IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

TITLE OF THE INVENTION

GESTURE IDENTIFICATION AND REPLICATION

Inventors:

　　　　ALON EFRAT, of:
　　　　Tucson, AZ

　　　　KOBUS BARNARD, of:
　　　　Tucson, AZ

Assignee:　　**ARIZONA BOARD OF REGENTS ON BEHALF OF THE UNIVERSITY OF ARIZONA**

Entity:　　　Small

Prepared by:　HAMILTON, DESANCTIS & CHA LLP
　　　　　　　Customer No.:　64128
　　　　　　　(303) 856-7155

GESTURE IDENTIFICATION AND REPLICATION

COPYRIGHT NOTICE

BACKGROUND

Field

[0002]   Embodiments of the present invention generally relate to interaction with and reproduction of video content. In particular, embodiments of the present invention relate to the application of robust algorithms to identify laser pointer gestures on a primary presentation screen and the reproduction of such gestures onto secondary presentation screens..

Description of the Related Art

[0003]   Commonly during conferences lectures are given to a large audience of hundreds, thousands or more attendees. Typically, during such lectures, the lecturer uses electronic slides (e.g., created in Microsoft PowerPoint slide presentation software, Apple Keynote, Google Presentation, Prezi, SlideRocket or the like) and/or video presentation materials, which are projected onto a screen or screens (referred to below as *primary screens*) by a projector. For such large audiences it is common to use several large screens (referred to below as *secondary screens*) to show the presentation materials at the same time. This enables more members of a large audience to view the presentation materials. In such an architecture, the video signal leaving the lecturer's computer is transmitted to several projectors rather than one.

[0004]   It is common for the lecturer to make use of gestures with a laser pointer during the talk to emphasize items and bullets on the presentation materials being projected onto the primary screen. These gestures can be an important part of the lecture, and can assist the learning process. The typical approach of simply reproducing the original video signal that is projected onto the primary screen onto the secondary screens deprives much of the audience of the benefit of the gestures as the lecturer directs the laser pointer spot only to the primary screen and such gestures are not reproduced on the secondary screens.

SUMMARY

[0005] Methods and systems are described for performing gesture identification and replication. According to one embodiment, a laser tracking routine running on a computer system, receives multiple video frames containing images of a primary screen on which a portion of a video presentation is displayed and onto which a laser pointer gesture is being made. The video presentation is associated with a first coordinate system and the images of the primary screen within the video frames are associated with a second coordinate system. A mapping routine running on the computer system determines a homography between the first coordinate system and the second coordinate system. Based on the plurality of video frames, the laser tracking routine identifies (i) the laser pointer gesture as a candidate gesture for reproduction onto one or more secondary screens and (ii) one or more coordinates of the laser pointer gesture in the second coordinate system. False gesture detection is reduced by a noise filtering routine running on the computer system by applying one or more noise filtering algorithms. When the noise filtering routine confirms the candidate gesture as a true laser pointer gesture, then the portion of the video presentation to be displayed on the one or more secondary screens is displayed concurrently with a synthetic representation of the laser pointer gesture by augmenting the plurality of video frames based on the one or more coordinates of the laser pointer gesture and the homography.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] Embodiments of the present invention are illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

[0007] **FIG. 1** conceptually illustrates a presentation architecture in accordance with an embodiment of the present invention.

[0008] **FIG. 2** is an example of a computer system with which embodiments of the present invention may be utilized.

[0009] **FIG. 3** is a high-level flow diagram illustrating gesture identification and reconstruction processing in accordance with an embodiment of the present invention.

[0010] **FIG. 4** is a flow diagram illustrating exemplary noise filtering processing in accordance with various embodiments of the present invention.

DETAILED DESCRIPTION

[0011]   Methods and systems are described for performing object oriented genetic programming (OOGP).   According to one embodiment, a method is provided for identifying a laser pointer gesture on a primary screen while accurately distinguishing between the laser pointer spot and possible other sources of light that might appear due to reflections caused by lighting internal or external to the presentation facility, changes in the presented frame caused by shadows and/or movement of the primary screen The identified laser pointer gesture can then be reproduced on one or more secondary screens by augmenting the video signal originating from the lecturer's computer system to produce an augmented video signal containing the identified laser pointer gesture for display on the one or more secondary screens.

[0012]   In the following description, numerous specific details are set forth in order to provide a thorough understanding of embodiments of the present invention. It will be apparent, however, to one skilled in the art that embodiments of the present invention may be practiced without some of these specific details. In other instances, well-known structures and devices are shown in block diagram form.

[0013]   Embodiments of the present invention include various steps, which will be described below. The steps may be performed by hardware components or may be embodied in machine-executable instructions, which may be used to cause a general-purpose or special-purpose processor programmed with the instructions to perform the steps.   Alternatively, the steps may be performed by a combination of hardware, software, firmware and/or by human operators.

[0014]   Embodiments of the present invention may be provided as a computer program product, which may include a machine-readable storage medium tangibly embodying thereon instructions, which may be used to program a computer (or other electronic devices) to perform a process. The machine-readable medium may include, but is not limited to, fixed (hard) drives, magnetic tape, floppy diskettes, optical disks, compact disc read-only memories (CD-ROMs), and magneto-optical disks, semiconductor memories, such as ROMs, PROMs, random access memories (RAMs), programmable read-only memories (PROMs), erasable PROMs (EPROMs), electrically

erasable PROMs (EEPROMs), flash memory, magnetic or optical cards, or other type of media/machine-readable medium suitable for storing electronic instructions (e.g., computer programming code, such as software or firmware). Moreover, embodiments of the present invention may also be downloaded as one or more computer program products, wherein the program may be transferred from a remote computer to a requesting computer by way of data signals embodied in a carrier wave or other propagation medium via a communication link (e.g., a modem or network connection).

[0015]   In various embodiments, the article(s) of manufacture (e.g., the computer program products) containing the computer programming code may be used by executing the code directly from the machine-readable storage medium or by copying the code from the machine-readable storage medium into another machine-readable storage medium (e.g., a hard disk, RAM, etc.) or by transmitting the code on a network for remote execution. Various methods described herein may be practiced by combining one or more machine-readable storage media containing the code according to the present invention with appropriate standard computer hardware to execute the code contained therein. An apparatus for practicing various embodiments of the present invention may involve one or more computers (or one or more processors within a single computer) and storage systems containing or having network access to computer program(s) coded in accordance with various methods described herein, and the method steps of the invention could be accomplished by modules, routines, subroutines, or subparts of a computer program product.

[0016]   Notably, while embodiments of the present invention may be described using modular programming terminology, the code implementing various embodiments of the present invention is not so limited. For example, the code may reflect other programming paradigms and/or styles, including, but not limited to object-oriented programming (OOP), agent oriented programming, aspect-oriented programming, attribute-oriented programming (@OP), automatic programming, dataflow programming, declarative programming, functional programming, event-driven programming, feature oriented programming, imperative programming, semantic-oriented programming,

functional programming, genetic programming, logic programming, pattern matching programming and the like.

<u>Terminology</u>

[0017]   Brief definitions of terms used throughout this application are given below.

[0018]   The terms "connected" or "coupled" and related terms are used in an operational sense and are not necessarily limited to a direct connection or coupling.

[0019]   The phrases "in one embodiment," "according to one embodiment," and the like generally mean the particular feature, structure, or characteristic following the phrase is included in at least one embodiment of the present invention, and may be included in more than one embodiment of the present invention. Importantly, such phases do not necessarily refer to the same embodiment.

[0020]   If the specification states a component or feature "may", "can", "could", or "might" be included or have a characteristic, that particular component or feature is not required to be included or have the characteristic.

[0021]   The term "responsive" includes completely or partially responsive.

[0022]   As discussed above, commonly during conferences, lectures are given to hundreds or thousands of attendees. If electronic slides are presented on a screen during the talk, their projections onto several screens is a necessity, since not all attendees would be able to see a single screen from close enough to enable readability. Thus, it is common to use several large secondary screens, all showing the slide at same time. Concurrently with the lecture, the lecturer may make laser pointer gestures. These gestures will appear on one screen (the primary screen) but not on the others. This reduces the quality of the media received by the audience viewing the secondary screens.

[0023]   Other existing solutions show, on a secondary screen, a video captured from a camera watching the primary screen. Such solutions do indeed capture and reproduce

the laser pointer spot, but they also add significant noise and blurriness to the original video content (and the primary screen). For example, text of slides presented on the primary screen likely to appear blurry to some extent, and colors used in pie charts could be altered and look unnatural. Illumination of the primary screen from other sources, shades and shakiness could also jeopardize the video quality in such prior solutions. Embodiments of the present invention seek to address various limitations of such previous presentation architectures.

[0024]   **FIG. 1** conceptually illustrates a presentation architecture 100 in accordance with an embodiment of the present invention.   According to the current example, presentation architecture 100 includes a lecturer computer 105, a system computer 110, a primary screen 120, a video camera 115 and one or more secondary screens 125a-n. The system components may be interconnected in various conventional manners (e.g., physically via custom or standard composite or component video cables, Universal Serial Bus (USB), Category 5 Ethernet cables, High-Definition Multimedia Interface (HDMI) cables or the like and/or wirelessly via wireless communications, including, but not limited to, wireless USB, Bluetooth or wireless networks (e.g., IEE 802.11 branded as Wi-Fi).

[0025]   According to one embodiment, lecturer computer 105 includes stored therein presentation materials desired to be displayed to the audience on primary screen 120 and secondary screens 125a-n via projectors (not shown) to facilitate their understanding of the subject matter being discussed. The presentation materials may be in the form of slides, video and/or multimedia. Lecturer computer 105 is configured to output a video signal 106 containing the content of the presentation materials, e.g., a currently selected slide, frames of video content or the like.

[0026]   System computer 110 is configured to receive video signal 106 from lecturer computer 105, analyze it, send an original signal to the primary screen 120 (typically in closest proximity to the lecturer) and send an augmented signal 111 to secondary screens 125a-n. In alternative embodiments, the functionality of lecturer computer 105 and system computer 110 may be combined into a single computer system.

[0027]   Video camera 115 monitors primary screen 120 and sends images of primary screen 120 in the form of a video of primary screen 116 to system computer 110 for analysis. As will be described in further detail below, in one embodiment of the present invention, a software program (not shown) executing on system computer 110, identifies in real time a laser pointer spot projected onto primary screen 120 by the lecturer.

[0028]   Notably, in accordance with one embodiment, the software program is capable of identifying candidate laser pointer gestures and filtering out false detections. Possible source of false detections include illuminations on primary screen 120 from other light sources (e.g., headlights, reflections of lamps in the classroom), bright spots in the original video content and/or changes in the presented frame caused by shadows and motion of primary screen 120 due to being bumped, for example. By distinguishing between true laser pointer gestures and other noise that might appear on primary screen 120, only true laser pointer gestures are reproduced on secondary screens 125a-n.

[0029]   In one embodiment, a gesture recognition algorithm implemented within the software program includes the ability to identify and characterize gestures performed by the lecturer with the laser pointer spot.   Recognized gesture patterns may include information regarding the trajectory performed by the pointer – for example, a circle around a bullet. Understanding these gestures facilitates the elimination of false detections.

[0030]   As described in further detail below, once the existence of a laser pointer gesture (e.g., a laser spot or pattern) on primary screen 120 is confirmed and validated, the software program identifies the current location of the laser pointer gesture in the captured video presentation coordinate system (e.g., the coordinate system associated with the captured video images of primary screen 120), and augments video signal 106 to create an augmented signal 111 containing a synthesized version of the laser pointer gesture to be displayed by or projected onto secondary screens 125a-n. For example, a synthetic circle of appropriate color (typically red) may be added to augmented signal 111 at the appropriate location based on a known mapping between the source video input coordinates (e.g., video signal 106 received by system computer 110 from lecturer

computer 105 – a/k/a slide coordinates) and the captured video presentation coordinate system. Those skilled in the art understand how to determine the mapping (a homography) between the captured video presentation coordinate system and the slide coordinates; however, for purposes of completeness, exemplary methodologies for calculating such a homography are provided in the papers attached hereto in Exhibit A.

[0031]   **FIG. 2** is an example of a computer system with which embodiments of the present invention may be utilized. Embodiments of the present invention include various steps, which will be described in more detail below. A variety of these steps may be performed by hardware components or may be tangibly embodied on a computer-readable storage medium in the form of machine-executable instructions, which may be used to cause a general-purpose or special-purpose processor programmed with instructions to perform these steps. Alternatively, the steps may be performed by a combination of hardware, software, and/or firmware. As such, **FIG. 2** is an example of a computer system 200, such as lecture computer 105 (in the form of a personal computer, a laptop computer, a tablet computer, a smartphone or the like) or system computer 110 (in the form of a server, a personal computer, a laptop computer or the like), upon which or with which embodiments of the present invention may be employed.

[0032]   According to the present example, the computer system includes a bus 230, one or more processors 205, one or more communication ports 210, a main memory 215, a removable storage media 240, a read only memory 220 and a mass storage 225.

[0033]   Processor(s) 205 can be any future or existing processor, including, but not limited to, an Intel® Itanium® or Itanium 2 processor(s), or AMD® Opteron® or Athlon MP® processor(s), or Motorola® lines of processors. Communication port(s) 210 can be any of an RS-232 port for use with a modem based dialup connection, a 10/100 Ethernet port, a Gigabit port using copper or fiber or other existing or future ports. Communication port(s) 210 may be chosen depending on a network, such a Local Area Network (LAN), Wide Area Network (WAN), or any network to which the computer system 200 connects.

[0034] Main memory 215 can be Random Access Memory (RAM), or any other dynamic storage device(s) commonly known in the art. Read only memory 220 can be any static storage device(s) such as Programmable Read Only Memory (PROM) chips for storing static information such as start-up or BIOS instructions for processor 205.

[0035] Mass storage 225 may be any current or future mass storage solution, which can be used to store information and/or instructions. Exemplary mass storage solutions include, but are not limited to, Parallel Advanced Technology Attachment (PATA) or Serial Advanced Technology Attachment (SATA) hard disk drives or solid-state drives (internal or external, e.g., having Universal Serial Bus (USB) and/or Firewire interfaces), such as those available from Seagate (e.g., the Seagate Barracuda 7200 family) or Hitachi (e.g., the Hitachi Deskstar 7K1000), one or more optical discs, Redundant Array of Independent Disks (RAID) storage, such as an array of disks (e.g., SATA arrays), available from various vendors including Dot Hill Systems Corp., LaCie, Nexsan Technologies, Inc. and Enhance Technology, Inc.

[0036] Bus 230 communicatively couples processor(s) 205 with the other memory, storage and communication blocks. Bus 230 can include a bus, such as a Peripheral Component Interconnect (PCI) / PCI Extended (PCI-X), Small Computer System Interface (SCSI), USB or the like, for connecting expansion cards, drives and other subsystems as well as other buses, such a front side bus (FSB), which connects the processor(s) 205 to system memory.

[0037] Optionally, operator and administrative interfaces, such as a display, keyboard, and a cursor control device, may also be coupled to bus 230 to support direct operator interaction with computer system 200. Other operator and administrative interfaces can be provided through network connections connected through communication ports 210.

[0038] Removable storage media 240 can be any kind of external hard-drives, floppy drives, IOMEGA® Zip Drives, Compact Disc – Read Only Memory (CD-ROM), Compact Disc – Re-Writable (CD-RW), Digital Video Disk – Read Only Memory (DVD-ROM).

[0039]   Components described above are meant only to exemplify various possibilities. In no way should the aforementioned exemplary computer system limit the scope of the invention.

[0040]   **FIG. 3** is a high-level flow diagram illustrating gesture identification and reconstruction processing in accordance with an embodiment of the present invention. According to the present example, after a predetermined or configurable number of frames (e.g., 2 to 15) are captured, the homography may be computed. When a candidate laser pointer gesture is identified, one or more noise filtering processes are performed to reduce the occurrence of false detections. If the candidate laser pointer gesture is confirmed, it is reproduced on the secondary screens by producing an augmented video signal containing a synthetic spot simulating the laser pointer spot and its position and motion mimics the gestures used by the lecturer.

[0041]   For purposes of simplicity and for sake of brevity, only a single pass of a video processing loop is described in the context of the present example. Those skilled in the art will appreciate the steps depicted in FIG. 3 may be placed in a loop (e.g., a polling loop that retrieves captured video content from video camera 115) or may be event driven (e.g., responsive to the availability of newly received video content from video camera 115).

[0042]   At decision block 310, it is determined if initialization processing has been completed. If so, then processing branches to block 330. If initialization processing has not been completed then processing continues with block 320.

[0043]   At block 320 a mapping (a homography) between the video input coordinates (e.g., video signal 106 received by system computer 110 from lecturer computer 105 – a/k/a slide coordinates) and the video presentation coordinates (e.g., the coordinate system associated with the captured video images of primary screen 120) is determined. Various methods exist for determining such a mapping. For purposes of completeness, a robust method for determining the mapping between these two frames of reference is provided in Exhibit A. In the context of the present example, this mapping is

assumed to be fixed for the duration of the lecture (and therefore determined only once). In alternative embodiments, the mapping may be performed on a periodic basis, on demand and/or responsive to a determination that the mapping is no longer valid.

[0044]   At block 330, as the software program running on system computer 110 receives images of primary screen 120 captured by video camera 115, a potential laser pointer gesture is identified based on a coarse analysis of pixel distribution differences between adjacent video frames. Those skilled in the art will be aware of various methods of identifying candidate laser pointer gestures contained within a video; however, for sake of completeness of the present disclosure exemplary approaches for identifying candidate laser pointer gestures are described in the papers attached as <u>Exhibit A</u> attached hereto.

[0045]   At block 340, false positives are reduced by putting the candidate laser pointer gesture identified in block 330 through one or more noise filtering processes, examples of which are described in further detail below with reference to **FIG. 4**.

[0046]   At decision block 350, it is determined whether the candidate laser pointer gesture has been confirmed. If so, processing continues with block 360; otherwise, gesture identification and reconstruction processing is complete

[0047]   **FIG. 4** is a flow diagram illustrating exemplary noise filtering processing in accordance with various embodiments of the present invention. According to the present example, four verification processes are applied to a candidate laser pointer gesture identified at block 330 of **FIG. 3**. The verification processes may be applied in parallel or serial; and more, fewer or different verification processes may be employed as desired for the particular implementation. According to one embodiment, the distance from the camera to the primary screen is either known or initially calculated. According to some embodiments, one or more of the discrete verification processes (e.g., 410-440) may be further dependent upon the results obtained from another. In other embodiments, the discrete verification processes may be independent from one another. In some embodiments, all verification processes must be satisfied to confirm a candidate laser

pointer gesture; however, in other embodiments, satisfying a super majority or a majority may be sufficient to confirm a candidate laser pointer gesture.

[0048]   At block 410, intensity verification is performed on the laser pointer spot (LPS) forming the candidate laser pointer gesture and appearing in the captured video frames. Depending on the intensity of the LPS and its distance from screen, the LSP might appear much brighter than nearby pixels in its vicinity. When its intensity is high enough, it tends to appear to the camera as a white spot, since all RGB sensors of the camera are saturated. Further enhancement may be obtained by reducing the average/median intensity compared to recently acquired frames.

[0049]   At block 420, size and shape verification are performed on the LPS. Most laser pointers generate a circular spot in the range of few millimeters to two to three centimeters. As such, size verification may filter out spots not falling within this range. Similarly, spots deviating too greatly from the expected circular shape may also be rejected. Notably, a fast moving LPS could be seen by the camera as a single elongated spot, rather than a circular spot. To accommodate for this, the size and shape verification process may accept such spots as a correct LPS, but only if the direction of motion is also consistent with the temporal filter of block 440.

[0050]   At block 430, color distribution verification is performed on the LPS. The LPS usually occupies several pixels of video camera 115. The intensity within the LPS (and registered by the different sensors of the camera) are different between the sensors. For example, commonly, the center of the spot is brighter than its outskirts. The intensity distribution profile (as compared to previously acquired profiles) is a strong indicator of whether the LPS of the candidate laser gesture is truly an LPS.

[0051]   At block 440, temporal smoothness is performed on the LPS over a number of frames. According to one embodiment, the LPS is meaningful only if it either stays in the same location on the primary screen or moves along a continuous curve which is piecewise smooth. Temporal smoothness may be defined as movement of the LPS in accordance with a sequence of segments, each of which is smooth (e.g., no sharp corners) and each of which is capable of being described, for example, using a low-degree polynomial, as is commonly done in computer graphics and solid modeling. For

computing this sequence, the computer system will use previously-acquired frames, and analyze the motion of the LPS spot. Moving in a smooth pattern over time is a good indicator of whether the LPS of the candidate laser gesture is truly an LPS.

[0052]	At decision block 450, a determination is made regarding the discrete verification processes in the aggregate support confirmation of the candidate laser pointer gesture. In one embodiment, all of the filters (e.g., 410-440) are required to be satisfied. In alternative embodiments, the filters may be weighted and their aggregate results compared to a predetermined or configurable threshold. In any event, if the results of the filtering in the aggregate support confirmation, then processing continues with block 470 wherein the LPS of the candidate laser pointer is approved; otherwise, processing branches to block 460 where the LPS of the candidate laser pointer gesture is rejected.

Alternative Embodiments:

[0053]	While embodiments of the present invention are described in the context of a particular usage model (e.g., a lecture or presentation to a large audience), it is to be understood that the gesture extraction and reproduction mechanisms have broad applicability in various other contexts involving interactions with and reproduction of video content. For purposes of illustrating the broad applicability of the gesture extraction methodologies described herein, a few alternative usage models are briefly described below.

[0054]	The inventors specifically contemplate uses within gaming (e.g., PC-based, console-based, mobile and/or online gaming), interactive voting/polling or group voting/tabulation and laser tag. In the context of gaming, for example, a laser pointer gesture made within a particular gamer's view presented in his/her television screen may be captured and reproduced within the view of another local or remote gamer.

[0055]	With reference to group voting/tabulation, while providers of interactive voting systems and audience response systems can supply voting application software and interactive audience participation equipment for purchase or lease, such as radio frequency (RF) or wireless keypad voting clickers, these systems may be cost prohibitive

in the context of a large audience due to the cost of the hardware. The inventors envision providing audience members with laser pointers instead of existing clickers. The audience can then participate in interactive voting, polling and the like by directing their pointers at the primary screen. The gesture extraction mechanisms described above can then be used to identify discrete laser spots to tally votes, gather and/or present statistics based thereon. Moreover, they could be used for move informative and involved gestures such as "drawing" around or "circling" around areas of interest and/or circling around regions of interest.

[0056]   In view of the foregoing, it will be clear that the invention is not limited to the specific embodiments described herein and that numerous modifications, changes, variations, substitutions, and equivalents will be apparent to those skilled in the art, without departing from the spirit and scope of the invention, as described in the claims.

<u>CLAIMS</u>

What is claimed is:

1.   A computer-implemented method comprising:

        receiving, by a laser tracking routine running on a computer system, a plurality of video frames containing images of a primary screen on which a portion of a video presentation is displayed and onto which a laser pointer gesture is being made, wherein the video presentation is associated with a first coordinate system and the images of the primary screen within the plurality of video frames are associated with a second coordinate system;

        determining, by a mapping routine running on the computer system, a homography between the first coordinate system and the second coordinate system;

        identifying, by the laser tracking routine, based on the plurality of video frames, (i) the laser pointer gesture as a candidate gesture for reproduction onto one or more secondary screens and (ii) one or more coordinates of the laser pointer gesture in the second coordinate system;

        reducing false gesture detection, by a noise filtering routine running on the computer system, applying one or more noise filtering algorithms; and

        when the noise filtering routine confirms the candidate gesture as a true laser pointer gesture, then causing the portion of the video presentation to be displayed on the one or more secondary screens concurrently with a synthetic representation of the laser pointer gesture by augmenting the plurality of video frames based on the one or more coordinates of the laser pointer gesture and the homography.

2.   The method of claim 1, wherein the noise filtering routine performs intensity verification on a laser pointer spot forming the candidate gesture.

3.   The method of claim 1, wherein the noise filtering routine performs size and shape verification on a laser pointer spot forming the candidate gesture.

4.   The method of claim 1, wherein the noise filtering routine performs color distribution verification on a laser pointer spot forming the candidate gesture.

5.   The method of claim 1, wherein the noise filtering routine performs temporal smoothness verification on a laser pointer spot forming the candidate gesture.

6.   The method of claim 1, wherein the noise filtering routine performs a plurality of verification processes on a laser pointer spot forming the candidate gesture selected from (i) an intensity verification process, (ii) a size and shape verification process, (iii) a color distribution verification process and (iv) a temporal smoothness verification process.

7.   The method of claim 6, wherein confirmation of the candidate gesture is achieved upon the laser pointer spot passing all of the plurality of verification processes.

ABSTRACT

[0057] Methods and systems for performing gesture identification and replication are provided. According to one embodiment, multiple video frames are received containing images of a primary screen on which a portion of a video presentation is displayed and onto which a laser pointer gesture is being made. The video presentation is associated with a first coordinate system and the images of the primary screen within the video frames are associated with a second coordinate system. A homography between the first coordinate system and the second coordinate system is determined. Based on the plurality of video frames, (i) the laser pointer gesture is identified as a candidate gesture for reproduction onto one or more secondary screens and (ii) one or more coordinates of the laser pointer gesture are identified in the second coordinate system. False gesture detection is reduced by applying one or more noise filtering algorithms. When the noise filtering algorithms confirm the candidate gesture as a true laser pointer gesture, then the portion of the video presentation to be displayed on the one or more secondary screens is displayed concurrently with a synthetic representation of the laser pointer gesture by augmenting the plurality of video frames based on the one or more coordinates of the laser pointer gesture and the homography.