

5 **A USER INTERFACE FOR A SYSTEM AND METHOD FOR HEAD SIZE
EQUALIZATION IN 360 DEGREE PANORAMIC IMAGES**

10 This is a continuation-in-part of Application No. 10/186,915 filed June 28,
2002 and entitled "REAL-TIME WIDE-ANGLE IMAGE CORRECTION SYSTEM
AND METHOD FOR COMPUTER IMAGE VIEWING".

15 **TECHNICAL FIELD**

The present invention relates in general to viewing images using a
computer, and more particularly to a user interface for a system and a method for
real-time correction of panoramic images captured by an omni-directional camera
to alleviate distortion and perception problems associated with such images.

20 **BACKGROUND OF THE INVENTION**

Wide-angle cameras, typically having a field of view greater than 60
degrees, often are used to capture large scenes. A wide-angle image taken by a
25 wide-angle camera can be viewed using at least two types of projections. Linear
perspective and cylindrical projections are the most typical projections used to
view wide-angle images.

30 A linear perspective projection is an image captured by a wide-angle lens
that is projected onto a flat piece of film. Linear perspective projection keeps
straight lines straight at the expense of maintaining shape. This causes
perspective deformations. The image would appear correct and not deformed if
the viewer of the image placed their eye at the projection center. However, when

viewing the wide-angle image with a smaller field-of-view the viewer expects smaller increases of image sizes as well as smaller amount of deformation on the image planes as the object rotates. This is why objects appear stretched at the edges. The larger than expected changes in size and deformation on the image plane make the user feel that the scene is not rigid, as if the scene were swimming around the viewer, particularly when viewing panoramic images.

A cylindrical projection is produced using a wide-angle camera having curved film and a rotating slit lens. Cylindrical projections are better at maintaining shape than linear projections. In addition, a cylindrical projection mitigates any apparent swimming motion. Even though the cylindrical projection is a viewing improvement over the linear projection, distortion and perception problems are still present. In particular, the cylindrical projection curves straight lines more than necessary. In addition, the cylindrical projection removes, almost completely, the illusion of turning one's head when viewing panoramic images.

Distortion and curvature problems in wide-angle images are caused by a field-of-view mismatch. In particular, due to limited viewing size on computer monitors and standard viewing distances, the angle subtended by the image when viewed is much smaller than the field-of-view of the camera that originally projected the scene to an image plane. This mismatch is the cause of many distortion and perception problems.

A related problem caused by viewing wide-angle images with a smaller field-of-view at viewing time is the misperception of depth. Wide-angle images exaggerate the depth disparity between near and far objects. One important visual cue of the depth of an object in a scene is the ratio between the image sizes of similar objects placed at near and far locations (called the depth foreshortening ratio). A smaller field-of-view results in a smaller foreshortening ratio. For example, assume a scene contains two persons standing next to each other with one person slightly further away from the camera than the other

person. If the scene is captured by a regular (approximately 60 degree) field-of-view camera and the same scene is captured with a wide-angle field-of-view camera, the size of the objects in the scene will appear different. In particular, with a regular field-of-view camera, the person farther away from the camera will appear slightly further away from the camera than the other person. However, with the wide-angle field-of-view camera, the person farther away from the camera will appear much smaller than the other person. Because of this exaggerated size difference, the person farther away from the camera appears much farther away than he really is. When a wide-angle image of a deep scene (such as a video conferencing scene) is viewed on a computer monitor, the viewer's field-of-view, in general, is much smaller than the field-of-view of the actual images. Therefore, the depth perceived by the viewer is much larger than the actual depth.

One application where these distortion and perception problems manifest themselves is in video conferencing systems. Wide-angle cameras often are used in video conferencing systems to capture and transmit an image containing all participants present in a meeting room. One problem, however, is that when viewing the video conference the wide-angle image exaggerates the depth of the room. This causes the people in the middle of the image (who are usually furthest away from the camera) to appear very small compared to others in the room due to the extreme foreshortening.

These distortion and perception problems typical with wide angle camera images are even more apparent when viewing panoramic images covering a 360 degree field of view. A panoramic image can be generated by aligning and "stitching" input images that are acquired from a single camera. Alternately, a panoramic image can be created by using multiple cameras to acquire the input images and stitching them together. For example, this is the case when multiple cameras are used to generate a panoramic video. Essentially, the creation of a panoramic video entails acquiring multiple videos depicting a full 360-degree

view of the surrounding scene. A camera rig, consisting of multiple cameras disposed in a back-to-back fashion, is sometimes employed for the purpose of capturing these videos. A stitching table which maps pixel positions in the input images to the pixel positions in the panoramic image is typically used to stitch the
5 images together in the panorama. Three hundred sixty degree omni-directional cameras have been used for meeting recording and video conferencing. The camera is typically placed in the center of a table with the meeting participants sitting around the table. If the meeting table has a rectangular shape, people's head sizes on the panoramic images will vary depending on their distances to the
10 camera. For example, the person who sits at the far end of the table will appear much smaller than the person who sits closer to the camera. Such head-size variations in the panoramic images do not look appealing visually and often it is difficult to recognize the people at the far end of the table or recognize their expressions.

15 Another problem with video conferencing employing panoramic images is that some omni-directional cameras generate a 3700x600 pixel panoramic image, but most users only have 1024x768 pixel displays which allow only a 1024x166 pixel panorama to be displayed. With this huge reduction in
20 resolution, the people on the far end of the tables are not viewable without head size normalization.

Accordingly, there exists a need for a system and a method for real-time correction of images captured by a panoramic camera to alleviate distortion and
25 perception problems associated with such images. This system and method should normalize head sizes in panoramic images and allow viewing of persons in the images with standard resolution monitors. Additionally, this system and method should be easy to set up and operate and should be computationally efficient.

SUMMARY OF THE INVENTION

The invention provides for a user interface for a system and method that provides for real-time correction of distortion and perception imperfections in an approximately 360 degree panoramic image. This real-time panoramic image correction system and method, employs parametric warping functions, namely Spatially Varying Uniform (SVU) scaling functions, to warp a panoramic image to equalize people's head sizes while not introducing discontinuity. The SVU function was previously described in patent application serial number 10/186,915 filed June 28, 2002 and entitled "REAL-TIME WIDE-ANGLE IMAGE CORRECTION SYSTEM AND METHOD FOR COMPUTER IMAGE VIEWING". However, the previous patent application used this technique only for up to 180 degree images. In this invention, the SVU scaling function is applied to approximately 360 degree panoramic images while taking horizontal features such as meeting table geometry and a plane touching the meeting participants' heads into account. In the aforementioned patent application, the SVU scaling function was used on up to 180 degree images which are usually generated by a wide-angle camera placed at one end of the meeting room table. In contrast, a 360 degree omni-directional camera is usually placed in the center of the table. Therefore, upper and lower curves used for computing the SVU scaling function for the wide-angle images and the approximately 360 degree panoramic images are different.

The real-time panoramic image correction system and method provides head size normalization, making far away objects, such as heads, appear closer and therefore larger. Close heads, which are big, are also made smaller. As a result people on the far end of a table are larger and are viewable even on 1024x166 pixel displays. In one embodiment the system and method of the invention, a user also selects the conference table shape and size and the real-time panoramic image correction system modifies the panoramic image stitching

table to match the meeting table so that the head sizes of meeting participants are normalized at the same time the stitching is computed.

The User Interface (UI) of the real-time panoramic image correction system and method simplifies the use of the system for the user. The UI includes a control panel that allows a user to enter meeting table size and shape and the amount of normalization desired (e.g. 0 to 100%). A speaker window can also be implemented on a display that displays the corrected panoramic image. In this speaker window, the normalized head from a meeting participant, particularly one that is speaking, can be extracted from the panoramic image and displayed in a separate window. Additionally, the corrected panoramic image can be displayed with extra pixels around its perimeter in order to make it conform to a standard size since corrected panoramic images will vary in size based on the amount of normalization applied. For example, a black bar may be placed across the top and bottom of the corrected panoramic image. This allows the corrected or normalized panoramic image to be transmitted over a network at a standard size and allows the image to adapt to any of the standard display sizes and resolutions.

20

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention can be further understood by reference to the following description and attached drawings that illustrate aspects of the invention. Other features and advantages will be apparent from the following detailed description of the invention, taken in conjunction with the accompanying drawings, which illustrate, by way of example, the principles of the present invention.

FIG. 1 shows an image taken by a 360 degree panoramic camera in a meeting room.

FIG. 2 shows the image in FIG. 1 corrected by the real-time panoramic image correction system and method of the invention.

5 FIG. 3 shows the upper and lower curves for the SVU scaling function (green curves) for the 180 degree images.

FIG. 4 shows the upper and lower curves for the SVU scaling function (green curves) for the 360 degree panoramic images.

10

FIG. 5A is a block diagram illustrating the general overview of one embodiment of the real-time panoramic image correction system of the invention.

15 FIG. 5B is a block diagram illustrating the general overview of another embodiment of the real-time panoramic image correction system of the invention which employs a concatenated stitching and warping table.

FIG. 6 is a block diagram illustrating a computing apparatus suitable for carrying out the invention.

20

FIG. 7 is a block diagram illustrating the details of the warp table generation module of the real-time panoramic image correction system shown in FIGs. 5A and 5B.

25 FIG. 8A is a general flow diagram illustrating the operation of the real-time panoramic image correction system shown in FIG. 5A.

FIG. 8B is a general flow diagram illustrating the operation of the real-time panoramic image correction system shown in FIG. 5B.

30

FIG. 9 is a flow diagram illustrating the details of the vertical scaling module shown in FIG. 7.

5 FIG. 10 illustrates the determination of the warping function by two sets of curves.

FIG. 11 is a diagram demonstrating the projection of the table edges onto cylindrical film (assuming the standard camera orientation).

10 FIG. 12 is a flow diagram illustrating the details of the horizontal scaling module shown in FIG. 7.

FIG. 13 is a flow diagram illustrating the details of the horizontal distortion correction module shown in FIG. 7.

15

FIG. 14 illustrates the vertical scaling function and a conceptual view of the division of the preliminary pixel coordinates into three sections.

20 FIG. 15 depicts a padded corrected panoramic image wherein extra pixels are added around the perimeter of the image so that it conforms to a standard size for network transmission or display.

DETAILED DESCRIPTION OF THE INVENTION

25

In the following description of the invention, reference is made to the accompanying drawings, which form a part thereof, and in which is shown by way of illustration a specific example whereby the invention may be practiced. It is to be understood that other embodiments may be utilized and structural changes may be made without departing from the scope of the present invention.

30

1.0 General Overview

FIG. 1 shows an image taken by a 360 degree omni-directional camera sitting in the middle of a table in a meeting room. This camera configuration is typical of what one might use in a video conferencing application. Referring to FIG. 1, one can see that the person in white appears much smaller than the other two people. The reason for this is that he sits further away from the camera. First of all, the person in white is much less visible and appears far away from the viewer thus affecting the experience of real time communication. Second, the image size is usually limited due to network bandwidth limitation, and due to the limited screen space. Many transmitted pixels are wasted if people's images are very small. Therefore it is desirable to equalize people's head sizes to maximally utilize the number of pixels being transmitted and displayed. FIG. 2 shows the image of FIG. 1 corrected using the real-time panoramic image correction system and method.

The real-time panoramic image correction system and method employs a technique to equalize or normalize people's head sizes while not creating image discontinuities. The technique is an extension of the SVU scaling function as described in patent application serial number 10/186,915 filed June 28, 2002 and entitled "REAL-TIME WIDE-ANGLE IMAGE CORRECTION SYSTEM AND METHOD FOR COMPUTER IMAGE VIEWING". In the previous patent application, a SVU scaling function was used on 180 degree images which are usually generated by a wide-angle camera placed at one end of the meeting room table. In contrast, as discussed above, a 360 degree camera is usually placed in the center of the table. Therefore, upper and lower curves for computing the SVU scaling function generated for the 360 degree panoramic images are different from those used in the wide-angle (e.g., up to 180 degrees) images. FIG. 3 shows the upper and lower source curves used to calculate the SVU scaling function (green curves) for the 180 degree images. FIG. 4 shows the upper and lower source curves used to calculate the SVU scaling function

(green curves) for the 360 degree image. Given the green source curves and the target curves (red curves), one can generate the SVU scaling functions in the same way as what is described in the aforementioned patent application.

5 Like in the previously mentioned co-pending patent application, correction of the panoramic image is achieved using a class of parametric warping functions called Spatially Varying Uniform (SVU) scaling functions that preserve local scaling perspectives and correct for depth misperception. The SVU scaling functions and scaling factors are used to perform vertical scaling and horizontal
10 scaling on the pixel coordinates of the image. This produces a preliminary warp table which maps the pixel locations of the preliminary corrected image to the original pixel coordinates. *This preliminary corrected image is a virtual image and is not actually constructed.* However, the preliminary corrected image maintains vertical lines as straight but distorts horizontal lines. This horizontal
15 distortion is corrected by performing a horizontal distortion correction using the SVU scaling functions and at least two different scaling factors. This processing yields a warp table. The warp table maps positions in a corrected panoramic image to the original pixel coordinates in the distorted panoramic image. In some embodiments of the invention wherein images are stitched to create the
20 panoramic image the warping table can be concatenated with the stitching table to increase computational efficiency. Using the warp table, or concatenated stitching and warp table, the corrected panoramic image is constructed in real time from the original panoramic image.

25 FIG. 5A is a block diagram illustrating the general overview of one embodiment of the real-time panoramic image correction system 500A of the invention. In general, the system 500A obtains horizontal feature geometry 505 (e.g., meeting table geometry and a plane touching the top of people's heads) and a panoramic image 510 and inputs the horizontal feature geometry 505 and
30 the panoramic image pixel coordinates 520 corresponding to the coordinates of each of the pixel within the panoramic image. The horizontal feature geometry

505 and panoramic image pixels coordinates 520 are processed by the real-time panoramic image correction system 500A and a corrected panoramic image 530 is the output.

5 The panoramic image 510 may be a single stitched image (such as from a still camera) or part of a sequence of images (such as from a video camera). The coordinates of each pixel within the panoramic image 510 are obtained to generate the panoramic image pixel coordinates 520. For example, in a rectilinear coordinate frame the pixel coordinates 520 are (x,y) locations of each
10 of the pixels within the panoramic image 510 which typically correspond to a series of images that were stitched to obtain the panoramic image. The real-time panoramic image correction module 500A includes a warp table generation module 540, a warp table 550 and a real-time image warping system 560.

15 The warp table generation module 540 is used to generate the warp table 550 using parametric image warping functions. The panoramic image pixel coordinates 520 are processed to create the warp table 550. The details of the warp table generation module 540 are discussed below. The warp table 550 contains the panoramic image pixel coordinates 520 arranged in such a way to
20 correspond to a location in the corrected panoramic image 530. Thus, the warp table 550 determines the new location in the corrected panoramic image 530 for the panoramic image pixel coordinates 520.

25 Once the warp table 550 is generated a panoramic image 510 is received as input to the real-time warping system 560. The real-time panoramic image warping system 560 applies the warp table 550 to the panoramic image 510 to create the corrected panoramic image 530. The real-time warping system 560 creates the corrected panoramic image 530 for each location in the corrected panoramic image 530 by obtaining the RGB values for the panoramic image 510
30 located at the pixel coordinates contained in the warp table 550. The RGB values at the panoramic pixel coordinates 520 are thus relocated in the corrected

panoramic image 530. The new location is determined using the warp table 550. It should be noted that any color space could be used, such as, for example, YUV.

5 In another embodiment of the invention 500B, shown in FIG. 5B, the warp table is concatenated with the stitching table used to stitch together the panoramic image. In this embodiment, a set of input images 512, are stitched into an imaginary panoramic image 515 and the image pixel coordinates for imaginary stitched panoramic image 522 are input into the system 500B. The
10 real-time panoramic image correction module 500B includes a warp table generation module 540, a warp table 550, a concatenated stitching and warping table 555 (discussed in greater detail in section 4.5), and a real-time image warping system 560.

15 The warp table generation module 540 is used to generate the warp table 550 using parametric image warping functions. The panoramic image pixel coordinates 522 of the imaginary panoramic image 515 are processed to create the warp table 550. The details of the warp table generation module 540 are discussed below. The warp table 550 contains the panoramic image pixel
20 coordinates 522 arranged in such a way to correspond to a location in the corrected panoramic image 530. Thus, the warp table 550 determines the new location in the corrected panoramic image 530 for the panoramic image pixel coordinates 522.

25 Once the warp table 550 is generated it is concatenated with the stitching table used to stitch the set of input images 512 to create a concatenated warping and stitching table 555. The set of input images 512 is received as input to the real-time image warping system 560. The real-time panoramic image warping system 560 applies the concatenated warping and stitching table 555 to the set
30 of input images 512 to create the corrected panoramic image 530. The real-time image warping system 560 creates the corrected panoramic image 530 for each

location in the corrected panoramic image 530 by obtaining the RGB values (or YUV, or other color space) for the set of input images 512 located at the pixel coordinates contained in the concatenated warping and stitching table 555. The RGB values at the imaginary panoramic pixel coordinates 522 are thus relocated
5 in the corrected panoramic image 530 during while the set of input images 512 are stitched. The new location is determined using the concatenated warping and stitching table 555.

The warping process occurs in real-time because the stitcher is static and
10 requires minimum computation. The generation of the warp table means that all that is required to create the corrected panoramic image 530 is to apply the warping table 550 (or the concatenated warp and stitching table) to the panoramic image 510 or set of input images 512, which can be done quickly.

15 2.0 Exemplary Operating Environment

The real-time panoramic image correction system 500A or 500B of the present invention is designed to operate in a computing environment. The following discussion is intended to provide a brief, general description of a
20 suitable computing environment in which the invention may be implemented.

FIG. 6 is a block diagram illustrating a computing apparatus suitable for carrying out the invention. Although not required, the invention will be described in the general context of computer-executable instructions, such as program
25 modules, being executed by a computer. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Moreover, those skilled in the art will appreciate that the invention may be practiced with a variety of computer system configurations, including personal computers, server
30 computers, hand-held devices, multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, minicomputers, mainframe

computers, and the like. The invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located on both local and remote computer storage media including memory storage devices.

With reference to FIG. 6, an exemplary system for implementing the invention includes a general-purpose computing device 600. In particular, the computing device 600 includes the processing unit 602, a system memory 604, and a system bus 606 that couples various system components including the system memory 604 to the processing unit 602. The system bus 606 may be any of several types of bus structures including a memory bus or memory controller, a peripheral bus, and a local bus using any of a variety of bus architectures. The system memory includes read only memory (ROM) 610 and random access memory (RAM) 612. A basic input/output system (BIOS) 614, containing the basic routines that help to transfer information between elements within the computing device 600, such as during start-up, is stored in ROM 610. The computing device 600 further includes a hard disk drive 616 for reading from and writing to a hard disk, not shown, a magnetic disk drive 618 for reading from or writing to a removable magnetic disk 620, and an optical disk drive 622 for reading from or writing to a removable optical disk 624 such as a CD-ROM or other optical media. The hard disk drive 616, magnetic disk drive 628 and optical disk drive 622 are connected to the system bus 606 by a hard disk drive interface 626, a magnetic disk drive interface 628 and an optical disk drive interface 630, respectively. The drives and their associated computer-readable media provide nonvolatile storage of computer readable instructions, data structures, program modules and other data for the computing device 600.

Although the exemplary environment described herein employs a hard disk, a removable magnetic disk 620 and a removable optical disk 624, it should be appreciated by those skilled in the art that other types of computer readable

media that can store data that is accessible by a computer, such as magnetic cassettes, flash memory cards, digital video disks, Bernoulli cartridges, random access memories (RAMs), read-only memories (ROMs), and the like, may also be used in the exemplary operating environment.

5

A number of program modules may be stored on the hard disk, magnetic disk 620, optical disk 624, ROM 610 or RAM 612, including an operating system 632, one or more application programs 634, other program modules 636 and program data 638. A user (not shown) may enter commands and information
10 into the computing device 600 through input devices such as a keyboard 640 and a pointing device 642. In addition, a camera 643 (such as a video camera) may be connected to the computing device 600 as well as other input devices (not shown) including, for example, a microphone, joystick, game pad, satellite dish, scanner, or the like. These other input devices are often connected to the
15 processing unit 602 through a serial port interface 644 that is coupled to the system bus 606, but may be connected by other interfaces, such as a parallel port, a game port or a universal serial bus (USB). The monitor 646 (or other type of display device) is also connected to the system bus 606 via an interface, such as a video adapter 648. In addition to the monitor 646, computing devices such
20 as personal computers typically include other peripheral output devices (not shown), such as speakers and printers.

The computing device 600 may operate in a networked environment using logical connections to one or more remote computers, such as a remote
25 computer 650. The remote computer 650 may be another personal computer, a server, a router, a network PC, a peer device or other common network node, and typically includes many or all of the elements described above relative to the computing device 600, although only a memory storage device 652 has been illustrated in FIG. 6. The logical connections depicted in FIG. 6 include a local
30 area network (LAN) 654 and a wide area network (WAN) 656. Such networking

environments are commonplace in offices, enterprise-wide computer networks, intranets and the Internet.

When used in a LAN networking environment, the computing device 600 is connected to the local network 654 through a network interface or adapter 658. When used in a WAN networking environment, the computing device 600 typically includes a modem 660 or other means for establishing communications over the wide area network 656, such as the Internet. The modem 660, which may be internal or external, is connected to the system bus 606 via the serial port interface 644. In a networked environment, program modules depicted relative to the computing device 600, or portions thereof, may be stored in the remote memory storage device 652. It will be appreciated that the network connections shown are exemplary and other means of establishing a communications link between the computers may be used.

3.0 Real-time Panoramic Image Correction System Details.

In general, the real-time panoramic image correction system 500 provides real-time correction of a panoramic image 510, and in particular normalizes or equalizes the head sizes of people sitting further away from the camera. The system 500 corrects for curvature and distortions in the vertical and the horizontal directions. In addition, the system 500 corrects for perception problems such as depth and swimming motions.

FIG. 7 is a block diagram illustrating the details of the warp table generation module 540 shown in FIGs. 5A or 5B. In particular, the warp table generation module 540 includes an input module 700 for inputting the panoramic image pixel coordinates 520 or the pixel coordinates of the imaginary panoramic image 522 and the horizontal feature geometry 505. These pixel coordinates 520, 522 represent a location of each pixel in the panoramic image 510 or the imaginary panoramic image 515. The module 540 also includes a vertical

scaling module 710. The vertical scaling module 710 uses warping functions to process each vertical scanline of the panoramic image pixel coordinates 520 or pixel coordinates of the imaginary panoramic image 522 such that the panoramic image pixel coordinates 520, 522 are scaled in a vertical direction. Conceptually, this can be thought of as “stretching” the panoramic image 510 (or the imaginary panoramic image 515) in the vertical direction. This vertical scaling corrects in part for exaggerated depth in the panoramic image 510 (or the imaginary panoramic image 515).

The warp table generation module 540 also includes a horizontal scaling module 720 for using parametric image warping functions to scale the image pixel coordinates 520, 522 in the horizontal direction. Similar to the vertical scaling, conceptually this can be thought of as “stretching” the panoramic image 510 or the imaginary panoramic image 515 in the horizontal direction. In addition, the horizontal scaling is performed individually on each vertical scanline such that aspect ratio of the panoramic image 510 or imaginary panoramic image 515 is preserved.

The vertical scaling module 710 and the horizontal scaling module 720 create a preliminary warp table which maps the preliminary pixel coordinates to the original pixel coordinates. This preliminary warp table could be used to create a preliminary corrected image. In practice this is not done, and the preliminary corrected image is a virtual image. The preliminary corrected image has the properties that vertical lines are maintained as vertical but horizontal lines are distorted.

In order to correct for this horizontal distortion, a horizontal distortion module 730 also is included in the warp table generation module 540. The horizontal distortion module 730 vertically scales the preliminary pixel coordinates in the preliminary warp table by different amounts depending on location. The output of the warp table generation module 540 is the warp table

550. The warp table maps the pixel coordinates in the corrected panoramic image 530 to the original panoramic image pixel coordinates 520 or the pixel coordinates of the imaginary stitched image 522. The warp table can be concatenated with the stitching table, creating a concatenated warping and stitching table, and can be used to stitch together the panoramic image from a set of input images.

4.0 Real-time Panoramic Image Correction Method.

10 The real-time panoramic image correction method of the present invention uses parametric image warping functions named Spatially Varying Uniform Scaling functions, or SVU scaling functions. Although SVU scaling functions may locally resemble a uniform scaling function to preserve aspect ratios, the scaling factor varies over the image to create warp. SVU scaling functions avoid
15 rotations and remove swimming motions when viewing panoramic images.

FIG. 8 is a general flow diagram illustrating the operation of one embodiment of the real-time panoramic image correction method of the real-time panoramic image correction system 500A shown in FIG. 5A. Panoramic image pixel coordinates 520 from the panoramic image 510 and horizontal feature geometry 505 are received as input (box 800). Next, the warp table 550 is
20 generated (box 810). This warp table 550 is used to determine a location of the corrected panoramic image pixel coordinates 530.

25 As discussed in detail below, the generation of warp table 550 includes vertical scaling, horizontal scaling, and horizontal distortion correction. Vertical scaling corrects the panoramic image 510 for exaggerated depth. In order to correct for exaggerated depth, the distant objects or people in the center of the panoramic image 510 need to be enlarged relative to those objects or people
30 who are close to the camera.

In order to maintain aspect ratio of the panoramic image 510 horizontal scaling is also performed. Specifically, parametric image warping functions (such as the SVU scaling functions) are used to perform horizontal scaling to generate a new width for each vertical scanline for the panoramic image 510. While the
5 SVU scaling functions maintain vertical lines as vertical, they induce a certain amount of vertical shear. Visually, this is perceived as slanting and distorted horizontal lines. Scenes often contain horizontal surfaces near the top or bottom of the image, such as a table and a ceiling on a room. for which the distortions may be noticeable.

10

In order to correct for this vertical shear, the real-time panoramic image correction method further includes performing horizontal distortion correction to correct the panoramic image 510 for horizontal distortion. In general, this is accomplished by relaxing the uniformity of the vertical scaling and nonlinearly
15 scaling at a vertical location in the panoramic image 510 (such as along a vertical scanline). After the vertical scaling, horizontal scaling, and horizontal distortion correction, the warp table 550 is generated that maps a pixel in the corrected panoramic image 530 to the panoramic image 510.

20

Once the warp table 550 has been generated, the panoramic image 510 is received as input (box 820). The panoramic image 510 then is warped in real time to create the corrected panoramic image 530. The warp table established where a pixel in the panoramic image 510 is to be found for each position in the corrected panoramic image 530 to produce a distortion-free panoramic image.

25

FIG. 8B is a general flow diagram illustrating the operation of one embodiment of the real-time panoramic image correction method of the real-time panoramic image correction system 500B shown in FIG. 5B, which operates in a manner very similar to that described with respect to FIG. 8A. Image pixel
30 coordinates 522 from the imaginary panoramic image 515 and horizontal feature geometry 505 are received as input (box 800). Next, the warp table 550 is

generated (box 810). The warp table is then concatenated with the warp table 550 to create a concatenated warping and stitching table (box 815). A set of input images used to generate the imaginary panoramic image is input (box 820). The concatenated warping and stitching table 555 is then used to determine a location of the corrected panoramic image pixel coordinates 530 while stitching the set of input images (box 830). The corrected panoramic image is then output (box 840).

The details of the various scaling procedures and modules are discussed in further detail below.

4.1 Vertical Scaling

FIG. 9 is a flow diagram illustrating the details of the vertical scaling module shown 710 in FIG. 7. The horizontal feature geometry 505 and panoramic pixel coordinates 520 or 522 are received as input (box 900). Next, parametric image warping functions (such as SVU scaling functions) are determined using a set of curves and scaling factors. In particular, source curves are specified (box 910) along with a warping factor (box 920). Using these source curves and the warping factor, target curves are computed (box 930).

FIG. 10 illustrates the determination the warping function by two sets of curves: (1) source curves; and (2) target curves. In one aspect of the invention, the source curves are input by a user via a simple user interface. These source curves provide the parameters for the parametric image warping function. Through the user interface a user is asked to define two cubic curves. These two *source* curves define common (real world) horizontal features, such as the tops of people's heads, and the edge of the table. As shown in FIG. 10, a top source curve S_t and a bottom source curve S_b are specified. It can be noted by comparison of FIG. 10 to FIG. 4 that in order to compute the source and target curves for the panorama four different parts of the upper and lower source curves

must be defined in order to define the source curves for the entire panoramic image.

- Instead of manually designating the source curves, they can be
- 5 automatically determined from the horizontal feature geometry such as, for example, the meeting table and a plane touching the top of meeting participants' heads. Suppose the meeting table is rectangular. The green source curves can be computed if one knows the table width W and length L , the camera height h , the focus length r of the camera, and the camera orientation θ . The focus
- 10 length r and camera height h are known once the camera is made. The table dimensions can be provided by the user during installation. FIG. 11 is a diagram demonstrating the projection of the table edges onto the cylindrical film (assuming the standard camera orientation). The following are the equations for the four parts of the lower source curves representing the table boundaries on
- 15 the cylindrical image.

$$z = h(1 - \frac{r \cos \theta}{W}) \quad \theta \in [-\alpha, \alpha], \alpha = \arccos(\frac{W}{\sqrt{W^2 + L^2}})$$

$$z = h(1 - \frac{r \sin \theta}{L}) \quad \theta \in [\alpha, \pi - \alpha]$$

$$z = h(1 + \frac{r \cos \theta}{W}) \quad \theta \in [\pi - \alpha, \pi + \alpha]$$

$$z = h(1 + \frac{r \sin \theta}{L}) \quad \theta \in [\pi + \alpha, 2\pi - \alpha]$$

Similarly, the following are the equations for the four portions of the upper source curves representing the plane touching the top of people's heads on the

cylindrical image, where h' represents the distance from the camera to the plane of the top of the meeting participants' heads .

$$z = -h'(1 - \frac{r \cos \theta}{W}) \quad \theta \in [-\alpha, \alpha], \alpha = \arccos(\frac{W}{\sqrt{W^2 + L^2}})$$

$$z = -h'(1 - \frac{r \sin \theta}{L}) \quad \theta \in [\alpha, \pi - \alpha]$$

$$z = -h'(1 + \frac{r \cos \theta}{W}) \quad \theta \in [\pi - \alpha, \pi + \alpha]$$

$$z = -h'(1 + \frac{r \sin \theta}{L}) \quad \theta \in [\pi + \alpha, 2\pi - \alpha]$$

5

The warping factor, α , which can also be chosen by the user, determines how much the image is warped. The warping factor, α , lies between 0 and 1, where $\alpha = 0$ is no warping and $\alpha = 1$ is the maximum warping. Conceptually, specifying $\alpha = 0$ will leave the image untouched, and specifying $\alpha = 1$ will pull
 10 pixels on source curves to the lines between the end points. Typically, the ideal warping factor lays approximately half way between 0 and 1.

Once the source curves and the warping factor are specified, the target curves can be computed (box 930). Let $y = S_t(x)$ and $y = S_b(x)$ be the equations
 15 of the top and bottom source curves respectively. Two *target* curves (where points on the source curve will move to) are determined by the source curves and α . Specifically, a top target curve T_t and a bottom target curve T_b are defined. If an equation of the line is defined between the end points of $S_t(x)$ as $y = y_t(x)$, and the equation of line connecting the bottom source ends as $y = y_b(x)$,

then the equation of the top target curve is $T_t(x) = (1 - \alpha)S_t(x) + \alpha y_t(x)$, and $T_b(x) = (1 - \alpha)S_b(x) + \alpha y_b(x)$.

5 A main scaling factor, $r(x)$, then is computed using the source curves and the target curves (box 940). In particular, given any vertical scanline x as shown in FIG. 10, let A, B denote its intersections with the source curves, and A', B' the intersections with the target curves. The SVU scaling function will scale AB to $A'B'$. Let

$$r(x) = \frac{\|A'B'\|}{\|AB\|}$$

$$r(x) = \frac{T_t(x) - T_b(x)}{S_t(x) - S_b(x)}$$

10

The ratio, $r(x)$, is a main scaling factor for a vertical scan line located at horizontal position x . The panoramic pixel coordinates 520 of the vertical scanline then are scaled in the vertical direction using the main scaling factor, $r(x)$ (box 950).

15

4.2 Horizontal Scaling

In order to preserve the aspect ratio, the panoramic pixel coordinates 520 also are scaled in the horizontal direction. This horizontal scaling uses the same main scaling factor, $r(x)$. In other words, just as for the vertical scaling a vertical scan line is also scaled horizontally by the main scaling factor, $r(x)$, to preserve aspect ratio.

20

FIG. 12 is a flow diagram illustrating the details of the horizontal scaling module shown in FIG. 7. The vertically-scaled panoramic pixel coordinates are received as input (box 1200). Using the main scaling factor, $r(x)$, the vertically-scaled panoramic pixel coordinates are scaled in the horizontal direction (box

25

1210). Once the vertically-scaled panoramic pixel coordinates 520 are scaled, the total width of a preliminary corrected image, w' , becomes,

$$w' = \int_0^w r(x) dx$$

5

where w is the width of the panoramic (or source) image 510.

Next, a preliminary warp table is generated (box 1220). The preliminary warp table contains preliminary pixel coordinates. The preliminary pixel coordinates are the panoramic pixel coordinates that have been vertically and horizontally scaled. Conceptually, the preliminary pixel coordinates can be used to construct a preliminary corrected image. Thus, for any pixel (x, y) in the panoramic image 510, let (x', y') denote its new position in the preliminary corrected image. This yields,

$$x' = \int_0^x r(x) dx$$

15

$$y' = T_t(x) + r(x) * (y - S_t(x))$$

The above equation is the forward mapping equation for the SVU scaling function. The SVU scaling function is not a perfect uniform scaling everywhere. It is easy to prove that the only function that is a perfect uniform scaling everywhere is a uniform global scaling function.

20

The SVU scaling function is similar to a projection onto a generalized cylindrical surface. However, such a simple projection does not produce locally uniform scaling. Local uniform scaling is desirable, and this lack of local uniform scaling causes objects in the preliminary corrected image to appear to be stretched.

25

4.3 Horizontal Distortion Correction

Once the panoramic pixel coordinates 520 have been scaled in the vertical and the horizontal directions, the resulting preliminary warp table is
5 corrected for horizontal distortion. Horizontal distortion correction is needed because, while the parametric class of image warping functions (such as a SVU scaling function) maintains vertical lines as vertical, they distort horizontal lines. To minimize this problem the uniformity of the scaling is relaxed and nonlinearly scaled at each vertical scan line.

10 FIG. 13 is a flow diagram illustrating the details of the horizontal distortion module 730 shown in FIG. 7. In general, the module 730 divides the preliminary pixel coordinates in the preliminary warp table into sections and scales each section in the vertical direction according to a specific scaling factor for that
15 section. The scaling factors may be different between sections, and are usually different for at least one section.

Specifically, the horizontal distortion correction module 730 initially divides the preliminary warp table into sections (box 1300). In one aspect of the
20 invention, the preliminary corrected image is conceptually divided as shown in FIG. 14. FIG. 14 illustrates the vertical scaling function and a conceptual view of the division of the preliminary corrected image into three sections. It should be noted that in FIG. 14 y is the vertical direction. Referring to FIG. 14, a first section 1400 is defined between the bottom source curve S_b and the top source
25 curve S_t . Other sections include section outside the source curves including a second section 1410 below the source curves and a third section 920 above the source curves.

The preliminary pixel coordinates located in the first section 1300 between
30 the source curves are scaled in the vertical (y) direction using the main scaling factor, $r(x)$ (box 1310). This scaling in the vertical direction is performed as

described above. The preliminary pixel coordinates located in the sections outside the source curves (the second section 1310 and the third section 1320) are scaled less in the vertical direction. This is performed by initially specifying a secondary scaling factor (box 1320). The secondary scaling factor, s , is shown in FIG. 14 below the main scaling factor, $r(x)$. Next, the preliminary pixel coordinates located in other sections 1410, 1420 outside of the source curves are scaled in the vertical direction using the secondary scaling factor (box 1330). It should be noted that the horizontal scaling remains the same (in other words, using the main scaling factor, $r(x)$) so that the straightness of vertical lines is maintained. In order to maintain continuity between sections, transitions between sections are made by applying a smoothing parameter, w (box 1340). The smoothing parameter smoothes the vertical scaling function and smoothly connects the disparate scaling factors as the vertical scaling function crosses the source curves.

Referring to FIGs. 10 and 14, consider the vertical line at x shown in FIG. 10. As shown in FIG. 14, $g(y)$ is the vertical scaling function that can be defined to be the vertical scaling factor at any point y on this vertical line. It should be noted that $g(y)$ is dependent on x . The function $g(y)$ is controlled by two parameters, the secondary scaling factor, s , and the smoothing parameter, w . The portion of the vertical scan line more than $w/2$ distance from the source curves is scaled by the main scaling factor, $r(x)$, between the source curves and by the secondary scaling factor, s , outside the source curves. The three constant segments are glued together by two cubic splines in $[S_t - 0.5w, S_t + 0.5w]$. Each cubic spine has ends with values s and $r(x)$ and a slope of 0 at both ends.

The smoothing parameter w controls the continuity at the source curves. For example, if the scene is discontinuous at the source curves, one can choose a very small w without noticeable artifacts. In the special case when $s = r(x)$, $g(y)$ becomes a constant, which is what was assumed in deriving the new pixel positions in the preliminary corrected image. Once the horizontal distortion

correction function has been applied to the preliminary warp table, the warp table 550 is generated that contains the locations of the panoramic pixel coordinates in the corrected panoramic image 530 (box 1350). The warp table 550 then is sent as output (box 1360). The warp table is used for constructing the corrected panoramic image 530 from the panoramic image 510.

4.4 Correction for Panoramic Image Viewing Misperception.

The SVU scaling function helps correct for the problems of depth misperception. This is achieved by using a symmetric SVU scaling function for viewing panoramic images. The bottom source curve is the mirror of the top source curve. The source curves pass through the image corners and have a minimum (or maximum) at the center vertical scan line. One parameter specifies the height of the center points and a second parameter is the warping factor, α , as discussed above.

4.5. Concatenated Warping/Stitching Table.

As mentioned previously, the warping table can be concatenated with the stitching table used to stitch images into the panoramic image. For example, in one embodiment, for a round meeting table, a cylindrical projection of the meeting room is created by stitching the images captured from multiple cameras together. The stitcher uses a stitching table that maps the camera images onto the panoramic image. The stitching table is typically created at the factory by the camera manufacturer during manufacturing calibration. The table contains the following information for each pixel (x,y) in the panorama:

(x,y) – panorama pixel coordinate

(u1,v1) – image pixel coordinate for camera 1

(u2,v2) – image pixel coordinate for camera 2

camera1 – main source camera

camera2 – second source camera if (x,y) is in an overlap region.

Let $S(x,y)$ denote the stitcher table entry at pixel (x,y) . That is, $S(x,y)=\{\text{camera1}, u1,v1, \text{camera2}, u2,v2\}$. To convert the stitcher from cylindrical to rectangular, one needs to concatenate the stitching table with the warping table. For each pixel (x,y) on the final panoramic image (after head-size normalization), let $W(x,y)$ denote the table entry which is the pixel coordinate on the cylindrical panoramic image and let T denote the concatenated table of S and W . Then $T(x,y) = S(W(x,y))$. This concatenated table can be pre-computed at installation time. At run time, the operation is the same as that for panorama stitching.

5.0 User Interface.

One embodiment of the real-time panoramic image correction system provides a User Interface (UI) to simplify the operation of the system for the user. The UI can include a control panel (e.g. a touch panel Liquid Crystal Display) for entering desired settings, as well as a display with a speaker window wherein a normalized head is extracted from the panoramic image and displayed. In addition, the UI can include a display for displaying the corrected or scaled panoramic images, the size of which will vary with the warping factor applied, with extra pixels around the perimeter of the image. This allows transmission of the corrected image over a network at a standard size without having to renegotiate the network stack in order to send images of different sizes and resolutions.

5.1 Control Panel.

The omni-directional camera used to capture the panoramic image may have a control panel, such as an touch LCD screen, which is used for setup and normal usage of the real-time panoramic image correction system of the invention. In one embodiment, the table shape and size, camera position and orientation and the head size normalization setting (0-100%) can be entered.

The table shape and size settings may include: round (preferably the default); small rectangular (e.g., 10x5'); large rectangular (e.g., 16x5'); and other (in this case the user is asked to specify length and width of the table). The default table setting (round) generates stitched images for non-round tables, but does not

5 normalize the head sizes for those table types.

As discussed above, in one embodiment of the invention, the omni-directional camera position and orientation can be set via the control panel. The default position is in the center of the table. The default orientation is with the

10 LCD readable near the long-end of the table (i.e., the text is parallel to the table's major axis of symmetry. The orientation of the camera is very important, since if it is incorrect, the warping can actually make the head sizes more different in size instead of equal in size. The camera orientation can vary +/- 15 degrees without noticeable effects. Various conventional methods can be used to maintain

15 camera orientation after setup. The camera table shape and size and omni-directional camera position and orientation can also be determined automatically.

Furthermore, in one embodiment of the invention, the head size normalization percentage can be set via the LCD. This parameter may be set

20 with a slider that allows continuous settings from 0% normalized (cylindrical) to 100% normalized (rectangular). The main reason to not have the setting at 100% is to make the normalization more robust to camera position and rotation changes, and also to provide some perspective distortion that some users may desire.

25

5.2 Speaker Window.

The UI can include a display pane with a speaker window wherein a head is extracted from the panoramic image by directly copying a sub-window from the

30 panoramic image. This sub-window is then displayed as a separate display pane. The extracted head can be non-normalized or normalized. This feature is

especially useful when a person is speaking in a video conference, especially when they are located relatively far away from the camera. By extracting the enlarged head the person's expressions are visible. This is especially beneficial when they are speaking.

5

5.3 Padded Panoramic Image and Anamorphic Panoramic image.

In addition, the UI may include a feature for displaying and transmitting a padded panoramic image. This padded panoramic image display adds extra
10 pixels around the corrected image to make it a standard size, since displaying images with different warp factors will be corrected to different sizes. For example, bars, such as black bars can be added on the top and the bottom of the image. An example of a padded corrected image is shown in FIG. 15. This padded panoramic image allows the system maintain a constant size and
15 resolution for each corrected panoramic image. This also allows for the transmission of each corrected panoramic image without renegotiating the network stack in order to send images of different sizes and resolutions. The party receiving the image can then display the padded panoramic image with or without the bars.

20

Another method to send head-size equalized or cylindrical panoramic video over a fixed resolution video stream is to use anamorphic video (i.e., pixel aspect ratios not equal to 1:1). For example, if a cylindrical panoramic image has a resolution of 1056x144 with square pixels, but the head-size normalized
25 panoramic image is 1056x120 with square pixels, then the head-size normalized video can be transmitted as 1056x144 with a pixel aspect ratio of height:width=120:144 (the pixels are slightly shorter than wide). When the receiver renders the video to a display, the image will be rendered as 1056x120. The pixel aspect ratio is typically transmitted in the video stream, but can also be
30 transmitted outside of the video stream (e.g., a SIP INFO message).

The foregoing description of the invention has been presented for the purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. It is intended that the scope of the invention be limited not by this detailed description of the invention, but rather by the claims appended hereto.

WHAT IS CLAIMED IS:

1. In a computer system comprising a display, user interface selection
5 device and user interface data entry device, a process for normalizing portions of
a 360 degree panoramic image, comprising:

displaying on the display a control panel for entering parameters required
to normalize portions of a 360 degree panoramic image; and

entering the parameters required to normalize portions of the 360 degree
10 panoramic image via said displayed control panel.

2. The process of claim 1 wherein the parameters required to
normalize the 360 degree panoramic image comprise camera position, camera
orientation, horizontal feature geometry and percentage of normalization desired.
15

3. The process of claim 2 wherein the horizontal feature geometry
comprises at least one of:
meeting table geometry on which a camera corresponding to the camera
position and orientation is set; and
20 geometry of a plane corresponding to the top of people's heads sitting
around the meeting table.

4. The process of Claim 2 wherein the meeting table geometry comprises the size and shape of the meeting table and wherein the camera orientation is specified relative to the meeting table geometry.

5 5. The process of Claim 1 wherein the control panel is displayed on a Liquid Crystal Display.

6. The process of Claim 1 wherein the control panel is accessible via a network user interface.

10

7. The process of Claim 2 wherein the percentage of normalization desired varies between 0% and 100%.

8. The process of Claim 1 wherein the portions of the panoramic
15 image normalized relate to head size of persons sitting around a table and captured by the camera.

9. The process of Claim 3 wherein the camera position and orientation is relative to the geometry of the table.

20

10. The process of Claim 1 wherein the computer system is embedded in a camera.

11. A graphical user interface for displaying portions of a panoramic image comprising:

a display pane for displaying an extracted portion of an approximately 360 degree panoramic image.

5

12. The graphical user interface of Claim 11 wherein the extracted portion of the panoramic image is the non-normalized head of a person depicted in the panoramic image.

10

13. The user interface of claim 11 wherein the extracted portion of the panoramic image is a normalized head of a person depicted in the panoramic image.

15

14. The user interface of claim 11 wherein the person is speaking.

15. A process for displaying and transmitting a scaled image over a network, comprising the process actions of:

scaling an image;

adjusting the scaled image with extra pixels around its perimeter so that

20 the scaled image fits a specified size; and

displaying the scaled image with extra pixels around its perimeter.

16. The process of claim 15, further comprising the process action of transmitting the adjusted scaled image.

17. The process of claim 16, further comprising the process actions of:
5 receiving the adjusted scaled image;
displaying the adjusted scaled image.

18. The process of claim 16, further comprising the process actions of:
receiving the adjusted scaled image;
10 displaying the adjusted scaled image without the extra pixels
around its perimeter.

19. The process of Claim 15 wherein the extra pixels are
displayed in black.
15

20. A process for displaying and transmitting a scaled image
over a network, comprising the process actions of:
scaling an image;
setting the pixel aspect ratio to account for the difference between
20 the standard resolution and and scaled resolution;
transmitting the image in the standard resolution; and
displaying the scaled image at the original scaled resolution.