Chapter 7

Super-resolution Using Sub-space Models

7.1 Introduction

In the previous chapter we demonstrated that generic MRF prior models can go a long way towards improving the performance of our super-resolution estimators. However, the generality of these priors is also their weakness, and in computing a super-resolution MAP estimate, there is a trade-off to be made between reduced noise and excessive smoothness.

In this chapter we introduce the use of compact image models that are tuned to particular classes of image. In certain cases, the range-space of these models may be learnt from training images. We consider in particular the case of super-resolution reconstruction of text and face images. In the first case, we show that a simple constrained model can give better results than the MAP-MRF estimators, whilst avoiding the need to impose spatial correlation. In the second case, we use a simple face model analogous to the “IdentiKit” method often used to compose images of police suspects. The model is composed from compact models of key facial features – the eyes, nose, mouth and cheek areas – which are learnt from training data using principal-components analysis. The power of this compact representation is demonstrated as we derive an ML estimator and two MAP estimators based on the model. In both cases, constraining the super-resolution estimate to lie in or near a low-dimensional, highly problem-specific sub-space greatly improves the conditioning of the problem. The MAP results typically exceed those possible with generic image priors.

The methods presented here are similar in spirit to those proposed by Baker and Kanade [10] who also examine the use of spatially varying priors for face images. The novelty over their approach is two-fold: first, the low resolution images need not be at the same resolution and indeed the resolution can vary across the image – this generalization is essential in the case that the low resolution images are related by a transformation more general than 2D pure translation; second, a generative model is used throughout.
In Section 7.2 we demonstrate the power of a simple constrained model applied to the reconstruction of text images. Section 7.3 describes the PCA-based face model, and Section 7.4 describes an ML estimator and two MAP estimators based on this model. In Section 7.5, the behaviour of the estimators is analysed using synthetic images. Finally, Section 7.6 presents results of the estimators applied to real image sequences.

7.2 Bound constraints

The motivation for this chapter comes from the success of a very simple constrained model applied to the super-resolution reconstruction of text image sequences. Although not properly explored in the super-resolution literature, it is widely recognized in the single-image restoration literature that placing hard constraints on the individual pixel intensities, which restrict the solution to some sub-space of the full image space, can give excellent results without the need for a spatial prior. Examples of such constraints are non-negativity and upper/lower bound constraints on pixel values.

As a motivating example, Figure 7.1 compares super-resolution estimates computed by both unconstrained and bound constrained ML estimators using synthetic images. The synthetic images are generated from the “Text” ground-truth images, using a $\sigma_{\text{psf}} = 0.7$ and down-sampling ratio $S = 3$. The super-resolution images are therefore constructed at $3 \times$ zoom. Three different levels of Gaussian noise is added to the synthetic images. Even for these tiny levels of noise, the unconstrained ML estimate is completely corrupted by reconstruction error. The bound-constrained estimate on the other hand is unaffected and of high-quality.

Figure 7.2 compares the performance of the bound constrained estimator with results obtained using MAP estimators. In this case the noise on the synthetic images is a much more realistic $\sigma = 5$ grey-levels. The constrained result is clearly superior to both MAP results. The lack of any imposed spatial correlation means that the constrained estimator is able to generate sharper results than the MAP estimators using spatial MRF priors.

Figure 7.3 compares ML, HMRF and bound-constrained reconstructions of a real image sequence – the “Czech” sequence seen in Section 5.13. The reconstruction is performed at $1.75 \times$ pixel zoom. The ML estimate is badly corrupted, whilst the bound-constrained estimate is comparable in quality to the MAP estimate using the first-derivative HMRF spatial prior.

Optimization The method used to compute the bound-constrained estimates is More and Toraldo’s “Gradient-projection conjugate gradient” (GPCG) algorithm [101]. This method is efficient for the solution of large-scale quadratic problems when the number of constraints is very large, as here. The number of iterations required is typically fewer than that required for convergence of the unconstrained ML estimator.