infrared (IR) light source illuminates a rear-side of a light-diffusive screen. As such, silhouette images of the input object can be generated based on a brightness contrast between the input object and IR illumination that is diffused through the light-diffusive screen. The silhouette images of the input object could be, for example, a plurality of matched pairs of images of the input object, such that each image of the matched pair corresponds to the input object from a different perspective at substantially the same time. A given matched pair of images can be employed to determine a location of the input object and the plurality of matched pairs of images can be employed to determine physical motion of the input object.

[0013] A controller can be configured to receive the plurality of images to determine three-dimensional location information associated with the input object. For example, the controller could apply an algorithm to determine the location of one or more features of the input object, such as the user's fingertips, in three-dimensional space. The controller could then translate the simulated inputs into device inputs based on the three-dimensional location information. For example, the controller could interpret gesture inputs based on motion associated with the one or more features of the input object and translate the gesture inputs into inputs to a computer or other device. The controller could also compare the motion associated with the one or more features of the input object with a plurality of predefined gestures stored in a memory, such that a match with a given predefined gesture could correspond with a particular device input.

[0014] FIG. **1** illustrates an example of a gesture recognition interface system **10** in accordance with an aspect of the invention. The gesture recognition interface system **10** includes a first camera **12** and a second camera **14**. The first camera **12** and the second camera **14** may each include an infrared (IR) filter, such that the respective camera may only be able to receive IR light. The gesture recognition interface system **10** also includes a gesture table **16**. As demonstrated in the example of FIG. **1**, the gesture table **16** includes IR light sources **18** that each illuminate an underside of a light-diffusive screen **20**. The light-diffusive screen **20** can be any of a variety of broadband diffuse materials. As such, IR light from the IR light sources **18** is diffused through the light-diffusive screen **20** at the top-surface of the light-diffusive screen **20**. Accordingly, an object that is placed above the light-diffusive screen **20** may appear to each of the first camera **12** and the second camera **14** as a silhouette image, such that it can appear as a substantially darker object in the foreground of an IR emitting light-diffusive screen **20**.

[0015] An input object **22** can provide simulated inputs over the light-diffusive screen **20**. In the example of FIG. **1**, the input object **22** is demonstrated as a user's hand, such that the simulated inputs can be provided through hand gestures. It is to be understood that the use of a hand to provide simulated inputs via hand gestures is but one example implementation of the gesture recognition interface system **10**. Examples of other types of input objects could include a stylus, wand, pointing stick, or any of a variety of devices that could provide gestures to simulate inputs. It is to be further understood that the input object need not be specially designed or suited for use in the gesture recognition interface system **10**.

[0016] In the example of FIG. **1**, the first camera **12** and the second camera **14** each receive separate silhouette images of the input object **22**, where each of the separate silhouette images received, respectively, by the first camera **12** and the second camera **14** are a matched pair. For example, each of the first camera **12** and the second camera **14** could rapidly take still photograph images at, for example, sixty times per second, such that each still photograph image taken by the first camera **12** is matched to a still photograph image taken by the second camera **14** at substantially the same time. The input object **22** can appear to be in a different location relative to the light-diffusive screen **20** in each silhouette image matched pair captured by each of the first camera **12** and the second camera **14**, respectively, due to parallax caused by the different mounted locations of each of the first camera **12** and the second camera **14**.

[0017] Based on the diffusive property of the light-diffusive screen **20**, the IR illumination that is provided from the IR light sources **18** is emitted from the top-surface of the light-diffusive screen **20** in a substantially more uniform manner. In other words, the IR illumination that is provided from the IR light sources **18** is more evenly distributed as it is emitted from the top-surface of the light-diffusive screen **20**. As a result, the input object **22** can be accurately detected in the images that are received by the cameras **12** and **14** regardless of the location of the input object **22** over the light-diffusive screen **20**. Furthermore, as opposed to the use of overhead illumination, the use of the IR illuminators **18** beneath the light-diffusive screen **20** mitigates the appearance of residual shadows that are cast on the top-surface of the light-diffusive screen **20**. Accordingly, outlines and features of the input object **22** over the light-diffusive screen **20**, and thus the resultant gestures, can be accurately determined in the gesture recognition interface system **10**.

[0018] The first camera **12** and the second camera **14** can each provide their respective separate silhouette images of the input object **22** to a controller **24**. The controller **24** could reside, for example, within a computer (not shown) for which the gesture recognition interface system **10** is designed to provide a gesture recognition interface. It is to be understood, however, that the hosting of a controller is not limited to a standalone computer, but could be included in embedded processors. The controller **24** can process the respective silhouette images associated with the input object **22** to generate three-dimensional location data associated with the input object **22**.

[0019] As an example, each of the first camera **12** and the second camera **14** could be mounted at a predetermined angle relative to the light-diffusive screen **20**. For a given matched pair of images of the input object **22**, if the predetermined angle of each of the cameras **12** and **14** is equal, then each point of the input object **22** in two-dimensional space in a given image from the camera **12** is equidistant from a corresponding point of the input object **22** in the respective matched image from the camera **14**. As such, the controller **24** could determine the three-dimensional physical location of the input object **22** based on a relative parallax separation of the matched pair of images of the input object **22** at a given time. In addition, using a computer algorithm, the controller **24** could also determine the three-dimensional physical location of features associated with the input object **22**, such as the inclination and orientation of fingers and/or the location of fingertips.

[0020] Furthermore, the controller **24** can include a calibration unit **26** that is configured to calibrate the gesture recognition interface system **10**. The calibration unit **26** can implement a calibration routine, such as interacting with software of the controller **24**, at power-up or setup of the gesture