Optimum Alignment of Panoramic Images for Stereoscopic Navigation in Image-Based Telepresence System

Luis E. Gurrieri and Eric Dubois
School of Electrical Engineering and Computer Science, University of Ottawa
Ottawa, Ontario K1N 6N5, Canada.
{Luis.Gurrieri,edubois}@uottawa.ca

Abstract

The addition of stereoscopic navigation to an image-based virtual environment is a desirable enhancement. This can be implemented by sampling the scene with a number of stereoscopic panoramas. In this regard, clusters of panoramas in a known spatial arrangement can be used to render omnistereoscopic views. However, slight misalignments between panoramas introduced by single-shot panoramic cameras must be corrected in order to capture the depth of the scene consistently in every direction. In this regard, a novel alignment correction method is proposed herein based on the dense disparity map between panoramas. This technique was successfully tested in the rendering of omnistereoscopic images in different scenarios and it is applicable to any off-the-shelf panoramic cameras. Unlike other panorama alignment methods, this is a featureless and uncalibrated solution to the alignment of closely taken panoramic snapshots. Furthermore, the omnistereoscopic rendering and aligning techniques proposed herein are computationally inexpensive alternatives for creating stereoscopic image-based virtual environments. [1]

1. Introduction

A pair of panoramic snapshots exhibits two zones of optimum stereoscopic depth registration located perpendicularly to their direction of alignment. In this context, the alignment occurs in a unique direction where, comparing the views, a *divergent* or *convergent* image flow behavior, i.e., a *zoom in/out* effect, is exhibited in the scene. Moreover, in a cluster of closely located cylindrical panoramas symmetrically distributed around a capture center, the area of stereoscopic usability is found in different azimuthal directions for different panoramic pairs. The correct geometric distribution of samples within a cluster enables to use these panoramic sections to compose two novel panoramic views corresponding to the left and right eyes. The key problem is to find the optimal aligning between any pair in order to maintain a consistent depth rendering in every direction.

The advent of single-shot panoramic cameras has made possible the efficient acquisition of panoramic clusters. Panoramic cameras based on multiple sensors have been increasingly used in image-based navigation applications [9]. This panoramic technology is based on mapping several overlapping wide-angle images on a 3D surface, i.e., cylinders, semi-spheres or cubes. However, the radial distortions in each image in combination with inherent parallax between sensors introduce distortions in the panoramic image in distinctive directions, which are more noticeable the closer the objects are to the camera [4]. This distortion is commonly neglected in single panoramic snapshots but becomes evident when comparing the scene registered in pairs of panoramas within a cluster. The key idea is to use the correct panoramic pair sections to compose an omnistereoscopic image, which is easier when the objects are far from the camera head. One problem is that panoramic pairs within a cluster are aligned in different directions with respect to a common reference of coordinates at the center of the cluster. Furthermore, each panorama exhibits geometric distortions in different directions, which is also dependent of how the panoramic camera rig changes its position while capturing the set. Therefore, a fine alignment correction must be done independently for each panoramic pair to minimize rendering errors in the omnistereoscopic composition.

The main focus of this article is to present a general panoramic alignment method applicable to any off-the-shelf panoramic camera in the context of the omnistereoscopic image composition. Our method relies on the local-

DOI: 10.1109/ICCVW.2011.6130263

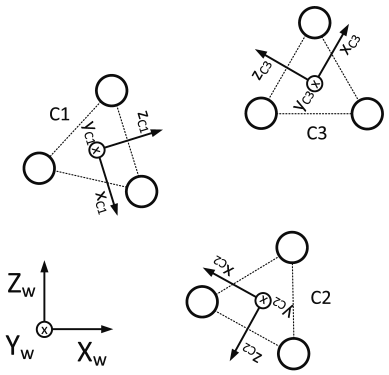


Figure 1. Stereoscopic panoramic sampling of a scene showing the clusters of regularly spaced panoramas (circles) in the *x*-*z* plane.

ization of specific disparity patterns by analyzing the dense image flow between panorama pairs. The proposed aligning algorithm is featureless, uncalibrated, and computationally inexpensive, making it attractive for the fast rendering of large sets of omnistereoscopic images.

1.1. Omnistereoscopic Sampling

In a telepresence context, the sampling of a scene requires a large number of distinct viewpoints, which makes the task challenging considering the acquisition time needed to create each stereoscopic panorama using conventional methods. Sequential methods for rendering high-quality omnistereoscopic images are time demanding and not practical when large numbers of viewpoints are involved [7][3]. Other methods that rely on panoramic video sequences [11][2] introduce new problems related to the consistency of the omnidirectional depth rendition and handling of the huge data overhead. Using a different approach, Vannijja et al. [10] proposed using clusters of cylindrical panoramas to create an omnistereoscopic image. In this case, the locations of the different panoramic samples are precisely known, enabling us to exploit the distinct viewpoints arising between panorama pairs. Although this strategy was proposed for a single viewpoint, it provides an efficient method for the stereoscopic sampling of a real-world scene, i.e., using clusters of panoramas instead of a single panorama at each viewpoint. The proposed approach is illustrated in Fig. 1 for the case of triangular clusters; however, other patterns with higher number of panoramas can be used. Notice that, unless the panoramic format is specified, a *cylindrical panorama* will be simply referred to as a *panorama*.

An example of a panoramic cluster is illustrated in Fig. 2, showing a triad of panoramas distributed in a triangular equidistant sampling pattern. The panoramic centers are coplanar and the baseline b is determined according to the illusion of depth to be perceived during the navigation of the virtual environment. Typically, a baseline of $b = 6.5\ cm$, corresponding to the average interocular distance in adults, is appropriate in most indoor situations. There is a local reference of coordinates at each panorama center and another for the cluster, located at the center of the arrangement, which is used as reference during the composition of left and right eye panoramas. Notice that $\theta \simeq \theta_1 \simeq \theta_2 \simeq \theta_3$ when objects in the scene are far from the panoramic cluster center.

In any given azimuthal direction, there is a unique panoramic pair which correctly renders a binocular stereoscopic view. The panoramic sections usable for stereoscopic rendering, indicated by the wide arrows in Fig. 2, are perpendicularly located with respect to the direction of optimal alignment for any given pair. A full omnistereoscopic view can be composed by determining the best panoramic pairs for stereoscopic composition in every azimuthal direction.

1.2. The Problem

The absolute location and relative orientation of the panoramas need to be known to exploit the spatial diversity between panoramic snapshots. The latter prior information, the optimum alignment between panorama pairs, must be known to identify the usable stereoscopic regions of interest as well as to keep a consistent depth rendering in every direction.

The panoramic clusters can be acquired by rotating a panoramic head around the symmetry center with a radius r, taking a number of snapshots during a circular trajectory of the camera. In this case, it is possible to determine the relative orientation between any two panoramas and, after compensating for the corresponding angle of rotation, find the right aligning direction between any pair of panoramas.

After compensation for the camera rotation between

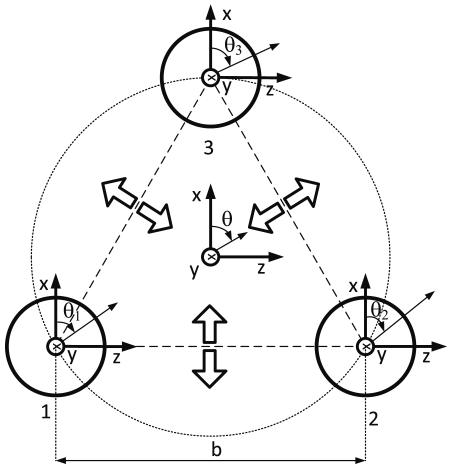


Figure 2. Cluster of panoramas: the panorama locations are indicated with the large circles while the wide arrows indicate the usable stereoscopic region of interest directions for each pair.

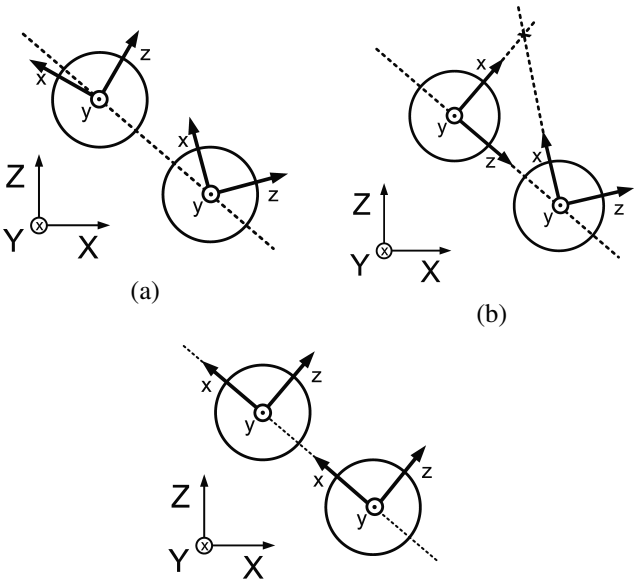
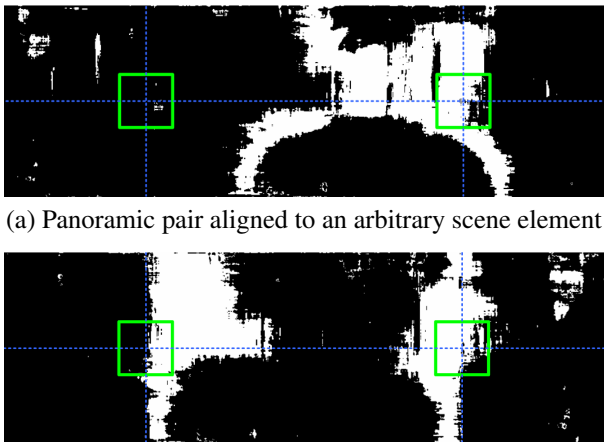


Figure 3. (a) Misaligned panoramic pair, (b) panoramic pair aligned to an arbitrary scene element, and (c) optimum panoramic alignment for stereoscopic mosaicing; the large circles indicate the panorama locations.

snapshots, panoramas require further pairwise alignment. After the panoramic sampling, the relative orientation is illustrated in Fig. 3-(a) where the x-axis is aligned with the azimuthal center of reference of each panorama, and each panoramic snapshot is taken in the x-z plane. As illustrated in Fig. 3-(b), panoramic pairs could be aligned to any arbitrary scene element. However, the optimum alignment for correct stereoscopic rendering occurs in a single azimuthal direction as illustrated in Fig. 3-(c). The horizontal component of the dense optical flow between panoramic images can be used to illustrate the optimum aligning case. For instance, the zero horizontal disparity regions in white in Fig. 4-(a) indicate the case of a panoramic pair aligned to an arbitrary scene element; this is the case shown in Fig. 3-(b). On the other hand, when the panorama pair is optimally aligned as illustrated in Fig. 3-(c), the target aligning pattern arises with a separation of $M/2$ pixels from each other around $\theta \pm 90^\circ$ as shown in Fig. 4-(b).

In recent years, methods for aligning cubic panoramas have been proposed, sharing the same goal of facilitating the navigation between viewpoints in image-based virtual environments [1][5][8]. These proposals are based on epipolar geometry, estimating the essential matrix using corresponding features between faces of different cubes. These techniques were focused on equalizing the viewing direction between largely spaced panoramic samples. Although these feature-based methods provide robust solutions to the registration problem, the alignment necessary



(a) Panoramic pair aligned to an arbitrary scene element

(b) Optimum alignment for omnistereoscopic rendering

Figure 4. Panoramic alignment cases: horizontal disparity map

to localize the stereoscopic regions of interest, which is the main focus of this work, is a different problem and, to the best knowledge of the authors, still unsolved. To propose a feature-less and computationally efficient method for finding the desired aligning pattern in panoramic pairs is the main goal of this article.

2. Panoramic Alignment

The panoramic images referred herein are cylindrical images unwrapped into a plane. After being converted these images to gray scale, they can be modeled as $I(n,m)$: $Z^2 \rightarrow Z$ defined for $n \in [-N/2, N/2)$ and $m \in [-M/2, M/2)$, where N and M are the vertical and horizontal dimensions, respectively. This is illustrated in Fig. 5. Each panoramic image exhibits periodicity in the horizontal dimension, i.e., $I(n,m) = I(n, m+M)$, a useful property that will be used in the development of the aligning algorithm.

Two panoramic images $(I_i, I_j) \in P_\xi$ are highly correlated when the distance between their cylindrical projection centers, b (Fig. 2), is small in comparison with the distance to the objects in the scene. In our case, three images per each sampled position $\xi = (X, Y, Z)$ defining a panoramic cluster $P_\xi = \{I_1, I_2, I_3\}$. As mentioned be-

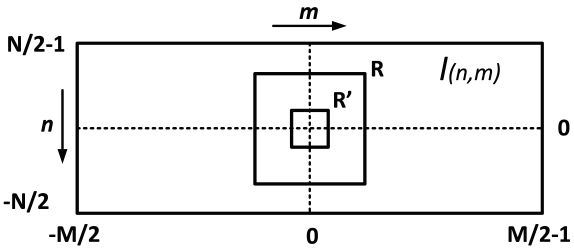


Figure 5. Planar panoramic image $I(n,m)$.

fore, alignment is necessary to be performed pairwise between elements of P_ξ as part of the omnistereoscopic rendering process. Hence, each cluster determines three pairs of images to be aligned independently: (I_i, I_j), where $i, j \in \{1, 2, 3\}$.

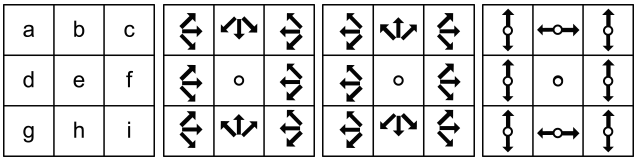
The analysis can be restricted to a square region **R** of W pixels in the center of each panorama (Fig. 5). It is important to note that, for a given horizontal shift λ, there is a unique subset of contiguous horizontal directions, $\Lambda \subset [0, M/2)$, where the pairwise image flow around the central region **R** exhibits a *convergence/alignment* pattern and at the same time a *divergence/alignment* pattern in the antipode direction $(\lambda + M/2)$. To find this pattern, the panoramic pair needs to be fine aligned first by introducing a correction shift m_α such as the shifted pair $(I_i(n, m - \lambda), I_j(n, m - \lambda - m_\alpha)$ exhibits the maximal image correlation in the central region **R**.

The optimum aligning process consists in, first, finding the first simultaneous occurrence of these patterns by shifting both panoramas horizontally a discreet amount $\lambda \in [0, M/2)$. This searching can be done sequentially or following another searching strategy. For each new λ, an optimization over the correction shift m_α is applied to the pair $(I_i(n, m-\lambda), I_j(n, m-\lambda-m_\alpha))$ before analyzing the patterns at 0° and 180°. Once obtaining the first occurrence, the next step consists in determining which $\lambda \in \Lambda$ offers the best aligning for omnistereoscopic rendering. This optimization is based on finding the horizontal shift $\lambda^* \in \Lambda$ which minimizes the Euclidean distance between matching points at **R**. The optimum shifting provides two aligned regions optimally separated by $M/2$ or 180° as illustrated in Fig. 3-(c) and 4-(b).

2.1. Disparity Operator

The dense optical flow analysis of each panoramic pair central region **R**, provides the disparity map between corresponding image elements at one pixel resolution. The algorithm proposed by Ogale et al. [6] was used to obtain the horizontal and vertical components of the dense disparity map in addition to the occlusion mapping, D_x, D_y, and O, respectively. This mapping is subject to errors due to poorly illuminated areas, flat textures, noise and occlusions. A better approach is to group pixels into square blocks and to estimate the characteristic disparity vector for each block, i.e., using a majority vote over the dense disparity map in the block **R'**.

In this case, the dense optical flow over square regions **R** of side length W centered on each panorama, I_i and I_j, is used to characterize the image displacement pattern. The dense disparity map between the pixels in the same region **R** of both panoramas (Fig. 5), at one pixel resolution, is returned in two matrices, D_x and D_y, and the occlusion map in O, each of the same size as **R**. The disparity operator $\Delta(D_x, D_y, O)_{R'}$ returns a disparity a vector (d_x, d_y) equivalent to the characteristic disparity of the block **R'**, which is indicated in Fig. 5. The disparity operator calculates the frequency of occurrence of each pixel disparity in **R'** in both dimensions and returns the most frequent value in each case. Occluded areas are eliminated and not considered in the majority vote estimation since the dense optical flow algorithm used returns unreliable values in those areas [6]. Hence, the occlusion map O is used to eliminate those occluded pixels from (D_x, D_y).



(a) (b) Ω_1 (c) Ω_2 (d) Ω_3

Figure 6. Pattern classification: (a) blocks defined on **R**, (b) focal convergence, (c) focal divergence, and (d) near alignment.

2.2. Alignment Patterns

In order to classify the characteristic disparity of each block, the central region **R** is divided in nine non-overlapping blocks (Fig. 6-(a)). There are three main image flow pattern of interest: *focal convergence* (Ω_1) indicated by the convergence of the image disparity vectors towards the center of reference (Fig. 6-(b)), *focal divergence* (Ω_2) indicated by the disparity vectors diverging from the center of the image (Fig. 6-(c)), and *near alignment* (Ω_3) between both images, usually when objects in the scene are far from the camera (Fig. 6-(d)).

The pair of overlapped panoramas (I_i,I_j) are illustrated in Fig. 7 where x-axis intersect both panoramas at $(n, m) = (0, 0)$ and the z-axis intersect them at $(0, M/2 - 1)$. An optimally aligned panoramic pair exhibits focal convergence and/or divergence at the panorama center, $\theta = 0^\circ$ $(m = 0)$, and its antipode, $\theta = 180^\circ$ $(m = M/2)$. This is illustrated in the **A** zones in Fig. 7. The optimum alignment direction determines the usable rendering areas or stereoscopic regions of interest centered around $\theta = \pm 90^\circ$ $(m \pm M/4)$, indicated as the **B** regions in Fig. 7.

2.3. Viewing Direction Equalization

The feature based approach to this problem for cubic panoramas as presented in [1], [5], and more recently in [8], can be adapted to align a panoramic pair to any arbitrary. However, a featureless approach provides a simpler solution in this case given the high correlation between images in **R** (Fig. 5).

First, a coarse viewing direction correction of I_j with respect to I_i is performed and, second, a fine alignment is achieved by using the characteristic image flow over the re-

gion **R'** (Fig. 5) to eliminate the horizontal disparity arising between panoramic centers. The first coarse alignment is done by finding the shift λ that minimizes

$$\Gamma(x) = \frac{1}{W^2} \sum_{i,j=-W/2}^{W/2-1} \| I_i(n,m) - I_j(n,m-x) \|^2, \quad (1)$$

where I_i and I_j are gray scale intensity panoramas, W is the size of the target central region **R**, and $x \in [-M/2, M/2)$. When the panoramas are already partially aligned, the searching region can be restricted to a smaller contiguous sequence. A value for W equivalent to $\theta = \pm 45°$ or $M/8$ is appropriate for the evaluation of $\Gamma(x)$. Initially, the horizontal shift that minimizes $\Gamma(x)$ is used to rotate I_j to the viewing direction of I_i. A final fine correction is performed by obtaining the horizontal disparity d_x from the disparity operator $\Delta(D_x, D_y, O)_{\mathbf{R'}}$, where **R'**, is a block region with $W' = M/18$. Hence, applying a correction shift, $m_\alpha = x + d_x$, to I_j aligns it with I_i, eliminating any horizontal disparity in **R'**. However, this is not necessarily correct at its antipodes as it will explained in the next section.

2.4. Disparity Pattern Classification

The optimal alignment, as is illustrated in Fig. 7, can be characterized by the image disparity patterns in the central regions, indicated as the square region **R** in Fig. 5 spanning $\theta = \pm 30°$ or, equivalently, $W = M/6$. This central region **R** is divided into nine blocks as shown in Fig. 6-(a). Each of these blocks is $W/3 = M/18$ pixels in length. The disparity operator, $\Delta(D_x, D_y, O)_\gamma$, where $\gamma \in \{a, b, c, d, e, f, g, h, i\}$, is applied to each block γ to obtain their characteristic disparity. Camera misalignments, such as roll, pitch, and/or tilt, lead to perturbations in the disparity patterns. However, this effect can be compensated up to a certain extent, i.e., a vertical disparity detected in the central block, $\gamma = e$, can be used to compensate camera misalignments during the capturing by subtracting a vertical bias from d_y in all blocks around the central one.

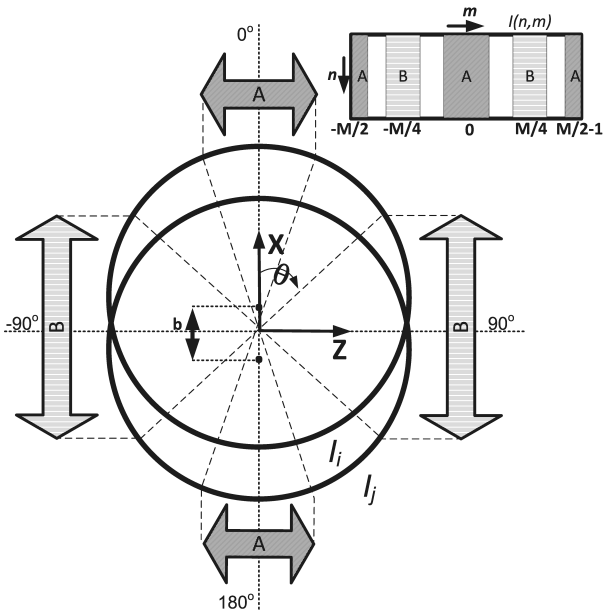


Figure 7. A pair of cylindrical panoramic images (large circles) optimally aligned, showing: (**A**), the directions of focal convergence and/or divergence, and (**B**), the stereoscopic regions of interest mapped onto the planar representation $I(n,m)$.

The pattern detection can be done by classifying the disparity pattern in the panorama center $(0,0)$, and its antipode $(0, M/2-1)$, into three mutually exclusive categories: focal convergence (Ω_1), focal divergence (Ω_2), and near alignment (Ω_3), as shown in Fig.6-(b), (c) and (d), respectively. The sign of the horizontal (d_x) and vertical (d_y) characteristic disparities in each block γ is used to determine to which class each block belongs. There is a limited number of categories in which a given block can be classified, i.e., the upper left block in Fig. 6-(a), $\gamma = a$, can be categorized as belonging to Ω_1 if the boolean event $A = 1$ ($d_x > 0$) is detected, or to Ω_2 if the event $C = 1$ ($d_x < 0$) is detected, or to Ω_3 if $O_x = 1$ ($d_x = 0$) is detected. The pattern encoding is a function of the block γ and the event α, where $\alpha \in \{A, B, C, D, O, O_x, O_y\}$. The class encoding as a function of α and γ is defined in Table 1, whereas the boolean events α as function of d_x and d_y are defined in Table 2. Notice that the central block, $\gamma = e$, must correspond to the boolean event $O = 1$ ($d_x = 0$ and $d_y = 0$), or perfect aligning, before starting to classify the other eight surrounding blocks.

Pattern	a	b	c	d	e	f	g	h	i
Ω_1	A	B	C	A	O	C	A	D	C
Ω_2	C	D	A	C	O	A	C	B	A
Ω_3	O_x	O_y	O_x	O_x	O	O_x	O_x	O_y	O_x

Table 1. Pattern encoding

Boolean Event	The event = 1 when
A	$d_x > 0$
B	$d_y < 0$
C	$d_x < 0$
D	$d_y > 0$
O	$d_x = 0$ and $d_y = 0$
O_x	$d_x = 0$
O_y	$d_y = 0$

Table 2. Events encoding

Finally, a test is performed over the events detected in all blocks but the central one to determine which disparity pattern occurs in that direction. Each block may be classified to belong to any of the three target classes given the random-

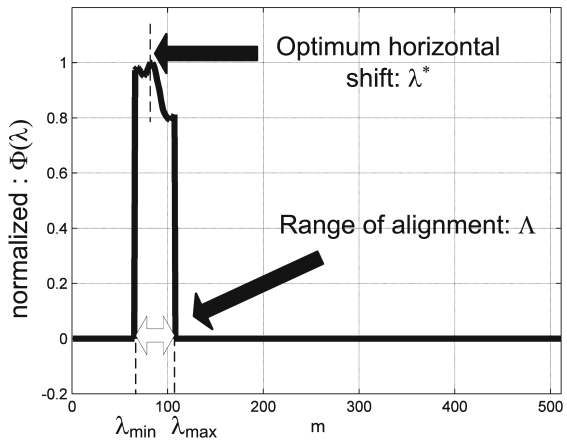


Figure 8. Optimum alignment occurrence as function of the horizontal shift.

ness on the scene. Hence, the most likely pattern is selected based on its frequency of occurrence, i.e., if the number of blocks classified as Ω_1 is larger than the number of them classified as Ω_2 or Ω_3, then that direction is classified as Ω_1 class. In order to formalize this process, a hard limit τ, where $\tau \geq 4$, can be used to define the decision threshold, i.e., after counting the number of blocks belonging to each class, the class with a number larger than τ is selected for a given azimuthal direction.

2.5. Optimum Azimuthal Direction

The previous section describes a method to classify the characteristic disparity of the central region **R** of a panoramic pair into three mutually exclusive patterns. However, it is still necessary to define which combination of these patterns in one azimuthal direction and its opposite direction identifies a $\lambda \in \Lambda$ as defined at the beginning of section 2. This happens when Ω_1 or Ω_3 are assigned to one direction and Ω_2 or Ω_3 to its antipodes. It is worth to mention that a perfect alignment, Ω_3, might be detected in any (or both) main directions in situations when the objects in the scene are substantially far from the panoramic cluster. The closer to the camera the scene elements are, the more likely it is to detect a focal convergence or divergence, Ω_1 or Ω_2, respectively.

There is a range of azimuthal directions where the desired patterns is manifested, referred herein as the *alignment range* Λ. The width of that region depends on the scene characteristics and varies from one location to another. At this point, the optimum azimuthal direction need to be identified using the criterion of best alignment. The proposed featureless approach is based on estimating the degree of alignment in any $\lambda \in \Lambda$ by calculating the normalized percentage of not occluded pixels whose horizontal disparity is zero. The region used for the evaluation is the central block, $\gamma = e$ (**R'** in Fig. 5), and the cost function can be defined as

$$\Phi(\lambda) = \Phi_{0^\circ}(\lambda) \cdot \Phi_{180^\circ}(\lambda), \tag{2}$$

where Φ_{0° and Φ_{180° operators return the percentage of not occluded pixels with zero horizontal disparity at 0° and 180°, respectively, and $\lambda \in \Lambda = [\lambda_{min}, \lambda_{max}]$. The optimum azimuthal direction λ^* maximizes the cost function $\Phi(\lambda)$ as shown in Fig. 8. The total shift applied to each panorama to achieve the optimum alignment, locating the regions **A** and the usable stereoscopic regions **B** locater in panoramic center and its antipode, is given by

$$I_i(n, m - \lambda^* - M/4), \tag{3}$$

$$I_j(n, m - m_\alpha - \lambda^* - M/4), \tag{4}$$

where m_α provides the fine alignment of I_j to I_i over **R** as defined in section 2.3.

An alternative feature based approach consists in selecting a direction in the alignment range and obtaining a set of N corresponding feature points U_1 and U_2 in the central regions of I_i and I_j at 0° and 180°. Having obtained these clouds of corresponding feature points, the optimum direction can be obtained by using an iterative optimization method such as the least mean squared to find the rotation λ^* that minimizes the Euclidean distance between feature points.

The featureless method was preferred over the feature based method given its simplicity and effectiveness.

3. Method

The entire optimum panoramic alignment algorithm proposed is summarized in this section.

The first step in the alignment process is to perform an image size scaling and gray scale conversion to the panoramic pair (I_i, I_j). The optimum alignment search begins by performing the viewing direction equalization of I_j with respect to I_i, which is used as a reference panorama. The viewing equalization can be done using the method described in section 2.3 given the similarity between panoramas within a cluster.

Once both panoramas are centered in the same scene element, the image disparity pattern is classified for both principal directions, λ and its antipodes $\lambda + M/2$, using the central region **R**. The dense image flow over **R** is calculated to obtain the horizontal and vertical disparity maps (D_x, D_y) and the occlusion map O [6]. Then, the block **R** is divided into nine blocks $\gamma \in \{a, b, c, d, e, f, g, h\}$ as described in section 2.4 (Fig. 6-(a)), and the characteristic disparity of each block is calculated.

The disparity operator $\Delta(D_x, D_y, O)_\gamma$, defined in section 2.1, is applied to the central block $\gamma = e$ in both principal directions. If the characteristic disparity of the central

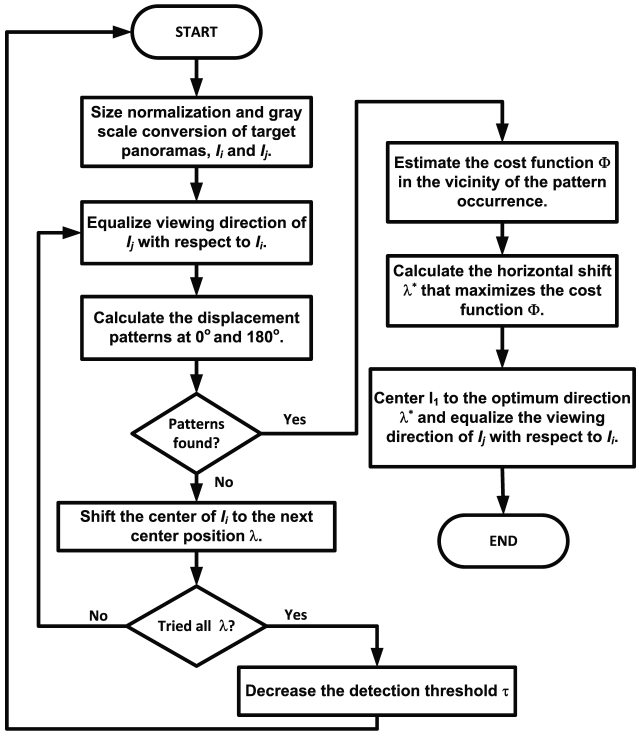


Figure 9. Optimal alignment algorithm.

block is different from the alignment case, then $(d_x, d_y) \neq (0, 0)$ simultaneously at $0°$ and $180°$, and the shift λ evaluated is not in the alignment region Λ and the process should be repeated for a new shift $\lambda \in [0, M/2)$. However, if the alignment is detected in both principal directions, the characteristic disparity is estimated for the eight remaining blocks surrounding the central one, and the pattern classification process described in section 2.4 is applied. As a result, the pattern in both principal directions is classified as one of the three target patterns, e.g., focal convergence (Ω_1), focal divergence (Ω_2), or near alignment (Ω_3).

The pattern classification may result in two events: *mismatch* or *match*, based on the target patterns for a positive detection defined in section 2.4. A *mismatch* event happens when the desired patterns are not found in that azimuthal direction. In that event, the next step is to center I_i, used as reference in the search, to the next azimuthal direction λ, where $\lambda \in [0, M/2)$, based on the search strategy used; if the search is intensive, then λ is incremented in one unit. After centering I_i in the new λ, I_j viewing direction is aligned with I_i and the pattern classification is performed again.

A *match* event is detected for the shift λ when the correct patterns were detected in both principal directions. In this case, an exhaustive evaluation of the cost function $\Phi(\lambda)$ for $\lambda \in \Lambda$ must be performed. The length in pixels of Λ varies from one viewpoint to another and also depends on the detection threshold τ defined in section 2.4. If all the possible directions were tested and no *match* event was detected, or if a *match* event was detected but in several distinct directions, then the search process must be repeated by lowering the threshold τ. The evaluation of the direction λ that maximizes the alignment likelihood defined by the cost function $\Phi(\lambda)$ in the neighborhood of the first *match* event results in the optimum alignment direction λ^*, as illustrated in Fig. 8.

Finally, the optimum shift is applied to the source panoramas I_i and I_j as defined in Eq. 3 and 4. The entire procedure is summarized in Fig. 9.

4. Results

A set of 68 panoramic snapshot pairs, acquired using a Ladybug 2 panoramic camera, were aligned with this technique demonstrating its effectiveness in diverse of indoor scenarios. From among these tests, two examples are presented in Fig. 10. Even though three panoramic pairs per triangular cluster were aligned independently, a single panorama pair per cluster is enough to illustrate the effectiveness of the proposed method. The initial horizontal disparity map for (I_1, I_2) is shown in Fig. 10-(a). A threshold has been applied to these maps to show only the areas where the image exhibits alignment based on the zero horizontal disparity criteria. After the method is applied, $I_1(n, m - \lambda^* - M/4)$ (Fig. 10-(b)) is used as a reference to illustrate the emergence of the alignment zones (Fig. 10-(c)). Notice that the panoramas in Fig. 10-(b) and (c) are centered in one of the directions of optimum stereoscopic composition, and therefore, showing the optimum alignment regions at $\theta = \pm 90°$. More results can be found on the companion website at http://www.LuisGurrieri.net/publications/cpia/.

5. Conclusions

A new approach to the fast acquisition and rendering of omnistereoscopic panoramas for image-based telepresence based on sampling the scene with a mesh of panoramic clusters has been introduced. In that context, a novel technique for the optimal alignment of panoramas specially designed for the identification of the region of best stereoscopic rendering potential has been proposed. This alignment technique does not rely on corresponding image features and does not require the camera parameters. The main advantage of this approach resides in its simplicity, making it suitable to the processing of the large numbers of panoramas needed in a stereoscopic scene sampling. Finally, techniques for the uncalibrated viewing direction equalization of closely located panoramas and methods for the image analysis and pattern classification were introduced as part of the proposed technique.

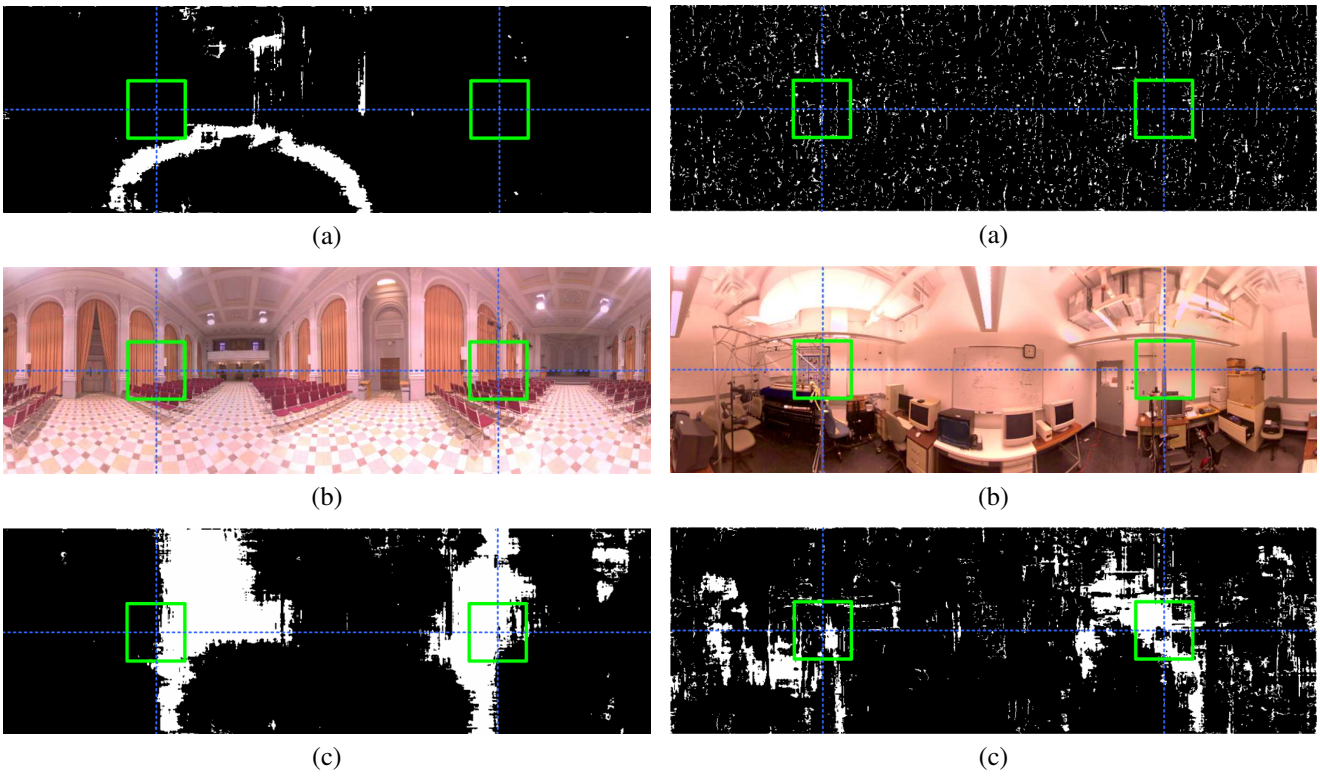


Figure 10. Two examples from distinctive locations: (a) initial horizontal disparity map arising between a pair $(I_1(n,m), I_2(n,m))$ where the zero disparity areas are represented in white and the square regions are located at $\theta = \pm 90^\circ$, (b) a reference panorama I_1 from the pair after aligning, i.e., $(I_1(n, m - \lambda^* - M/4), I_2(n, m - m_\alpha - \lambda^* - M/4))$, and (c) the horizontal disparity map of the aligned pair.

6. Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors would like to thank Quyen Sy for her help.

References

[1] M. Fiala and G. Roth. Automatic alignment and graph map building of panoramas. In *IEEE Int. Workshop on Haptic Audio Visual Environments and their Applications*, pages 103–108, 2005. 3, 4

[2] M. Hori, M. Kanbara, and N. Yokoya. Arbitrary stereoscopic view generation using multiple omnidirectional image sequences. In *Proc. IEEE Int. Conf. Pattern Recognition*, pages 286–289, 2010. 2

[3] F. Huang, R. Klette, and K. Scheibe. *Panoramic Imaging: Sensor-Line Cameras and Laser Range-Finders*. Wiley, 2008. 2

[4] S. B. Kang and R. Weiss. Characterization of errors in composing cylindrical panoramic images. *Panoramic vision: sensors, theory and applications, Ch. 12*, pages 205–226, 2001. 1

[5] F. Kangni and R. Laganiere. Epipolar geometry for the rectification of cubic panoramas. In *The 3rd Canadian Conference on Computer and Robot Vision*, pages 70–77, 2006. 3, 4

[6] A. S. Ogale and Y. Aloimonos. A roadmap to the integration of early visual modules. In *Proc. IEEE Int. Conf. Computer Vision: Special Issue on Early Cognitive Vision*, volume 72, pages 9–25, 2007. 4, 6

[7] S. Peleg and M. Ben-Ezra. Stereo panorama with a single camera. In *Proc. IEEE Conf. Computer Vision Pattern Recognition*, volume 1, pages 395–401, June 1999. 2

[8] S. Salehi and E. Dubois. Alignment of Cubic-Panorama image datasets using epipolar geometry. In *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, pages 1545–1548, Prague, Czech Republic, May 2011. 3, 4

[9] M. Uyttendaele, A. Criminisi, S. B. Kang, S. Winder, R. Szeliski, and R. Hartley. Image-based interactive exploration of real-world environments. *IEEE Computer Graphics and Applications*, pages 52–63, 2004. 1

[10] V. Vanijja and S. Horiguchi. A stereoscopic image-based approach to virtual environment navigation. *The Computer, the Internet and Management*, 14(2):68–81, Aug. 2006. 2

[11] K. Yamaguchi, H. Takemura, K. Yamazawa, and N. Yokoya. Real-time generation and presentation of view-dependent binocular stereo images using a sequence of omnidirectional images. In *Proc. IEEE Int. Conf. Pattern Recognition*, volume 4, pages 589–593, 2000. 2