Fotopanoramen – Squaring the Circle in Panoramas

Projektleitung
Prof. Pietro Perona, California Institute of Technology, Computational Vision Lab
Prof. Dr. Gabriele Peters

Zeitraum
2005 – 2007

Kontakt
Prof. Dr. Gabriele Peters
Fachbereich Informatik
Fachhochschule Dortmund
Emil-Figge-Straße 42
44227 Dortmund
Tel.: (0231) 755-6796
E-Mail: gabriele.peters@fh-dortmund.de

Abstract

Pictures taken by a rotating camera cover the viewing sphere surrounding the center of rotation. Having a set of images registered and blended on the sphere what is left to be done, in order to obtain a flat panorama, is projecting the spherical image onto a picture plane. This step is unfortunately not obvious -- the surface of the sphere may not be flattened onto a page without some form of distortion. The objective of this paper is discussing the difficulties and opportunities that are connected to the projection from viewing sphere to image plane. We then show that multiple projections may coexist successfully in the same mosaic: these projections are chosen locally and depend on what is present in the pictures. We show that such multi-view projections can produce more compelling results than the global projections.

1 Introduction

As we explore a scene we turn our eyes and head and capture images in a wide field of view. For millennia painters and (more recently) photographers have grappled with the problem of creating pictures that render the visual impression of \`being there‘. Recent advances in storage, computation and display technology have made it possible to develop \`virtual reality‘ environments where the user feels \`immersed‘ in a virtual scene and can explore it by moving within it. However, the humble still picture, painted or printed on a flat surface, is still a popular medium: it is inexpensive to reproduce, easy and convenient to carry, store and display. Even more importantly, it has unrivaled size, resolution and contrast. Furthermore, the advent of inexpensive digital cameras, their seamless integration with computers, and recent progress in detecting and matching informative image features together with the development of good blending techniques have made it possible for any amateur photographer to produce automatically mosaics of photographs covering very wide fields of view and conveying the vivid visual impression of large panoramas, something that so far was the exclusive preserve of the artist. The geometry of single view point panoramas has long been well understood. This has been used, e.g., for mosaicing of video sequences.
By contrast when the point of view changes the mosaic is \`impossible‘ unless the structure of the scene is very special. Let‘s explore for a moment the \`easy‘ case, where all pictures share the same center of projection C.

If we consider the viewing sphere, i.e. the unit sphere centered in C, we may identify each pixel in each picture with the ray connecting C with that pixel and passing through the surface of the viewing sphere, as well as through the physical point in the scene that is imaged by that pixel. By detecting and matching visual features in different images we may register automatically the images with respect to each other. We may then map every pixel of every images we collected to the corresponding point of the viewing sphere and obtain a spherical image that summarizes all our information on the scene. This spherical image is the most natural representation: we may represent this way a scene of arbitrary angular width and if we place our head in C, the center of the sphere, we may rotate it around and capture the same images as if we were in the scene.

What is left to be done, in order to obtain our panorama-on-a-page, is projecting the spherical image onto a picture plane. This step is unfortunately not obvious -- the surface of the sphere may not be flattened onto a page without some form of distortion. The choice of projection from the sphere to the plane has been dealt with extensively by painters and cartographers.

The best known projection is linear perspective (also called \`gnomonic‘ and \`rectilinear‘). It may be obtained by projecting the relevant points of the viewing sphere onto a tangent plane, by means of rays emanating from the center of the sphere C. Linear perspective became popular amongst painters during the Renaissance. Brunelleschi is credited with being the first to use correct linear perspective. It is believed by many to be the only \`correct‘ projection because it maps lines in 3D space to lines on the 2D image plane and because when the picture is viewed from one special point, the \`center of projection‘ of the picture, the retinal image that is obtained is the same as when observing the original scene. A further, somewhat unexpected, virtue is that perspective pictures look \`correct‘ even if the viewer moves away from the center of projection, a very useful phenomenon called \`robustness of perspective’.

Unfortunately, linear perspective has a number of drawbacks. First of all: it may only represent scenes that are at most 180° wide: as the field of view becomes wider, the area of the tangent plane dedicated to representing one degree of visual angle in the peripheral portion of the picture becomes very large compared to the center, and eventually becomes unbounded. Second, there is an even more stringent limit to the size of the visual field that may be represented successfully using linear perspective: beyond widths of 30°-40° architectural structures (parallelepipeds) appear to be distorted, despite the fact that their edges are straight. Furthermore, spheres that are not in the center of the viewing field project to ellipses onto the image plane and appear unnatural and distorted. A similar phenomenon affects cylinders. Renaissance painters knew of these shortcomings and adopted a number of corrective measures, some of which we will discuss later.

The objective of this paper is discussing the difficulties and opportunities that are connected to the projection from viewing sphere to image plane, in the context of digital image mosaics. We first explore a number of alternatives to linear perspective which were developed by painters and cartographers. These are `global‘ projections and do not depend on image content. We explore experimentally the tradeoffs of these projections: how they distort architecture and people and how well do they tolerate wide fields of view. This is described in [ICCV 2005]. We then show that multiple projections may coexist successfully in the same mosaic: these projections are chosen locally and depend on what is seen in the pictures that form the mosaic. We conclude with a discussion of the work that lies ahead.

2 Multi View Projection

In [ICCV 2005] we discuss pros and cons of global projections such as perspective, Transverse Mercator, Mercator, Stereographic, and Geographic projection (figure 1). They are `global‘, in that once a tangent point or a tangent line is chosen, the projection is completely determined by this parameter. This is by no means a necessary property for a good projection. We may instead tailor the projection locally to the content of the images in order to improve the final effect. We next explore a few options for such multi-view projections.

2.1 Multi-Plane Perspective Projection

A global projection of wide panoramas bends lines, which is unpleasant to the eye. To obtain both a rectilinear appearance and a large field of view we suggest using a multi-plane perspective projection. Rather than projecting the sphere onto a single plane, multiple tangent planes to the sphere are used. Each projection is linear perspective. The tangent planes have to be arranged so that they may be unfolded into a flat surface without distortion, e.g., the points of tangency belong to a maximal circle. One may think of the intersections of the tangent planes being fitted with hinges that allow flattening. The projection onto each plane is perspective and covers only a limited field of view, thus it is pleasant to the eye.

Figure 1. Spherical projections. There are many spherical projections. Each has its pros and cons.

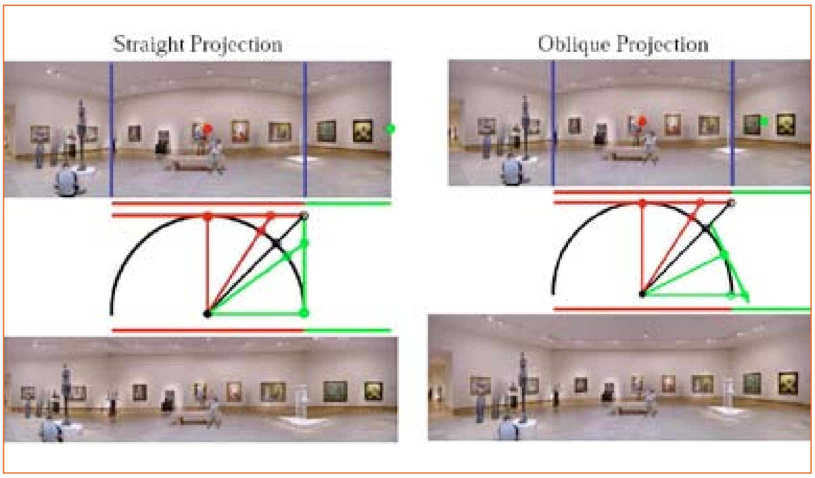

Figure 2. In each panel the top figure displays the geographic projection and the interaction required by the user – definition of the interaction lines between the tangent planes (marked in blue) and the center of projection for each tangent plane (marked in green and red). The middle panel displays a top view of the projection. The bottom panel displays the final result.

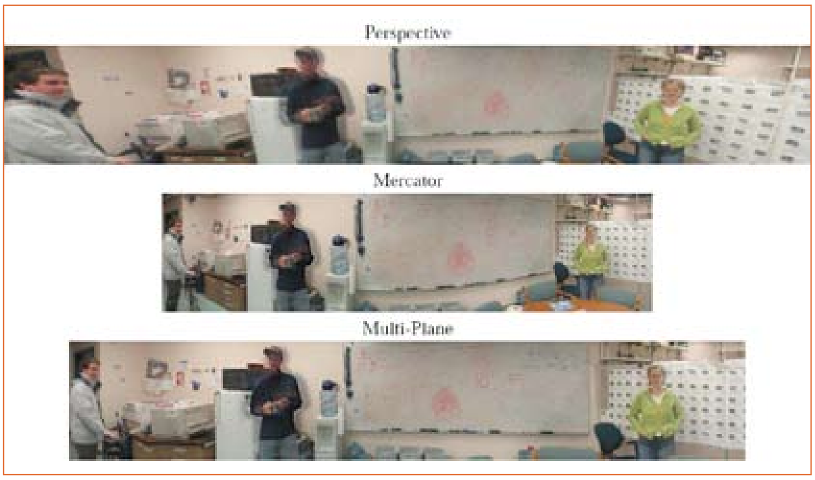

Figure 3. Architecture vs. spherical objects. The perspective projection distorts people at large viewing angles. The Mercator projection keeps the people undistorted, but distorts the wall and white-board at the background. The Multi-Plane projection provides the most compelling result with no noticeable distortions in both background and people.

Figure 4. Muti-Plane Multi-View. The multi-plane projection rectified the background but the chair on the right got distorted. Using the Multi-View approach the chair is undistorted.

This process introduces large orientation discontinuities at the intersection between the projection planes, however, in many man-made environment these discontinuities will not be noticed if they occur along natural discontinuities.
The tangent planes must therefore be chosen in a way that fits the geometry of the scene, e.g. so that the vertical edges of a room project onto the seams and each projection plane corresponds to a single wall. Orientation discontinuities caused by the projection this way co-occur with orientation discontinuities in the scene and therefore they are visually unnoticeable (see figure 1, “Multi-Plane”). Sometimes no seam may be found that completely corresponds to discontinuities in the scene: for example in figure 4 the chair on the right is clearly distorted. Another caveat is that some arrangements will cause a loss in the impression of depth: for example, when projecting a panorama of a standard room onto a square prism (see left panel of figure 2). Most often the sensation of depth can be maintained by appropriate choice of the projection planes (see right panel of figure 2).

We have currently implemented a simple user interface to allow choosing the position of the multiple tangent planes. We assume that the hinges between tangent planes are either associated to vertical or horizontal lines: the user is presented with the Geographic projection of the panorama and clicks once anywhere on a single vertical line to choose a seam and once again to choose the point of tangency of each projection plane. Automating this operation is an interesting exercise which we leave for the future.

2.2 Preserving Foreground Objects

The multi-plane perspective projection takes us back to the second challenge presented in Section 1. Recall, that even for small fields of view nearby (foreground) objects are often perceived as distorted. Our solution to this problem draws its inspiration from the Renaissance artists.

During the Renaissance the rules of perspective were understood, and linear perspective was used to produce pictures that had a realistic look. Painters noticed earlier on, that spheres and cylinders (and therefore people) would appear distorted if they were painted according to the rules of a global perspective projection (a sphere will project to an ellipse). It thus became common practice to paint people, spheres and cylinders by using linear perspective centered around each object. (see for example the The School of Athens by

Raphael). This results in paintings with multiple view points. There is one global view point used for the background and an additional view point for each foreground person/object.

Renaissance paintings look good precisely because they are constructed using a multiplicity of projections. Each projection is chosen in order to minimize the apparent distortion of either the ambient architecture, or of a specific person/object. We follow this example and adopt the multi-view point approach to construct realistic looking panoramas. We first separate the background and foreground objects. A panorama is constructed from the background by using a global projection: perspective for fields of view that are narrower than, say, 40° and Multi-Plane otherwise. The foreground objects are projected using a `local' perspective projection, with a central line of sight going through the center of each object, and then they are pasted onto the background. More in detail:

(1) Obtain a foreground-background segmentation for each image and cut out the foreground objects.

(2) Fill in the holes in the background caused by cutting out the foreground objects using a texture propagation technique.

(3) Construct a panorama of the filled background images.

(4) Overlay foreground objects on top of the background panorama. For each foreground object, find its bounding box in the original image and in the panorama if it were projected along with the background. Rescale the cut-out object to have the same height as its projection (note, that the width will be different). Paste the object so that the centers of the bounding boxes align.

This process is illustrated in figure 5. Five frames were taken out of a video sequence showing a child walking from right to left, while facing the camera. The child was cut-out from each image, texture propagation was used to fill in the holes and a perspective panorama of the background was constructed (see figure 5 top). The cut-outs of the child were then pasted onto the background in two ways. Once applying the same perspective projection used for the background, which resulted in distorting the child's head into a variety of ellipsoidal shapes (see figure 5 middle). Then using the multi-view approach described above which produced a significantly better looking result, removing all the head distortions, see figure 5 bottom. figure 4 displays our full solution including both multi-plane projection for the background and multi-view projection to correct the chair in the foreground. Another example is depicted in figure 3.

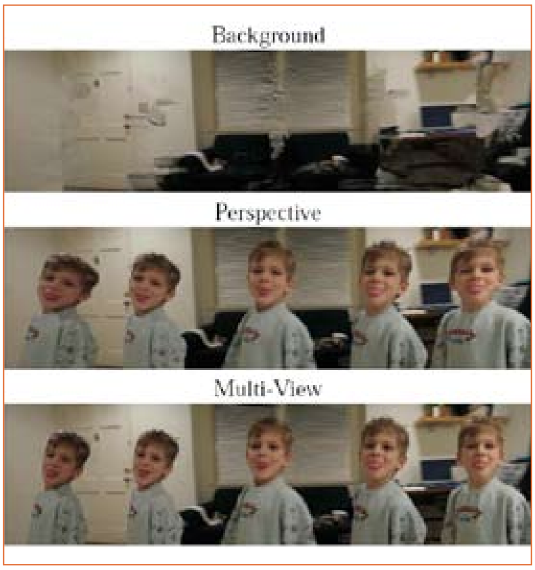

Figure 5. Correcting perspective distortions. Top: Panorama of the background only. Artifacts in the hole filling are visible, but are inessential as they wil be eventually covered by the foreground object. Center: A global perspective projection of both background and foreground. The child's head appears distorted. Bottom: A multi-view point panorama providing the most compelling look with no head distortions.

References

• Lihi Zelnik-Manor, Gabriele Peters, and Pietro Perona, Squaring the Circle in Panoramas, 10th IEEE International Conference on Computer Vision (ICCV 2005), 2005.