

Image Mosaicing and Super-resolution

Abstract

This book investigates the problem of how information contained in multiple, overlapping images of the same scene may be combined to produce images of superior quality. This area, generically titled *frame fusion*, offers the possibility of reducing noise, extending the field of view, removal of moving objects, removing blur, increasing spatial resolution and improving dynamic range. As such, this research has many applications in fields as diverse as forensic image restoration, computer generated special effects, video image compression, and digital video editing.

An essential enabling step prior to performing frame fusion is *image registration*, by which an accurate estimate of the point-to-point mapping between views is computed. A robust and efficient algorithm is described to automatically register multiple images using only information contained within the images themselves. The accuracy of this method, and the statistical assumptions upon which it relies, are investigated empirically.

Two forms of frame-fusion are investigated. The first is *image mosaicing*, which is the alignment of multiple images into a single composition representing part of a 3D scene. Various methods of presenting the composite image are demonstrated, and in particular, a novel algorithm is developed for automatically choosing an optimal viewing transformation for certain cases. In addition, a new and efficient method is demonstrated for the matching of point features across multiple views.

The second frame-fusion method is *super-resolution*, which aims to restore poor-quality video sequences by removing the degradations inherent in the imaging process. The framework presented here uses a generative model of the imaging process, which is discussed in detail and an efficient implementation presented. An algorithm is developed which seeks a maximum likelihood estimate under this model, and the factors affecting its performance are investigated analytically and empirically.

The use of “generic” prior image models in a Bayesian framework is described and shown to produce dramatically improved super-resolution results. Finally, super-resolution algorithms are developed which make use of image models which are tuned to specific classes of image. These algorithms are shown to produce results of comparable or better quality than those using generic priors, while having a lower computational complexity. The technique is applied to images of text and faces.

Throughout this work, the performance of the algorithms is evaluated using real image sequences. The applications demonstrated include the separation of latent marks from cluttered, non-periodic backgrounds in forensic images; the automatic creation of full 360° panoramic mosaics; and the super-resolution restoration of various scenes, including text and faces in low-quality video.

Contents

1	Introduction	1
1.1	Background	1
1.2	Modelling assumptions	4
1.3	Applications	5
1.4	Principal contributions	5
2	Literature Survey	7
2.1	Image registration	7
2.1.1	Registration by a geometric transformation	7
2.1.2	Ensuring global consistency	9
2.1.3	Other parametric surfaces	10
2.2	Image mosaicing	10
2.3	Super-resolution	12
2.3.1	Simple super-resolution schemes	12
2.3.2	Methods using a generative model	13
2.3.3	Super-resolution using statistical prior image models	14
3	Registration: Geometric and Photometric	17
3.1	Introduction	17
3.2	Imaging geometry	18
3.3	Estimating homographies	20
3.3.1	Linear estimators	20
3.3.2	Non-linear refinement	21
3.3.3	The maximum likelihood estimator of H	21
3.4	A practical two-view method	22
3.5	Assessing the accuracy of registration	26
3.5.1	Assessment criteria	26
3.5.2	Obtaining a ground-truth homography	26
3.6	Feature-based vs. direct methods	34
3.7	Photometric registration	37
3.7.1	Sources of photometric difference	37
3.7.2	The photometric model	37
3.7.3	Estimating the parameters	37
3.7.4	Results	38

3.8	Application: Recovering latent marks in forensic images	40
3.8.1	Motivation	40
3.8.2	Method	40
3.8.3	Further examples	42
3.9	Summary	45
4	Image Mosaicing	47
4.1	Introduction	47
4.2	Basic method	47
4.2.1	Outline	48
4.2.2	Practical considerations	49
4.3	Rendering from the mosaic	51
4.3.1	The reprojection manifold	51
4.3.2	The blending function	54
4.3.3	Eliminating seams by photometric registration	57
4.3.4	Eliminating seams due to vignetting	57
4.3.5	A fast alternative to median filtering	58
4.4	Simultaneous registration of multiple views	59
4.4.1	Motivation	59
4.4.2	Extending the two-view framework to N-views	63
4.4.3	A novel algorithm for feature-matching over N-views	66
4.4.4	Results	68
4.5	Automating the choice of reprojection frame	70
4.5.1	Motivation	70
4.5.2	Synthetic camera rotations	73
4.6	Applications of image mosaicing	75
4.7	Mosaicing non-planar surfaces	76
4.8	Mosaicing “user’s guide”	76
4.9	Summary	77
4.9.1	Further examples	78
5	Super-resolution: Maximum Likelihood and Related Approaches	81
5.1	Introduction	81
5.2	What do we mean by “resolution”?	82
5.3	Single-image methods	83
5.4	The multi-view imaging model	84
5.4.1	A note on the assumptions made in the model	86
5.4.2	Discretization of the imaging model	86
5.4.3	Related approaches	87
5.4.4	Computing the elements in M_n	88
5.4.5	Boundary conditions	94
5.5	Justification for the Gaussian PSF	95
5.6	Synthetic test images	96
5.7	The average image	97
5.7.1	Noise robustness	100
5.8	Rudin’s forward-projection method	105

5.9	The maximum-likelihood estimator	109
5.10	Predicting the behaviour of the ML estimator	110
5.11	Sensitivity of the ML estimator to noise sources	114
5.11.1	Observation noise	114
5.11.2	Poorly estimated PSF	115
5.11.3	Inaccurate registration parameters	115
5.12	Irani and Peleg’s method	123
5.12.1	Least-squares minimization by steepest descent	123
5.12.2	Irani and Peleg’s algorithm	124
5.12.3	Relationship to the ML estimator	126
5.12.4	Convergence properties	126
5.13	Gallery of results	127
5.14	Summary	136
6	Super-resolution Using Bayesian Priors	137
6.1	Introduction	137
6.2	The Bayesian framework	138
6.2.1	Markov random fields	139
6.2.2	Gibbs priors	139
6.2.3	Some common cases	140
6.3	The optimal Wiener filter as a MAP estimator	140
6.4	Generic image priors	142
6.5	Practical optimization	146
6.6	Sensitivity of the MAP estimators to noise sources	147
6.6.1	Exercising the prior models	147
6.6.2	Robustness to image noise	148
6.7	Hyper-parameter estimation by cross-validation	156
6.8	Gallery of results	164
6.9	Super-resolution “user’s guide”	167
6.10	Summary	167
7	Super-resolution Using Sub-space Models	169
7.1	Introduction	169
7.2	Bound constraints	170
7.3	Learning a face model using PCA	171
7.4	Super-resolution using the PCA model	175
7.4.1	An ML estimator (FS-ML)	175
7.4.2	MAP estimators	176
7.5	The behaviour of the face model estimators	177
7.6	Examples using real images	181
7.7	Summary	188

8	Conclusions and Extensions	189
8.1	Summary	189
8.2	Extensions	190
8.2.1	Application to digital video	191
8.2.2	Model-based super-resolution	192
8.3	Final observations	194
A	Large-scale Linear and Non-linear Optimization	195
A.1	Notation	195
A.2	Quadratic functions	195
A.2.1	Steepest descent	196
A.2.2	Conjugate gradient descent	196
A.3	General non-linear functions	199
A.3.1	Newton methods	199
A.3.2	Computing a Newton step	201
A.4	Non-linear least-squares	201
	References	203
	Index	217