Panoramic mosaicing optimization

Lionel Robinault¹, ², Stéphane Bres³, Serge Miguet¹

¹ LIRIS - Lyon2 University
² FOXSTREAM - LIRIS, Lyon2 University
³ LIRIS, INSA Lyon

Abstract

Motorized dome-type cameras, also called PTZ camera, allow the creation of panoramas. These panoramas represent the whole of the scene seen by the camera. In the case of a PTZ camera and with certain constraints, the scene seen by the camera can be considered as a sphere. The creation of a panorama consists in traversing a sphere in an exhaustive way. The acquired images are then projected on unspecified support which can be a cylinder, a cube or others. The projection of the rectangular images onto a sphere inevitably involves partial overlap between images. These overlaps lead to useless calculations. In order to limit the number of images we propose the calculation of an optimal trajectory for the camera according to intrinsic and extrinsic constraints.

Keywords: panoramic image, mosaic, environment mapping

Introduction and Motivation

The fall of the manufacturing cost of the motorized dome cameras allowed a democratization of their use. Until recently, the use of these cameras was limited to visualization controlled by an operator, via a joystick. The power of the calculators makes it possible today to automate certain operations of detection and following. Among the various possible approaches, one consists of creating a basic model facilitating segmentation. In the case of a PTZ camera, this implies the construction of a panorama. Creation of a panorama or a mosaic of images in an automatic way and by possibly controlling the viewing conditions is a challenge which motivates the industrialists like the world of research.

The developments of the techniques of visualization of these ten last years and in particular work of Chen [1] and Szeliski [2] allowed the creation and the exploitation of virtual panorama starting from micro-computers. There are several commercial programs available which permit the creation of panorama from simple photographs in a few clicks of the mouse. The applications are varied. Chen [1] was one of the first to use a cylindrical panorama with 360° to represent a scene and to allow navigation in real time in an application of virtual reality. His approach was used in the commercial product QuickTimeVR. For Lee [3] the construction of the panorama is used as support with the compression and the transmission of a video flow. Hsu and Anandan [4] describe what they call Mosaic-Based Compression (MBC) and describe several types of representations which make it possible to avoid the redundancy of information in the video data. Douze and Al [5] propose a method of robust tracking in a video sequence by using a panorama. Kang et al. [6] use a PTZ camera to carry out to follow people. Other applications are more marginal and outside of the framework of visualization. Peer and Al [7] present a system making it possible to map depth in panoramic. For Sato [8] the image mosaics can be used to support the analysis of documents.

One can note even in this very rapid review of the state of the art, that the applications are numerous and
various. However, even before being able to exploit a panorama, it first needs to be built. We can imagine several possible approaches to this construction depending on the concerned application. We are particularly interested in the creation of a panorama comprising only the static components of the scene and we use a camera PTZ for that. Briefly, the solution we choose consists in moving the camera to a position for a few seconds, depending on the intensity of the movement in the image, and building a basic background image starting from mixture of Gaussians [9]. This approach leads us to optimize the number of frames in order to limit useless overlap and thus to limit the number of images to be built. If the optical center of the camera is superposed with the axes of rotation, the world seen by the camera can be regarded as a sphere whose center is precisely the optical center. Indeed, if the optical center is superposed with the rotation center, it does not move during the movement of the camera. So, if two points are aligned with the optical center, these three points remainders aligned whatever the movement of the camera. This implies that whatever the swing angle, only one radius passes by these two points and that thus these two points are projected on a single point in the image. Consequently, when the optical center of the camera is confused with the center (or centers) of rotation, it is not possible to obtain a notion the depth even by multiplying the images taken at different angles. Since no concept of depth is possible under these conditions, we can thus consider that all the points are at the same distance from the center i.e. on surface of a sphere of unknown radius. Our problems is thus to traverse the all points of a sphere starting from the projection of the rectangular image.

A first solution consists with of a sweeping of the sphere by sectors. That is, for a given angle in the panorama, we carry out several acquisitions while varying the pitch angle. A simple optimization consists in varying the pitch angle in ascending then descending order with each shift of panorama. This first solution offers some advantages. The first is that the trajectory is simple to calculate since it is sufficient to assure that at the equator, the number of steps of angle in the panorama ensures a partial overlap of adjacent images. In an application of video surveillance, the second advantage is that the nature of the mobile objects (primarily of the people) have a vertically lengthened form. If the images are taken relatively quickly, the cutting of these objects will be limited. The principal disadvantage is that when one approaches the poles, image overlap is increasingly important. Information to be treated is thus duplicated unnecessarily. We propose a solution allowing the definition of the trajectory to limit the number of still images needed.

The article is built in the following way. In a first part we will study the surface occupied by an image according to the angles in the panorama and pitch angle. We will devote particular attention to the equations which define the projection of the image boundary on the sphere. These equations we will allow more simple calculation of the intersection between the images and the height of the band covered for a given pitch angle. In the second part we propose a search algorithm for optimal camera trajectory and we present results of the application of the method.

Equations

In the continuation of this article, θ represents the angle in the panorama and ϕ the pitch angle. The equations which delimit the low and high image boundaries are simple to calculate since they are large circles of slope (ϕ₀ ± ϕ₁), shifted angle θ₀ where ϕ₀ and θ₀ is the camera angle and ϕ the angle corresponding to the height of the image,

\[ ϕ₀ = tan \left( \frac{H}{2f} \right) \]

with H = height of the image and f = focal distance expressed in pixels.

The equation of the high and low image boundary is given by

\[ ϕ = cos^2 \left( \frac{cos(ϕ₀ ± ϕ₁)}{\sqrt{cos^2(θ) + cos^2(ϕ₀ ± ϕ₁) sin^2(θ)}} \right) \]

The equation of the left and right image boundary is a little more delicate to express. It is also a large circle which corresponds to the intersection of the sphere and a plane perpendicular to the xoz plan to which one applies a rotation of angle (θ₀ - θ₁) around z axis and of angle ϕ₀ around y axis.

The equation of the left and right image boundary is given by:

\[ ϕ = tan \left( \sqrt{A cos^2(θ) + B sin^2(θ) + C sin(θ) cos(θ) - 1} \right) \]

with \[ A = \frac{1}{sin^2(ϕ₀); \ B = 1 + \frac{cos(ϕ₀ ± ϕ₁)}{sin^2(θ₀ ± ϕ₁) sin^2(ϕ₀)} \]
and \[ C = -2 \frac{\cos \phi \cos(\theta \pm \Delta \theta)}{\sin(\theta \pm \Delta \theta) \sin^2 \theta} \]

This expression enables us to simply calculate the absolute value of \( \phi \). In order to resolve the problem of sign introduced into this expression, we note that the \( \phi \) and \( Z \) have the same sign. The equation of \( Z \) being given by:

\[ z = \frac{-\cos \phi}{\sin \theta} \cos \theta \cos \Delta \theta + \frac{\cos(\theta \pm \Delta \theta)}{\sin(\theta \pm \Delta \theta) \sin \theta} \cos \phi \sin \theta \]

The change of sign takes place for \( Z = 0 \), i.e. for \( \theta = \tan^{-1}(\cos \phi \tan \Delta \theta) \).

![Figure 1](image1.png)

**Figure 1.1**: \( \phi_0 = 0^\circ \)  
**Figure 1.2**: \( \phi_0 = 16^\circ \)  
**Figure 1.3**: \( \phi_0 = 45^\circ \)  
**Figure 1.4**: \( \phi_0 = 73^\circ \)

The four equations of the image boundary enable us to more finely analyze the projections of the solid angles of each image according to the conditions of the still image. We now will study an approach by bands in which, for a value \( \phi_0 \), one varies the angle in the panorama. The step of \( \Delta \theta \) variation will be thus optimized for each \( \phi_0 \) in order to limit covering and thus to limit the number of image to be acquired to sweep the whole of the scene.

For \( \phi_0 \) and \( \Delta \theta \) we obtain a band as represented below.

![Figure 2](image2.png)

**Figure 2**: Representation of the high and low boundary for images 640x480 with \( \phi_0 = 45^\circ \), \( f = 830 \) and \( \Delta \theta = 45^\circ \) (top) et \( \Delta \theta = 55^\circ \) (bottom)

By seeking the minimal value \( \phi_{\text{hi}} \) of the high boundary and the maximum value \( \phi_{\text{lo}} \) of the low boundary, we obtain the maximum height of the band for which one is certain to cover the surface without leaving of hole. The step value must be judiciously selected. Too small a \( \Delta \theta \) produces useless overlap whereas too large a \( \Delta \theta \) is likely to strongly decrease the height of the band, at the risk of not even ensuring a connectivity between the images. The parameters \( (\phi_1, \Delta \theta_1) \) of the following band are then given so that for \( \phi_1 > \phi_0 \), \( \phi_{\text{lo}} \leq \phi_{\text{hi}} \).

The height of the band depends on \( \phi_0 \), but especially on the step \( \Delta \theta \) which conditions the overlap of the images on the same tape. For the high image boundary, the analysis is rather simple since as long as there is \( 0 \leq \Delta \theta \leq \Delta \theta_1 \), the minimal value of the high boundary, corresponding to the intersection of the two images, is always located on the high image boundary for a value \( \theta = \theta_0 + \Delta \theta_2 \).

For the low boundary, the things are a little more complicated. We can define 4 fields. For \( 0 \leq \Delta \theta \leq \Delta \theta_1 \) (see figure 3.1), the maximum value of the low boundary corresponds to the intersection between the low boundary of image 1 and the left side of image 2. The \( \Delta \theta_1 \) value corresponds to the angle for which the intersection between the two images is located on the axis of symmetry of image 1. In this interval the function \( \phi_{\text{lo}} = \phi_{\text{olo}} + \frac{\Delta \theta}{2} \) is increasing and evolves of

\[ \phi_{\text{lo}} = \phi_0 - \tan^{-1} \left( \frac{h}{\sqrt{f^2 + f^2}} \right) \text{ for } \Delta \theta = 0 \text{ to } \phi_{\text{lo}} = \phi_0 - \tan^{-1} \left( \frac{h}{2f} \right) \text{ for } \Delta \theta = \Delta \theta_1. \]
If $\Delta \theta_1 < \Delta \theta \leq \Delta \theta_2$ (see figure 3.2), the function is continuous. Indeed the intersection between the two images is always located on the lower boundary of image 1 on the axis of symmetry. In this interval the function of the low image boundary is decreasing. The maximum value reached thus remains $\widetilde{\phi_l} = \phi_l - \tan\left(\frac{h}{2f}\right)$. At $\Delta \theta = \Delta \theta_2$, the corner lower right of image 1 is confused with the left lower corner of image 2. If one increases the value of $\Delta \theta$ further, the two images are intersected along their vertical boundary.

However if $\Delta \theta_2 < \Delta \theta \leq \Delta \theta_1$ (see figure 3.3), the value of the angle $\phi$ remains lower than $\phi_l = \phi_l - \tan\left(\frac{h}{2f}\right)$.

The maximum value is unchanged. Then, $\Delta \theta_1 < \Delta \theta \leq \Delta \theta_2$ (see figure 3.4), the value is higher east increases regularly until reaching value $\phi_l = \phi_l + \tan\left(\frac{h}{\sqrt{l^2 + f^2}}\right)$ for $\Delta \theta = \Delta \theta_1$.

If $\Delta \theta > \Delta \theta_1$, there is no intersection between the images. We cannot thus ensure sweeping without any holes in the panorama. Finally, one should note that this study is valid when $\phi_l$ and $\phi_0$ are all two strictly the higher than zero. If their values are strictly lower, it is enough to consider their absolute value to fall down in the case describes higher. When the values of $\phi_l$ and $\phi_0$ are not same sign, i.e. when $\phi_l > 0$ and $\phi_0 < 0$, the function $\phi_0 = f(\Delta \theta)$ is analyzed like the absolute value of a high boundary.

The following graph represents the functions and $\phi_{0}$ for an image 640x480 with $\phi_0 = 45^\circ$ and $f = 830$. We can note that the height of the covered band is maximum for $\Delta \theta = 0$, which does not have much interest, and that this height decreases when $\Delta \theta$ increases. We point out that $\Delta \theta$ is inversely proportional to the number of images necessary to cover a complete band.

**Research of the optimal trajectory in an approach by band**

We have now all the tools allowing us to seek the optimal trajectory in an approach by band. Starting from a value $\phi_0$, which corresponds to one of the limits of the material used, we can calculate the limits $\phi_{0,n}$ for several numbers of images $n$. These limits correspond to the maximum heights which we are ensured to cover for the number of image used, i.e. of the step $\Delta \theta$ between each image. For each value of $\phi_{0,n}$, we can calculate the values $\phi_{1,n}$, for which $\phi_{1,n} \leq \phi_{0,n}$. We thus obtain a tree representing the whole of the possible solutions.

In practice, we will place limits on the possible number of images in the solution. Firstly, the number of images must sufficient being to allow partial overlap of the images, which comes down to saying that $\Delta \theta \leq \theta_1$. Furthermore, if the number of images becomes high, the advantages one can obtain in band height becomes negligible. This is all the more important as one is seeking to reduce the number of images. Finally, the optimal number of images for each band occurs when $\Delta \theta$ in the close to $\theta_2$. It is tempting to think that the optimal number of images is that for which $\Delta \theta$ is closest to $\theta_2$. This is not necessarily the case for two reasons. First is that the number of images is an integer, and there is little chance of finding a number such that $\Delta \theta = \theta_2$. The second reason is that it is possible that under certain conditions, increasing slightly $\Delta \theta$, which inevitably involves a reduction in...
the high limit, can save us an image on the tape. On the other hand, decreasing $\Delta \theta$ leads to increasing the number of images but also makes it possible to increase the high limit slightly. This small increase might in the end allow us to remove a band.

In the graph below we represented by crosses, the whole of the possible solutions, i.e. the end of the branches of the tree. This graph makes it possible to determine the number of images necessary to reach a certain value of $\varphi$. The layout in blue indicates the optimal solution expressed as the number of images to reach a certain value of $\varphi$. The graph below is calculated for the filming material which we use and which does not allow one to go down below - 30°. We can notice that in certain cases, in particular between 11 and 15 images, the fact of increasing the number of image does not make it possible significantly to increase the extent of the $\varphi$ covered.

![Graph showing possible solutions](image.png)

**Figure 5 :** $\varphi$ covered according to the number of image.

We can also note that, although our material does not make it possible to control the pitch angle to values lower than - 30°, the limit $\varphi_0$ is lower than this value. This is with a particularly unfavourable case for which the selected number of image ($n = 7$) led to a value of $\Delta \theta$ lain between $\Delta \theta_3$ and $\Delta \theta_4$ and obviously, nearer to $\Delta \theta_4$ than of $\Delta \theta_3$.

Each sheets of the tree (represented by a cross in the graph), thus corresponds to the value of $\varphi$ reached for a total number of images and for one given trajectory. While following the branches which lead to these sheets, we can reconstitute the whole of the stages to follow, i.e. find the parameters of each band.

**Conclusion and future works**

This preliminary work, enabled us to calculate an optimized trajectory making it possible to acquire the whole of the scene with a minimal number of images and for a given focal distance. We made the choice of a regular spacing per band. It is the calculation of spacing which is optimal. Under certain limiting condition, with larger focal distances, it can be necessary to increase spacing between $\Delta \theta_3$ and $\Delta \theta_4$. We can gain an image on the band by decreasing the delta of this band without having to add an additional band. However, on the whole of the scene seen by the camera, there can there be particular regions one would like to visualize with a higher resolution, i.e. with a larger focal distance. It can be the case for the distant regions or zones of major interest. On contrary, there can be also regions of few interest like the sky or close objects. In this case, a weak focal distance is sufficient. Our next work will be thus to study the calculation of the trajectory by controlling the zoom factor according to the regions of interests. Initially, it will be with the operator to select the regions of interest and the value of the focal distance to be applied. In a second time, we would like the application itself to be able to select the optimal focal distance. The automatic calculation of the resolution can be a function of the movement in the image, texture or the distance. For that we will have to be able to automatically calculate an approximation of the distance between the camera and the concerned point. Several solutions are possible. Simplest consists in using a second camera in order to have a binocular vision. This second camera is not required to have an important resolution. We could also use an omnidirectional camera. Another solution consists in using a particularity of our camera. Indeed, on our PTZ camera, the geometrical center is not superposed with the optical center. Peer and Al [7] precisely use this eccentricity to calculate a chart of depth. A last solution consists in controlling the focus. We can estimate the distance according to fuzziness/sharpness of the regions on a function of the focus.

**References**


Figure 6 : position of the images on a cylindrical panorama

Figure 7 : Construction of a cubic panorama

Figure 8 : Face $\varphi = 0, \theta = 0$ of a cubic panorama