Image Alignment and Stitching: A Tutorial
Image Alignment and Stitching: A Tutorial

Richard Szeliski

Microsoft Research, USA
szeliski@microsoft.com

now
the essence of knowledge
Boston – Delft
Editorial Scope

Foundations and Trends® in Computer Graphics and Vision will publish survey and tutorial articles in the following topics:

- Rendering: Lighting models; Forward rendering; Inverse rendering; Image-based rendering; Non-photorealistic rendering; Graphics hardware; Visibility computation
- Shape: Surface reconstruction; Range imaging; Geometric modelling; Parameterization;
- Mesh simplification
- Animation: Motion capture and processing; Physics-based modelling; Character animation
- Sensors and sensing
- Image restoration and enhancement
- Segmentation and grouping
- Feature detection and selection
- Color processing
- Texture analysis and synthesis
- Illumination and reflectance modeling
- Shape Representation
- Tracking
- Calibration
- Structure from motion
- Motion estimation and registration
- Stereo matching and reconstruction
- 3D reconstruction and image-based modeling
- Learning and statistical methods
- Appearance-based matching
- Object and scene recognition
- Face detection and recognition
- Activity and gesture recognition
- Image and Video Retrieval
- Video analysis and event recognition
- Medical Image Analysis
- Robot Localization and Navigation

Information for Librarians
Foundations and Trends® in Computer Graphics and Vision, 2006, Volume 2, 4 issues. ISSN paper version 1572-2740. ISSN online version 1572-2759. Also available as a combined paper and online subscription.
Image Alignment and Stitching: A Tutorial

Richard Szeliski

Abstract

This tutorial reviews image alignment and image stitching algorithms. Image alignment algorithms can discover the correspondence relationships among images with varying degrees of overlap. They are ideally suited for applications such as video stabilization, summarization, and the creation of panoramic mosaics. Image stitching algorithms take the alignment estimates produced by such registration algorithms and blend the images in a seamless manner, taking care to deal with potential problems such as blurring or ghosting caused by parallax and scene movement as well as varying image exposures. This tutorial reviews the basic motion models underlying alignment and stitching algorithms, describes effective direct (pixel-based) and feature-based alignment algorithms, and describes blending algorithms used to produce seamless mosaics. It ends with a discussion of open research problems in the area.
# Contents

1 Introduction 1

2 Motion Models 5

2.1 2D (planar) Motions 6

2.2 3D Transformations 8

2.3 Cylindrical and Spherical Coordinates 15

2.4 Lens Distortions 18

3 Direct (pixel-based) Alignment 21

3.1 Error Metrics 22

3.2 Hierarchical Motion Estimation 25

3.3 Fourier-Based Alignment 27

3.4 Incremental Refinement 31

3.5 Parametric Motion 37

4 Feature-Based Registration 43

4.1 Keypoint Detectors 43

4.2 Feature Matching 47

4.3 Geometric Registration 52

4.4 Direct vs. Feature-Based Alignment 59

5 Global Registration 63
5.1 Bundle Adjustment 63
5.2 Parallax Removal 68
5.3 Recognizing Panoramas 70

6 Compositing 75
6.1 Choosing a Compositing Surface 75
6.2 Pixel Selection and Weighting 78
6.3 Blending 84

7 Extensions and Open Issues 91

References 95
Introduction

Algorithms for aligning images and stitching them into seamless photo-mosaics are among the oldest and most widely used in computer vision. Frame-rate image alignment is used in every camcorder that has an “image stabilization” feature. Image stitching algorithms create the high-resolution photo-mosaics used to produce today’s digital maps and satellite photos. They also come bundled with most digital cameras currently being sold, and can be used to create beautiful ultra wide-angle panoramas.

An early example of a widely used image registration algorithm is the patch-based translational alignment (optical flow) technique developed by Lucas and Kanade [123]. Variants of this algorithm are used in almost all motion-compensated video compression schemes such as MPEG and H.263 [113]. Similar parametric motion estimation algorithms have found a wide variety of applications, including video summarization [20,203,111],93 video stabilization [81], and video compression [95,114]. More sophisticated image registration algorithms have also been developed for medical imaging and remote sensing – see [29,226,71] for some previous surveys of image registration techniques.
In the photogrammetry community, more manually intensive methods based on surveyed ground control points or manually registered tie points have long been used to register aerial photos into large-scale photo-mosaics [181]. One of the key advances in this community was the development of bundle adjustment algorithms that could simultaneously solve for the locations of all of the camera positions, thus yielding globally consistent solutions [207]. One of the recurring problems in creating photo-mosaics is the elimination of visible seams, for which a variety of techniques have been developed over the years [135, 136, 148, 50, 1].

In film photography, special cameras were developed at the turn of the century to take ultra wide-angle panoramas, often by exposing the film through a vertical slit as the camera rotated on its axis [131]. In the mid-1990s, image alignment techniques were started being applied to the construction of wide-angle seamless panoramas from regular hand-held cameras [124, 193, 43, 194]. More recent work in this area has addressed the need to compute globally consistent alignments [199, 167, 178], the removal of “ghosts” due to parallax and object movement [50, 178, 210, 1], and dealing with varying exposures [124, 210, 116, 1]. (A collection of some of these papers can be found in [19].) These techniques have spawned a large number of commercial stitching products [43, 168], for which reviews and comparison can be found on the Web.

While most of the above techniques work by directly minimizing pixel-to-pixel dissimilarities, a different class of algorithms works by extracting a sparse set of features and then matching these to each other [227, 35, 38, 7, 129, 30]. Feature-based approaches have the advantage of being more robust against scene movement and are potentially faster, if implemented the right way. Their biggest advantage, however, is the ability to “recognize panoramas,” i.e., to automatically discover the adjacency (overlap) relationships among an unordered set of images, which makes them ideally suited for fully automated stitching of panoramas taken by casual users [30].

What, then, are the essential problems in image alignment and stitching? For image alignment, we must first determine the appropriate mathematical model relating pixel coordinates in one image to pixel coordinates in another. Section 2 reviews these basic motion
models. Next, we must somehow estimate the correct alignments relating various pairs (or collections) of images. Section 3 discusses how direct pixel-to-pixel comparisons combined with gradient descent (and other optimization techniques) can be used to estimate these parameters. Section 4 discusses how distinctive features can be found in each image and then efficiently matched to rapidly establish correspondences between pairs of images. When multiple images exist in a panorama, techniques must be developed to compute a globally consistent set of alignments and to efficiently discover which images overlap one another. These issues are discussed in Section 5.

For image stitching, we must first choose a final compositing surface onto which to warp and place all of the aligned images (Section 6). We also need to develop algorithms to seamlessly blend overlapping images, even in the presence of parallax, lens distortion, scene motion, and exposure differences (Section 6). In the last section of this survey, additional applications of image stitching and open research problems were discussed.
References


References


References


References


References


References


References


106 References


Full text available at: http://dx.doi.org/10.1561/0600000009


References


