

[0110] It has been proposed by Sigel and others to utilize the natural image of the finger for this purpose and certain US patents address this in the group referenced above. Copending applications by one of the inventors (Tim Pryor) also describe finger related activity.

[0111] As disclosed in said co-pending application, it is however, often desirable to use retro-reflective material on the finger, disclosed herein as either temporarily attached to the finger as in jewelry or painted on the finger using retro-reflective coating "nail polish" or adhered to the finger such as with adhesive tape having a retro-reflective coating. Such coatings are typically those of Scotchlite 7615 and its equivalent that have high specific reflectivity, contrasting well to their surroundings to allow easy identification. The brightness of the reflection allows dynamic target acquisition and tracking at lowest cost.

[0112] The camera system employed for the purposes of low cost desirable for home use is typically that used for Internet video conferencing and the like today. These cameras are CCD's and more recently CMOS, cameras having low cost (25-100 dollars) yet relatively high pixel counts and densities. It is considered that within a few years these will be standard on all computers, for all intents and purposes, "free" to the applications here proposed, and interfaced via "fire wire" (IEEE 1394) or USB (universal serial bus).

[0113] The use of retroreflective and/or highly distinctive targets (eg bright orange triangles) allows reliable acquisition of the target in a general scene, and does not restrict the device to pointing on a desktop application under controlled lighting as shown in Sigel or others. Active (self luminous) targets such as LEDS also allow such acquisition, but are more costly, cumbersome and obtrusive and generally less preferable.

[0114] If we consider camera system 10 sitting on top of the screen 7 and looking at the user or more particularly, the user's hand, in a normal case of Internet telephony there is a relatively large field of view so that the user's face can also be seen. This same field of view can be used for this invention but it describes a relatively large volume. For higher precision, add-on lenses or zoom lenses on the camera may be used to increase the resolution.

[0115] Or it is possible according to the invention to have a plurality of cameras, one used for the Internet and the other used for the input application here described. Indeed with the ever dropping prices, the price of the actual camera including the plastic lens on the CMOS chip is so low, it is possible perhaps even to have multiple cameras with fixed magnifications, each having a separate chip!

[0116] These can easily be daisy chained with either fire wire or USB such that they can either be selected at will electronically in fact by the different magnifications or pointing directions desired

[0117] Let us now return now to the question of determining location or orientation of a human portion such as typically a hand, or finger—in this case, a finger. In order to make this invention operate in the lowest possible cost it is desirable that the lighting available be low cost as well. Indeed if the camera units are shared with telephony using the natural lighting of the object, then the cost of specialized lighting required for the retro-reflectors adds cost to the system. The power for the lighting, such as LEDs can

generally be conveyed over the USB or 1394 bus however. The user can also point or signal with an object such as 15 having datum 16 on it, such as a retroreflective dot 16 or line target 17.

[0118] It is possible to expand the sensing of 2D positions described above into 3, 4, 5 and 6 dimensions.(x,y plus z, pitch, yaw, roll). Two sensing possibilities of the many possible, are described in various embodiments here in.

[0119] 1. The first, illustrated in FIGS. 1a and b is to utilize a single camera, but multiple discrete features or other targets on the object which can provide a multidegree of freedom solution. In one example, the target spacing on the object is known apriori and entered into the computer manually or automatically from software containing data about the object, or can be determined through a taught determining step.

[0120] 2. The second is a dual camera solution shown in FIGS. 1c and d that does not require a priori knowledge of targets and in fact can find the 3D location of one target by itself, useful for determining finger positions for example. For 6-degree freedom of information, at least three point, targets are required, although line targets, and combinations of lines and points can also be used.

[0121] FIG. 1b illustrates a 3-D (3 Dimensional) sensing embodiment using single camera stereo with 3 or more datums on a sensed object, or in another example, the wrist of the user.

[0122] As shown the user holds in his right hand 29, object 30 which has at least 3 visible datums 32, 33, and 34 which are viewed by TV camera 40 whose signal is processed by computer 41 which also controls projection display 42. TV camera 40 also views 3 other datums 45, 46 and 47, on the wrist 48 of the users left hand, in order to determine its orientation or rough direction of pointing of the left hand 51, or its position relative to object 30, or any other data (eg relation to the screen position or other location related to the mounting position of the TV camera, or to the users head if viewed, or what ever. The position and orientation of the object and hand can be determined from the 3 point positions in the camera image using known photogrammetric equations (see Pinckney, reference U.S. Pat. No. 4,219,847 and other references in papers referenced).

[0123] Alternatively to the 3 discrete point target, a colored triangular target for example can be used in which the intersections of lines fitted to its sides define the target datums, as discussed below

[0124] It is also possible to use the camera 40 to see other things of interest as well. For the direction of pointing of the user at an object 55 represented on display 42 is determine for example datum 50 on finger 52 of users left hand 51 (whose wrist position and attitude can be also determined).

[0125] Alternatively, the finger can be detected just from its general gray level image, and can be easily identified in relation to the targeted wrist location (especially if the user, as shown, has clenched his other fingers such that the finger 52 is the only one extended on that hand).

[0126] The computer can process the gray level image using known techniques, for example blob and other algorithms packaged with the Matrox brand Genesis image processing board for the PC, and determine the pointing