Optimizing and Learning for Super-resolution

Lyndsey C. Pickup, Stephen J. Roberts and Andrew Zisserman
Information Engineering Building, Dept. of Engineering Science,
Parks Road, Oxford, OX1 3PJ, UK
{elle,sjrob,az}@robots.ox.ac.uk

Abstract

In multiple-image super-resolution, a high resolution image is estimated from a number of lower-resolution images. This involves computing the parameters of a generative imaging model (such as geometric and photometric registration, and blur) and obtaining a MAP estimate by minimizing a cost function including an appropriate prior.

We consider the quite general geometric registration situation modelled by a plane projective transformation, and make two novel contributions: (i) in previous approaches the MAP estimate has been obtained by first computing and fixing the registration, and then computing the super-resolution image with this registration. We demonstrate that superior estimates are obtained by optimizing over both the registration and image; (ii) the parameters of the edge preserving prior are learnt automatically from the data, rather than being set by trial and error.

We show examples on a number of real sequences including multiple stills, digital video, and DVDs of movies.

1 Introduction

Multi-frame image super-resolution refers to the process where a group of images of the same scene are fused to produce an image or images with a higher spatial resolution, or with more visible detail in the high spatial frequency features [9]. The limits on the resolution of the original imaging device can be improved by exploiting the relative sub-pixel motion between the scene and the imaging plane. Such problems are common, with everything from holiday snaps and DVD frames to satellite terrain imagery providing collections of low-resolution images to be enhanced, for instance to produce a more aesthetic image for media publication [17], or for higher-level vision tasks such as object recognition or localization [7]. Figure 1 shows two examples of multi-frame super-resolution.

The problem is often broken down into several distinct stages: image registration or motion estimation, low-resolution image blur estimation, selection of a suitable prior, and super-resolution image estimation. However, these stages are seldom truly independent, and this is too often ignored in current super-resolution techniques [1, 2, 4, 9, 14].

In this work we introduce an algorithm to estimate a super-resolution image at the same time as finding the low-resolution image registrations, and show that this *simultaneous* approach offers visible benefits on results obtained from real data sequences. The registration model we handle is fully projective, and we also incorporate a photometric

BMVC 2006 doi:10.5244/C.20.46

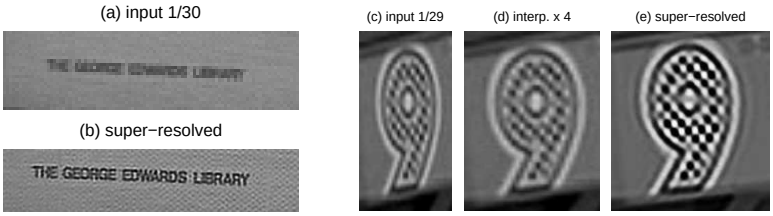

Figure 1: (a) One of 30 close-ups from a digital camera sequence; (b) automatic super-resolution output. (c) Close-up from one of 29 DVD movie frames; (d) interpolated to correct aspect ratio (for comparison only); (e) automatic super-resolution output.

model to handle brightness changes often present in images captured in a temporal sequence. This makes the model far more general than most super-resolution approaches. Additionally, the algorithm learns parameters for the high-resolution image prior (or regularizer). Each scene will have different underlying image statistics, so by adapting to these, the algorithm preserves as much richness and detail as possible from the original scene without encountering problems with conditioning.

1.1 Background

The vast majority of current super-resolution methods either pre-register the inputs using standard registration techniques, or assume that a perfect registration is given *a priori* [4, 9]. In most cases, the selection of values for the prior and blur functions are deferred to the user, rather than chosen with reference to the input data. A few methods assume no scene motion, and use other cues such as lighting or varying zoom [10].

There are two notable methods which learn a registration using the super-resolution model. Hardie *et al.* [7] use their high-resolution estimate to reconsider the registrations, but limit themselves to shifts on a $\frac{1}{4}$-pixel-spaced grid, so registration is a search across grid locations, which would quickly become infeasible with more degrees of freedom. Tipping and Bishop [18] marginalize out the high-resolution image to learn a Euclidean registration directly, but with such a high computational cost that their inputs are restricted to 9×9 pixels. Both these approaches also rely on Gaussian image priors.

The Generalized Cross Validation (GCV) work of Nguyen *et al.* [14] also learns a blur width (though not registrations), then learns a regularization coefficient based on the data, though also restricted to a Gaussian (*i.e.* not edge-preserving) prior. Taken together, these techniques motivate the development of an approach capable of registering the images at the same time as super-resolving, but without the need for the Gaussian form of prior.

2 The anatomy of multi-frame super-resolution

A high-resolution scene $\mathbf{x}$, with N pixels, is assumed to have generated a set of K low-resolution images $\mathbf{y}^{(k)}$, each with M pixels. For each image, the warping, blurring and subsampling of the scene is modelled by an $M \times N$ sparse matrix $\mathbf{W}^{(k)}$ [4, 18], and a global affine photometric correction results from addition and multiplication across all

pixels by scalars λ_α and λ_β respectively [4]. Thus the generative model is

$$\mathbf{y}^{(k)} \quad = \quad \lambda_\alpha^{(k)} \mathbf{W}^{(k)} \mathbf{x} + \lambda_\beta^{(k)} \mathbf{1} + \boldsymbol{\epsilon}^{(k)}, \tag{1}$$

where $\boldsymbol{\epsilon}^{(k)}$ represents noise on the low-resolution image, and consists of *i.i.d.* samples from a zero-mean Gaussian with *std* σ_N, and images $\mathbf{x}$ and $\mathbf{y}^{(k)}$ are represented as vectors. Given $\left\{\mathbf{y}^{(k)}\right\}$, the goal is to recover $\mathbf{x}$, without any prior knowledge of $\left\{\mathbf{W}^{(k)}, \boldsymbol{\lambda}^{(k)}, \sigma_N\right\}$. The problem is almost always poorly conditioned, so a prior over $\mathbf{x}$ is usually required to avoid solutions which are subjectively very implausible to the human viewer.

The *Maximum a Posteriori* (MAP) equation is derived from the combination of the generative model of equation (1) with an image prior based on the Huber function, $\rho(.)$, which has previously been shown to be a good selection for image super-resolution [3, 4]. The objective function takes the form

$$\mathcal{F} \quad = \quad \underbrace{\sum_{k=1}^{K} \left|\left| \mathbf{y}^{(k)} - \lambda_\alpha^{(k)} \mathbf{W}^{(k)} \mathbf{x} - \lambda_\beta^{(k)} \mathbf{1} \right|\right|_2^2}_{\text{generative model}} \quad + \quad \underbrace{\nu \sum_{g \in \mathcal{G}(\mathbf{x})} \rho\left(g, \alpha\right)}_{\text{prior}}, \tag{2}$$

where the prior term will be explained in section 2.3. In the main body of the algorithm proposed here, $\mathcal{F}$ is optimized explicitly with respect to $\mathbf{x}$, the set of geometric registration parameters $\boldsymbol{\phi}$ (which parameterize $\mathbf{W}$), and the set of λ_α and λ_β values, $\boldsymbol{\lambda}$, at the same time. This is alternated with a cross-validation-driven update of the prior parameters ν and α, and the algorithm is run with a selection of point-spread function (PSF) kernels. A discussion of the registration, point-spread function and image prior concerns is given below, followed by implementation details for our approach in section 3.

2.1 Image Registration

Standard approaches to super-resolution first determine the registration, then fix it and optimize a function like $\mathcal{F}$ with respect only to $\mathbf{x}$ to obtain the final super-resolution estimate. However, if the set of input images is assumed to be noisy, it is reasonable to expect the registration to be adversely affected by the noise. In contrast, we make use of the high-resolution image estimate common to all the low-resolution images.

The registration problem itself is not convex, and repeating textures can cause naïve intensity-based registration algorithms to fall into a local minimum, though when initialized sensibly, very accurate results are obtained. The pathological case where the footprints of the low-resolution images fail to overlap in the high-resolution frame can be avoided by adding an extra term to $\mathcal{F}$ to penalise large deviations in the registration parameters from the initial registration estimate (see section 3.3).

Errors in either geometric or photometric registration in the low-resolution dataset have consequences for the estimation of other super-resolution components. The uncertainty in localization can give the appearance of a larger PSF because the effects of a scene point on the low-resolution image set is more dispersed. Uncertainty in photometric registration increases the variance of intensity values at each spatial location, giving the appearance of more low-resolution image noise, because low-resolution image values will tend to lie further from the values of the back-projected estimate. Increased noise in turn is an indicator that a change in the prior weighting is required, thus lighting parameters can have a knock-on effect on the image edge appearances.

2.2 The point spread function

By far the most difficult component of most super-resolution systems to determine is the point-spread function (PSF), which is of crucial importance, because it describes how each pixel in $\mathbf{x}$ influences pixels in the observed images. Resulting from optical blur in the camera, artifacts in the sensor medium (film or a CCD array), and potentially also through motion during the image exposure, the PSF is almost invariably modelled either as an isotropic Gaussian or a uniform disk in super-resolution, though some authors suggest other function derived from assumptions on the camera optics [2, 3]. The exact shape of the kernel depends on the entire process from photon to pixel.

Identifying and reversing the blur process is the domain of *Blind Image Deconvolution*. Approaches based on Generalized Cross-Validation [16] or Maximum Likelihood [12] are less sensitive to noise than other available techniques [11], and both have direct analogues in current super-resolution work [14, 18]. Because of the parametric nature of both sets of algorithms, neither is truly capable of recovering an arbitrary point-spread function. With this in mind, we choose a few sensible forms of PSF and concentrate on super-resolution which handles mismatches between the true and assumed PSF as gracefully as possible.

2.3 The super-resolution image prior

Super-resolution is an ill-posed problem, and ML solutions are usually corrupted by "chequer-board" patterns resulting from fitting to noise on the input dataset. The Huber potential function has been shown to be a good choice for an edge-preserving prior [3, 4], and is defined over pair-wise image gradient estimates in the horizontal, vertical and two diagonal directions, leading to an eight-neighbour graph structure over the image, denoted $\mathcal{G}(\mathbf{x})$. The Huber function is

$$\rho(z,\alpha) \quad = \quad \begin{cases} z^2 & \text{if } |z| < \alpha \\ 2\alpha|z| - \alpha^2 & \text{otherwise} \end{cases} \tag{3}$$

where α is the single parameter.

Many authors favour a Gaussian form for the prior over $\mathbf{x}$ for its simplicity and conjugacy with the data likelihood term of the optimization, though it tends to have the undesirable effect of softening black-white transitions at image edges. In super-resolution images, edges are preserved better with functions like the Huber potential, or Bilinear Total Variation [5], which are more tolerant of outliers, or even patch-based texture priors [15, 20]. These approaches have partition functions which are very costly to compute directly, and so it can be expensive to learn the prior parameters directly.

3 Super-resolution with motion and prior estimation

In this section, we fill out the remaining details of the simultaneous super-resolution approach, which consists of three distinct components. First, there are convenient initializations for the registrations and the estimate of $\mathbf{x}$, which by themselves even give a quick and reasonable super-resolution estimate. The second and third algorithm components form the body of the iterative loop: the MAP estimation, and the cross-validation regularizer update. Convergence is defined to be the point at which all parameters change by

less than a preset threshold in successive iterations. The loop is repeated till this point, typically taking 3-10 iterations. The algorithm is summarized in figure 2.

3.1 Initialization

Input images are assumed to be pre-registered by a standard algorithm [8] so that points at the image centres correspond to within a small number of low-resolution pixels.

A candidate PSF is selected in order to compute the *average image*, $\mathbf{a}$, which is a stable though excessively smooth approximation to $\mathbf{x}$. Each pixel in $\mathbf{a}$ is a weighted combination of pixels in $\mathbf{y}$, such that a_i depends strongly on y_j if y_j depends strongly on x_i, according to the weights in $\mathbf{W}$. Lighting changes must also be taken into consideration, so

$$\mathbf{a} \quad = \quad \mathbf{S}^{-1}\mathbf{W}^T\mathbf{\Lambda}_\alpha^{-1}(\mathbf{y} - \mathbf{\Lambda}_\beta), \tag{4}$$

where $\mathbf{W}$, $\mathbf{y}$, $\mathbf{\Lambda}_\alpha$ and $\mathbf{\Lambda}_\beta$ are the stacks of the K groups of $\mathbf{W}^{(k)}$, $\mathbf{y}^{(k)}$, $\lambda_\alpha^{(k)}\mathbf{I}$, and $\lambda_\beta^{(k)}\mathbf{1}$ respectively, and $\mathbf{S}$ is a diagonal matrix whose elements are the column sums of $\mathbf{W}$. Notice that both inverted matrices are diagonal, so $\mathbf{a}$ is simple to compute. Using $\mathbf{a}$ in place of $\mathbf{x}$, we optimize the first term of $\mathcal{F}$ with respect to $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$ only. This provides a good estimate for the registration parameters, without requiring $\mathbf{x}$ or the prior parameters.

To initailise $\mathbf{x}$, the Scaled Conjugate Gradients algorithm is applied to the ML solution, but terminated after around $\frac{K}{4}$ steps, before the instabilities dominate. This gives a sharper result than initializing with $\mathbf{a}$ as in [4]. When only a few images are available, a more stable ML solution can be found by using a constrained optimization to bound the pixel values so they must lie in the permitted image intensity range.

The prior parameters are initialized to around $\alpha = 0.01$ and $\nu = 0.1$; as these are both strictly positive quantities, logs of the values are used. For the PSF, a Gaussian with *std* ≈ 0.45 low-resolution pixels is reasonable for in-focus images, and a disk of radius upwards of 0.8 is suitable for slightly defocused scenes.

3.2 Learning the prior parameters with unknown registration

It is necessary to determine ν and α while still in the process of converging on the estimates of $\mathbf{x}$, $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$. This is done by removing some *individual low-resolution pixels*

1. Initialize PSF, image registrations, super-resolution image and prior parameters according to section 3.1.
2. (a) (Re)-sample the set of validation pixels
   (b) Update α and ν (prior parameters) using cross-validation-style Gradient Descent (section 3.2). This includes a few steps of a sub-optimization of $\mathcal{F}$ *w.r.t.* $\mathbf{x}$.
   (c) Optimize $\mathcal{F}$ (equation 2) jointly with respect to $\mathbf{x}$ (super-resolution image), λ (photometric transform) and $\boldsymbol{\phi}$ (geometric transform).
3. If the maximum absolute change in α, ν, or any element of $\mathbf{x}$, $\boldsymbol{\lambda}$ or $\boldsymbol{\phi}$ is above preset convergence thresholds, return to 2.

Figure 2: Basic structure of our multi-frame super-resolution algorithm.

from the problem, solving for $\mathbf{x}$ using the remaining pixels, then projecting this back into the original image frames to determine its quality using the withheld validation pixels using the robust L1 norm. The selected α and ν should minimize this cross-validation error.

This defines a subtly different cross-validation approach to those used previously for image super-resolution, because validation pixels are selected at random from the collection of $K \times M$ *individual linear equations* comprising the overall problem, rather than from the K *images*. This distinction is important when uncertainty in the registrations is assumed, since validation *images* can be misregistered in their entirety. Assuming independence of the registration error on each frame given $\mathbf{x}$, the pixel-wise validation approach has a clear advantage.

In determining a search direction in (ν, α)-space, $\mathcal{F}$ can be optimized *w.r.t.* $\mathbf{x}$, starting with the current $\mathbf{x}$ estimate, for *just a few steps* to determine whether the parameter combination improves the estimate. This intermediate optimization does not need to run to convergence in order to provide a gradient direction worthy of exploration. This is much faster than the usual approach of running a complete optimization for a number of parameter combinations, especially useful if the initial estimate is poor. An arbitrary 5% of pixels are used for validation, ignoring regions within a few pixels of edges, to avoid boundary complications, and because inputs are centred on the region of interest.

3.3 Optimization and Implementation Details

This problem closely resembles the well-studied problem of Bundle Adjustment [19], in that the camera parameters and image features are found simultaneously. Because most high-resolution pixels are observed in most frames, the super-resolution problem is closest to the "strongly convergent camera geometry" setup, and conjugate gradient methods are expected to converge rapidly [19]. Using the Scaled Conjugate Gradients (SCG) implementation from Netlab [13], rapid convergence is observed up to a point, beyond which a slow steady decrease in $\mathcal{F}$ gives no subjective improvement in the solution, but this can be avoided by specifying sensible convergence criteria.

The elements of $\mathbf{x}$ are scaled to lie in the range $[-\frac{1}{2}, \frac{1}{2}]$, and the geometric registration is decompose into a "fixed" component, which is the initial mapping from $\mathbf{y}^{(k)}$ to $\mathbf{x}$, and a projective correction term, which is itself decomposed into constituent shifts, rotations, axis scalings and projective parameters, which are the $\boldsymbol{\phi}$ parameters, then concatenated with $\boldsymbol{\lambda}$ to give one parameter vector. This is then "whitened" to be zero-mean and have a *std* of 0.35 units, which is approximately the standard deviation of $\mathbf{x}$. The prior over registration values suggested in section 2.1 is achieved simply by penalising large values in this registration vector.

Boundary conditions are treated as in [18], making the super-resolution image big enough so that the PSF kernel associated with any low-resolution pixel under any expected registration is adequately supported. Gradients with respect to $\mathbf{x}$ and $\boldsymbol{\lambda}$ can be found analytically, and those with respect to $\boldsymbol{\phi}$ are found numerically.

4 Experimental Results

The performance of simultaneous registration, super-resolution and prior updating, is evaluated using real data from a variety of sources. This is contrasted with a "registration-

fixing" approach, whereby registrations between the inputs are found then fixed before the super-resolution process. This fixed registration is also initialised as in section 3.1, then refined using an intensity-based scheme. Finally $\mathcal{F}$ is optimized *w.r.t.* $\mathbf{x}$ *only* to obtain a high-resolution estimate.

Experiments are first performed on synthetic data, generated using (1) applied to 256×256-pixel images at a zoom factor of 4. Values for $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$ are sampled randomly from a Gaussian with *std* of 0.25 units. For both algorithms, registrations for the set of images usually agree with each other to within a few tenths of a pixel in the high-resolution frame, but this variance is smaller in the simultaneous super-resolution method than for registrations found without using the fixed model.

Surrey Library Sequence: An area of interest is highlighted in the 30-frame Surrey Library sequence from `http://www.robots.ox.ac.uk/~vgg/data4.html`. The camera motion is a slow pan through a small angle, and the sign on a wall is illegible given any one of the inputs alone. Gaussian PSFs with *std* $= 0.375, 0.45, 0.525$ are selected, and used in both algorithms. There are 77003 elements in $\mathbf{y}$, and $\mathbf{x}$ has 45936 elements with a zoom factor of 4. $\mathbf{W}$ has around 3.5×10^9 elements, of which around 0.26% are nonzero with the smallest of these PSF kernels, and 0.49% with the largest. Most instances of the simultaneous algorithm converge in 2 to 5 iterations. Results appear in figure 3, showing that while both algorithms perform well with the middle PSF size, the simultaneous-registration algorithm handles the worse PSF estimates more gracefully.

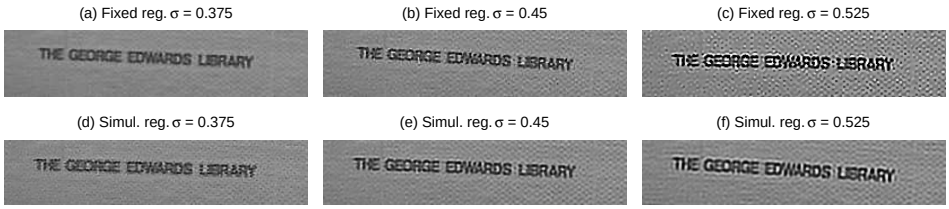

Figure 3: Surrey Library sequence. (a,b,c) Super-resolution found using fixed registrations. (d,e,f) Super-resolution images using our algorithm. One of the 30 low-resolution images can be seen in figure 1 (a).

Eye-test Card Sequence: The second experiment uses just 10 images of an eye-test card, captured using a webcam. The card is tilted and rotated slightly, and image brightness varies as the lighting and camera angles change. Gaussian PSFs with *std* $= 0.3, 0.375, 0.45$ are used in both super-resolution algorithms. The results appear in the left portion of figure 4.

Camera "9" Sequence: The model is adapted to handle DVD input, where the aspect ratio of the input images is 1.25:1, but they represent 1.85:1 video. The correction in the horizontal scaling is incorporated into the "fixed" part of the homography representation, and the PSF is assumed to be anisotropic. This avoids an undesirable interpolation of the inputs prior to super-resolving, which would lose high-frequency information, or working with squashed images throughout the process, which would violate the assumption of an isotropic prior on $\mathbf{x}$. The sequence consists of 29 I-frames[1] from the movie *Groundhog Day*. An on-screen hand-held TV camera moves independently of the real camera, and

[1] I-Frames are encoded as complete images, rather than requiring nearby frames in order to render them.

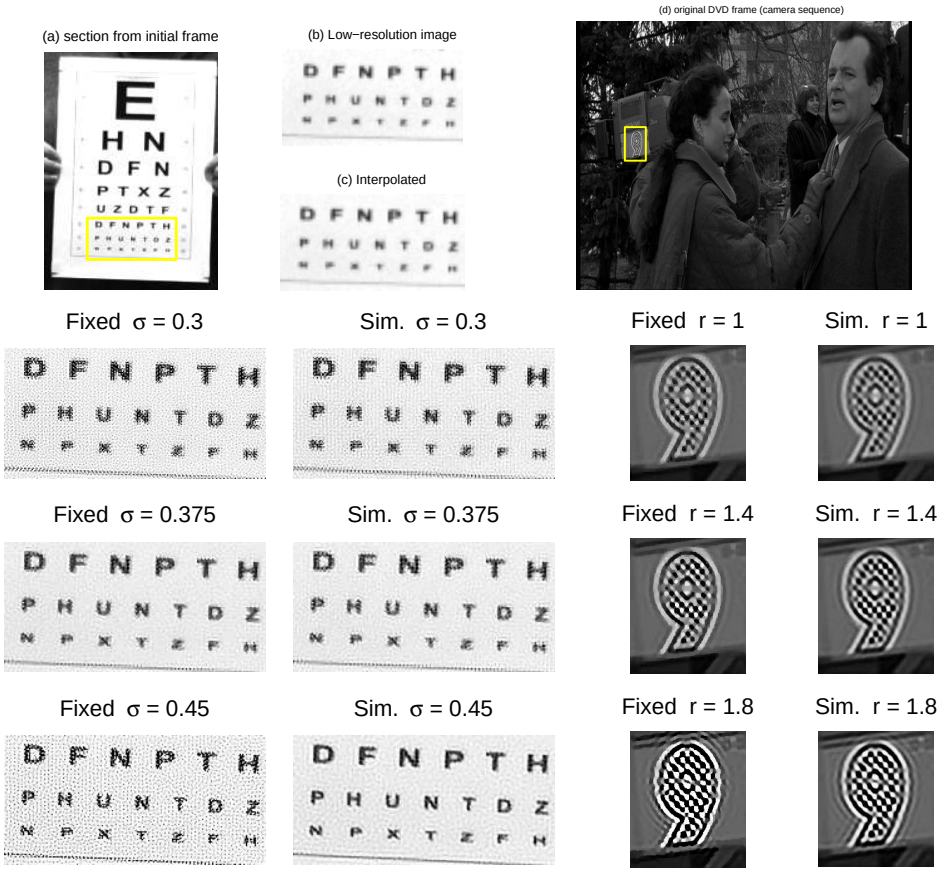

Figure 4: (a,b,c) Input and interest region (raw and interpolated) for 10-frame eye-test card sequence; (d) raw DVD frame for Camera "9" sequence (see figure 1 for interest region); Lower section, first and third columns: results obtained by fixing registration prior to super-resolution; Lower section, second and fourth columns: results obtained using the simultaneous approach.

the logo on the side is chosen as the interest region. Disk-shaped PSFs with radii of 1, 1.4, and 1.8 pixels are used. In both the eye-test card and camera "9" sequences, the simultaneously-optimized super-resolution images again appear subjectively better to the human viewer, and are more consistent across different PSFs.

Finally, a selection of results obtained from difficult DVD input sequences take from the movie *Lola Rennt* is show in figure 5. In the "cars" sequence, there are just 9 I-frames showing a pair of cars, and the areas of interest are the car number plates. The "badge" sequence shows the badge of a bank security officer. Seven I-frames are available, but all are dark, making the noise level proportionally very high. Significant improvements at a zoom factor of 4 (in each direction) can be seen.

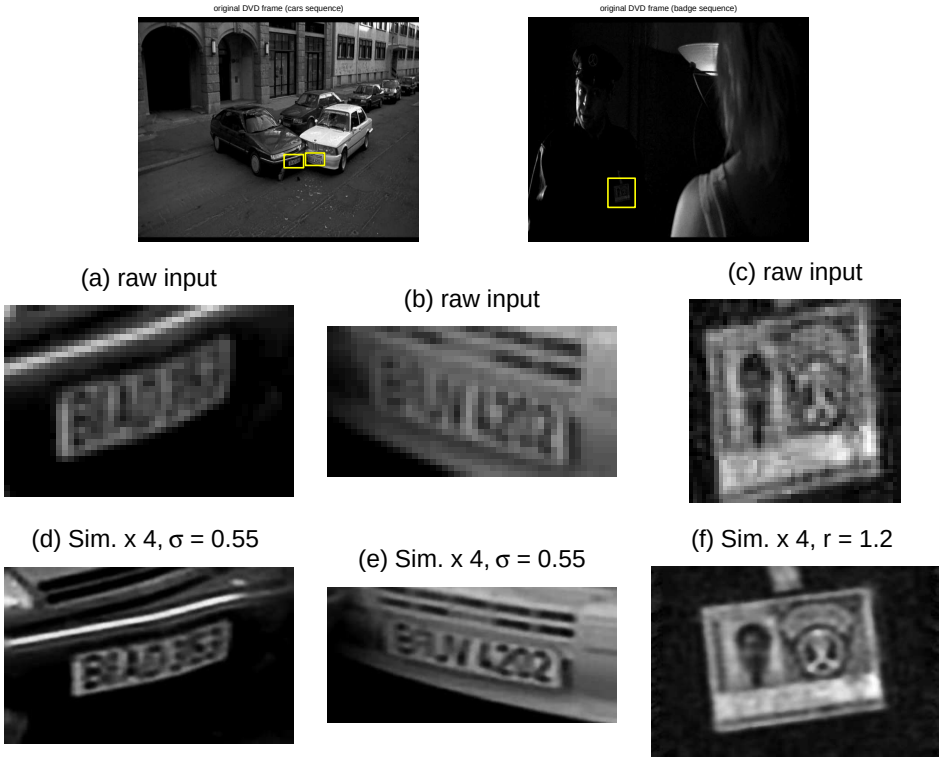

Figure 5: Results from the movie *Lola Rennt* on DVD. Top row: two raw DVD frames. Middle Row: the selected interest regions, shown at the DVD aspect ratio. Bottom Row: The same interest regions super-resolved using the simultaneous method. (a, d) black car's number plate, (b, e) white car's number plate, (c, f) security guard's ID badge (intensities have been scaled for ease of viewing).

5 Conclusion

A novel method for combining super-resolution with image registration and the learning of a Huber edge-preserving image prior has been presented. Results on real data from several sources show this approach to be superior to the practice of fixing the registration prior to the super-resolution process. Future work directions include methods for selecting the parametric family of point-spread function kernels, extending the model to handle nonplanar registrations, *e.g.* with the probabilistic optic flow framework [6], and incorporating a better model for the lossy DVD compression.

Acknowledgements

This work was funded in part by EC Network of Excellence PASCAL and by the EPSRC.

References

[1] Y. Altunbasak, A. Patti, and R. Mersereau. Super-resolution still and video reconstruction from mpeg-coded video. *IEEE Trans. Circuits And Syst. Video Technol.*, 12:217–226, 2002.

[2] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE PAMI*, 24(9):1167–1183, 2002.

[3] S. Borman. *Topics in Multiframe Superresolution Restoration*. PhD thesis, University of Notre Dame, Notre Dame, Indiana, May 2004.

[4] D. P. Capel. *Image Mosaicing and Super-resolution*. PhD thesis, University of Oxford, 2001.

[5] S. Farsiu, M. Elad, and P. Milanfar. A practical approach to super-resolution. In *Proc. of the SPIE: Visual Communications and Image Processing*, San-Jose, 2006.

[6] R. Fransens, C. Strecha, and L Van Gool. A probabilistic approach to optical flow based super-resolution. In *Proc. Workshop on Generative Model Based Vision*, 2004.

[7] R. C. Hardie, K. J. Barnard, and E. A. Armstrong. Joint map registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633, 1997.

[8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[9] M. Irani and S. Peleg. Super resolution from image sequences. *ICPR*, 2:115–120, Jun 1990.

[10] M. V. Joshi, S. Chaudhuri, and R. Panuganti. A learning-based method for image super-resolution from zoomed observations. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 35(3):527–537, 2005.

[11] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43—46, May 1996.

[12] R. L. Lagendijk, A. M. Tekalp, and J. Biemond. Maximum likelihood image and blur identification: A unifying approach. *Optical Engineering*, 29:422—435, 1990.

[13] I. Nabney. *Netlab algorithms for pattern recognition*. Springer, 2002.

[14] N. Nguyen, P. Milanfar, and G. Golub. Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. *IEEE Transactions on Image Processing*, 10(9):1299–1308, Sep 2001.

[15] L. C. Pickup, S. J. Roberts, and A. Zisserman. A sampled texture prior for image super-resolution. In *NIPS 16*, 2003.

[16] S. J. Reeves and R. M. Mersereau. Blur identification by the method of generalized cross-validation. *IEEE Transactions on Image Processing*, 1(3):301—311, 1992.

[17] Salient Stills. http://www.salientstills.com/.

[18] M. E. Tipping and C. M. Bishop. Bayesian image super-resolution. In *NIPS 15*, 2002.

[19] W. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle adjustment: A modern synthesis. In W. Triggs, A. Zisserman, and R. Szeliski, editors, *Vision Algorithms: Theory and Practice*, LNCS, pages 298–375. Springer Verlag, 2000.

[20] Q. Wang, X. Tang, and H. Shum. Patch based blind image super resolution. In *Proc. ICCV*, volume 1, pages 709—716, 2005.