Versatile Imaging System

Frank Edughom Ekpar

Abstract—This paper introduces a versatile imaging system that encompasses a wide variety of methods, devices and systems. The versatile imaging system or device introduced in this paper is designed to facilitate the synthesis of representations of stimuli covering substantially all directions or only a subset of directions around a given reference or viewpoint, comprises at least one grid of one or more focusing elements disposed on an N-dimensional and arbitrarily shaped surface, at least one grid of one or more sensor elements disposed on an N-dimensional and arbitrarily shaped surface, and optionally, at least one grid of one or more stimulus guide elements disposed on an N-dimensional and arbitrarily shaped surface, where N can be chosen to be 1, 2, 3, or any other suitable quantity. A sampling of contemporary imaging systems that fall within the scope of the system is described. Pointers to architectures that take greater advantage of the features of the system are presented.

Index Terms—Versatile Imaging System, Representations of Stimuli, Sensor Grid, Omnidirectional Imaging System, Lensless Imaging.

I. INTRODUCTION AND LITERATURE SURVEY

Images and video streams constitute a vital aspect of communications in human society. Through the ages, humans have used images and video in various forms (drawings and paintings, printed letters, photographs, x-ray images, ultrasound images, to mention but a few) to convey concepts, give expression to artistic and aesthetic creations, diagnose and treat diseases, and more generally to solve a wide variety of communication problems. It is no wonder then that the study of the capture, characterization, manipulation, transmission and consumption of images and video has occupied some of the brightest minds. Digital cameras with a limited field of view are used to capture a very high number of images and video today. In spite of their limitations, these images suffice for a vast array of applications. Nevertheless, there exist many other application domains including space exploration, tele-presence, virtual reality and surveillance where a substantially wider field of view confers significant advantages and could sometimes prove indispensable. These application domains are better served via the use of panoramic images. Panoramic imaging techniques date back as far as the 18th century [1] when early works of art featuring panoramic projections were created.

A wide variety of techniques for creating, storing, transmitting and visualizing panoramic images have been reported in the scientific literature. More recently, techniques for creating and visualizing very large data sets including panoramic image data sets have been introduced by Frank et al in [2]. One key challenge for a panoramic imaging system is the encoding of enough visual information to make the system practical and cost-effective.

The use of rotating cameras to build panoramic mosaics, as described in [3] and [4], while facilitating the capture of panoramas of static scenes, is of limited practical value in more dynamic scenes. Catadioptric image capture systems generally comprising both reflective and refractive elements provide a more robust technique for capturing omnidirectional images without the use of moving parts [5]. Since they generally have no moving parts and can generate a panoramic view in a single image frame, catadioptric systems can be used to capture both static and dynamic scenes and video in real-time. Aspects of the geometrical characteristics of catadioptric systems were described by Geyer et al [6]. Yagi et al [7] and Chahl et al [8] proposed other panoramic image capture systems that are amenable to real-time operation. The popular Apple QuickTime Virtual Reality Authoring System [9] allows the generation of cylindrical panoramic images via stitching from several overlapping still image segments captured using a conventional camera and provides rendering and multimedia integration features.

Robust methods for navigating panoramic images and for correcting distortions in panoramic images have been introduced by Frank et al in [10] and [11].

The ideal panoramic imaging system should provide means of capturing substantially all visual stimuli around a given reference point in any number of desired dimensions and be capable of providing perspective-corrected views of any selected portion of the visual space around the selected reference point in real-time.

This paper presents a versatile imaging system that accommodates a very wide variety of designs and architectures and alleviates many of the problems associated with contemporary imaging systems. In particular, the system presented in this paper is based on principles that permit the construction of a versatile device for creating representations of stimuli covering substantially all directions or only a subset of directions around a given reference or viewpoint, comprising at least one grid of one or more focusing elements disposed on an N-dimensional and arbitrarily shaped surface, at least one grid of one or more sensor elements disposed on an N-dimensional and arbitrarily shaped surface, and optionally, at least one grid of one or more stimulus guide elements disposed on an N-dimensional and arbitrarily shaped surface, where N can be chosen to be 1, 2, 3, or any other suitable quantity. Aspects of this imaging system are disclosed in some of the issued claims of United States Patent Number 7567274 [12] granted to the author.

A survey of the literature has been presented in the introduction in Section I. The remainder of this paper is...
organized as follows: Section II describes the problem definition, methodology/approach in the form of the core design principles of the system introduced in this paper. Examples of results, concrete implementations and a discussion of related issues appear in Sections III to VII. Image Stitching as a method of capturing images that could substantially cover all directions around a selected reference point in 3D space is summarized in Section III. One of the methods employed in the remediation of the problems associated with image stitching namely, the use of the multi-resolution spline technique to effect the blending of overlapping image segments, appears in Section IV. Catadioptric imaging, another imaging technique capable of synthesizing panoramic images, is described in Section V. Perspective correction of panoramic images is discussed in Section VI. Lens-less imaging systems that fall within the scope of the system presented here is mentioned in Section VII. Future scope and the desirable features of higher-performance embodiments of the versatile imaging system introduced in this paper are highlighted in Section VII. Concluding remarks appear in Section IX.

II. PROBLEM DEFINITION/METHODOLOGY/APPROACH: CORE DESIGN PRINCIPLES

Efficiently capturing image and video data streams that substantially cover all view points around a selected reference or view point is the ultimate goal of an immersive imaging system. Image and video data could be combined with other data types such as audio and positioning data to create interactive virtual tours [12]. The system presented in this paper is adapted to capture a wide variety of image and video data types. The data could be multi-dimensional (for example 3D or stereoscopic panoramic video with a time dimension, 3D or 2D panoramic scenes comprising still images, standard video, standard or directional audio, global positioning system (GPS) data, holographic data, infrared data, ultrasonic data, ultraviolet data, etc) and could be acquired using any panoramic imaging system or could be created from the combination of a plurality of conventional images. In particular, panoramic data could be acquired using at least one arbitrarily shaped image acquisition device which could be embodied in a spherical image/video acquisition unit comprising at least one grid of one or more photosensitive elements on a surface with a spherical geometry or an approximation thereto and at least one enclosing concentric grid of one or more focusing elements on a surface with a spherical geometry or an approximation thereto. The photosensitive unit could comprise a grid of CCD sensors, CMOS sensors, a combination of these and other sensors (for example acoustical sensors) or any other suitable set or combination of sensors. The focusing element unit could comprise a grid of one or more lenses capable of focusing images and video from substantially all directions or only a subset of directions surrounding the device. Other arrangements, including, but not limited to, an arrangement requiring optical or alternative transmission of the light rays corresponding to the images or video from the focusing unit to the sensor or detector unit, are also possible. Panoramic data and virtual tour data could also be acquired using at least one versatile device for creating representations of stimuli covering substantially all directions or only a subset of directions around a given reference or view point, said versatile device for creating representations of stimuli comprising at least one grid of one or more focusing elements disposed on an N-dimensional and arbitrarily shaped surface, at least one grid of one or more sensor elements disposed on an N-dimensional and arbitrarily shaped surface, and optionally, at least one grid of one or more stimulus guide elements disposed on an N-dimensional and arbitrarily shaped surface, wherein said focusing element grid is adapted to focus stimuli covering substantially all directions around a given reference or view point onto the sensor element grid and may optionally do so via the stimulus guide unit if necessary, wherein each focusing element or group of focusing elements is associated with and focuses a subset (typically the subset impinging on it) of the entire stimulus space onto a sensor element or group of sensor elements responsive to the stimuli, wherein when an optional stimulus guide element grid is provided, the focusing element grid can be adapted to focus the stimuli onto the stimulus guide element grid for suitable formatting and onward transmission to the sensor element grid, wherein when used, each stimulus guide element or group of elements is associated with and receives stimuli from a focusing element or group of focusing elements and is in turn associated with and transmits stimuli to a sensor element or group of sensor elements. N can be chosen to be 1, 2, 3, or any other suitable quantity. The use of the versatile data acquisition device or method described here permits easier and more flexible manufacturing or assembly of data acquisition systems. Using the versatile data acquisition device, panoramic imaging systems that eliminate the spatial and temporal discontinuities and many other shortcomings of existing panoramic imaging systems could easily be built. Furthermore, other types of sensors could be incorporated into the versatile device to simultaneously capture other useful data such as audio for the creation of super-rich and truly immersive virtual tours.

III. IMAGE STITCHING: RESULTS AND DISCUSSION

Although the preferred implementation of the adaptation of the versatile imaging system to the capture of panoramic images is a system capable of acquiring a complete panoramic image or video frame at once, practical limitations sometimes necessitate the use of a plurality of overlapping image segments to compose a complete panorama. Fig. 1 illustrates the concept of image stitching for the purpose of generating panoramas. The source image segments (typically overlapping) are represented as S1, S2, \( S_{\text{N-1}} \) and here \( N \) is the number of separate image segments required to compose the panorama. O(P) conceptually represents the set of operations and transformations applied to the source image segments to synthesize the panorama – which is indicated as P in Fig. 1.
Fig. 2 depicts a sample 360-degree panorama generated by stitching a sequence of overlapping image segments. The relevant literature – for example [3, 4, 9] contains a detailed analysis of image stitching practices.

The separate image segments used in image stitching could be viewed as the sub-images generated by the corresponding partitions on the arbitrarily-shaped surface of the sensor elements of the versatile imaging system.

IV. IMAGE BLENDING USING GAUSSIAN AND LAPLACIAN PYRAMIDS: RESULTS AND DISCUSSION

Although in principle it is possible to capture panoramic images and video of the type described here in a single image frame using a single imaging device, it is often more practical to capture a series of overlapping high-resolution segments of the image and then stitch these together to form a single image mosaic.

Techniques for stitching overlapping image segments are well known. One significant problem with image stitching is how to make the seams between overlapping segments invisible. A wide variety of image blending techniques exist. Generally, the choice of a specific image blending technique depends on the requirements of the specific application. The multi-resolution spline technique proposed by Peter J. Burt et al [13] gives satisfactory results albeit requiring large amounts of memory. Peter J. Burt et al first decompose the images to be splined into a set of band-pass filtered component images with the component images in each frequency band assembled into a corresponding band-pass mosaic. Component images are joined using a weighted average within a transition zone proportional in size to the wavelengths that comprise the band. Ultimately, summation of the band-pass mosaic images is used to derive the output mosaic image.

Peter J. Burt et al [13] utilize a combination of Gaussian-filtered and Laplacian-filtered pyramids of images at distinct resolutions or levels (as illustrated in Fig. 3) to accomplish the blending of the images. In Fig. 3, G0, G1,...,Gl represent the Gaussian pyramid levels whilst L0, L1,...,Ll represent the Laplacian pyramid levels.

In Fig. 4 a segment of an image mosaic obtained by stitching overlapping image segments without the application of image blending is depicted. Seams are visible in the image of Fig. 4. Note that in Fig. 5, the same image has been manipulated with the multi-resolution spline technique used to blend the overlapping regions together. The seams are no longer visible in Fig. 5.

V. CATADIOPTRIC IMAGING SYSTEMS: RESULTS AND DISCUSSION

Panoramic images can be acquired using a wide variety of
techniques. One of the simplest methods of acquiring panoramic images involves the capture of a sequence of overlapping image segments around a unique effective viewpoint using a conventional digital camera. The overlapping segments are then stitched together and blended (using the multi-resolution spline technique [13] or any other appropriate technique) to produce the complete panoramic mosaic. Rotating cameras could also be used to create panoramic mosaics as described in [3] and [4]. Stitching and the use of rotating cameras might suffice for the capture of panoramas of static scenes but are of limited practical value in more dynamic scenes. To ameliorate this problem, catadioptric systems comprising both reflective and refractive elements are traditionally used to capture omnidirectional images in a manner that satisfies the unique effective viewpoint constraint. Catadioptric imaging systems are designed to be capable of capturing a complete 360-degree field of view in a single image frame with no moving parts. Since they generally have no moving parts and can generate a panoramic view in a single image frame, catadioptric systems can be used to capture both static and dynamic scenes and video in real-time. Aspects of the geometrical characteristics of catadioptric systems were described by Geyer et al [6]. Yagi et al [7] and Chahl et al [8] proposed other panoramic image capture systems that are amenable to real-time operation. Furthermore a detailed treatment of image formation by catadioptric imaging systems was carried out by Simon Baker and Shree K. Nayar in [5]. Catadioptric systems typically generate annular images. Fig. 6 is a sample annular image captured using a catadioptric imaging system.

VI. PERSPECTIVE CORRECTION: RESULTS AND DISCUSSION

For convenient manipulation, panoramic images could first be re-projected onto a spherical surface to create a spherical environment map. This permits the application of a uniform perspective correction algorithm during the navigation of the image. Fig. 7 is a conceptual illustration of the projection or re-projection of a panoramic image onto the surface of a sphere. Portions of the sphere for which panoramic image data is not available (for example, when the panorama is acquired using a system with a vertical field of view that is less than \( \pi \) radians) can be replaced with user-supplied data or simply filled with a uniform color. The region of interest or view window is depicted on the \( U \) \( -V \) coordinate plane as a perspective projection of the region of the panorama indicated by the viewing parameters.

In addition to the correction of any substantial geometric distortions, views extracted from panoramic images require perspective correction for comfortable viewing by a human observer. Fig. 7 and Fig. 8 illustrate the generation of a perspective-corrected view and its representation on the \( U \) \( -V \) coordinate plane. Suppose that the ray from the center, \( O_0 \), of the sphere to the center of the \( U \) \( -V \) coordinate plane (the viewing direction for the indicated view) intersects the surface of the sphere at the point with angular coordinates \( (\Theta, \varphi) \). Then the point with coordinates \( (u, v) \) on the perspective-corrected \( U \) \( -V \) image plane corresponds to the point with coordinates \( (\Theta, \varphi) \) on the surface of the sphere. The perspective correction is via the projection of the region of the sphere indicated by the selected viewing parameters unto the \( U \) \( -V \) image plane. The viewing parameters include a lateral angle \( \Theta \), the vertical angle (elevation or azimuth) \( \varphi \) and a magnification coefficient.

These viewing parameters together with a view window onto which the perspective-corrected views of the spherical

---

DOI: [http://dx.doi.org/10.24018/ejers.2019.4.12.1668](http://dx.doi.org/10.24018/ejers.2019.4.12.1668)
environment map are rendered specify a region of interest on the spherical environment map that is projected onto a perspective-corrected object plane as shown in Fig. 7 and Fig. 8 in which the spherical environment map is presented in 3-dimensional X-Y-Z space. The magnification coefficient is directly related to $\lambda$, the angular span of the width of the perspective-corrected view window as measured from the center of the sphere. When the viewing is implemented in software, the viewing parameters can be entered using the mouse, keyboard, joystick or any other suitable input device. The perspective transformation is achieved by projecting the selected region of interest on the surface of the sphere onto the perspective-corrected object plane. Equations (1) through (8) describe the transformations required to create the perspective-corrected view. In equations (1) through (8) as in Fig. 7 and Fig. 8, $u$ and $v$ are the coordinates of a point on the perspective-corrected U-V object plane (the point P($u$, $v$) in Fig. 7), $\theta$ and $\phi$ are the corresponding lateral and azimuth angles on the spherical environment map for the viewing parameters $\theta$ and $\phi$ and the magnification coefficient-related angular span $\lambda$; $d$ is the distance of the center of the object plane from the center, $O$, of the sphere while W and H are, respectively, the width and height of the perspective-corrected view window on the U-V object plane. 

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_0 & \sin \phi_0 \\ 0 & -\sin \phi_0 & \cos \phi_0 \end{bmatrix}$$  \hspace{1cm} (1)$$

$$B = \begin{bmatrix} \cos \theta_0 & 0 & -\sin \theta_0 \\ 0 & 1 & 0 \\ \sin \theta_0 & 0 & \cos \theta_0 \end{bmatrix}$$  \hspace{1cm} (2)$$

$$C = AB$$  \hspace{1cm} (3)$$

$$d = \frac{W}{2\tan \frac{\lambda}{2}}$$  \hspace{1cm} (4)$$

$$D = [u \quad v \quad d]$$  \hspace{1cm} (5)$$

$$E = DC = [\mathcal{E}_1 \quad \mathcal{E}_2 \quad \mathcal{E}_3]$$  \hspace{1cm} (6)$$

$$\theta = a \tan 2(\mathcal{E}_1, \mathcal{E}_2)$$  \hspace{1cm} (7)$$

$$\phi = a \tan 2\left(\mathcal{E}_2 \sqrt{\mathcal{E}_1^2 + \mathcal{E}_3^2}\right)$$  \hspace{1cm} (8)$$

In equations (7) and (8), $\tan(\text{argument1}, \text{argument2})$ is a trigonometric function that gives the arctangent of argument1/argument2. Once the $\theta$ and $\phi$ angles have been computed using equations (1) through (8), the image/video data contained in the spherical environment map at the locations corresponding to the $\theta$ and $\phi$ angles is then copied onto the locations indicated by the $u$ and $v$ coordinates on the perspective-corrected view window on the U-V object plane - possibly using interpolation to account for non-integral values of the parameters and to improve the quality of the resultant image/video. Compared with the approaches used by the contemporary systems, equations (1) through (8) provide a faster means of constructing perspective-corrected views from spherical environment maps while providing full pan, tilt and zoom controls. Creating a look-up table containing a subset of points from the perspective-corrected window, using equations (1) through (8) on the points in the look-up table only and then using bilinear or other suitable forms of interpolation to correct all other points not in the look-up table leads to faster albeit less accurate perspective correction. It is also possible to use a table of pre-computed trigonometric values to speed up the calculations without any appreciable loss in image/video quality. Fig. 9 shows a portion of a 360-degree panoramic image of an office with visible perspective distortion at the top of the window and ceiling. In Fig. 10, the perspective correction algorithm described here has been applied to the image of Fig. 9.
VII. LENS-LESS IMAGING SYSTEMS: RESULTS AND DISCUSSION

Although most imaging systems in use today employ lenses to focus stimuli onto appropriate sensors, there exists a class of imaging systems that can acquire images without the use of lenses. In these lens-less imaging systems, the focusing and stimulus guide units described in the system presented in this paper could be embodied in the diffraction and other stimulus-modulating properties of the air, diffraction gratings and other suitable materials utilized in the design and construction of these systems. Khademhosseini et al. [14] describe a Lens-free on-chip imaging system using nano-structured surfaces in which the detected far-field diffraction pattern is used for rapid reconstruction of the object distribution on the chip at the sub-pixel level using a compressive sampling algorithm. Furthermore, Sungkyu Seo, Serhan O. Isikman, Ikbal Sencan, Onur Mudanyali, Ting-Wei Su, Waheb Bishara, Anthony Erlinger and Aydogan Ozcan also present a detailed investigation of the performance of lens-free holographic microscopy toward high-throughput on-chip blood analysis in [15].

VIII. HIGHER PERFORMING EMBODIMENTS: FUTURE SCOPE

Constructing an imaging system with a substantial subset of the components of the system presented here would permit unprecedented imaging capabilities. The use of the versatile data acquisition system described in Section 2 permits easier and more flexible manufacturing or assembly of data acquisition systems. Using the versatile data acquisition device, panoramic imaging systems that eliminate the spatial and temporal discontinuities and many other shortcomings of existing panoramic imaging systems could easily be built. Furthermore, other types of sensors could be incorporated into the versatile device to simultaneously capture other useful data such as audio for the creation of super-rich and truly immersive virtual tours. In principle, the sensor elements that comprise the versatile imaging system presented here could conform to an arbitrary shape. Other aspects of the system with respect to the capture of holographic and stereoscopic image and video data are highlighted in [12].

IX. CONCLUSION

In this paper, a versatile imaging system that encompasses a wide variety of methods, devices and systems was presented. The design principles of the system were highlighted. Contemporary approximations to subsets of the capabilities in systems such as multi-image sequence stitching, catadioptric imaging systems and lens-free imaging systems were also described. These illustrate aspects of imaging architectures that take greater advantage of aspects of the features of the system. Future systems could be designed to take advantage of more of the features described in this paper as explained while describing higher performing embodiments in the previous section.

REFERENCES

[1] R. Benosman, “Panoramic imaging from 1767 to the present”, Proc. International Conference on Advanced Robotics, Workshop, Volume 1, 2001, pp. 9-10.
[2] F. Ekpar, H. Hase, and M. Yoneda, “On the Interactive Visualization of Very Large Data Sets”, Proceedings. 7th International IEEE Conference on Computer and Information Processing, 2007, pp. 627-632.
[3] R. Szeliski, “Video mosaics for virtual environments”, Computer Graphics and Applications, 16(3), 1996, pp. 23-30.
[4] H. Sawhney, S. Hsu, and R. Kumar, “Robust video mosaicing through topology inference and local to global alignment”, Proc. Fifth European Conference on Computer Vision, Volume 2, 1998, pp. 103-119.
[5] Simon Baker and Shree K. Nayar, “A Theory of Single-Viewpoint Catadioptric Image Formation”, International Journal of Computer Vision, Volume 35, Issue 2, 1999, pp. 175 - 196.
[6] C. Geyer, and K. Daniilidis, “Catadioptric projective geometry”, Proc. International Conference on Advanced Robotics, Volume 1, 2001, pp. 17-30.
[7] Y. Yagi, S. Kawato, and S. Tsuji, “Real-time omnidirectional image sensor (copis) for vision-guided navigation”, IEEE Transactions on Robotics and Automation, 10 (1), 1994, pp. 11-22.
[8] J. Chahl, and M. Srivinasan, “Reflective surfaces for panoramic imaging”, Applied Optics, Volume 36, 1997, pp. 8275-8285.
[9] Shenchang Eric Chen, “QuickTime VR: an image-based approach to virtual environment navigation”, Computer Graphics (Proc. SIGGRAPH), 1995, pp. 29-38.
[10] F. Ekpar, H. Hase, and M. Yoneda, “Constructing arbitrary perspective-corrected views from panoramic images using neural networks”, Proc. 7th International Conference on Neural Information Processing, Volume 1, 2000, pp. 156-160.
[11] F. Ekpar, H. Hase, and M. Yoneda, “Correcting distortions in panoramic images using constructive neural networks”, International Journal of Neural Systems, Volume 1, 2003, pp. 239-250.
[12] F. Ekpar, “Method and apparatus for creating interactive virtual tours”, United States Patent Number 7567274, Issued 2009.
[13] Peter J. Burt and Edward H. Adelson, “A Multiresolution Spline with Application to Image Mosaics”, ACM Transactions on Graphics, Volume 2, Number 4, 1983, pp. 217-236.
[14] Khademhosseini B, Sencan I, Biener G, Su TW, Coskun AF, Tseng D, and Ozcan A, “Lensfree on-chip imaging using nanostructured surfaces”, Applied Physics Letters, 2010, Volume 96, Issue 17, 171106.
[15] Sungkyu Seo, Serhan O. Isikman, Ikbal Sencan, Onur Mudanyali, Ting-Wei Su, Waheb Bishara, Anthony Erlinger and Aydogan Ozcan, “High Throughput Lens-free Blood Analysis on a chip”, Analytical Chemistry, 2010, Volume 82, Issue 11, pp. 4621-4627.