Novel Views from Non-calibrated Stereo

M. A. Akhloufi, P. Cohen and V. Polotski
Groupe de Recherche en Perception et Robotique
Ecole Polytechnique de Montreal
P.O. Box 6079, Station Centre-Ville,
Montreal (Qc), Canada, H3C 3A7.
(akhloufi, cohen, polotski)@ai.polymtl.ca

Abstract

In this paper we present a new approach to the synthesis of novel views from two images given by an uncalibrated stereo system. Unlike methods based on inferring the 3D structure of the scene or on using dense correspondence between source images to produce a new synthesized view, we use epipolar constraints associated with two cameras configuration and represented by a ***fundamental matrix*** *to reproject corresponding features in the image plane of the view to be synthesized. This requires only sparse correspondence between features in the source images.* ***Perspective image warping*** *is used to render the remaining dense set of image points via texture mapping. This new approach allows interactive view synthesis in applications such as: immersive telepresence systems, virtual and augmented reality and telerobotics. Only an initialization process which consists in matching features between the source views is needed. The efficiency of the method is illustrated on images of synthetic and real scenes.*

1 Introduction

In recent years, researchers from the traditionally separate fields of computer vision and computer graphics have been working on a common problem, namely the development of tools that permit a realistic rendering from a sparse set of images.

In classical approaches novel views are rendered from appropriate reconstructed 3D models. In computer graphics the object is first represented using a 3D modeler. Then a texture is mapped to add realism to the scene [22, 8]. In computer vision the classical approach is similar. First, a 3D model is reconstructed from a set of input images. Then, a texture extracted from the images is mapped onto the model [6]. Novel views are obtained from a final model at a given viewpoint. Many techniques have been proposed for 3D model reconstruction [13, 12, 6]. However, it is known that 3D model reconstruction is complex, time consuming and prone to errors [11]. The second step of the procedure, the rendering from the 3D models, may also involve intensive computation for achieving a visually realistic image.

To overcome these problems, an emerging field of research, *image based rendering*[17, 20] has been recently introduced in computer vision and computer graphics. It is currently gaining relevent interest. One of the first works in image-based rendering was proposed by Chen and Williams [5], who developed the *QuickTime VR* system. It consists mainly of mosaicking a set of images taken from a camera rotating about the axis passing through its principal point. These views are then stitched together prior to be reprojected, via cylindrical mapping, on a common cylindrical reference frame. The user can interactively move within the cylinder to display different views in the panoramic sequence captured by the rotating camera. The system suffers from the restrictions associated with the camera motion used to capture the views. Other techniques use view interpolation between source images to synthesize the novel view [4], but do not achieve a perspective rendering and need a dense set of correspondences between input images.

Introducing geometrical constraints leads to methods which produce geometrically valid pixel reprojections and, hence, synthesize views that are close to the desired real ones. In this category, we may distinguish three techniques: *(1)* Faugeras and *al.* [19, 7, 16] use the epipolar constraint captured by the *fundamental matrix* to reproject the corresponding points in the new image plane from dense correspondence between the source images. *(2)* Avidan and Shashua [3] use *trilinear tensors* to synthesize a novel view, also from dense correspondence. *(3)* In [21] the author propose