took the idea further [8]. The results indicate that the mouse was twice as fast as the other two

[1] To avoid confusion in this paper, the term *pointer* describes the laser pointer and the term *cursor*, the "mouse" pointer on the screen.

input devices but the control variables such as the update rate (and thus latency) of the different input devices were significantly different. The laser pointer was disadvantaged by its low update rate (7 Hz compared to an estimated 40 Hz for a standard mouse), as well as the typical training bias induced by subjects experience with the mouse. In fact, in another study where the latency was comparable between both devices, the laser pointer was found to perform as well as the standard mouse [1]. It is important to note, though, that this later user study was conducted with only four subjects.

Target acquisition, i.e. pointing and selecting, by using a laser pointer with an extended laser-on (similar to a mouse-over) technique has been tested and proven inaccurate. This was mainly due to jitter in the laser spot caused by the user's hand wiggle [6, 9]. To reduce jitter in the measured position, an averaging filter was used to estimate pointing position over ½ second. This approach slightly reduced the jitter but at the expense of adding ½ second to selection time.

A potentially better way to use laser pointers is to use semantic snarfing [6, 7]. This technique allows the user to point at an area of interest, which is then transferred to a secondary device such as a Portable Data Assistant (PDA) so that the user can point more precisely. According to these papers, semantic snarfing is faster but prone to errors because the interface is scaled down due to the PDA's small screen size. A proposed solution to avoid the latency introduced by the extended laser-on technique is to add a push-button on the laser pointer or use other techniques such as *hotspots* or split laser beams [2, 12]. A laser pointer equipped with a push-button can also be used as a pen when the screen is translucent and the camera located behind it [11]. Concerning the laser pointer, one study compared the performance of a visible and invisible laser and found that the response time was better with the visible one [1]. This result was explained by the fact that the user focused on the laser spot rather than on the cursor, thus eliminating computer display latency and providing an ideal feedback loop between the motor and perceptual systems. In this paper, a method to implement an auto-calibrated laser pointer system using planar homographies is presented. A similar idea was also put forward but not detailed for the purpose of 2D presentations [10, 14].

Finally, we measure the performance of a system implemented with this technique and compare the perceived system responsiveness when using visible (red) vs. invisible (infrared) laser pointers. The proposed thesis is that a visible pointer will make system latencies more perceptible because of the visible gap between the laser spot and the cursor, especially during fast movement.

3. Tracking algorithm

Our system is quite flexible, allowing for a large range of screen sizes, as well as multiple configuration of the camera-screen set.

3.1. Assumptions

The system will work as long as the three following conditions are filled:

- The screen is entirely visible (and in focus) inside the field of view of the camera
- There is no relative movement between the screen and the camera once the system is started
- The laser spot is brighter than the rest of the screen

3.2. Preprocessing

The first step is to estimate the pose of the screen in order to accurately compute the pointer position. In order to achieve that, we detect he position of the four corners of the computer screen and store them for future frame. To help their detection, we maximize the contrast with the surroundings by projecting a white blank screen. Also, to compensate for the phase difference between the projector and the camera, the binary image of the screen is accumulated over ten frames to get a more accurate estimate of the corners.

The homography matrix describing the mapping between the camera plane P and the display screen plane P' is then computed from the four detected corners in image coordinates and the four image screen corners in screen coordinates. A 3×3 homography matrix is a mathematical tool used to compute the projection from a plane P onto another plane P' when both planes are defined in a 3D space, such that:

$$\begin{bmatrix} x'w \\ y'w \\ w \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

w being the homogenous scale factor.

Finding the homography that correlates two planes is done by solving a set of linear equations in a similar way to that proposed in [10] but assuming that the last factor of the matrix is equal to 1, therefore reducing the system to equations needed to solve the problem using LU decomposition [16].

Each of the equations is written such that:

$$AC^t = B^t$$

Where:

$$A = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1'x_1 & -x_1'y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y_1'x_1 & -y_1'y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2'x_2 & -x_2'y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y_2'x_2 & -y_2'y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3'x_3 & -x_3'y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y_3'x_3 & -y_3'y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4'x_4 & -x_4'y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y_4'x_4 & -y_4'y_4 \end{bmatrix}$$

$$C = \begin{bmatrix} a & b & c & d & e & f & g & h \end{bmatrix}$$

$$B = \begin{bmatrix} x_1' & y_1' & x_2' & y_2' & x_3' & y_3' & x_4' & y_4' \end{bmatrix}$$

In the case of this system, a homography is used to map the camera image coordinates (P) to computer screen coordinates (P‘).

The edges of the displayed image's footprint in the camera image are computed as 2D vectors to test whether the laser spot is inside or outside the projected screen.

3.3. Processing algorithm

Once the preprocessing phase completed, the laser spot is tracked on the display screen in two steps.

First, find the laser spot in the image by finding the centroid of the bounding box of the brightest spot on the screen. The laser spot is assumed to be the only part of the image with a high light intensity. Each pixel's intensity is compared to a threshold and, if higher, added to the bounding box of the assumed laser spot. The center of that bounding box is assumed to be the centroid of the laser spot.

Second, to eliminate laser spots outside of the display screen, the laser spot's centroid is compared to each of the four previously computed corners to determine if it is inside the display screen. If that is the case, the homography is applied to the detected laser spot coordinates to map them in pattern space, and therefore in screen coordinate space. This leads to a position on the screen that is transferred to a cursor position.

4. Implementation

The system (see Figure 1) is made of a monochrome video camera with a resolution of 640 × 480 pixels and a horizontal field-of-view (FOV) of 37o, an XGA (1024 × 768) projector and a standard front projection screen. Both the camera and the projector were connected to a dual Pentium 4 computer running at 2.0 GHz with 512 Mbytes of RAM for processing. Two laser pointers have been tested: a visible (red) laser pointer for the user to see the actual laser spot, and an invisible (infrared) laser pointer that creates a laser spot outside of the human's perceptible light spectrum.

All the auto-adjustments of the camera were turned off. Also, during the test with the infrared (IR) pointer, an IR interferential filter is mounted on the camera after the preprocessing phase to increase robustness to external lighting.

In this experiment the projector and the camera were located at about 2m of the screen to test robustness with an extremely bright condition in terms of screen illumination compared to normal use. This was to test whether or not the laser pointer could be discriminated in this condition.

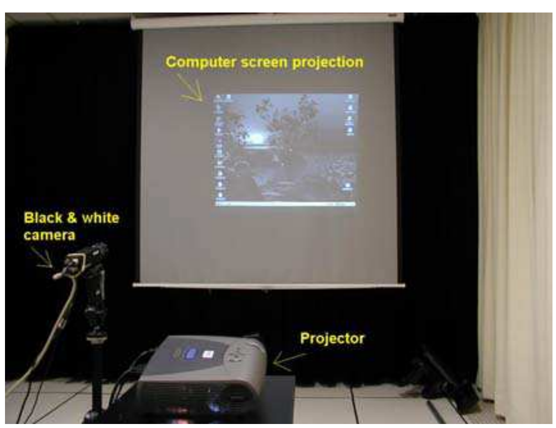

Figure 1. The laser pointing system

5. Evaluation method

The implemented system was evaluated under three ambient lighting conditions: dark room, artificially lit room and sunlit room. Also, the system was tested with two laser pointer wavelengths: 680nm (red) and 820nm (infrared).

The metrics used to evaluate the performance of the system are: range, accuracy, update rate, jitter and latency.

The accuracy was calculated by pointing at different regions of the screen and recording the difference in pixel from the tip of the cursor to the laser spot.

The update rate was calculated by moving constantly the laser spot on the screen and by measuring the time difference between cursor updates in a sequence of images. That video sequence of the scene was recorded by a high-speed (250 fps) video camera.

To test for jitter, the pointer was placed on a stable surface and pointed at the screen while the displacements were recorded by a video camera.

To measure latency, the cursor was placed at an arbitrary position in the computer screen. The pointer was then activated to point at an arbitrary region within the screen area. The time taken by the cursor to reach the static laser spot was then measured from the resulting video by counting the number of frames between the two events. The video of the scene was again recorded by the high speed camera.

6. Results

The system performed well under two of the three lighting conditions tested. Under direct sunlight, the system could not detect the screen because of a lack of contrast between the projected image and the rest of the screen. Under the other two lighting conditions (normal indoor and dark room), the performance results are summarized in Table 1:

Table 1. Performance results

Parameter	**Result**
Range	Line of sight
X accuracy	±2 screen pixels typical
Y accuracy	±3 screen pixels typical
Latency	52ms ± 8 ms average
Update	60 Hz ± 1 Hz (Overall) 605 Hz ± 5 Hz (Algo. only)
Jitter	1Hz ±1 screen pixel

It is interesting to note that from informal testing, the system works even when the camera is off by up to 30^{o} from perpendicular axis of the screen.

6.1. Range

The only things that limit the range of the system are: the need for a line of sight between the pointer and the screen, between the camera and the screen and the need for good contrast between the laser spot and the display.

In practice however, range is basically limited by factors such as user's pointing performance, and depends on its performance. As long as the projector projects an image that is bright enough to contrast with the environment and as long as the footprint of that image in the camera is good enough to get the desired resolution and that the user can point accurately, the system will work.

For most uses, a range of 10 m is achievable and sufficient.

6.2. Accuracy

Accuracy of the laser pointer device is relative to the size (in pixels) of the laser spot on the screen and captured by the camera. If the camera is closer, the laser will be integrated over several pixels on the camera's retinal plane, thus providing a better estimate of the laser spot's centroid.

Also, laser diodes have spreading factors. Commercial laser pointers tend to leak more on the sides of the spot than higher quality diodes. This influences centroid computation.

Thirdly, the angle between the projected screen and the camera will have influence on the accuracy and uniformity of the results. If the angle is great, one pixel in the camera will map to several pixels on the screen.

Fourthly, resolution of the system has an influence on accuracy performance as pixels on the projected screen have to be differentiable if one is to point at them.

The camera's resolution and the screen to camera distance limit resolution.

Resolution causes inaccuracies as long as the effective image resolution, i.e. resolution taking into account the sub-pixel factor, is below the projected image's resolution. On the other hand, if the resolution is too high, movement detection does not necessarily result in cursor movement and useless cursor move events are generated.

Detecting the centroid of the bounding box of the laser pointer enables to achieve quarter-pixel resolution, increasing the effective camera resolution fourfold over native resolution.

To achieve optimal resolution, the camera should be placed so that the retinal plane is parallel projected image plane, and that the four boundaries of the auto-detected projected image are parallel to the corresponding camera image boundaries. Then, since the detection algorithm permits twice the resolution along both axes, the horizontal and vertical camera resolutions should be half the size in pixels of the detected projected image.

Under the experimental conditions described above, the 640×480 camera provided enough resolution to detect one screen pixel movements. Finally, from testing the system, it seems that radial distortion in the camera lens has a noticeable effect on the measurement. Accuracy is better in the center of the camera image than on the sides. Also, inaccuracies tend to shift towards the screen sides. For example, the cursor tends to be a little higher than the laser spot when in the upper portion of the screen and a little under when in the lower portion of it.

In summary, the accuracy depends of five factors:

- Distance between the camera and the screen
- Light leaking of the laser pointer
- Angle between the camera and the projector
- Effective camera resolution
- Radial distortion of the camera lens

6.3. Update rate

At 60 Hz the update rate is good enough for interaction, and induces a latency that does not hinder user performance. Clearly here, the update rate is limited by the camera speed (in frames per second), since less than 10% of the computing power is used to achieve it, as indicated by the internal (algorithmic) update rate of 605 Hz.

6.4. Jitter

From informal tests, we found that at ±1 screen pixel, the jitter of the system is limited as compared to natural human hand wiggle. This jitter is mainly be explained by the fact that the camera resolution (640x480 pixels) is slightly inferior to the projector resolution (1024x768), thus causing a ±1 pixel incertitude on the laser spot position. This inaccuracy, however, is in part diminished by the fact that the laser spot position is computed at a sub-camera pixel accuracy by finding the centroid of it on the camera image.

6.5. Latency

With an overall system latency of 52 ms ± 8 ms, the system is well under the 80 ms limit, which has been shown to hamper user performance in tracking tasks with position control [14].

6.6. Infrared vs. visible laser pointer

Tests were made using a visible and an invisible laser pointer. Although the laser wavelength used doesn't affect the system's performance in terms of criteria listed above. For the user, however, there is a noticeable difference in terms of adaptation. When using the red laser pointer. The user sees the red laser spot on the screen and expects the cursor to follow directly, thus yielding a high sensitivity to latency and update rate. When using the IR laser pointer, the laser spot is not visible and the human sensory motor system adapts to the latency. The system seems more fluid though it reacts at the same speed. It is important to note though that one previous study [1] indicates that the IR pointer increased the mean response time for pointing tasks. It is therefore possible that a visible pointer could be performing better albeit giving the impression of greater latency.

Another important observation is that, in a 3D stereo display environment, a visible laser pointer cannot be used because humans will generally perceive it as a double spot due to the nature of stereo.

7. Discussion & Future Work

This paper presented and evaluated a new computer vision based algorithm that can be used to track a laser spot on a projected screen, in order to use laser pointers as an input device. One application that generates as interest in this system is presentations, but virtual reality and augmented reality systems could benefit from such an interface.

We found that performance of the pointing system is sufficient to be used in most applications. Resolution could become a hardware problem if the ratio display resolution over effective camera resolution is not kept close to 1. This leads to an over sensitivity of the system if the displayed screen's footprint in the camera image is in higher resolution than the actual displayed image. The system then reacts to movements in the laser pointer that would not lead to an actuation of the cursor. On the other hand, if the displayed screen is represented by less effective pixels in the system than in reality, interaction will be rather jaggy and imprecise. Latency and update rate can be improved by using a faster camera (>60 fps), although current performance seems more than adequate for those criteria.

In order for the system to work at its full capacity, the camera must be more sensitive in the wavelengths range of the laser spot; otherwise, the identification of the laser spot could fail. Also, better contrast between the display image and the rest of the scene will result in better preprocessing, and therefore a better mapping.

More research is needed to formally compare the performance of an auto calibrated laser pointer tracking system and a manually calibrated system of the same kind. This is a necessary step in assessing if an auto-calibrated pointing interface is a good alternative to manually calibrated systems.

Acknowledgements

We would like to thank our colleagues Guy Godin and Mark Fiala for their help during the development of the system.

References

[1] D. Cavens, F. Vogt, S. Fels, and M. Meitner. Interacting with the big screen:Pointers to ponder. In Proceedings of ACM Conference on Computer Human Interaction (CHI 2002), pages 678–679, Minneapolis, Minnesota, USA, April 20- 25 2002.

[2] K. Cheng, K, and Pulo, K. Direct Interaction with Large-Scale Display Systems using Infrared Laser tracking Devices. In Proceedings of the Australian Symposium on Information Visualisation (InVis.Au'03), pages 67-74, Adelaide, Australia, 2003.

[3] J. Davis and X. Chen. LumiPoint: Multi-user laser-based interaction on large tiled displays. Displays, 23(5):205–211, Elsevier Science 2002.

[4] R. R. Eckert and J. A. Moore. The classroom of the 21st century: The interactive learning wall. SIGCHI Bulletin, 32(2):33–40, April 2000.

[5] C. Kirstein and H. M¨uller. Interaction with a projection screen using a camera tracked laser pointer. In Proceedings of International Conference on Multimedia Modeling (MMM 98), pages 191–192, Lausanne, Switzerland, October 12-15 1998.

[6] B. A. Myers, R. Bhatnagar, J. Nichols, C. H. Peck, D. Kong, R. Miller, and A. C. Long. Interacting at a distance: Measuring the performance of laser pointers and other devices. In Proceedings of ACM Conference on Computer Human Interaction (CHI 2002), pages 33–40, Minneapolis, Minnesota, USA, April 20-25 2002.

[7] B. A. Myers, C. H. Peck, J. Nichols, D. Kong, and R. Miller. Interacting at a distance using semantic snarfing. In Proceedings of Ubicomp 2001, pages 305–314, Atlanta, Georgia, USA, September 30 - October 2 2001.

[8] D. R. Olsen and T. Nielsen. Laser pointer interaction. In Proceedings of ACM Conference on Computer Human Interaction (CHI 2001), pages 17–22, Seattle, WA, USA, March 31 - April 4 2001.

[9] C. H. Peck. Useful parameters for the design of laser pointer interaction techniques. In Proceedings of ACM Conference on Computer Human Interaction (CHI 2001), pages 461–462, Seattle, WA, 31 March-5 April 2001.

[10] R. Sukthankar, R. G. Stockton, and M. D. Mullin. Self-Calibrating Camera-Assisted Presentation Interface. In Proceedings of the Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore, December 5-8 2000.

[11] T. Winograd and F. Guimbretiere. Visual instruments for an interactive mural. In ACM SIGCHI CHI99 Extended Abstracts, pages 234–235, Pittsburgh, PA, 1999.

[12] S. V. Matveyev and M. Göbel. Direct Interaction based on a Two-Point Laser Pointer Technique. In Proceedings of ACM SIGGRAPH 2003, San Diego, CA, 2003, DVD only.

[13] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.

[14] R. Sukthankar, R. G. Stockton, and M. D. Mullin. Smarter presentation: Exploiting homography in camera-projector systems. In Proceedings of International Conference on Computer Vision, pages 247–253, Vancouver, Canada, July 9-12 2001.

[15] C. Wickens. The effects of control dynamics on performance. In K. B. et al., editor, Handbook of Human Perception and Performance, volume II, pages 39–1 to 39–60, New York, 1986. Wiley.

[16] W. H. Press, W. T. Vettering, S. A. Teukolsky, , and B. P. Flannery. Numerical Recipes in C++. Cambridge University Press, Cambridge, 2002.