living room. To scroll down a list, or to move to a different menu, the user would point to different parts of the room, or make various hand gestures. If the room inhabitants are engaged in a conversation, they are likely to make hand gestures that look similar to those used for menu control, without necessarily intending to communicate with the virtual control. The popup menu system does not know the intent of the gestures, and may misinterpret them and perform undesired actions in response.

[0024] As another example, a person watching television in a living room may be having a conversation with someone else, or be moving about to lift a glass, grasp some food, or for other purposes. If a gesture-based television remote control were to interpret every user motion as a possible command, it would execute many unintended commands, and could be very ineffective.

[0025] A third limitation of camera-based input systems is that they cannot determine the force that a user applies to a virtual control, such as a virtual key. In musical applications, force is an important parameter. For instance, a piano key struck gently ought to produce a softer sound than one struck with force. Furthermore, for virtual keyboards used as text input devices, a lack of force information can make it difficult or impossible to distinguish between a finger that strikes the typing surface intentionally and one that approaches it or even touches it without the user intending to do so.

[0026] Systems based on analyzing sound information related to user input gestures can address some of the above problems, but carry other disadvantages. Extraneous sounds that are not intended as commands could be misinterpreted as such. For instance, if a virtual keyboard were implemented solely on the basis of sound information, any unintentional taps on the surface providing the keyboard guide, either by the typist or by someone else, might be interpreted as keystrokes. Also, any other background sound, such as the drone of the engines on an airplane, might interfere with such a device.

[0027] What is needed is a virtual control system and methodology that avoids the above-noted limitations of the prior art. What is further needed is a system and method that improves the reliability of detecting, classifying, and interpreting input events in connection with a virtual keyboard. What is further needed is a system and method that is able to distinguish between intentional user actions and unintentional contact with a virtual keyboard or other electronic device.

SUMMARY OF THE INVENTION

[0028] The present invention combines stimuli detected in two or more sensory domains in order to improve performance and reliability in classifying and interpreting user gestures. Users can communicate with devices by making gestures, either in the air, or in proximity with passive surfaces or objects, and not especially prepared for receiving input. By combining information from stimuli detected in two or more domains, such as auditory and visual stimuli, the present invention reduces the ambiguity of perceived gestures, and provides improved determination of time and location of such user actions. Sensory input are correlated in time and analyzed to determine whether an intended command gesture or action occurred. Domains such as vision and sound are sensitive to different aspects of ambient interference, so that such combination and correlation substantially increases the reliability of detected input.

[0029] In one embodiment, the techniques of the present invention are implemented in a virtual keyboard input system. A typist may strike a surface on which a keyboard pattern is being projected. A virtual keyboard, containing a keystroke detection and interpretation system, combines images from a camera or other visual sensor with sounds detected by an acoustic sensor, in order to determine with high accuracy and reliability whether, when, and where a keystroke has occurred. Sounds are measured through an acoustic or piezoelectric transducer, intimately coupled with the typing surface. Detected sounds may be generated by user action such as, for example, taps on the typing surface, fingers or other styluses sliding on the typing surface, or by any other means that generate a sound potentially having meaning in the context of the device or application.

[0030] Detected sounds (signals) are compared with reference values or waveforms. The reference values or waveforms may be fixed, or recorded during a calibration phase. The sound-based detection system confirms keystrokes detected by the virtual keyboard system when the comparison indicates that the currently detected sound level has exceeded the reference signal level. In addition, the sound-based detection system can inform the virtual keyboard system of the exact time of occurrence of the keystroke, and of the force with which the user's finger, stylus, or other object hit the surface during the keystroke. Force may be determined, for example, based on the amplitude, or by the strength of attack, of the detected sound. In general, amplitude, power, and energy of sound waves sensed by the sound-based detection system are directly related to the energy released by the impact between the finger and the surface, and therefore to the force exerted by the finger. Measurements of amplitude, power, or energy of the sound can be compared to each other, for a relative ranking of impact forces, or to those of sounds recorded during a calibration procedure, in order to determine absolute values of the force of impact.

[0031] By combining detected stimuli in two domains, such as a visual and auditory domain, the present invention provides improved reliability and performance in the detection, classification, and interpretation of input events for a virtual keyboard.

[0032] In addition, the present invention more accurately determines the force that the user's finger applies to a typing surface. Accurate measurement of the force of the user input is useful in several applications. In a typing keyboard, force information allows the invention to distinguish between an intentional keystroke, in which a finger strikes the typing surface with substantial force, and a finger that approaches the typing surface inadvertently, perhaps by moving in sympathy with a finger that produces an intentional keystroke. In a virtual piano keyboard, the force applied to a key can modulate the intensity of the sound that the virtual piano application emits. A similar concept can be applied to many other virtual instruments, such as drums or other percussion instruments, and to any other interaction device where the force of the interaction with the typing surface is of interest. For operations such as turning a device on or off, force information is useful as well, since requiring a certain