Appendix A

Notation

This is a brief guide to the notational conventions used in this text.

Scalars, vectors and matrices

We denote scalars by either small or capital letters a, A, α. We denote column vectors by bold small letters $\mathbf{a}, \boldsymbol{\phi}$. When we need a row vector we usually present this as the transpose of a column vector $\mathbf{a}^T, \boldsymbol{\phi}^T$.

We represent matrices by bold capital letters $\mathbf{B}, \boldsymbol{\Phi}$. The i^{th} row and j^{th} column of matrix $\mathbf{A}$ is written as a_{ij}. The j^{th} column of matrix $\mathbf{A}$ is written as $\mathbf{a}_j$. When we need to refer to the i^{th} row of a matrix, we write this as $\mathbf{a}_{i\bullet}$ where the bullet $\bullet$ indicates that we are considering all possible values of the column index.

We concatenate two $D \times 1$ column vectors horizontally as $\mathbf{a} = [\mathbf{b}, \mathbf{c}]$ to form the $D \times 2$ matrix $\mathbf{A}$. We concatenate two $D \times 1$ column vectors vertically as $\mathbf{a} = [\mathbf{b}^T, \mathbf{c}^T]^T$ to form the $2D \times 1$ vector $\mathbf{a}$. Although this notation is cumbersome, it allows us to represent vertical concatenations within a single line of text.

Variables and parameters

We denote variables with Roman letters $\mathbf{a}, \mathbf{b}$. The most common examples are the observed data which is always denoted by $\mathbf{x}$ and the state of the world which is always denoted by $\mathbf{w}$. However, other hidden or latent variables are also represented by Roman letters. We denote parameters of the model by Greek letters $\boldsymbol{\mu}, \boldsymbol{\Phi}, \sigma^2$. These are distinguished from variables in that there is usually a single set of parameters that explains the relation between many sets of variables.

Functions

We write functions as a name, followed by square brackets that contain the arguments of the function. For example, $\log[x]$ returns the logarithm of the scalar variable x. Sometimes we will write a function with bullets $\bullet$ as arguments (e.g., $\text{atan2}[\bullet, \bullet]$) to focus the interest on the function itself rather than the arguments.

When the function returns one or more vector or matrix arguments, it is written in bold. For example, the function $\mathbf{aff}[\mathbf{x}, \boldsymbol{\Phi}, \boldsymbol{\tau}]$ applies an affine transformation to the 2D point $\mathbf{x}$ with parameters $\boldsymbol{\Phi}, \boldsymbol{\tau}$ and returns a new 2D vector output. When a

function returns multiple outputs, we write this in Matlab notation so $[\mathbf{U}, \mathbf{L}, \mathbf{V}] = \mathbf{svd}[\mathbf{X}]$ returns the three parts $\mathbf{U}, \mathbf{L}, \mathbf{V}$ of the singular value decomposition of $\mathbf{X}$.

Some functions are used repeatedly throughout the text. These include:

- $\min_x f[x]$, which returns the minimum possible value of the function $f[x]$ as we vary x over its entire valid range,
- $\text{argmin}_x f[x]$, which returns the value of the argument x that minimizes $f[x]$,
- $\max_x$ and argmax_x, which fulfill the same roles as $\min_x$ and argmin_x but where we are maximizing the function,
- $\mathbf{diag}[\mathbf{A}]$, which returns a column vector containing the elements on the diagonal of matrix $\mathbf{A}$,
- $\delta[x]$ for continuous x, which is a Dirac delta function and has the key property $\int f[x]\delta[x - x_0]dx = f[x_0]$,
- $\delta[x]$ for discrete x, which returns 1 when the argument x is 0 and returns 0 otherwise, and
- heaviside$[x]$, which represents the Heaviside step function. It returns 0 when the argument $x < 0$ and returns 1 otherwise.

Probability distributions

We write the probability of a random variable x as $Pr(x)$. We write the joint probability of two variables a, b as $Pr(a, b)$ and the conditional probability of a given b as $Pr(a|b)$. Sometimes, we wish to specify the exact value b^* that a is conditioned upon and here we write $Pr(a|b = b^*)$. Occasionally, we denote that variables a and b are independent by writing $a \perp\!\!\!\perp b$. Similarly we indicate that a and b are conditionally independent given c by writing $a \perp\!\!\!\perp b|c$.

Probability distributions are written in the style $Pr(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \text{Norm}_{\mathbf{x}}[\boldsymbol{\mu}, \boldsymbol{\Sigma}]$. This returns the value of the multivariate normal distribution for data $\mathbf{x}$ when the distribution has mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. In this way, we always distinguish the argument of the distribution (here $\mathbf{x}$) from the parameters (here $\boldsymbol{\mu}, \boldsymbol{\Sigma}$).

Sets

We denote sets with calligraphic letters $\mathcal{S}$. The notation $\mathcal{S} \subset \mathcal{T}$ indicates that $\mathcal{S}$ is a subset of $\mathcal{T}$. The notation $x \in \mathcal{S}$ indicates that x is a member of the set $\mathcal{S}$. The notation $\mathcal{A} = \mathcal{B} \cup \mathcal{C}$ indicates that set $\mathcal{A}$ is the union of sets $\mathcal{B}$ and $\mathcal{C}$. The notation $\mathcal{A} = \mathcal{B} \setminus \mathcal{C}$ indicates that set $\mathcal{A}$ consists of all of the elements of $\mathcal{B}$ except those that are in $\mathcal{C}$.

Often we write out a set explicitly in terms of the elements and for this we use curly brackets so that $\mathcal{A} = \{x, y, z\}$ indicates that the set $\mathcal{A}$ contains x, y, and z and nothing else. When a set is empty, we write $\mathcal{A} = \{\}$. We use the notation $\{x_i\}_{i=1}^I$ as shorthand to represent the set $\{x_1, x_2, \ldots, x_I\}$, and we may write the same set in the compact form $x_{1\ldots I}$ if it is part of an equation.

Appendix B

Optimization

Throughout this book, we have used iterative nonlinear optimization methods to find the maximum likelihood or MAP parameter estimates. We now provide more details about these methods. It is impossible to do full justice to this topic in the space available; many entire books have been written about nonlinear optimization. Our goal is merely to provide a brief introduction to the main ideas.

B.1 Problem statement

Continuous nonlinear optimization techniques aim to find the set of parameters $\hat{\boldsymbol{\theta}}$ that minimize a function $f[\bullet]$. In other words, they try to compute,

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \left[f[\boldsymbol{\theta}] \right], \tag{B.1}$$

where $f[\bullet]$ is termed a *cost function* or *objective function*.

Although optimization techniques are usually described in terms of minimizing a function, most optimization problems in this book involve *maximizing* an objective function based on log probability. To turn a maximization problem into a minimization we multiply the objective function by minus one. In other words, instead of maximizing the log probability, we minimize the negative log probability.

B.1.1 Convexity

The optimization techniques that we consider here are iterative: they start with an estimate $\boldsymbol{\theta}^{[0]}$ and improve it by finding successive new estimates $\boldsymbol{\theta}^{[1]}, \boldsymbol{\theta}^{[2]}, \ldots, \boldsymbol{\theta}^{[\infty]}$, each of which is better than the last, until no more improvement can be made. The techniques are purely local in the sense that the decision about where to move next is based on only the properties of the function at the current position. Consequently, these techniques cannot guarantee the correct solution: they may find an estimate $\boldsymbol{\theta}^{[\infty]}$ from which no local change improves the cost. However, this

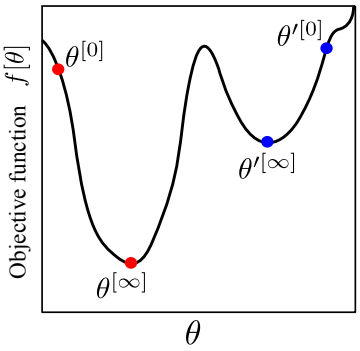

Figure B.1 Local minima. Optimization methods aim to find the minimum of the objective function $f[\theta]$ with respect to parameters θ. Roughly, they work by starting with an initial estimate $\theta^{[0]}$ and moving iteratively downhill until no more progress can be made (final position represented by $\theta^{[\infty]}$). Unfortunately, it is possible to terminate in a local minimum. For example, if we start at $\theta'^{[0]}$ and move downhill, we wind up in position $\theta'^{[\infty]}$.

does not mean there is not a better solution in some distant part of the function that has not yet been explored (figure B.1). In optimization parlance, they can only find *local minima*. One way to mitigate this problem is to start the optimization from a number of different places and choose the final solution with the lowest cost.

In the special case where the function is *convex*, there will only be a single minimum, and we are guaranteed to find it with sufficient iterations (figure B.2). For a 1D function, it is possible to establish the convexity by looking at the second derivative of the function; if this is positive everywhere (i.e., the slope is continuously increasing) then the function is convex and the global minimum can be found. The equivalent test in higher dimensions is to examine the *Hessian matrix* (the matrix of second derivatives of the cost function with respect to the parameters). If this is positive definite everywhere (see appendix C.2.6), then the function is convex and the global minimum will be found. Some of the cost functions in this book are convex, but this is unusual; most optimization problems found in vision do not have this convenient property.

B.1.2 Overview of approach

In general the parameters $\boldsymbol{\theta}$ over which we search are multidimensional. For example, when $\boldsymbol{\theta}$ has two dimensions, we can think of the function as a two dimensional surface (figure B.3). With this in mind, the principles behind the methods we will discuss are simple. We alternately

- choose a search direction $\mathbf{s}$ based on the local properties of the function, and
- search to find the minimum along the chosen direction. In other words, we seek the distance λ to move such that

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmin}} \left[f[\boldsymbol{\theta}^{[t]} + \lambda \mathbf{s}] \right], \tag{B.2}$$

  and then set $\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} + \hat{\lambda}\mathbf{s}$. This is termed a *line search*.

We now consider each of these stages in turn.

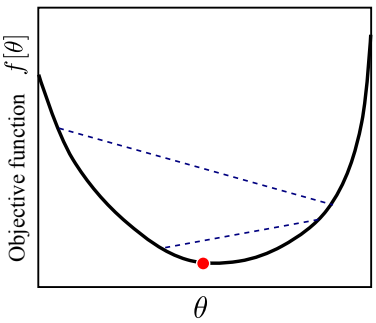

Figure B.2 Convex functions. If the function is convex, then the global minimum can be found. A function is convex if no chord (line between two points on the function) intersects the function. The figure shows two example chords (blue dashed lines). The convexity of a function can be established algebraically by considering the matrix of second derivatives. If this is positive definite for all values of $\boldsymbol{\theta}$, then the function is convex.

B.2 Choosing a search direction

We will describe two general methods for choosing a search direction (*steepest descent* and *Newton's method*) and one method which is specialized for least squares problems (the *Gauss-Newton method*). All of these methods rely on computing derivatives of the function with respect to the parameters at the current position. To this end, we are relying on the function being smooth so that the derivatives are well behaved.

For most models, it is easy to find a closed form expression for the derivatives. If this is not the case, then an alternative is to approximate them using finite differences. For example, the first derivative of $f[\bullet]$ with respect to the j^{th} element of $\boldsymbol{\theta}$ can be approximated by

$$\frac{\partial f}{\partial \theta_j} \approx \frac{f\left[\boldsymbol{\theta} + a\mathbf{e}_j\right] - f\left[\boldsymbol{\theta}\right]}{a}, \tag{B.3}$$

where a is a small number and $\mathbf{e}_j$ is the unit vector in the j^{th} direction. In principle as a tends to zero, this estimate becomes more accurate. However, in practice the calculation is limited by the floating point precision of the computer, so a must be chosen with care.

B.2.1 Steepest descent

An intuitive way to choose the search direction is to measure the gradient and select the direction which moves us downhill fastest. We could move in this direction until the function no longer decreases, then recompute the steepest direction and move again. In this way, we gradually move toward a local minimum of the function (figure B.3a). The algorithm terminates when the gradient is zero and the second derivative is positive, indicating that we are at the minimum point and any further local changes would not result in further improvement. This approach is termed *steepest descent*. More precisely, we choose

$$\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} - \lambda \left.\frac{\partial f}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}^{[t]}}, \tag{B.4}$$

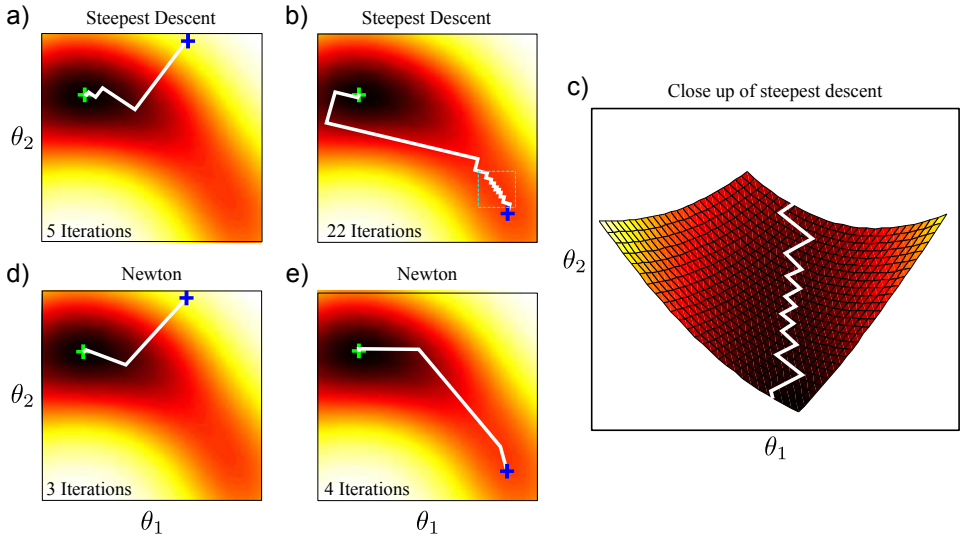

Figure B.3 Optimization on a two dimensional function (color represents height of function). We wish to find the parameters that minimize the function (green cross). Given an initial starting point $\boldsymbol{\theta}^0$ (blue cross), we choose a direction and then perform a local search to find the optimal point in that direction. a) One way to chose the direction is steepest descent: at each iteration, we head in the direction where the function changes the fastest. b) When we initialize from a different position, the steepest descent method takes many iterations to converge due to oscillatory behavior. c) Close-up of oscillatory region (see main text). d) Setting the direction using Newton's method results in faster convergence. e) Newton's method does not undergo oscillatory behavior when we initialize from the second position.

where the derivative $\partial f/\partial\boldsymbol{\theta}$ is the *gradient vector*, which points uphill, and λ is the distance moved downhill in the opposite direction $-\partial f/\partial\boldsymbol{\theta}$. The line search procedure (section B.3) selects the value of λ.

Steepest descent sounds like a good idea but can be very inefficient in certain situations (figure B.3b). For example, in a descending valley, it can oscillate ineffectually from one side to the other rather than proceeding straight down the center: the method approaches the bottom of the valley from one side, but overshoots because the valley itself is descending, so the minimum along the search direction is not exactly in the valley center (figure B.3c). When we re-measure the gradient and perform a second line search, we overshoot in the other direction. This is not an unusual situation: it is guaranteed that the gradient at the new point will be perpendicular to the previous one, so the only way to avoid this oscillation is to hit the valley at exactly right angles.

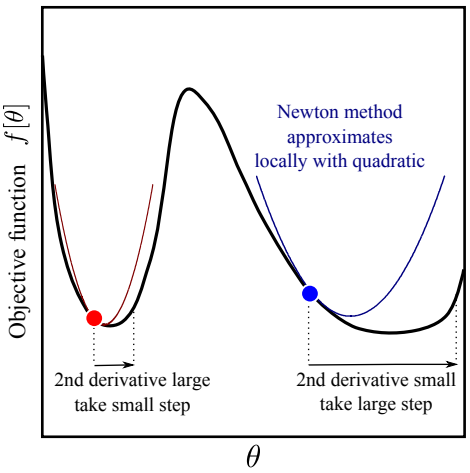

Figure B.4 Use of second derivatives. The gradient at the red and blue points is the same, but the magnitude of the second derivative is larger at the red point than the blue point: the gradient is changing faster at the red point than the blue point. The distance we move should be moderated by the second derivative: if the gradient is changing fast, then the minimum may be nearby and we should move a small distance. If it is changing slowly, then it is safe to move further. Newton's method takes into account the second derivative: it uses a Taylor expansion to create a quadratic approximation to the function and then moves toward the minimum.

B.2.2 Newton's method

Newton's method is an improved approach that also exploits the second derivatives at the current point: it considers both the gradient of the function and how that gradient is changing.

To motivate the use of second derivatives, consider a one-dimensional function (figure B.4). If the magnitude of the second derivative is low, then the gradient is changing slowly. Consequently, it will probably take a while before it completely flattens out and becomes a minimum, and so it is safe to move a long distance. Conversely, if the magnitude of the second derivative is high, then things are changing rapidly, and we should move only a small distance.

Now consider the same argument in two dimensions. Imagine we are at a point where the gradient is identical in both dimensions. For steepest descent, we would move equally in both dimensions. However, if the magnitude of the second derivative in the first direction is much greater than that in the second, we would nonetheless wish to move further in the second direction.

To see how to exploit the second derivatives algebraically, consider a truncated Taylor expansion around the current estimate $\boldsymbol{\theta}^{[t]}$:

$$f[\boldsymbol{\theta}] \approx f[\boldsymbol{\theta}^{[t]}] + (\boldsymbol{\theta}-\boldsymbol{\theta}^{[t]})^T \left.\frac{\partial f}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}^{[t]}} + \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}^{[t]})^T \left.\frac{\partial^2 f}{\partial \boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}^{[t]}} (\boldsymbol{\theta}-\boldsymbol{\theta}^{[t]}), \tag{B.5}$$

where $\boldsymbol{\theta}$ is a $D \times 1$ variable, the first derivative vector is of size $D \times 1$, and the Hessian matrix of second derivatives is $D \times D$. To find the local extrema, we now take derivatives with respect to $\boldsymbol{\theta}$ and set the result to zero

$$\frac{\partial f}{\partial \boldsymbol{\theta}} \approx \left.\frac{\partial f}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}^{[t]}} + \left.\frac{\partial^2 f}{\partial \boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}^{[t]}} (\boldsymbol{\theta}-\boldsymbol{\theta}^{[t]}) = 0. \tag{B.6}$$

By re-arranging this equation, we get an expression for the minimum $\hat{\boldsymbol{\theta}}$,

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{[t]} - \left(\frac{\partial^2 f}{\partial \boldsymbol{\theta}^2}\right)^{-1} \frac{\partial f}{\partial \boldsymbol{\theta}}, \tag{B.7}$$

where the derivatives are still taken at $\boldsymbol{\theta}^{[t]}$, but we have stopped writing this for clarity. In practice we would implement Newton's method as a series of iterations

$$\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} - \lambda \left(\frac{\partial^2 f}{\partial \boldsymbol{\theta}^2}\right)^{-1} \frac{\partial f}{\partial \boldsymbol{\theta}}, \tag{B.8}$$

where the λ is the step size. This can be set to one, or we can find the optimal value using line search.

One interpretation of Newton's method is that we have locally approximated the function as a quadratic. On each iteration, we move toward its extremum (or move exactly to it if we fix $\lambda = 1$). Note that we are assuming that we are close enough to the correct solution that the nearby extremum *is* a minimum and not a saddle point or maximum. In particular, if the Hessian is not positive definite then a direction that is not downhill may be chosen. In this sense Newton's method is not as robust as steepest descent.

Subject to this limitation, Newton's method converges in fewer iterations than steepest descent (figure B.3d-e). However, it requires more computation per iteration as we have to invert the $D \times D$ Hessian matrix at each step. Choosing this method usually implies that we can write the Hessian in closed form; approximating the Hessian from finite derivatives requires many function evaluations and so is potentially very costly.

B.2.3 Gauss-Newton method

Cost functions in computer vision often take the special form of a least squares problem

$$f[\boldsymbol{\theta}] = \sum_{i=1}^{I} (\mathbf{x}_i - \mathbf{g}[\mathbf{w}_i, \boldsymbol{\theta}])^T (\mathbf{x}_i - \mathbf{g}[\mathbf{w}_i, \boldsymbol{\theta}]), \tag{B.9}$$

where $\mathbf{g}[\bullet, \bullet]$ is a function that transfers the variables $\{\mathbf{w}_i\}$ into the space of the variables $\{\mathbf{x}_i\}$, and is parameterized by $\boldsymbol{\theta}$. In other words, we seek the values of $\boldsymbol{\theta}$ that most closely map $\{\mathbf{w}_i\}$ to $\{\mathbf{x}_i\}$ in a least squares sense. This cost function is a special case of the more general form $f[\boldsymbol{\theta}] = \mathbf{z}^T\mathbf{z}$, where

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}_1 - \mathbf{g}[\mathbf{w}_1, \boldsymbol{\theta}] \\ \mathbf{x}_2 - \mathbf{g}[\mathbf{w}_2, \boldsymbol{\theta}]| \\ \vdots \\ \mathbf{x}_I - \mathbf{g}[\mathbf{w}_I, \boldsymbol{\theta}] \end{bmatrix}. \tag{B.10}$$

The *Gauss-Newton method* is an optimization technique that is used to solve least squares problems of the form

$$\hat{\boldsymbol{\theta}} = \text{argmin}\,[f[\boldsymbol{\theta}]] \qquad \text{where } f[\boldsymbol{\theta}] = \mathbf{z}[\boldsymbol{\theta}]^T \mathbf{z}[\boldsymbol{\theta}]. \tag{B.11}$$

To minimize this objective function, we approximate the term $\mathbf{z}[\boldsymbol{\theta}]$ with a Taylor series expansion around the current estimate $\boldsymbol{\theta}^{[t]}$ of the parameters:

$$\mathbf{z}[\boldsymbol{\theta}] \approx \mathbf{z}[\boldsymbol{\theta}^{[t]}] + \mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}), \tag{B.12}$$

where $\mathbf{J}$ is the Jacobian matrix. The entry j_{mn} at the m^{th} row and the n^{th} column of $\mathbf{J}$ contains the derivative of the m^{th} element of $\mathbf{z}$ with respect to the n^{th} parameter so that

$$j_{mn} = \frac{\partial z_m}{\partial \theta_n}. \tag{B.13}$$

Now we substitute the approximation for $\mathbf{z}[\boldsymbol{\theta}]$ into the original cost function $f[\boldsymbol{\theta}] = \mathbf{z}^T\mathbf{z}$ to yield

$$\begin{aligned} f[\boldsymbol{\theta}] &\approx (\mathbf{z}[\boldsymbol{\theta}^{[t]}] + \mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}))^T(\mathbf{z}[\boldsymbol{\theta}^{[t]}] + \mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]})) \\ &= \mathbf{z}[\boldsymbol{\theta}^{[t]}]^T\mathbf{z}[\boldsymbol{\theta}^{[t]}] + 2(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]})^T\mathbf{J}^T\mathbf{z}[\boldsymbol{\theta}^{[t]}] + (\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]})^T\mathbf{J}^T\mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}). \end{aligned} \tag{B.14}$$

Finally, we take derivatives of this expression with respect to the parameters $\boldsymbol{\theta}$ and equate to zero to get the relation

$$\frac{\partial f}{\partial \boldsymbol{\theta}} \approx 2\mathbf{J}^T\mathbf{z}[\boldsymbol{\theta}^{[t]}] + 2\mathbf{J}^T\mathbf{J}(\boldsymbol{\theta} - \boldsymbol{\theta}^{[t]}) = 0. \tag{B.15}$$

Re-arranging, we get the update rule:

$$\boldsymbol{\theta} = \boldsymbol{\theta}^{[t]} - (\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T\mathbf{z}[\boldsymbol{\theta}^{[t]}]. \tag{B.16}$$

We can rewrite this by noting that

$$\left.\frac{\partial f}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}^{[t]}} = \left.\frac{\partial \mathbf{z}^T\mathbf{z}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}^{[t]}} = 2\mathbf{J}^T\mathbf{z}[\boldsymbol{\theta}^{[t]}], \tag{B.17}$$

to give the final Gauss-Newton update

$$\boldsymbol{\theta}^{[t+1]} = \boldsymbol{\theta}^{[t]} - \lambda(\mathbf{J}^T\mathbf{J})^{-1}\frac{\partial f}{\partial \boldsymbol{\theta}}, \tag{B.18}$$

where the derivative is taken at $\boldsymbol{\theta}^{[t]}$ and λ is the step size.

Comparing with the Newton update (equation B.8), we see that we can consider this update as approximating the Hessian matrix as $\mathbf{H} \approx \mathbf{J}^T\mathbf{J}$. It provides better results than gradient descent without ever computing second derivatives. Moreover, the term $\mathbf{J}^T\mathbf{J}$ is normally positive definite resulting in increased stability.

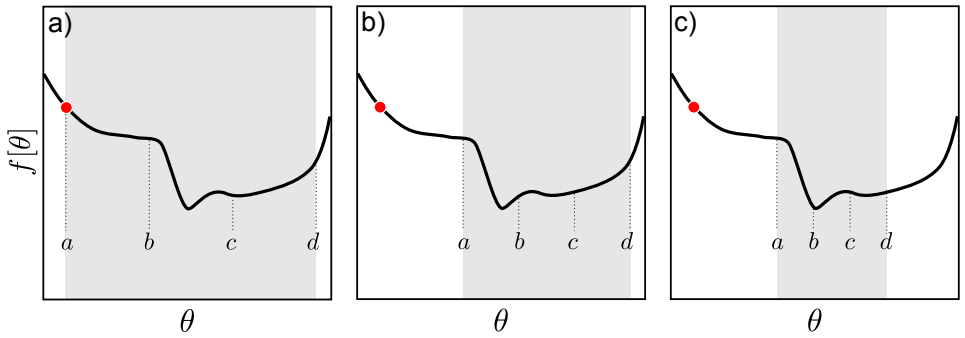

Figure B.5 Line search over region $[a, d]$ using bracketing approach. Gray region indicates current search region. a) We define two points b, c that are interior to the search region and evaluate the function at these points. Here $f[b] > f[c]$ so we eliminate the range $[a, b]$. b) We evaluate two points $[b, c]$ interior to the new range and compare their values. This time we find that $f[b] < f[c]$ so we eliminate the range $[c, d]$. c) We repeat this process until the minimum is closely bracketed.

B.2.4 Other methods

There are numerous other methods for choosing the optimization direction. Many of these involve approximating the Hessian in some way with the goal of either ensuring that a downhill direction is always chosen or reducing the computational burden. For example, if computation of the Hessian is prohibitive, a practical approach is to approximate it with its own diagonal. This usually provides a better direction than steepest descent.

Quasi-Newton methods such as the *Broyden Fletcher Goldfarb Shanno (BFGS) method* approximate the Hessian with information gathered by analyzing successive gradient vectors. The *Levenberg-Marquardt* algorithm interpolates between the Gauss-Newton algorithm and steepest descent with the aim of producing a method that requires few iterations and is also robust. *Damped Newton* and *trust-region methods* also attempt to improve the robustness of Newton's method. The nonlinear *conjugate gradient algorithm* is another valuable method when only first derivatives are available.

B.3 Line search

Having chosen a sensible direction using steepest descent, Newton's method or some other approach, we must now decide how far to move: we need an efficient method to find the minimum of the function in the chosen direction. Line search methods start by determining the range over which to search for the minimum.

This is usually guided by the magnitude of the second derivative along the line, which provides information about the likely search range (see figure B.4).

There are many heuristics to find the minimum, but we will discuss only the direct search method (figure B.5). Consider searching over the region $[a, d]$. We compute the function at two internal points b and c where $a < b < c < d$. If $f[b] < f[c]$, we eliminate the range $[c, d]$ and search over the new region $[a, c]$ at the next iteration. Conversely, if $f[b] > f[c]$, we eliminate the range $[a, b]$ and search over a new region $[b, d]$.

This method is applied iteratively until the minimum is closely bracketed. It is typically not worth exactly locating the minimum; the line search direction is rarely optimal and so the minimum of the line search is usually far from the overall minimum of the function. Once the remaining interval is sufficiently small, an estimate of the minimum position can be computed by making a parabolic fit to the three points that remain after eliminating one region or the other and selecting the position of the minimum of this parabola.

B.4 Reparameterization

Often in vision problems, we must find the best parameters $\boldsymbol{\theta}$ subject to one or more constraints. Typical examples include optimizing variances σ^2 which must be positive, covariance matrices, which must be positive definite, and matrices that represent geometric rotations, which must be orthogonal. The general topic of constrained optimization is beyond the scope of this volume, but we briefly describe a trick that can be used to convert constrained optimization problems into unconstrained ones, which can be solved using the techniques already described.

The idea of reparameterization is to represent the parameters $\boldsymbol{\theta}$ in terms of a new set of parameters $\boldsymbol{\phi}$, which do not have any constraints on them, so that

$$\boldsymbol{\theta} = \mathbf{g}[\boldsymbol{\phi}], \tag{B.19}$$

where $\mathbf{g}[\bullet]$ is a carefully chosen function.

Then we optimize with respect to the new unconstrained parameters $\boldsymbol{\phi}$. The objective function becomes $f[\mathbf{g}[\boldsymbol{\phi}]]$ and the derivatives are computed using the chain rule so that the first derivative would be

$$\frac{\partial f}{\partial \boldsymbol{\phi}} = \sum_{k=1}^{K} \frac{\partial f}{\partial \theta_k} \frac{\partial \theta_k}{\partial \boldsymbol{\phi}}. \tag{B.20}$$

where θ_k is the k^{th} element of $\boldsymbol{\theta}$. This strategy is easier to understand with some concrete examples.

Parameters that must be positive

When we optimize a variance parameter $\theta = \sigma^2$ we must ensure that the final answer is positive. To this end, we use the relation:

$$\theta = \exp[\phi], \tag{B.21}$$

and now optimize with respect to the new scalar parameter ϕ. Alternatively, we can use the square relation:

$$\theta = \phi^2, \tag{B.22}$$

and again optimize with respect to the parameter ϕ.

Parameters that must lie between 0 and 1

To ensure that a scalar parameter θ lies between zero and one, we use the logistic sigmoid function:

$$\theta = \frac{1}{1+\exp[-\phi]}, \tag{B.23}$$

and optimize with respect to the new scalar parameter ϕ.

Parameters that must be positive and sum to one

To ensure that the elements of a $K \times 1$ multivariable parameter $\boldsymbol{\theta}$ sum to one and are all positive we use the softmax function:

$$\theta_k = \frac{\exp[\phi_k]}{\sum_{j=1}^{K}\exp[\phi_j]}, \tag{B.24}$$

and optimize with respect to the new $K \times 1$ variable $\boldsymbol{\phi}$.

3D rotation matrices

A 3×3 rotation matrix contains three independent quantities spread throughout its nine entries. A number of nonlinear constraints exist between the entries: the norm of each column and row must be one, each column is perpendicular to the other columns, each row is perpendicular to the other rows, and the determinant is one.

One way to enforce these constraints is to re-parameterize the rotation matrix as a *quaternion* and optimize with respect to this new representation. A quaternion $\mathbf{q}$ is a 4D quantity $\mathbf{q} = [q_0, q_1, q_2, q_3]$. Mathematically speaking, they are a four dimensional extension of complex numbers, but the relevance for vision is that they can be used to represent 3D rotations. We use the relation:

$$\boldsymbol{\Theta} = \frac{1}{q_0^2+q_1^2+q_2^2+q_3^2}\begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2q_1q_2-2q_0q_3 & 2q_1q_3+2q_0q_2 \\ 2q_1q_2+2q_0q_3 & q_0^2-q_1^2+q_2^2-q_3^2 & 2q_2q_3-2q_0q_1 \\ 2q_1q_3-2q_0q_2 & 2q_2q_3+2q_0q_1 & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix}. \tag{B.25}$$

Although the quaternion contains four numbers, only the ratios of those numbers are important (giving 3 degrees of freedom): each element of equation B.25

consists of squared terms, which are normalized by the squared amplitude constant, and so any constant that multiplies the elements of $\mathbf{q}$ is canceled out when we convert back to a rotation matrix.

Now we optimize with respect to the quaternion $\mathbf{q}$. The derivatives with respect to the k^{th} element of $\mathbf{q}$ can be computed as

$$\frac{\partial f}{\partial q_k} = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\partial f}{\partial \boldsymbol{\Theta}_{ij}} \frac{\partial \boldsymbol{\Theta}_{ij}}{\partial q_k}. \tag{B.26}$$

The quaternion optimization is stable as long as we do not approach the singularity at $\mathbf{q} = \mathbf{0}$. One way to achieve this is to periodically re-normalize the quaternion to length 1 during the optimization procedure.

Positive definite matrices

When we optimize over a $K \times K$ covariance matrix $\boldsymbol{\Theta} = \boldsymbol{\Sigma}$, we must ensure that the result is positive definite. A simple way to do this is to use the relation:

$$\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Phi}^T, \tag{B.27}$$

where $\boldsymbol{\Phi}$ is an arbitrary $K \times K$ matrix.

Appendix C

Linear algebra

C.1 Vectors

A vector is a geometric entity in D dimensional space that has both a direction and a magnitude. It is represented by a $D \times 1$ array of numbers. In this book, we write vectors as bold, small, Roman or Greek letters (e.g., $\mathbf{a}, \boldsymbol{\phi}$). The transpose $\mathbf{a}^T$ of vector $\mathbf{a}$ is a $1 \times D$ array of numbers where the order of the numbers is retained.

C.1.1 Dot product

The *dot product* or *scalar product* between two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$c = \mathbf{a}^T\mathbf{b} = \sum_{d=1}^{D} a_d b_d, \tag{C.1}$$

where $\mathbf{a}^T$ is the transpose of $\mathbf{a}$ (i.e., $\mathbf{a}$ converted to a row vector) and the returned value c is a scalar. Two vectors are said to be *orthogonal* if the dot product between them is zero.

C.1.2 Norm of a vector

The *magnitude* or *norm* of a vector is the square root of the sum of the square of the D elements so that,

$$\text{norm}[\mathbf{a}] = |\mathbf{a}| = \left(\sum_{d=1}^{D} a_d^2\right)^{1/2} = (\mathbf{a}^T\mathbf{a})^{1/2}. \tag{C.2}$$

C.1.3 Cross product

The *cross product* or *vector product* is specialized to three dimensions. The operation $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ is equivalent to the matrix multiplication (section C.2.1):

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{C.3}$$

or for short

$$\mathbf{c} = \mathbf{A}_{\times}\mathbf{b}, \tag{C.4}$$

where $\mathbf{A}_{\times}$ is the 3×3 matrix from equation C.3 that implements the cross product.

It is easily shown that the result $\mathbf{c}$ of the cross product is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. In other words,

$$\mathbf{a}^T(\mathbf{a} \times \mathbf{b}) = \mathbf{b}^T(\mathbf{a} \times \mathbf{b}) = 0. \tag{C.5}$$

C.2 Matrices

Matrices are used extensively throughout the book and are written as bold, capital, Roman or Greek letters (e.g., $\mathbf{A}$,$\mathbf{\Phi}$). We categorize matrices as *landscape* (more columns than rows), *square* (the same number of columns and rows) or *portrait* (more rows than columns). They are always indexed by row first and then column, so a_{ij} denotes the element of matrix $\mathbf{A}$ at the i^{th} row and the j^{th} column.

A *diagonal* matrix is a square matrix with zeros everywhere except on the diagonal (i.e., elements a_{ii}) where the elements may take any value. An important special case of a diagonal matrix is the *identity* matrix $\mathbf{I}$. This has zeros everywhere except for the diagonal, where all the elements are ones.

C.2.1 Matrix multiplication

To take the matrix product $\mathbf{C} = \mathbf{AB}$, we compute the elements of $\mathbf{C}$ as

$$c_{ij} = \sum_{k=1}^{K} a_{ik}b_{kj}. \tag{C.6}$$

This can only be done when the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$. Matrix multiplication is associative so that $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{ABC}$. However it is not commutative so that in general $\mathbf{AB} \neq \mathbf{BA}$.

C.2.2 Transpose

The transpose of a matrix $\mathbf{A}$ is written as $\mathbf{A}^T$ and is formed by reflecting it around the principal diagonal, so that the k^{th} column becomes the k^{th} row and vice-versa. If we take the transpose of a matrix product $\mathbf{AB}$, then we take the transpose of the original matrices but reverse the order so that

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T. \tag{C.7}$$

C.2.3 Inverse

A square matrix $\mathbf{A}$ may or may not have an inverse $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. If a matrix does not have an inverse, it is called *singular*.

Diagonal matrices are particularly easy to invert: the inverse is also a diagonal matrix, with each diagonal value d_{ii} replaced by $1/d_{ii}$. Hence, any diagonal matrix that has non-zero values on the diagonal is invertible. It follows that the inverse of the identity matrix is the identity matrix itself.

If we take the inverse of a matrix product $\mathbf{AB}$, then we can equivalently take the inverse of each matrix individually, and reverse the order of multiplication

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{C.8}$$

C.2.4 Determinant and trace

Every square matrix $\mathbf{A}$ has a scalar associated with it called the determinant, denoted by $|\mathbf{A}|$ or $\det[\mathbf{A}]$. It is (loosely) related to the scaling applied by the matrix. As a rule of thumb, matrices where the magnitude of the determinant is small tend to make vectors smaller upon multiplication and matrices where the magnitude of the determinant is large tend to make them larger. If a matrix is singular, the determinant will be zero and there will be at least one direction in space that is mapped to the origin when the matrix is applied. For a diagonal matrix, the determinant is the product of the diagonal values. It follows that the determinant of the identity matrix is 1. Determinants of matrix expressions can be computed using the following rules:

$$\begin{aligned} |\mathbf{A}^T| &= |\mathbf{A}| && \text{(C.9)} \\ |\mathbf{AB}| &= |\mathbf{A}||\mathbf{B}| && \text{(C.10)} \\ |\mathbf{A}^{-1}| &= 1/|\mathbf{A}|. && \text{(C.11)} \end{aligned}$$

The trace of a matrix is a second number associated with a square matrix $\mathbf{A}$. It is the sum of the diagonal values (the matrix itself need not be diagonal). The traces of compound terms are bound by the following rules:

$$\begin{aligned}
\mathrm{tr}[\mathbf{A}^{\mathbf{T}}] &= \mathrm{tr}[\mathbf{A}] && \text{(C.12)}\\
\mathrm{tr}[\mathbf{AB}] &= \mathrm{tr}[\mathbf{BA}] && \text{(C.13)}\\
\mathrm{tr}[\mathbf{A}+\mathbf{B}] &= \mathrm{tr}[\mathbf{A}]+\mathrm{tr}[\mathbf{B}] && \text{(C.14)}\\
\mathrm{tr}[\mathbf{ABC}] &= \mathrm{tr}[\mathbf{BCA}] = \mathrm{tr}[\mathbf{CAB}], && \text{(C.15)}
\end{aligned}$$

where in the last relation, the trace is invariant for cyclic permutations only, so that in general $\mathrm{tr}[\mathbf{ABC}] \neq \mathrm{tr}[\mathbf{BAC}]$.

C.2.5 Orthogonal and rotation matrices

An important class of square matrix is the orthogonal matrix. Orthogonal matrices have the following special properties:

1. Each column has norm one, and each row has norm one.
2. Each column is orthogonal to every other column, and each row is orthogonal to every other row.

The inverse of an orthogonal matrix $\mathbf{\Omega}$ is its own transpose, so $\mathbf{\Omega}^T\mathbf{\Omega} = \mathbf{\Omega}^{-1}\mathbf{\Omega} = \mathbf{I}$; orthogonal matrices are easy to invert! When this class of matrix pre-multiplies a vector, the effect is to rotate it around the origin and possibly reflect it.

Rotation matrices are a sub-class of orthogonal matrices that have the additional property that the determinant is one. As the name suggests, when this class of matrix pre-multiplies a vector, the effect is to rotate it around the origin with no reflection.

C.2.6 Positive definite matrices

A $D \times D$ real symmetric matrix $\mathbf{A}$ is *positive definite* if $\mathbf{x}^T\mathbf{Ax} > 0$ for all non-zero vectors $\mathbf{x}$. Every positive definite matrix is invertible and its inverse is also positive definite. The determinant and trace of a symmetric positive definite matrix are always positive. The covariance matrix $\mathbf{\Sigma}$ of a normal distribution is always positive definite.

C.2.7 Null space of a matrix

The right null space of a matrix $\mathbf{A}$ consists of the set of vectors $\mathbf{x}$ for which

$$\mathbf{Ax} = \mathbf{0}. \tag{C.16}$$

Similarly, the left null space of a matrix $\mathbf{A}$ consists of the set of vectors $\mathbf{x}$ for which

$$\mathbf{x}^T\mathbf{A} = \mathbf{0}^T. \tag{C.17}$$

A square matrix only has a non-trivial null space (i.e., not just $\mathbf{x} = \mathbf{0}$) if the matrix is singular (non-invertible) and hence the determinant is zero.

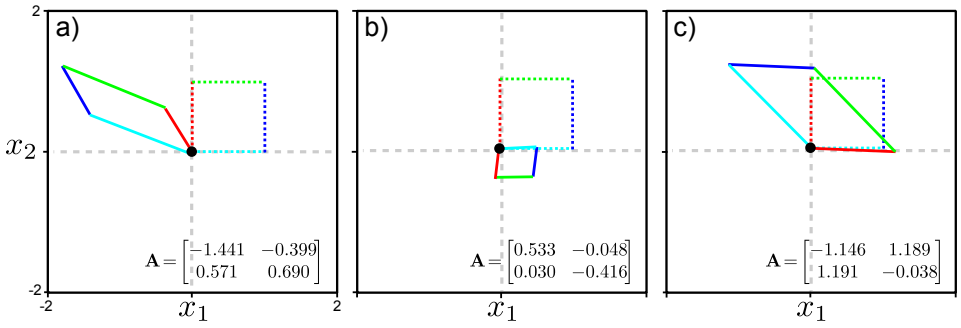

Figure C.1 Effect of applying three linear transformations to a unit square. Dashed square is before transformation. Solid square is after. The origin is always mapped to the origin. Co-linear points remain co-linear. Parallel lines remain parallel. The linear transformation encompasses shears, reflections, rotations and scalings.

C.3 Tensors

We will occasionally have need for $D > 2$ dimensional quantities that we shall refer to as D-dimensional tensors. For our purposes a matrix can be thought of as the special case of a 2 dimensional tensor, and a vector as the special case of a 1D tensor.

The idea of taking matrix products generalizes to higher dimensions and is denoted using the special notion $\times_n$ where n is the dimension over which we take the product. For example, the l^{th} element f_l of the tensor product $\mathbf{f} = \mathbf{A}\times_2 \mathbf{b}\times_3 \mathbf{c}$ is given by

$$f_l = \sum_m \sum_n A_{lmn} b_m c_n, \tag{C.18}$$

where l, m and n index the 3D tensor $\mathbf{A}$, and $\mathbf{b}$ and $\mathbf{c}$ are vectors.

C.4 Linear transformations

When we pre-multiply a vector by a matrix this is called a linear transformation. Figure C.1 shows the results of applying several different 2D linear transformations (randomly chosen 2×2 matrices) to the 2D vectors that represent the points of the unit square. We can deduce several things from this figure. First, the point (0,0) at the origin in always mapped back onto itself. Second, collinear points remain collinear. Third, parallel lines are always mapped to parallel lines. Viewed as a geometric transformation, pre-multiplication by a matrix can account for shearing, scaling, reflection and rotation around the origin.

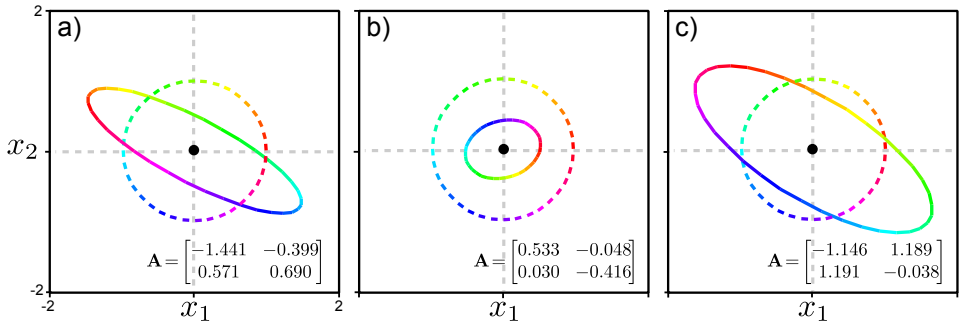

Figure C.2 Effect of applying three linear transformations to a circle. Dashed circle is before transformation. Solid ellipse is after. After the transformation, the circle is mapped to an ellipse. This demonstrates that there is one special direction that is expanded the most (becomes the major axis of the ellipse), and one special direction that is expanded the least (becomes the minor axis of the ellipse).

A different perspective on linear transforms comes from applying different transformations to points on the unit circle (figure C.2). In each case, the circle is transformed to an ellipse. The ellipse can be characterized by its major axis (most elongated axis) and its minor axis (most compact axis), which are perpendicular to one another. This tells us something interesting: in general, there is a special direction in space (position on the original circle) that gets stretched the most (or compressed the least) by the transformation. Likewise there is a second direction that gets stretched the least or compressed the most.

C.5 Singular value decomposition

The singular value decomposition (SVD) is a factorization of a $M \times N$ matrix $\mathbf{A}$, such that

$$\mathbf{A} = \mathbf{U}\mathbf{L}\mathbf{V}^T, \tag{C.19}$$

where $\mathbf{U}$ is a $M \times M$ orthogonal matrix, $\mathbf{L}$ is a $M \times N$ diagonal matrix, and $\mathbf{V}$ is a $N \times N$ orthogonal matrix. It is always possible to compute this factorization, although a description of how to do so is beyond the scope of this book.

The best way to get the flavor of the SVD is to consider some examples. First let us consider a square matrix:

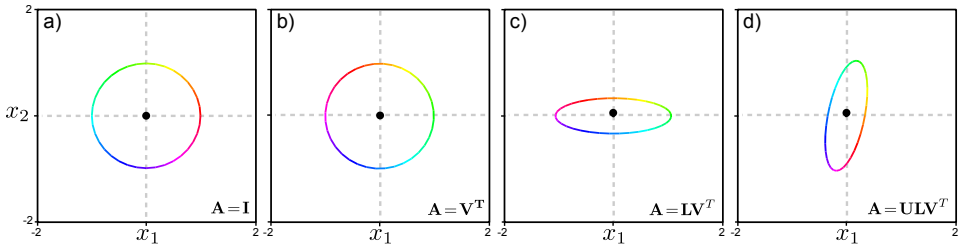

Figure C.3 Cumulative effect of SVD components for matrix $\mathbf{A}_3$. a) Original object. b) Applying matrix $\mathbf{V}^T$ rotates and reflects the object around the origin. c) Subsequently applying $\mathbf{L}$ causes stretching/compression along the coordinate axes. d) Finally, applying matrix $\mathbf{U}$ rotates and reflects this distorted structure.

$$\begin{aligned}\mathbf{A}_1 &= \begin{bmatrix}0.183 & 0.307 & 0.261\\ -1.029 & 0.135 & -0.941\\ 0.949 & 0.515 & -0.162\end{bmatrix} = \mathbf{U}\mathbf{L}\mathbf{V}^T \\ &= \begin{bmatrix}-0.204 & -0.061 & -0.977\\ 0.832 & -0.535 & -0.140\\ -0.514 & -0.842 & 0.160\end{bmatrix}\begin{bmatrix}1.590 & 0 & 0\\ 0 & 0.856 & 0\\ 0 & 0 & 0.303\end{bmatrix}\begin{bmatrix}-0.870 & -0.302 & 0.389\\ -0.135 & -0.613 & -0.778\\ -0.474 & 0.729 & -0.492\end{bmatrix}.\end{aligned} \tag{C.20}$$

Notice that by convention, the non-negative values on the principal diagonal of $\mathbf{L}$ decrease monotonically as we move from top-left to bottom-right. These are known as the *singular values*.

Now consider the singular value decomposition of a portrait matrix:

$$\begin{aligned}\mathbf{A}_2 &= \begin{bmatrix}0.537 & 0.862\\ 1.839 & 0.318\\ -2.258 & -1.307\end{bmatrix} = \mathbf{U}\mathbf{L}\mathbf{V}^T \\ &= \begin{bmatrix}-0.263 & 0.698 & 0.665\\ -0.545 & -0.676 & 0.493\\ 0.795 & -0.233 & 0.559\end{bmatrix}\begin{bmatrix}3.273 & 0\\ 0 & 0.76\\ 0 & 0\end{bmatrix}\begin{bmatrix}-0.898 & -0.440\\ -0.440 & 0.898\end{bmatrix}.\end{aligned} \tag{C.21}$$

For this rectangular matrix, the orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$ are different sizes and the diagonal matrix $\mathbf{L}$ is the same size as the original matrix. The singular values are still found on the diagonal, but the number is determined by the smallest dimension. In other words, if the original matrix was $M \times N$ then there will be $\min[M, N]$ singular values.

To further understand the SVD, let us consider a third example:

$$\mathbf{A}_3 = \begin{bmatrix}-0.147 & 0.357\\ -0.668 & 0.811\end{bmatrix} = \begin{bmatrix}0.189 & 0.981\\ 0.981 & -0.189\end{bmatrix}\begin{bmatrix}1.068 & 0\\ 0 & 0.335\end{bmatrix}\begin{bmatrix}-0.587 & 0.8091\\ 0.809 & 0.587\end{bmatrix}. \tag{C.22}$$

Figure C.3 illustrates the cumulative effect of the transformations in the decomposition $\mathbf{A}_3 = \mathbf{ULV}^T$. The matrix $\mathbf{V}^T$ rotates and reflects the original points. The matrix $\mathbf{L}$ scales the result differently along each dimension. In this case it is stretched along the first dimension and shrunk along the second. Finally, the matrix $\mathbf{U}$ rotates the result.

Figure C.4 provides a second perspective on this process. Each pair of panels depicts what happens when we modify a different part of the SVD but keep the remaining parts the same. When we change $\mathbf{V}$, the shape of the final ellipse is the same, but the mapping from original directions to points on the ellipse changes (observe the color change along the major axis). When we modify the first element of $\mathbf{L}$, the length of the major axis changes. When we change the other non-zero element of $\mathbf{L}$, the length of the minor axis changes. When we change the matrix $\mathbf{U}$, the orientation of the ellipse changes.

C.5.1 Analyzing the singular values

We can learn a lot about a matrix by looking at the singular values. We saw in the previous section, that as we decrease the smallest singular value the minor axis of the ellipse becomes progressively smaller. When it actually becomes zero, both sides of the unit circle collapse into one another (as do points from circles of all radii). Now there is a many-to-one mapping from the original points to the transformed ones, and the matrix is no longer invertible. In general, a matrix is only invertible if all of the singular values are non-zero.

The number of non-zero singular values is called the rank of the matrix. The ratio of the smallest to the largest singular values is known as the condition number: it is roughly a measure of how 'invertible' the matrix is. As it becomes close to zero, our ability to invert the matrix decreases.

The singular values scale the different axes of the ellipse by different amounts (figure C.3c). Hence, the area of a unit circle is changed by a factor that is equal to the product of the singular values. In fact this scaling factor is the determinant (section C.2.4). When the matrix is singular, at least one of the singular values is zero and hence the determinant is also zero. The right null space consists of all of the vectors that can be reached by taking a weighted sum of those columns of $\mathbf{V}$ whose corresponding singular values are zero. Similarly, the left null space consists of all of the vectors that can be reached by taking a weighted sum of those columns of $\mathbf{U}$ whose corresponding singular values are zero.

Orthogonal matrices only rotate and reflect points, and rotation matrices just rotate them. In either case, there is no change in area to the unit circle: all the singular values are one for these matrices and the determinant is also one.

C.5.2 Inverse of a matrix

We can also see what happens when we invert a square matrix in terms of the singular value decomposition. Using the rule $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, we have

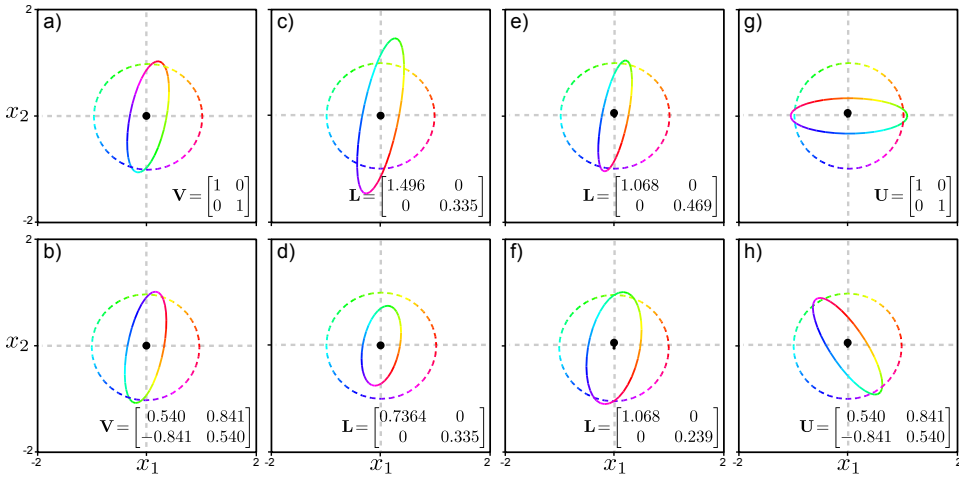

Figure C.4 Manipulating different parts of the SVD of $\mathbf{A}_3$. a-b) Changing matrix $\mathbf{V}$ does not affect the final ellipse, but changes which directions (colors) are mapped to the minor and major axes. c-d) Changing the first diagonal element of $\mathbf{L}$ changes the length of the major axis of the ellipse. e-f) Changing the second diagonal element of $\mathbf{L}$ changes the length of the minor axis. g-h) Changing $\mathbf{U}$ affects the final orientation of the ellipse.

$$\mathbf{A}^{-1} = (\mathbf{U}\mathbf{L}\mathbf{V}^T)^{-1} = (\mathbf{V}^T)^{-1}\mathbf{L}^{-1}\mathbf{U}^{-1} = \mathbf{V}\mathbf{L}^{-1}\mathbf{U}^T, \tag{C.23}$$

where we have used the fact that $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices so $\mathbf{U}^{-1} = \mathbf{U}^T$ and $\mathbf{V}^{-1} = \mathbf{V}^T$. The matrix $\mathbf{L}$ is diagonal so $\mathbf{L}^{-1}$ will also be diagonal with new non-zero entries that are the reciprocal of the original values. This also shows that the matrix is not invertible when any of the singular values are zero: we cannot take the reciprocal of 0.

Expressed in this way, the inverse has the opposite geometric effect to that of the original matrix: if we consider the effect on the transformed ellipse in figure C.3d, it first rotates by $\mathbf{U}^T$ so its major and minor axis are aligned with the coordinate axes (figure C.3c). Then it scales these axes (using the elements of $\mathbf{L}^{-1}$), so that the ellipse becomes a circle (figure C.3b). Finally, it rotates the result by $\mathbf{V}$ to get back to the original position (figure C.3a).

C.6 Matrix calculus

We are often called upon to take derivatives of compound matrix expressions. The derivative of a function f[$\mathbf{a}$] that takes a vector as its argument and returns a scalar is a vector $\mathbf{b}$ with elements

$$b_i = \frac{\partial f}{\partial a_i} \tag{C.24}$$

The derivative of a function f[$\mathbf{A}$] that returns a scalar, with respect to an $M \times N$ matrix $\mathbf{A}$ will be a $M \times N$ matrix $\mathbf{B}$ with elements

$$b_{ij} = \frac{\partial f}{\partial a_{ij}}. \tag{C.25}$$

The derivative of a function $\mathbf{f}[\mathbf{a}]$ that returns a vector with respect to vector $\mathbf{a}$ is a matrix $\mathbf{B}$ with elements

$$b_{ij} = \frac{\partial f_i}{\partial a_j}. \tag{C.26}$$

where f_i is the i^{th} element of the vector returned by the function $\mathbf{f}[\mathbf{a}]$.

We now provide several commonly used results for reference.

1. Derivative of linear function:

$$\begin{aligned}
\frac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}} &= \mathbf{a} && \text{(C.27)}\\
\frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} &= \mathbf{a} && \text{(C.28)}\\
\frac{\partial \mathbf{a}^T\mathbf{X}\mathbf{b}}{\partial \mathbf{X}} &= \mathbf{a}\mathbf{b}^T && \text{(C.29)}\\
\frac{\partial \mathbf{a}^T\mathbf{X}^T\mathbf{b}}{\partial \mathbf{X}} &= \mathbf{b}\mathbf{a}^T. && \text{(C.30)}
\end{aligned}$$

2. Derivative of quadratic function:

$$\begin{aligned}
\frac{\partial \mathbf{b}^T\mathbf{X}^T\mathbf{X}\mathbf{c}}{\partial \mathbf{X}} &= \mathbf{X}(\mathbf{b}\mathbf{c}^T + \mathbf{c}\mathbf{b}^T) && \text{(C.31)}\\
\frac{\partial (\mathbf{B}\mathbf{x}+\mathbf{b})^T\mathbf{C}(\mathbf{D}\mathbf{x}+\mathbf{d})}{\partial \mathbf{x}} &= \mathbf{B}^T\mathbf{C}(\mathbf{D}\mathbf{x}+\mathbf{d}) + \mathbf{D}^T\mathbf{C}^T(\mathbf{B}\mathbf{x}+\mathbf{b}) && \text{(C.32)}\\
\frac{\partial \mathbf{x}^T\mathbf{B}\mathbf{x}}{\partial \mathbf{x}} &= (\mathbf{B}+\mathbf{B}^T)\mathbf{x} && \text{(C.33)}\\
\frac{\partial \mathbf{b}^T\mathbf{X}^T\mathbf{D}\mathbf{X}\mathbf{c}}{\partial \mathbf{X}} &= \mathbf{D}^T\mathbf{X}\mathbf{b}\mathbf{c}^T + \mathbf{D}\mathbf{X}\mathbf{c}\mathbf{b}^T && \text{(C.34)}\\
\frac{\partial (\mathbf{X}\mathbf{b}+\mathbf{c})^T\mathbf{D}(\mathbf{X}\mathbf{b}+\mathbf{c})}{\partial \mathbf{X}} &= (\mathbf{D}+\mathbf{D}^T)(\mathbf{X}\mathbf{b}+\mathbf{c})\mathbf{b}^T. && \text{(C.35)}
\end{aligned}$$

3. Derivative of determinant:

$$\frac{\partial \det[\mathbf{Y}]}{\partial x} = \det[\mathbf{Y}]\mathrm{tr}\left[\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\right], \tag{C.36}$$

which leads to the relation

$$\frac{\partial \det[\mathbf{Y}]}{\partial \mathbf{Y}} = \det[\mathbf{Y}]\mathbf{Y}^{-T}. \tag{C.37}$$

4. Derivative of log determinant:

$$\frac{\partial \log[\det[\mathbf{Y}]]}{\partial \mathbf{Y}} = \mathbf{Y}^{-T}. \tag{C.38}$$

5. Derivative of inverse:

$$\frac{\partial \mathbf{Y}^{-1}}{\partial x} = -\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}. \tag{C.39}$$

6. Derivative of trace:

$$\frac{\partial \mathrm{tr}[\mathbf{F}[\mathbf{X}]]}{\partial \mathbf{X}} = \left(\frac{\partial \mathbf{F}[\mathbf{X}]}{\partial \mathbf{X}}\right)^{T}. \tag{C.40}$$

More information about matrix calculus can be found in Petersen *et al.* (2006).

C.7 Common problems

In this section, we discuss several standard linear algebra problems that are found repeatedly in computer vision.

C.7.1 Least squares problems

Many inference and learning tasks in computer vision result in least squares problems. The most frequent context is when we use maximum likelihood methods with the normal distribution. The least squares problem may be formulated in a number of ways. We may be asked to find the vector $\mathbf{x}$ that solves the system

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{C.41}$$

in a least squares sense. Alternatively, we may be given i of smaller sets of equations of the form

$$\mathbf{A}_i\mathbf{x} = \mathbf{b}_i, \tag{C.42}$$

and again asked to solve for $\mathbf{x}$. In this latter case, we form the compound matrix $\mathbf{A} = [\mathbf{A}_1^T, \mathbf{A}_2^T \ldots \mathbf{A}_I^T]^T$ and compound vector $\mathbf{b} = [\mathbf{b}_1^T, \mathbf{b}_2^T \ldots \mathbf{b}_I^T]^T$, and the problem is the same as in equation C.41.

We may equivalently see the same problem in an explicit least-squares form,

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmin}} \left[(\mathbf{A}\mathbf{x} - \mathbf{b})^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \right]. \tag{C.43}$$

Finally, we may be presented the problem as a sum of smaller terms

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left[\sum_{i=1}^{I} \left(\mathbf{A}_i\mathbf{x} - \mathbf{b}_i\right)^T \left(\mathbf{A}_i\mathbf{x} - \mathbf{b}_i\right) \right], \tag{C.44}$$

in which case we form compound matrices $\mathbf{A}$ and $\mathbf{b}$, which changes the problem back to that in equation C.43.

To make progress, we multiply out the terms in equation C.43

$$\begin{aligned}
\hat{\mathbf{x}} &= \underset{\mathbf{x}}{\operatorname{argmin}} \left[\left(\mathbf{A}\mathbf{x} - \mathbf{b}\right)^T \left(\mathbf{A}\mathbf{x} - \mathbf{b}\right) \right]. \\
&= \underset{\mathbf{x}}{\operatorname{argmin}} \left[\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{A}\mathbf{x} - \mathbf{x}^T\mathbf{A}^T\mathbf{b} + \mathbf{b}^T\mathbf{b} \right] \\
&= \underset{\mathbf{x}}{\operatorname{argmin}} \left[\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{x}^T\mathbf{A}^T\mathbf{b} + \mathbf{b}^T\mathbf{b} \right],
\end{aligned} \tag{C.45}$$

where we have combined two terms in the last line by noting that they are both the same: they are transposes of one another, but they are also scalars, so they equal their own transpose. Now we take the derivative with respect to $\mathbf{x}$ and equate the result to zero to give

$$2\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{A}^T\mathbf{b} = 0, \tag{C.46}$$

which we can re-arrange to give the standard least squares result

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{C.47}$$

This result can only be computed if there are at least as many rows in $\mathbf{A}$ as there are unknown values in $\mathbf{x}$ (i.e., if the matrix $\mathbf{A}$ is square or portrait). Otherwise, the matrix $\mathbf{A}^T\mathbf{A}$ will be singular. For implementations in Matlab, it is better to make use of the backslash operator '\' rather than explicitly implement equation C.47.

C.7.2 Principal direction / minimum direction

We define the principal and minimal directions as

$$\begin{aligned}
\hat{\mathbf{b}} &= \underset{\mathbf{b}}{\operatorname{argmax}} \left[\mathbf{A}\mathbf{b}\right] \qquad \text{subject to } |\mathbf{b}| = 1 \\
\hat{\mathbf{b}} &= \underset{\mathbf{b}}{\operatorname{argmin}} \left[\mathbf{A}\mathbf{b}\right] \qquad \text{subject to } |\mathbf{b}| = 1,
\end{aligned} \tag{C.48}$$

respectively. This problem has exactly the geometric form of figure C.2. The constraint that $|\mathbf{b}| = 1$ means that $\mathbf{b}$ has to lie on the circle (or sphere or hypersphere in higher dimensions). In the principal direction problem, we are hence seeking the direction that is mapped to the major axis of the resulting ellipse/ellipsoid. In the minimum direction problem, we seek the direction that is mapped to the minor axis of the ellipsoid.

We saw in figure C.4 that it is the matrix $\mathbf{V}$ from the singular value decomposition of $\mathbf{A}$ that controls which direction is mapped to the different axes of the ellipsoid. To solve the principal direction problem, we hence compute the SVD $\mathbf{A} = \mathbf{U}\mathbf{L}\mathbf{V}^T$, and set $\mathbf{b}$ to be the first column of $\mathbf{V}$. To solve the minimum direction problem, we set $\mathbf{b}$ to be the last column of $\mathbf{V}$.

C.7.3 Orthogonal Procrustes problem

The orthogonal Procrustes problem is to find the closest linear mapping $\boldsymbol{\Omega}$ between one set of vectors $\mathbf{A}$ and another $\mathbf{B}$ such that $\boldsymbol{\Omega}$ is an orthogonal matrix. In layman's terms, we seek the best Euclidean rotation (possibly including mirroring) that maps points $\mathbf{A}$ to points $\mathbf{B}$.

$$\hat{\boldsymbol{\Omega}} = \underset{\boldsymbol{\Omega}}{\operatorname{argmin}} \left[|\boldsymbol{\Omega}\mathbf{A} - \mathbf{B}|_F \right], \tag{C.49}$$

where $|\bullet|_F$ denotes the Frobenius norm of a matrix – the sum of the square of all of the elements. To make progress, we recall that the trace of a matrix is the sum of its diagonal entries and so $|\mathbf{X}|_F = \text{tr}[\mathbf{X}^T\mathbf{X}]$ which gives the new criterion

$$\begin{aligned}
\hat{\boldsymbol{\Omega}} &= \underset{\boldsymbol{\Omega}}{\operatorname{argmin}} \Big[\text{tr}[\mathbf{A}^T\mathbf{A}] + \text{tr}[\mathbf{B}^T\mathbf{B}] - 2\text{tr}[\mathbf{A}^T\boldsymbol{\Omega}^T\mathbf{B}] \Big] \\
&= \underset{\boldsymbol{\Omega}}{\operatorname{argmax}} \Big[\text{tr}[\mathbf{A}^T\boldsymbol{\Omega}^T\mathbf{B}] \Big] \\
&= \underset{\boldsymbol{\Omega}}{\operatorname{argmax}} \Big[\text{tr}[\boldsymbol{\Omega}^T\mathbf{B}\mathbf{A}^T] \Big],
\end{aligned} \tag{C.50}$$

where we have used relation C.15 between the last two lines. We now compute the SVD $\mathbf{B}\mathbf{A}^T = \mathbf{U}\mathbf{L}\mathbf{V}^T$ to get the criterion

$$\begin{aligned}
\hat{\boldsymbol{\Omega}} &= \underset{\boldsymbol{\Omega}}{\operatorname{argmax}} \Big[\text{tr}[\boldsymbol{\Omega}^T\mathbf{U}\mathbf{L}\mathbf{V}^T] \Big] \\
&= \underset{\boldsymbol{\Omega}}{\operatorname{argmax}} \Big[\text{tr}[\mathbf{V}^T\boldsymbol{\Omega}^T\mathbf{U}\mathbf{L}] \Big],
\end{aligned} \tag{C.51}$$

and notice that

$$\text{tr}[\mathbf{V}^T\boldsymbol{\Omega}^T\mathbf{U}\mathbf{L}] = \text{tr}[\mathbf{Z}\mathbf{L}] = \sum_{i=1}^{I} z_{ii} l_{ii}, \tag{C.52}$$

where we defined $\mathbf{Z} = \mathbf{V}^T\boldsymbol{\Omega}^T\mathbf{U}$, and used the fact that the $\mathbf{L}$ is a diagonal matrix, so each entry scales the diagonal of $\mathbf{Z}$ on multiplication.

We note that the matrix $\mathbf{Z}$ is orthogonal (it is the product of three orthogonal matrices). Hence, every value on the diagonal of the orthogonal matrix $\mathbf{Z}$ must be less than or equal to one (the norms of each column are exactly one), and so we maximize the criterion in equation C.52 by choosing $\mathbf{Z} = \mathbf{I}$ when the diagonal

values are equal to one. To achieve this, we set $\boldsymbol{\Omega}^T = \mathbf{V}\mathbf{U}^T$ so that the overall solution is

$$\hat{\boldsymbol{\Omega}} = \mathbf{U}\mathbf{V}^T. \tag{C.53}$$

A special case of this problem is to find the closest orthogonal matrix $\boldsymbol{\Omega}$ to a given square matrix $\mathbf{B}$ in a least squares sense. In other words, we seek to optimize

$$\hat{\boldsymbol{\Omega}} = \underset{\boldsymbol{\Omega}}{\operatorname{argmin}} \left[|\boldsymbol{\Omega} - \mathbf{B}|_F \right]. \tag{C.54}$$

This is clearly equivalent to optimizing the criterion in equation C.49 but with $\mathbf{A} = \mathbf{I}$. It follows that the solution can be found by computing the singular value decomposition $\mathbf{B} = \mathbf{U}\mathbf{L}\mathbf{V}^T$ and setting $\boldsymbol{\Omega} = \mathbf{U}\mathbf{V}^T$.

C.8 Tricks for inverting large matrices

Inversion of a $D \times D$ matrix has a complexity of $\mathcal{O}(D^3)$. In practice, this means it is difficult to invert matrices whose dimension is larger than a few thousand. Fortunately, matrices are often highly structured, and we can exploit that structure using a number of tricks to speed up the process.

C.8.1 Diagonal and block-diagonal matrices

Diagonal matrices can be inverted by forming a new diagonal matrix, where the values on the diagonal are reciprocal of the original values. Block diagonal matrices are matrices of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_N \end{bmatrix}. \tag{C.55}$$

The inverse of a block-diagonal matrix can be computed by taking the inverse of each block separately so that

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_1^{-1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^{-1} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_N^{-1} \end{bmatrix}. \tag{C.56}$$

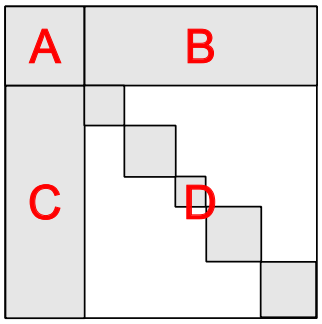

Figure C.5 Inversion relation #1. Gray regions indicate parts of matrix with non-zero values, white regions represent zeros. This relation is suited to the case where the matrix can be divided into four sub-matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ and the bottom right block is easy to invert (e.g., diagonal, block diagonal or structured in another way that means that inversion is efficient). After applying this relation, the remaining inverse is the size of sub-matrix $\mathbf{A}$.

C.8.2 Inversion relation #1: Schur complement identity

The inverse of a matrix with sub-blocks $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ in the top-left, top right, bottom-left, and bottom right positions, respectively, can easily be shown to be

$$\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix}^{-1} = \begin{bmatrix}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1}\\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1}+\mathbf{D}^{-1}\mathbf{C}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1}\end{bmatrix}, \tag{C.57}$$

by multiplying the original matrix with the right hand side and showing that the result is the identity matrix.

This result is extremely useful when the matrix $\mathbf{D}$ is diagonal or block-diagonal (figure C.5). In this circumstance, $\mathbf{D}^{-1}$ is fast to compute, and the remaining inverse quantity $(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$ is much smaller and easier to invert than the original matrix. The quantity $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}$ is known as the Schur complement.

C.8.3 Inversion relation #2

Consider the $d \times d$ matrix $\mathbf{A}$, the $k \times k$ matrix $\mathbf{C}$ and the $k \times d$ matrix $\mathbf{B}$ where $\mathbf{A}$ and $\mathbf{C}$ are symmetric, positive definite matrices. The following equality holds:

$$(\mathbf{A}^{-1}+\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}^{-1} = \mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{A}\mathbf{B}^T+\mathbf{C})^{-1}. \tag{C.58}$$

Proof:

$$\begin{aligned}\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A}\mathbf{B}^T+\mathbf{B}^T &= \mathbf{B}^T+\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A}\mathbf{B}^T\\ \mathbf{B}^T\mathbf{C}^{-1}(\mathbf{B}\mathbf{A}\mathbf{B}^T+\mathbf{C}) &= (\mathbf{A}^{-1}+\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})\mathbf{A}\mathbf{B}^T.\end{aligned} \tag{C.59}$$

Taking the inverse of both sides we get

$$(\mathbf{A}^{-1}+\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}^{-1} = \mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{A}\mathbf{B}^T+\mathbf{C})^{-1}, \tag{C.60}$$

as required.

This relation is very useful when $\mathbf{B}$ is a landscape matrix with many more columns C than rows R. On the left-hand side, the term that we must invert is of size $C \times C$, which might be very costly. However, on the right-hand side, the inversion is only of size $R \times R$, which might be considerably more efficient.

C.8.4 Inversion relation #3: Sherman-Morrison-Woodbury

Consider the $d \times d$ matrix $\mathbf{A}$, the $k \times k$ matrix $\mathbf{C}$ and the $k \times d$ matrix $\mathbf{B}$ where $\mathbf{A}$ and $\mathbf{C}$ are symmetric, positive definite matrices. The following equality holds:

$$(\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{A}\mathbf{B}^T + \mathbf{C})^{-1}\mathbf{B}\mathbf{A}. \tag{C.61}$$

This is sometimes known as the *matrix inversion lemma*.

Proof:

$$\begin{aligned}
&(\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1} \\
&\qquad = (\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}(\mathbf{I} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A} - \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A}) \\
&\qquad = (\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}\left((\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})\mathbf{A} - \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A}\right) \\
&\qquad = \mathbf{A} - (\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A}.
\end{aligned} \tag{C.62}$$

Now, applying inversion relation #2 to the term in brackets:

$$\begin{aligned}
(\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1} &= \mathbf{A} - (\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}\mathbf{A} \\
&= \mathbf{A} - \mathbf{A}\mathbf{B}^T(\mathbf{B}\mathbf{A}\mathbf{B}^T + \mathbf{C})^{-1}\mathbf{B}\mathbf{A},
\end{aligned} \tag{C.63}$$

as required.

C.8.5 Matrix determinant lemma

The matrices that we need to invert are often the covariances in the normal distribution. When this is the case we sometimes also need to compute the determinant of the same matrix and this may be prohibitively expensive. Fortunately, there is a direct analogy of the matrix inversion lemma for determinants which can reduce the cost when the covariance is structured.

Consider the $d \times d$ matrix $\mathbf{A}$, the $k \times k$ matrix $\mathbf{C}$ and the $k \times d$ matrix $\mathbf{B}$ where $\mathbf{A}$ and $\mathbf{C}$ are symmetric, positive definite covariance matrices. The following equality holds:

$$|\mathbf{A}^{-1} + \mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}\mathbf{B}^T|\mathbf{C}|^{-1}|\mathbf{A}|^{-1}. \tag{C.64}$$

Bibliography

Aeschliman, C., Park, J., & Kak, A. C. (2010) A novel parameter estimation algorithm for the multivariate t-distribution and its application to computer vision. In *European Conference on Computer Vision*, pp. 594–607. 134, 140

Agarwal, A., & Triggs, B. (2006) Recovering 3D human pose from monocular images. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (1): 44–48. 144, 166, 169

Agarwal, S., Snavely, N., Seitz, S. M., & Szeliski, R. (2010) Bundle adjustment in the large. In *European Conference on Computer Vision*, pp. 29–42. 453

Agarwal, S., Snavely, N., Simon, I., Seitz, S. M., & Szeliski, R. (2009) Building Rome in a day. In *IEEE International Conference on Computer Vision*, pp. 72–79. 453

Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S. M., Colburn, A., Curless, B., Salesin, D., & Cohen, M. F. (2004) Interactive digital photomontage. *ACM Transactions on Graphics* **23** (3): 294–302. 316

Aghajanian, J., Warrell, J., Prince, S. J. D., Li, P., Rohn, J. L., & Baum, B. (2009) Patch-based within-object classification. In *IEEE International Conference on Computer Vision*, pp. 1125–1132. 595

Ahonen, T., Hadid, A., & Pietikäinen, M. (2004) Face recognition with local binary patterns. In *European Conference on Computer Vision*, pp. 469–481. 534

Alahari, K., Kohli, P., & Torr, P. H. S. (2008) Reduce, reuse & recycle: Efficiently solving multi-label MRFs. In *IEEE Computer Vision & Pattern Recognition*. 316

Allen, B., Curless, B., & Popovic, Z. (2003) The space of human body shapes: reconstruction and parameterization from range scans. *ACM Transactions on Graphics* **22** (3): 587–594. 500

Aloimonos, J. Y. (1990) Perspective approximations. *Image and Vision Computing* **8** (3): 177–192. 385

Amberg, B., Blake, A., & Vetter, T. (2009) On compositional image alignment, with an application to active appearance models. In *IEEE Computer Vision & Pattern Recognition*, pp. 1714–1721. 499

Amini, A., Weymouth, T., & Jain, R. (1990) Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **12** (9): 855–867. 275, 499

Amit, Y., & Geman, D. (1997) Shape quantization and recognition with randomized trees. *Neural Computation* **9** (7): 1545–1588. 209

Amit, Y., & Kong, A. (1996) Graphical templates for model registration. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **18** (3): 225–236. 274

Andriluka, M., Roth, S., & Schiele, B. (2009) Pictorial structures revisited: people detection and articulated pose estimation. In *IEEE Computer Vision & Pattern Recognition*. 274

Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., & Davis, J. (2005) SCAPE: shape completion and animation of people. *ACM Transactions on Graphics* **24** (3): 408–416. 496, 497, 498, 500

Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing* **50** (2): 174–188. 567

Avidan, S., & Shamir, A. (2007) Seam carving for content-aware image resizing. *ACM Transactions on Graphics* **26** (3): 10. 274

Ayache, N. (1991) *Artificial vision for mobile robots: stereo vision and multisensory perception*. MIT Press. 567

Babalola, K., Cootes, T., Twining, C., Petrovic, V., & Taylor, C. (2008) 3D brain segmentation using active appearance models and local regressors. In *Medical Image Computing and Computer-Assisted Intervention 2008*, ed. by D. Metaxas, L. Axel, G. Fichtinger, & G. Székely, volume 5241 of *Lecture Notes in Computer Science*, pp. 401–408. Springer Berlin / Heidelberg. 481

Bailey, T., & Durrant-Whyte, H. (2006) Simultaneous localization and mapping (SLAM): Part II. *Robotics & Automation Magazine, IEEE* **13** (3): 108–117. 568

Baker, H. H., & Binford, T. O. (1981) Depth from edge and intensity-based stereo. In *International Joint Conference on Artificial Intelligence*, pp. 631–636. 274

Ballan, L., & Cortelazzo, G. M. (2008) Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes. In *3D Data Processing, Visualization and Transmission*. 385

Baluja, S., & Rowley, H. A. (2003) Boosting sex identification performance. *International Journal of Computer Vision* **71** (1): 111–119. 210

Barber, D. (2012) *Bayesian Reasoning and Machine Learning*. Cambridge University Press. 226, 239, 275

Bartlett, M. S., Lades, H. M., & Sejnowski, T. J. (1998) Independent component representations for face recognition. In *Proceedings of the SPIE Symposium on Electronic Imaging: Science and Technology: Conference on Human Vision and Electronic Imaging III*, pp. 528–539. 533

Basri, R., Costa, L., Geiger, D., & Jacobs, D. (1998) Determining the similarity of deformable shapes. *Vision Research* **38** (15–16): 2365–2385. 274

Batlle, J., Mouaddib, E., & Salvi, J. (1998) Recent progress in coded structured light as a technique to solve the correspondence problem: a survey. *Pattern Recognition* **31** (7): 963–982. 385

Baumgart, B. G. (1974) *Geometric modeling for computer vision*. Stanford University PhD dissertation. 385

Bay, H., Ess, A., Tuytelaars, T., & Gool, L. J. V. (2008) Speeded-up robust features (SURF). *Computer Vision and Image Understanding* **110** (3): 346–359. 352

Beardsley, P. A., Reid, I. D., Zisserman, A., & Murray, D. W. (1995) Active visual navigation using non-metric structure. In *IEEE International Conference on Computer Vision*, pp. 58–65. 567

Bekios-Calfa, J., Buenaposada, J. M., & Baumela, L. (2011) Revisiting linear discriminant techniques in gender recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **33** (4): 858–864. 209

Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997) Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **19** (7): 711–720. 533

Belkin, M., & Niyogi, P. (2001) Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems*, pp. 585–591. 352

Belongie, S., Carson, C., Greenspan, H., & Malik, J. (1998) Color- and texture-based image segmentation using EM and its application to content based image retrieval. In *IEEE International Conference on Computer Vision*, pp. 675–682. 140

Belongie, S., Malik, J., & Puzicha, J. (2002) Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (4): 509–522. 352

Bengio, Y., & Delalleau, O. (2009) Justifying and generalizing contrastive divergence. *Neural Computation* **21** (6): 1601–1621. 239

Bergtholdt, M., Cremers, D., & Schörr, C. (2005) Variational segmentation with shape priors. In *Handbook of Mathematical Models in Computer Vision*, ed. by Y. C. N. Paragios & O. Faugeras, 131–144. 499

Beymer, D., & Konolige, K. (1999) Real-time tracking of multiple people using continuous detection. In *IEEE Frame Rate Workshop*. 567

Birchfield, S., & Tomasi, C. (1998) Depth discontinuities by pixel-to-pixel stereo. In *IEEE International Conference on Computer Vision*, pp. 1073–1080. 274

Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*. Springer, 2nd edition. 19, 33, 45, 77, 97, 159, 209, 239, 275

Blake, A. (2006) Visual tracking: a short research roadmap. In *Handbook of Mathematical Models in Computer Vision*, ed. by N. Paragios & Y. C. and. O. Faugeras, pp. 293–307. Springer. 567

Blake, A., Curwen, R. W., & Zisserman, A. (1993) A framework for spatiotemporal control in the tracking of visual contours. *International Journal of Computer Vision* **11** (2): 127–145. 567

Blake, A., & Isard, M. (1996) The CONDENSATION algorithm – conditional density propagation and applications to visual tracking. In *Advances in Neural Information Processing Systems*, pp. 361–367. 567

Blake, A., & Isard, M. (1998) *Active Contours*. Springer. 499, 538, 565, 566, 567

Blake, A., Isard, M., & Reynard, D. (1995) Learning to track the visual motion of contours. *Artificial Intelligence* **78** (1–2): 179–212. 567

Blake, A., Kohli, P., & Rother, C. eds. (2011) *Advances in Markov Random Fields for Vision and Image Processing*. MIT Press. 316

Blanz, V., Basso, C., Poggio, T., & Vetter, T. (2003) Reanimating faces in images and video. *Comput. Graph. Forum* **22** (3): 641–650. 500

Blanz, V., Grother, P., Phillips, P. J., & Vetter, T. (2005) Face recognition based on frontal views generated from non-frontal images. In *IEEE Computer Vision & Pattern Recognition*, pp. 454–461. 495, 500, 533

Blanz, V., & Vetter, T. (1999) A morphable model for the synthesis of 3D faces. In *SIGGRAPH*, pp. 187–194. 494, 495, 500

Blanz, V., & Vetter, T. (2003) Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **25** (9): 1063–1074. 494, 495, 500

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research* **3**: 993–1022. 595

Bleyer, M., & Chambon, S. (2010) Does color really help in dense stereo matching? In *International Symposiumon 3D Data Processing, Visualization and Transmission*. 318

Bor Wang, S., Quattoni, A., Morency, L. P., Demirdjian, D., & Darrell, T. (2006) Hidden conditional random fields for gesture recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 1521–1527. 274

Bosch, A., Zisserman, A., & Munoz, X. (2007) Image classification using random forests and ferns. In *IEEE International Conference on Computer Vision*. 209

Bouwmans, T., Baf, F. E., & Vachon, B. (2010) Statistical background modeling for foreground detection: A survey. In *Handbook of Pattern Recognition and Computer Vision*, ed. by C. H. Chen, L. F. Pau, & P. S. P. Wang, pp. 181–199. World Scientific Publishing. 97

Boykov, Y., & Funka Lea, G. (2006) Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision* **70** (2): 109–131. 317

Boykov, Y., & Jolly, M.-P. (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *IEEE International Conference on Computer Vision*, pp. 105–112. 317

Boykov, Y., & Kolmogorov, V. (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **26** (9): 1124–1137. 316

Boykov, Y., & Veksler, O. (2006) Graph cuts in vision and graphics: theories and applications. In *Handbook of Mathematical Models in Computer Vision*, ed. by Y. C. N. Paragios & O. Faugeras, pp. 79–96. Springer. 316

Boykov, Y., Veksler, O., & Zabih, R. (1999) Fast approximate energy minimization via graph cuts. In *IEEE International Conference on Computer Vision*, pp. 377–384. 307, 308

Boykov, Y., Veksler, O., & Zabih, R. (2001) Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **23** (11): 1222–1239. 316

Brand, J., & Mason, J. (2000) A comparative assessment of three approaches to pixel-level skin-detection. In *International Conference on Pattern Recognition*, pp. 1056–1059. 97

Brand, M. (2002) Charting a manifold. In *Advances in Neural Information Processing Systems*, pp. 961–968. 352

Bregler, C., & Malik, J. (1998) Tracking people with twists and exponential maps. In *IEEE Computer Vision & Pattern Recognition*, pp. 8–15. 500

Breiman, L. (2001) Random forests. *Machine Learning* **45**: 5–32. 209

Brishnapuram, B., Figueiredo, M., Carin, L., & Hartemink, A. (2005) Sparse multinomial logistic regression: fast algorithms and generalization bounds. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **27** (6): 957–968. 209

Broida, T. J., & Chellappa, R. (1986) Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **8** (1): 90–99. 567

Brown, M., Hua, G., & Winder, S. A. J. (2011) Discriminative learning of local image descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **33** (1): 43–57. 352

Brown, M., & Lowe, D. G. (2005) Unsupervised 3D object recognition and reconstruction in unordered datasets. In *3D Digital Imaging and Modeling*, pp. 56–63. 453

Brown, M., & Lowe, D. G. (2007) Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision* **74** (1): 59–73. 421

Brown, M. Z., Burschka, D., & Hager, G. D. (2003) Advances in computational stereo. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **25** (8): 993–1008. 318

Brubaker, M. A., Fleet, D. J., & Hertzmann, A. (2010) Physics-based person tracking using the anthropomorphic walker. *International Journal of Computer Vision* **87** (1–2): 140–155. 500

Brunelli, R., & Poggio, T. (1993) Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **15** (10): 1042–1052. 533

Buchanan, A. M., & Fitzgibbon, A. W. (2005) Damped Newton algorithms for matrix factorization with missing data. In *IEEE Computer Vision & Pattern Recognition*, pp. 316–322. 453

Burgess, C. J. C. (2010) Dimension reduction: a guided tour. *Foundations and Trends in Machine Learning* **2** (4): 275–365. 352

Byröd, M., & Åström, K. (2010) Conjugate gradient bundle adjustment. In *European Conference on Computer Vision*, pp. 114–127. 453

Cai, D., He, X., Hu, Y., Han, J., & Huang, T. S. (2007) Learning a spatially smooth subspace for face recognition. In *IEEE Computer Vision & Pattern Recognition*. 533

Canny, J. (1986) A computational approach to edge detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **8** (6): 679–698. 352

Carreira-Perpiñán., M. Á., & Hinton, G. E. (2005) On contrastive divergence learning. In *Artificial Intelligence and Statistics*, volume 2005, p. 17. 239

Caselles, V., Kimmel, R., & Sapiro, G. (1997) Geodesic active contours. *International Journal of Computer Vision* **22** (1): 61–79. 499

Chellappa, R., Sinha, P., & Phillips, P. J. (2010) Face recognition by computers and humans. *IEEE Computer* **43** (2): 46–55. 533

Chen, L.-F., Liao, H.-Y. M., Ko, M.-T., Lin, J.-C., & Yu, G.-J. (2000) A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognition* **33** (10): 1713–1726. 533

Chen, R., & Liu, J. S. (2000) Mixture Kalman filters. *Journal of the Royal Statistical Society B.* **62** (3): 493–508. 567

Chen, S. E. (1995) QuickTime VR: an image-based approach to virtual environment navigation. In *SIGGRAPH*, pp. 29–38. 421

Cheung, G. K. M., Baker, S., & Kanade, T. (2004) Shape-from-silhouette across time part I: Theory and algorithms. *International Journal of Computer Vision* **62** (3): 221–247. 385

Chia, A. Y. S., Rahardja, S., Rajan, D., & Leung, M. K. H. (2010) Object recognition by discriminative combinations of line segments and ellipses. In *IEEE Computer Vision & Pattern Recognition*, pp. 2225–2232. 499

Chittajallu, D. R., Shah, S. K., & Kakadiaris, I. A. (2010) A shape-driven MRF model for the segmentation of organs in medical images. In *IEEE Computer Vision & Pattern Recognition*, pp. 3233–3240. 317

Cho, Y., Lee, J., & Neumann, U. (1998) A multi-ring color fiducial system and an intensity-invariant detection method for scalable fiducial-tracking augmented reality. In *International Workshop on Augmented Reality*, pp. 147–165. 420

Choi, S., Kim, T., & Yu, W. (2009) Performance evaluation of RANSAC family. In *British Machine Vision Conference*. 420

Chum, O., Philbin, J., Sivic, J., Isard, M., & Zisserman, A. (2007) Total recall: Automatic query expansion with a generative feature model for object retrieval. In *IEEE International Conference on Computer Vision*, pp. 1–8. 595

Chum, O., Werner, T., & Matas, J. (2005) Two-view geometry estimation unaffected by a dominant plane. In *IEEE Computer Vision & Pattern Recognition*, pp. 772–779. 420

Claus, D., & Fitzgibbon, A. W. (2005) A rational function lens distortion model for general cameras. In *IEEE Computer Vision & Pattern Recognition*, pp. 213–219. 385

Cohen, L. (1991) On active contour models and balloons. *CGVIP: Image Understanding* **53** (2): 211–218. 499

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001) Active appearance models. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **23** (6): 681–685. 499

Cootes, T. F., & Taylor, C. J. (1997) A mixture model for representing shape variation. In *British Machine Vision Conference*. 499

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995) Active shape models – their training and application. *Computer Vision & Image Understanding* **61** (1): 38–59. 499

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001) *Introduction to Algorithms*. MIT Press, 2nd edition. 270, 316

Coughlan, J., Yuille, A., English, C., & Snow, D. (2000) Efficient deformable template detection and localization without user interaction. *Computer Vision & Image Understanding* **78** (3): 303–319. 274

Cristianini, M., & Shawe-Taylor, J. (2000) *An Introduction to Support Vector Machines*. Cambridge University Press. 209

Csurka, G., Dance, C., Fan, L., Willamowski, J., & Bray, C. (2004) Visual categorization with bags of keypoints. In *ECCV International Workshop on Statistical Learning in Computer Vision*. 209, 574, 575, 595

Cuzzolin, F. (2006) Using bilinear models for view-invariant action and identity recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 1701–1708. 534

Dalal, N., & Triggs, B. (2005) Histograms of oriented gradients for human detection. In *IEEE Computer Vision & Pattern Recognition*, pp. 886–893. 352

Davison, A. J., Reid, I. D., Molton, N., & Stasse, O. (2007) MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (6): 1052–1067. 564, 565, 567

de Aguiar, E., Stoll, C., Theobalt, C., Ahmed, N., Seidel, H.-P., & Thrun, S. (2008) Performance capture from sparse multi-view video. *ACM Transactions on Graphics* **27** (3). 385

de La Gorce, M., Paragios, N., & Fleet, D. J. (2008) Model-based hand tracking with texture, shading and self-occlusions. In *IEEE Computer Vision & Pattern Recognition*.

De La Torre, F. (2011) A least-squares framework for component analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence* . 352, 533

De Ridder, D., & Franc, V. (2003) Robust subspace mixture models using t-distributions. In *British Machine Vision Conference*, pp. 319–328. 140

Delaitre, V., Laptev, I., & Sivic, J. (2010) Recognizing human actions in still images: a study of bag-of-features and part-based representations. In *British Machine Vision Conference*. 596

Delong, A., & Boykov, Y. (2009) Globally optimal segmentation of multi-region objects. In *IEEE International Conference on Computer Vision*, pp. 285–292. 317

Dempster, A. P., Laird, M. N., & Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39** (1): 1–38. 140

Deng, J., Berg, A., Li, K., & Fei-Fei, L. (2010) What does classifying more than 10,000 image categories tell us? *European Conference on Computer Vision* pp. 71–84. 595

Deng, Y., & Lin, X. (2006) A fast line segment based stereo algorithm using tree dynamic programming. In *European Conference on Computer Vision*, pp. 201–212. 274

Deutscher, J., Blake, A., & Reid, I. D. (2000) Articulated body motion capture by annealed particle filtering. In *IEEE Computer Vision & Pattern Recognition*, pp. 2126–2133. 500

Devernay, F., & Faugeras, O. D. (2001) Straight lines have to be straight. *Machine Vision Applications* **13** (1): 14–24. 385

Dollár, P., Tu, Z., & Belongie, S. (2006) Supervised learning of edges and object boundaries. In *IEEE Computer Vision & Pattern Recognition*, pp. 1964–1971. 209, 352

Domke, J., Karapurkar, A., & Aloimonos, Y. (2008) Who killed the directed model? In *IEEE Computer Vision & Pattern Recognition*. 317

Duda, R. O., Hart, P. E., & Stork, D. G. (2001) *Pattern Classification*. John Wiley and Sons, 2nd edition. 97

Durrant-Whyte, H., & Bailey, T. (2006) Simultaneous localization and mapping (SLAM): part I. *Robotics & Automation Magazine, IEEE* **13** (2): 99–110. 568

Durrant-Whyte, H., Rye, D., & Nebot, E. (1996) Localisation of automatic guided vehicles. In *International Symposium on Robotics Research*, pp. 613–625. 567

Durrant-Whyte, H. F. (1988) Uncertain geometry in robotics. *IEEE Transactions on Robot Automation* **4** (1): 23–31. 567

Edwards, G. J., Taylor, C. J., & Cootes, T. F. (1998) Interpreting face images using active appearance models. In *IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 300–305. 533

Efros, A. A., & Freeman, W. T. (2001) Image quilting for texture synthesis and transfer. In *SIGGRAPH*, pp. 341–346. 311, 312, 317

Efros, A. A., & Leung, T. K. (1999) Texture synthesis by non-parametric sampling. In *IEEE International Conference on Computer Vision*, pp. 1033–1038. 317

Eichner, M., & Ferrari, V. (2009) Better appearance models for pictorial structures. In *British Machine Vision Conference*. 274

Elder, J. H. (1999) Are edges incomplete? *International Journal of Computer Vision* **34** (2–3): 97–122. 336, 352

Elgammal, A. (2011) Figure-ground segmentation – pixel-based. In *Guide to visual analysis of humans: looking at people*, ed. by T. Moeslund, A. Hilton, Krüger, & L. Sigal. Springer. 97

Elgammal, A., Harwood, D., & Davis, L. (2000) Non-parametric model for background subtraction. In *European Conference on Computer Vision*, pp. 751–767. 97

Elgammal, A. M., & Lee, C.-S. (2004) Separating style and content on a nonlinear manifold. In *IEEE Computer Vision & Pattern Recognition*, pp. 478–485. 534

Engels, C., Stewénius, H., & Nistér, D. (2006) Bundle adjustment rules. *Photogrammetric Computer Vision* . 453

Everingham, M., Sivic, J., & Zisserman, A. (2006) Hello! My name is ... Buffy – automatic naming of characters in TV video. In *British Machine Vision Conference*, pp. 889–908. 274, 533

Everingham, M., Van Gool, L., Williams, C., Winn, J., & Zisserman, A. (2010) The PASCAL visual object classes (VOC) challenge. *International Journal of Computer Vision* **88** (2): 303–338. 595

Faugeras, O. (1993) *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press. 385

Faugeras, O., Luong, Q., & Papadopoulo, T. (2001) *The Geometry of Multiple Images*. MIT PRESS. 385, 453

Faugeras, O. D. (1992) What can be seen in three dimensions with an uncalibrated stereo rig. In *European Conference on Computer Vision*, pp. 563–578. 453

Faugeras, O. D., & Keriven, R. (1998) Variational principles, surface evolution, PDEs, level set methods, and the stereo problem. *IEEE Transactions on Image Processing* **7** (3): 336–344. 454

Faugeras, O. D., Luong, Q.-T., & Maybank, S. J. (1992) Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision*, pp. 321–334. 453

Felzenszwalb, P., & Zabih, R. (2011) Dynamic programming and graph algorithms in computer vision. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **33** (4): 721–740. 272, 274, 316

Felzenszwalb, P. F., Girshick, R. B., McAllester, D. A., & Ramanan, D. (2010) Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **32** (9): 1627–1645. 274

Felzenszwalb, P. F., & Huttenlocher, D. P. (2005) Pictorial structures for object recognition. *International Journal of*

Computer Vision **61** (1): 55–79. 270, 271, 272, 274, 500

Felzenszwalb, P. F., & Veksler, O. (2010) Tiered scene labeling with dynamic programming. In *IEEE Computer Vision & Pattern Recognition*. 274, 317

Ferencz, A., Learned-Miller, E. G., & Malik, J. (2008) Learning to locate informative features for visual identification. *International Journal of Computer Vision* **77** (1–3): 3–24. 534

Fischler, M., & Bolles, R. (1981) Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM* **24** (6): 381–395. 420

Fischler, M. A., & Erschlager, R. A. (1973) The representation and matching of pictorial structures. *IEEE Transactions on Computers* **22** (1): 67–92. 274

Fitzgibbon, A. W., & Zisserman, A. (1998) Automatic camera recovery for closed or open image sequences. In *European Conference on Computer Vision*, pp. 311–326. 453

Ford, L., & Fulkerson, D. (1962) *Flows in Networks*. Princeton University Press. 316

Forssén, P.-E., & Lowe, D. G. (2007) Shape descriptors for maximally stable extremal regions. In *IEEE International Conference on Computer Vision*, pp. 1–8. 352

Förstner, W. (1986) A feature-based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing* **26** (3): 150–166. 352

Forsyth, D. A., Arikan, O., Ikemoto, L., O'Brien, J., & Ramanan, D. (2006) Computational studies of human motion: Part 1, Tracking and motion synthesis. *Foundations and Trends in Computer Graphics and Computer Vision* **1** (2/2): 77–254. 500

Frahm, J.-M., Georgel, P. F., Gallup, D., Johnson, T., Raguram, R., Wu, C., Jen, Y.-H., Dunn, E., Clipp, B., & Lazebnik, S. (2010) Building Rome on a cloudless day. In *European Conference on Computer Vision*, pp. 368–381. 450

Frahm, J.-M., & Pollefeys, M. (2006) RANSAC for (quasi-)degenerate data (QDEGSAC). In *IEEE Computer Vision & Pattern Recognition*, pp. 453–460. 420

Franco, J.-S., & Boyer, E. (2005) Fusion of multi-view silhouette cues using a space occupancy grid. In *IEEE International Conference on Computer Vision*, pp. 1747–1753. 385

Freeman, W. T., Pasztor, E. C., & Carmichael, O. T. (2000) Learning low-level vision. *International Journal of Computer Vision* **40**: 25–47. 275, 311, 316

Freifeld, O., Weiss, A., Zuffi, S., & Black, M. J. (2010) Contour people: A parameterized model of 2D articulated human shape. In *IEEE Computer Vision & Pattern Recognition*, pp. 639–646. 499

Freiman, M., Kronman, A., Esses, S. J., Joskowicz, L., & Sosna, J. (2010) Non-parametric iterative model constraint graph min-cut for automatic kidney segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, pp. 73–80. 317

Freund, Y., & Schapire, R. (1996) Experiments with a new boosting algorithm. In *International conference on machine learning*, pp. 148–156. 209

Freund, Y., & Schapire, R. E. (1995) A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory: Eurocolt '95*, pp. 23–37. 209

Frey, B., Kschischang, F., Loeliger, H., & Wiberg, N. (1997) Factor graphs and algorithms. In *Allerton Conference on Communication, Control and Computing*. 275

Frey, B. J., & Jojic, N. (1999a) Estimating mixture models of images and inferring spatial transformations using the EM algorithm. In *IEEE Computer Vision & Pattern Recognition*, pp. 416–422. 140

Frey, B. J., & Jojic, N. (1999b) Transformed component analysis: joint estimation of spatial transformations and image components. In *IEEE Computer Vision & Pattern Recognition*, pp. 1190–1196,. 140

Friedman, J., Hastie, T., & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Annals of Statistics* **28** (2): 337–407. 209

Friedman, J. H. (1999) Greedy function approximation: A gradient boosting machine. Technical report, Department of Statistics, Stanford University. 209

Friedman, N., & Russell, S. J. (1997) Image segmentation in video sequences: a probabilistic approach. In *Uncertainty in Artificial Intelligence*, pp. 175–181. 97

Fua, P., & Leclerc, Y. G. (1995) Object-centered surface reconstruction: Combining multi-image stereo and shading. *International Journal of Computer Vision* **16** (1): 35–56. 454

Fusiello, A., Trucco, E., & Verri, A. (2000) A compact algorithm for rectification of stereo pairs. *Mach. Vis. Appl.* **12** (1): 16–22. 453

Gao, X.-S., Hou, X., Tang, J., & Cheng, H.-F. (2003) Complete solution classification for the perspective-three-point problem. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **25** (8): 930–943. 385

Gavrila, D., & Davis, L. S. (1996) 3-D model-based tracking of humans in action: a multi-view approach. In *IEEE Computer Vision & Pattern Recognition*, pp. 73–80. 500

Geiger, B., Ladendorf, B., & Yuille, A. (1992) Occlusions and binocular stereo. In *European Conference on Computer Vision*, pp. 425–433. 274

Geiger, D., Gupta, A., Costa, L. A., & Vlontzos, J. (1995) Dynamic-programming for detecting, tracking and matching deformable contours. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **17** (3): 294–302. 275

Gelb, A. (1974) *Applied Optimal Estimation*. MIT Press. 567

Gelman, A., Carlin, J. B., Stern, H. S. ., & Rubin, D. B. (2004) *Bayesian Data Analysis*. Chapman and Hall / CRC. 45, 65

Geman, S., & Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **6** (6): 721–741. 316

Geyer, C., & Daniilidis, K. (2001) Catadioptric projective geometry. *International Journal of Computer Vision* **45** (3): 223–243. 385

Ghahramani, Z. (2001) An introduction to hidden Markov models and Bayesian networks. In *Hidden Markov models: applications in computer vision*, ed. by B. H. Juang, pp. 9–42. World Scientific Publishing. 274

Ghahramani, Z., & Hinton, G. (1996a) Parameter estimation for linear dynamical systems. Technical Report CRG–TR–96–2, Department of Computer Science, University of Toronto. 567

Ghahramani, Z., & Hinton, G. (1996b) Switching state-space models. Technical Report CRG–TR–96–3, Department of Computer Science, University of Toronto. 567

Ghahramani, Z., & Hinton, G. E. (1996c) The EM algorithm for mixtures of factor analyzers. Technical Report CRG–TR–96–1, University of Toronto. 140

Goldberg, A., & Tarjan, R. (1988) A new approach to the maximum flow problem. *Journal of the Association for Computing Machinery* **35** (4): 921–940. 316

Golomb, B. A., Lawrence, D. T., & Sejnowski, T. (1990) SEXNET: a neural network identifies sex from human faces. In *Advances in Neural Information Processing Systems*, pp. 572–579. 209

Gong, M., & Yang, Y. H. (2005) Near real-time reliable stereo matching using programmable graphics hardware. In *IEEE Computer Vision & Pattern Recognition*, pp. 924–931. 274

Gonzalez, R. C., & Woods, R. E. (2002) *Digital Image Processing*. Prentice Hall, 2nd edition. 352

Gower, J. C., & Dijksterhuis, G. B. (2004) *Procrustes Problems*. Oxford University Press. 420

Grady, L. (2006) Random walks for image segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (11): 1768–1783. 317

Grauman, K., & Darrell, T. (2005) The pyramid match kernel: Discriminative classification with sets of image features. In *IEEE International Conference on Computer Vision*, pp. 1458–1465. 595

Grauman, K., Shakhnarovich, G., & Darrell, T. (2003) A Bayesian approach to image-based visual hull reconstruction. In *IEEE Computer Vision & Pattern Recognition*, pp. 187–194. 385

Greig, D. M., Porteous, B. T., & Seheult, A. H. (1989) Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B* **51** (2): 271–279. 316

Grimes, D. B., Shon, A. P., & Rao, R. P. N. (2003) Probabilistic bilinear models for appearance-based vision. In *IEEE International Conference on Computer Vision*, pp. 1478–1485. 534

Gross, R., Matthews, I., & Baker, S. (2002) Eigen light-fields and face recognition across pose. In *FGR*, pp. 3–9. 533

Guesebroek, J. M., Boughouts, G. J., & Smeulders, A. W. M. (2005) A.W.M.: The Amsterdam library of object images. *International Journal of Computer Vision* **61** (1): 103–112. 134

Guillaumin, M., Verbeek, J. J., & Schmid, C. (2009) Is that you? metric learning approaches for face identification. In *IEEE International Conference on Computer Vision*, pp. 498–505. 534

Guivant, J. E., & Nebot, E. M. (2001) Optimization of the simultaneous localization and map-building algorithm for real-time implementation. *IEEE Transactions on Robotics* **17** (3): 242–257. 567

Handa, A., Chli, M., Strasdat, H., & Davison, A. J. (2010) Scalable active matching. In *IEEE Computer Vision & Pattern Recognition*, pp. 1546–1553. 567

Harris, C. (1992) Tracking with rigid objects. In *Active Vision*, ed. by A. Blake & A. L. Yuille, pp. 59–73. 420

Harris, C., & Stephens, M. J. (1988) A combined corner and edge detector. In *Alvey Vision Conference*, pp. 147–152. 352, 453

Harris, C. G., & Pike, J. M. (1987) 3D positional integration from image sequences. In *Alvey Vision Conference*, pp. 233–236. 567

Hartley, R. I. (1992) Estimation of relative camera positions for uncalibrated cameras. In *European Conference on Computer Vision*, pp. 579–587. 453

Hartley, R. I. (1994) Projective reconstruction from line correspondence. In *IEEE Computer Vision & Pattern Recognition*, pp. 579–587. 453

Hartley, R. I. (1997) In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **19** (6): 580–593. 434, 453

Hartley, R. I., & Gupta, R. (1994) Linear pushbroom cameras. In *European Conference on Computer Vision*, pp. 555–566. 385

Hartley, R. I., & Zisserman, A. (2004) *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, 2nd edition. 19, 369, 385, 406, 420, 431, 440, 453

He, X., Yan, S., Hu, Y., Niyogi, P., & Zhang, H. (2005) Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **27** (3): 328–340. 533

He, X., Zemel, R. S., & Carreira-Perpiñán, M. Á. (2004) Multiscale conditional random fields for image labeling. In *IEEE Computer Vision & Pattern Recognition*, pp. 695–702. 209, 210

He, X., Zemel, R. S., & Ray, D. (2006) Learning and incorporating top-down cues in image segmentation. In *European Conference on Computer Vision*, pp. 338–351. 210

Heap, T., & Hogg, D. (1996) Towards 3D hand tracking using a deformable model. In *IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 140–145. 500

Heeger, D., & Bergen, J. (1995) Pyramid-based texture analysis/synthesis. In *Computer graphics and interactive techniques*, pp. 229–238. ACM. 317

Heess, N., Williams, C. K. I., & Hinton, G. E. (2009) Learning generative texture models with extended fields of experts. In *British Machine Vision Conference*. 317

Hernández, C., & Schmitt, F. (2004) Silhouette and stereo fusion for 3D object modeling. *Computer Vision & Image Understanding* **96** (3): 367–392. 454

Hinton, G. E. (2002) Training products of experts by minimizing contrastive divergence. *Neural Computation* **14** (8): 1771–1800. 239

Hirschmüller, H. (2005) Accurate and efficient stereo processing by semi-global matching and mutual information. In *IEEE Computer Vision & Pattern Recognition*, pp. 807–814. 318

Hirschmüller, H., & Scharstein, D. (2009) Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **31** (9): 1582–1599. 318

Hofmann, T. (1999) Probabilistic latent semantic analysis. In *Uncertainty in artificial intelligence*, pp. 289–296. 595

Hogg, D. (1983) Model-based vision: a program to see a walking person. *Image and Vision Computing* **1** (1): 5–20. 500

Hoiem, D., Efros, A., & Hebert, M. (2005) Automatic photo pop-up. *ACM Transactions on Graphics (SIGGRAPH)* **24** (3): 577–584. 206

Hoiem, D., Efros, A. A., & Hebert, M. (2007) Recovering surface layout from an image. *International Journal of Computer Vision* **75** (1): 151–172. 205, 206, 209

Horn, B. K. P. (1990) Relative orientation. *International Journal of Computer Vision* **4** (1): 59–78. 453

Horn, E., & Kiryati, N. (1999) Toward optimal structured light patterns. *Image and vision computing* **17** (2): 87–97. 385

Horprasert, T., Harwood, D., & Davis, L. S. (2000) A robust background subtraction and shadow detection. In *Asian Conference on Computer Vision*, pp. 983–988. 97

Hsu, R. L., Abdel-Mottaleb, M., & Jain, A. K. (2002) Face detection in color images. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (5): 696–707. 97

Huang, C., Ai, H., Li, Y., & Lao, S. (2007a) High-performance rotation invariant multi-view face detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (4): 671–686. 210

Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007b) Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report Technical Report 07–49, University of Massachusetts, Amherst. 529, 534

Huang, T. S., & Faugeras, O. D. (1989) Some properties of the E matrix in two-view motion estimation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **11** (12): 1310–1312. 453

Huang, Y., Liu, Q., & Metaxas, D. N. (2011) A component-based framework for generalized face alignment. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* **41** (1): 287–298. 491, 499

Huber, P. J. (2009) *Robust Statistics*. John Wiley and Sons, 2nd edition. 420

Humayun, A., Mac Aodha, O., & Brostow, G. J. (2011) Learning to find occlusion regions. In *CVPR*. 209

Ioffe, S. (2006) Probabilistic linear discriminant analysis. In *European Conference on Computer Vision*, pp. 531–542. 533

Isack, H., & Boykov, Y. (2012) Energy-based geometric multi-model fitting. *International Journal of Computer Vision* **97** (2): 123–147. 316, 416, 420

Isard, M., & Blake, A. (1998) A mixed-state CONDENSATION tracker with automatic model-switching. In *IEEE International Conference on Computer Vision*, pp. 107–112. 567

Ishikawa, H. (2003) Exact optimization for Markov random fields with convex priors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **25** (10): 1333–1336. 316

Ishikawa, H. (2009) Higher order clique reduction in binary graph cut. In *IEEE Computer Vision & Pattern Recognition*. 317

Jazwinski, A. H. (1970) *Stochastic Processes and Filtering Theory*. Academic Press. 567

Jegou, H., Douze, M., & Schmid, C. (2008) Recent advances in large scale image search. In *Emerging Trends in Visual Computing*, pp. 305–326. 595

Jégou, H., Douze, M., & Schmid, C. (2009) Packing bag-of-features. In *IEEE International Conference on Computer Vision*, pp. 2357–2364. 595

Jeong, Y., Nistér, D., Steedly, D., Szeliski, R., & Kweon, I.-S. (2010) Pushing the envelope of modern methods for bundle adjustment. In *IEEE Computer Vision & Pattern Recognition*, pp. 1474–1481. 453

Jiang, H., & Martin, D. R. (2008) Global pose estimation using non-tree models. In *IEEE Computer Vision & Pattern Recognition*. 274

Jojic, N., & Frey, B. J. (2001) Learning flexible sprites in video layers. In *IEEE Computer Vision & Pattern Recognition*, pp. 199–206. 140

Jojic, N., Frey, B. J., & Kannan, A. (2003) Epitomic analysis of appearance and shape. In *IEEE International Conference on Computer Vision*, pp. 34–41. 141

Jones, E., & Soatto, S. (2005) Layered active appearance models. In *IEEE International Conference on Computer Vision*, pp. 1097–1102. 499

Jones, M. J., & Rehg, J. M. (2002) Statistical color models with application to skin detection. *International Journal of Computer Vision* **46** (1): 81–96. 97, 140

Jordan, M. I. (2004) Graphical models. *Statistical science* **19** (1): 140–155. 239

Jordan, M. I., & Jacobs, R. A. (1994) Hierarchical mixtures of experts and the EM algorithm. *Neural Computation* **6** (2): 181–214. 211

Juan, O., & Boykov, Y. (2006) Active graph cuts. In *IEEE Computer Vision & Pattern Recognition*, pp. 1023–1029. 316

Julier, S., & Uhlmann, J. (1997) A new extension of the Kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, volume 3, p. 26. 567

Kadir, T., & Brady, M. (2001) Saliency, scale and image description. *International Journal of Computer Vision* **45** (2): 83–105. 352

Kakumanu, P., Makrogiannis, S., & Bourbakis, N. G. (2007) A survey of skin-colour modeling and detection methods. *Pattern Recognition* **40** (3): 1106–1122. 97

Kalman, R. E. (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* **82** (1): 35–45. 567

Kalman, R. E., & Bucy, R. S. (1961) New results in linear filtering and prediction theory. *Transactions of the American Society for Mechanical Engineering D* **83** (1): 95–108. 567

Kanade, T., Rander, P., & Narayanan, P. J. (1997) Virtualized reality: Constructing virtual worlds from real scenes. *IEEE MultiMedia* **4** (1): 34–47. 385

Kass, M., Witkin, A., & Terzopolous, D. (1987) Snakes: Active contour models. *International Journal of Computer Vision* **1** (4): 321–331. 275, 499

Kato, H., & Billinghurst, M. (1999) Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *International Workshop on Augmented Reality*, pp. 85–94. 420

Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., & Tachibana, K. (2000) Virtual object manipulation on a table-top AR environment. In *International Symposium on Augmented Reality*, pp. 111–119. 420

Kendall, D. G. (1984) Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society* **16** (2): 81–121. 462

Khan, Z., & Dellaert, F. (2004) Robust generative subspace modelling: the subspace t-distribution. Technical Report GIT–GVU–04–11, Georgia Institute of Technology. 140

Kim, J. C., Lee, K. M., Choi, B., & Lee, S. U. (2005) A dense stereo matching using two pass dynamic programming with generalized control points. In *IEEE Computer Vision & Pattern Recognition*, pp. 1075–1082. 274

Klein, G., & Murray, D. (2007) Parallel tracking and mapping for small AR workspaces. In *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*. 567

Kohli, P., Kumar, M. P., & Torr, P. H. S. (2009a) P3 & beyond: Move making algorithms for solving higher order functions. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **31** (9): 1645–1656. 317

Kohli, P., Ladicky, L., & Torr, P. H. S. (2009b) Robust higher order potentials for enforcing label consistency. *International Journal of Computer Vision* **82** (3): 302–324. 317

Kohli, P., & Torr, P. H. S. (2005) Efficiently solving dynamic Markov random fields using graph cuts. In *IEEE International Conference on Computer Vision*, pp. 922–929. 316

Kolev, K., & Cremers, D. (2008) Integration of multiview stereo and silhouettes via convex functionals on convex domains. In *European Conference on Computer Vision*, pp. 752–765. 454

Koller, D., & Friedman, N. (2009) *Probabilistic graphical models*. MIT Press. 220, 226, 239, 275

Koller, D., Klinker, G., Rose, E., Breen, D., Whitaker, R., & Tuceryan, M. (1997) Real-time vision-based camera tracking for augmented reality applications. In *ACM Symposium on Virtual Reality Software and Technology*, pp. 87–94. 420

Kolmogorov, V. (2006) Convergent tree-reweighted message passing for energy minimization. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (10): 1568–1583. 317

Kolmogorov, V., Criminisi, A., Blake, A., Cross, G., & Rother, C. (2006) Probabilistic fusion of stereo with color and contrast for bi-layer segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (9): 1480–1492. 316

Kolmogorov, V., & Rother, C. (2007) Minimizing non-submodular graph functions with graph-cuts – a review. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (7): 1274–1279. 316, 317

Kolmogorov, V., & Zabih, R. (2001) Computing visual correspondence with occlusions via graph cuts. In *IEEE International Conference on Computer Vision*, pp. 508–515. 308, 316, 317

Kolmogorov, V., & Zabih, R. (2002) Multi-camera scene reconstruction via graph cuts. In *European Conference on Computer Vision*, pp. 82–96. 316

Kolmogorov, V., & Zabih, R. (2004) What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis & Machine Intelligence* **26** (2): 147–159. 316

Komodakis, N., Tziritas, G., & Paragios, N. (2008) Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. *Computer Vision & Image Understanding* **112** (1): 14–29. 316

Kotz, S., & Nadarajah, S. (2004) *Multivariate t distributions and their applications*. Cambridge University Press. 140

Kschischang, F. R., Frey, B., & Loeliger, H. A. (2001) Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory* **47** (2): 498–519. 275

Kumar, M. P., Torr, P., & Zisserman, A. (2004) Extending pictorial structures for object recognition. In *British Machine Vision Conference*. 274

Kumar, M. P., Torr, P. H. S., & Zisserman, A. (2005) OBJ CUT. In *IEEE Computer Vision & Pattern Recognition*, pp. 18–25. 317

Kumar, M. P., Veksler, O., & Torr, P. H. S. (2011) Improved moves for truncated convex models. *Journal of Machine Learning Research* **12**: 31–67. 316

Kumar, N., Belhumeur, P., & Nayar, S. K. (2008) Face tracer: A search engine for large collections of images with faces. In *European Conference on Computer Vision*. 209, 210

Kumar, N., Berg, A. C., Belhumeur, P. N., & Nayar, S. K. (2009) Attribute and simile classifiers for face verification. In *IEEE International Conference on Computer Vision*, pp. 365–372. 534

Kumar, S., & Hebert, M. (2003) Discriminative random fields: A discriminative framework for contextual interaction in classification. In *IEEE International Conference on Computer Vision*, pp. 1150–1159. 316

Kutulakos, K. N., & Seitz, S. M. (2000) A theory of shape by space carving. *International Journal of Computer Vision* **38** (3): 199–218. 454

Kwatra, V., Schödl, A., Essa, I., Turk, G., & Bobick, A. (2003) Graphcut textures: Image and video synthesis using graph cuts. *ACM Transactions on Graphics (SIGGRAPH 2003)* **22** (3): 277–286. 316, 317

Lanitis, A., Taylor, C. J., & Cootes, T. F. (1997) Automatic interpretation and coding of face images using flexible models. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **19** (7): 743–756. 499

Laptev, I., Marszałek, M., Schmid, C., & Rozenfeld, B. (2008) Learning realistic human actions from movies. In *CVPR*. 592, 593, 595

Laurentini, A. (1994) The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **16** (2): 150–162. 385

Lawrence, N. D. (2004) Probabilistic non-linear principal component analysis with Gaussian process latent variable models. Technical Report CS–04–08, University of Sheffield. 140, 352

Lawrence, N. D. (2005) Probabilistic non-linear principal component analysis with Gaussian process latent variable models. *Journal of Machine Learning Research* **6**: 1783–1816. 499

Lazebnik, S., Schmid, C., & Ponce, J. (2006) Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *IEEE Computer Vision & Pattern Recognition*, pp. 2169–2178. 595

Lee, J., Moghaddam, B., Pfister, H., & Machiraju, R. (2005) A bilinear illumination model for robust face recognition. In *IEEE International Conference on Computer Vision*, pp. 1177–1184. 534

Lempitsky, V., Blake, A., & Rother, C. (2008) Image segmentation by branch-and-mincut. In *European Conference on Computer Vision*, pp. 15–29. 317

Lempitsky, V., Rother, C., Roth, S., & Blake, A. (2010) Fusion moves for Markov random field optimization. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **32** (8): 1392–1405. 316

Leonard, J. J., & Feder, H. J. S. (2000) A computational efficient method for large-scale concurrent mapping and localisation. In *International Symposium on Robotics Research*, pp. 169–176. 567

Leordeanu, M., Hebert, M., & Sukthankar, R. (2007) Beyond local appearance: Category recognition from pairwise interactions of simple features. In *IEEE Computer Vision & Pattern Recognition*. 499

Leotta, M. J., & Mundy, J. L. (2011) Vehicle surveillance with a generic, adaptive, 3D vehicle model. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **33** (7): 1457–1469. 500

Lepetit, V., & Fua, P. (2005) Monocular model-based 3D tracking of rigid objects: A survey. *Foundations and Trends in Computer Graphics and Vision* **1** (1): 1–89. 420

Lepetit, V., & Fua, P. (2006) Keypoint recognition using randomized trees. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (9): 1465–1479. 420

Lepetit, V., Lagger, P., & Fua, P. (2005) Randomized trees for real-time keypoint recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 775–781. 209, 417

Lepetit, V., Moreno-Noguer, F., & Fua, P. (2009) EPnP: an accurate $O(n)$ solution to the PnP problem. *International Journal of Computer Vision* **81** (2): 155–166. 385

Leventon, M. E., Grimson, W. E. L., & Faugeras, O. D. (2000) Statistical shape influence in geodesic active contours. In *IEEE Computer Vision & Pattern Recognition*, pp. 1316–1323. 499

Levin, A., Lischinski, D., & Weiss, Y. (2004) Colorization using optimization. *ACM Transactions on Graphics* **23** (3): 689–694. 316

Lhuillier, M., & Quan, L. (2002) Match propagation for image-based modeling and rendering. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (8): 1140–1146. 318

Li, F.-F., & Perona, P. (2005) A Bayesian hierarchical model for learning natural scene categories. In *IEEE Computer Vision & Pattern Recognition*, pp. 524–531. 595, 596

Li, J., & Allinson, N. M. (2008) A comprehensive review of current local features for computer vision. *Neurocomputing* **71** (10-12): 1771–1787. 352

Li, P., Fu, Y., Mohammed, U., Elder, J., & Prince, S. J. D. (2011) Probabilistic models for inference about identity. *IEEE Transactions on Pattern Analysis & Machine Intelligence* . 516, 528, 529, 533, 534

Li, P., Warrell, J., Aghajanian, J., & Prince, S. (2010) Context-based additive logistic model for facial keypoint localization. In *British Machine Vision Conference*, pp. 1–11. 533

Li, S. Z. (2010) *Markov Random Field Modeling in Image Analysis*. Springer, 3rd edition. 316

Li, S. Z., & Jain, A. K. eds. (2005) *Handbook of face recognition*. Springer. 533

Li, S. Z., & Zhang, Z. (2004) Floatboost learning and statistical face detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **26** (9): 1112–1123. 210

Li, S. Z., Zhang, Z. Q., Shum, H. Y., & Zhang, H. J. (2003) Floatboost learning for classification. In *Advances in Neural Information Processing Systems*, pp. 993–1000. 209

Li, S. Z., Zhuang, Z., Blake, A., Zhang, H., & Shum, H. (2002) Statistical learning of multi-view face detection. In *European Conference on Computer Vision*, pp. 67–82. 210

Li, Y., Du, Y., & Lin, X. (2005) Kernel-based multifactor analysis for image synthesis and recognition. In *IEEE International Conference on Computer Vision*, pp. 114–119. 534

Li, Y., Sun, J., Tang, C.-K., & Shum, H.-Y. (2004) Lazy snapping. *ACM Transactions on Graphics* **23** (3): 303–308. 317

Lienhart, R., Kuranov, A., & Pisarevsky, V. (2003) Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In *Deutsche Arbeitsgemeinschaft für Mustererkennung*, pp. 297–304. 210

Liu, C., & Rubin, D. B. (1995) ML estimation of the t distribution using EM and its extensions ECM and ECME. *Statistica Sinica* **5** (1): 19–39. 140

Liu, C., & Shum, H. Y. (2003) Kullback-Leibler boosting. In *IEEE Computer Vision & Pattern Recognition*, pp. 407–411. 209

Liu, C., & Wechsler, H. (1998) Probabilistic reasoning models for face recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 827–832. 533

Liu, J., Sun, J., & Shum, H.-Y. (2009) Paint selection. *ACM Transactions on Graphics* **28** (3). 317

Longuet-Higgins, H. C. (1981) A computer algorithm for reconstructing a scene from two projections. *Nature* **293**: 133–135. 453

Loop, C. T., & Zhang, Z. (1999) Computing rectifying homographies for stereo vision. In *IEEE Computer Vision & Pattern Recognition*, pp. 1125–1131. 453

Lourakis, M. I. A., & Argyros, A. A. (2009) SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software* **36** (1). 453

Lowe, D. G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60** (2): 91–110. 352, 453, 534

Loxam, J., & Drummond, T. (2008) Student-t mixture filter for robust, real-time visual tracking. In *European Conference on Computer Vision*, pp. 372–385. 140

Lu, S., Metaxas, D. N., Samaras, D., & Oliensis, J. (2003) Using multiple cues for hand tracking and model refinement. In *IEEE Computer Vision & Pattern Recognition*, pp. 443–450. 500

Lucas, B. D., & Kanade, T. (1981) An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pp. 647–679. 453

Lucey, S., & Chen, T. (2006) Learning patch dependencies for improved pose mismatched face verification. In *IEEE Computer Vision & Pattern Recognition*, pp. 909–915. 533

Ma, Y., Derksen, H., Hong, W., & Wright, J. (2007) Segmentation of multivariate mixed data via lossy data coding and compression. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (9): 1546–1562. 140

Ma, Y., Soatto, S., & Kosecká, J. (2004) *An Invitation to 3-D Vision*. Springer. 385, 453

Mac Aodha, O., Brostow, G. J., & Pollefeys, M. (2010) Segmenting video into classes of algorithm-suitability. In *IEEE Computer Vision & Pattern Recognition*. 209

Mackay, D. J. (2003) *Information theory, learning and inference algorithms*. Cambridge University Press. 65

Makadia, A. (2010) Feature tracking for wide-baseline image retrieval. In *European Conference on Computer Vision*, pp. 310–323. 595

Mäkinen, E., & Raisamo, R. (2008a) Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (3): 541–547. 210

Mäkinen, E., & Raisamo, R. (2008b) An experimental comparison of gender classification methods. *Pattern recognition methods* **29** (10): 1544–1556. 210

Malcolm, J. G., Rathi, Y., & Tannenbaum, A. (2007) Graph cut segmentation with nonlinear shape priors. In *IEEE International Conference on Image Processing*, pp. 365–368. 317

Malladi, R., Sethian, J. A., & Vemuri, B. C. (1994) Evolutionary fronts for topology-independent shape modeling and recovery. In *European Conference on Computer Vision*, pp. 3–13. 499

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2002) Robust wide baseline stereo from maximally stable extremal regions. In *British Machine Vision Conference*. 352

Matthews, I., & Baker, S. (2004) Active appearance models revisited. *International Journal of Computer Vision* **60** (2): 135–164. 499

Matthews, I., Cootes, T. F., Bangham, J. A., Cox, S., & Harvey, R. (2002) Extraction of visual features for lipreading. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (2): 198–213. 499

Matthews, I., Xiao, J., & Baker, S. (2007) 2D vs. 3D deformable face models: Representational power, construction, and real-time fitting. *International Journal of Computer Vision* **75** (1): 93–113. 496, 499, 500

Matthies, L., Kanade, T., & Szeliski, R. (1989) Kalman filter-based algorithms for estimating depth from image sequences. *International Journal of Computer Vision* **3** (3): 209–238. 567

Matusik, W., Buehler, C., Raskar, R., Gortler, S. J., & McMillan, L. (2000) Image-based visual hulls. In *SIGGRAPH*, pp. 369–374. 385

Maybank, S. J. (1998) *Theory of reconstruction from image motion*. Springer-Verlag. 453

Maybeck, P. S. (1990) The Kalman filter: an introduction to concepts. In *Autonomous Robot Vehicles*, ed. by I. J. Cox & G. T. Wilfong, pp. 194–204. Springer-Verlag. 567

McLachlan, G. J., & Krishnan, T. (2008) *The EM algorithm and extensions*. Wiley, 2nd edition. 140

Mei, C., Sibley, G., Cummins, M., Newman, P., & Reid, I. (2009) A constant time efficient stereo SLAM system. In *British Machine Vision Conference*. 567

Meir, R., & Mätsch, G. (2003) An introduction to boosting and leveraging. In *Advanced Lectures on Machine Learning*, ed. by S. Mendelson & A. Smola, pp. 119–184. 209

Messer, K., Matas, J., Kittler, J., Luettin, J., & Maitre, G. (1999) XM2VTS: The extended M2VTS database. In *Conference on Audio and Video-base Biometric Personal Verification*, pp. 72–77. 534

Mikolajczyk, K., & Schmid, C. (2002) An affine invariant interest point detector. In *European Conference on Computer Vision*, pp. 128–142. 352

Mikolajczyk, K., & Schmid, C. (2004) Scale & affine invariant interest point detectors. *International Journal of Computer Vision* **60** (1): 63–86. 352

Mikolajczyk, K., & Schmid, C. (2005) A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **27** (10): 1615–1630. 352

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., & Gool, L. J. V. (2005) A comparison of affine region detectors. *International Journal of Computer Vision* **65** (1-2): 43–72. 352

Mikulík, A., Perdoch, M., Chum, O., & Matas, J. (2010) Learning a fine vocabulary. In *European Conference on Computer Vision*, pp. 1–14. 595

Minka, T. (2002) Bayesian inference in dynamic models: an overview. Technical report, Carnegie Mellon University. 567

Moeslund, T. B., Hilton, A., & Krüger, V. (2006) A survey of advances in vision-based human motion capture and analysis. *Computer Vision & Image Understanding* **104** (2–3): 90–126. 500

Moghaddam, B., Jebara, T., & Pentland, A. (2000) Bayesian face recognition. *Pattern Recognition* **33** (11): 1771–1782. 533, 535

Moghaddam, B., & Pentland, A. (1997) Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **19** (7): 696–710. 140, 533

Moghaddam, B., & Yang, M. H. (2002) Learning gender with support faces. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (5): 707–711. 209

Mohammed, U., Prince, S. J. D., & Kautz, J. (2009) Visio-lization: generating novel facial images. *ACM Transactions on Graphics (SIGGRAPH)* **28** (3). 314, 315

Moni, M. A., & Ali, A. B. M. S. (2009) HMM based hand gesture recognition: a review on techniques and approaches. In *Int. Conf. on Computer Science and Information Technology*, pp. 433–437. 274

Montemerlo, M., Thrun, S., Koller, D., & Wegbreit, B. (2002) FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *AAAI/IAAI*, pp. 593–598. 567

Montemerlo, M., Thrun, S., Koller, D., & Wegbreit, B. (2003) FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *International Joint Conference on Uncertainty in Artificial Intelligence*, pp. 1151–1156. 567

Moons, T. (1998) A guided tour through multiview relations. In *SMILE*, ed. by R. Koch & L. J. V. Gool, volume 1506 of *Lecture Notes in Computer Science*, pp. 304–346. Springer. 453

Moons, T., Van Gool, L. J., & Vergauwen, M. (2009) 3D reconstruction from multiple images: Part 1 – principles. *Foundations and Trends in Computer Graphics and Vision* **4** (4): 287–404. 453

Moore, A. P., Prince, S. J. D., & Warrell, J. (2010) "Lattice Cut" – constructing superpixels using layer constraints. In *IEEE Computer Vision & Pattern Recognition*, pp. 2117–2124. 316, 317

Moore, A. P., Prince, S. J. D., Warrell, J., Mohammed, U., & Jones, G. (2008) Superpixel lattices. In *IEEE Computer Vision & Pattern Recognition*. 274

Moosmann, F., Nowak, E., & Jurie, F. (2008) Randomized clustering forests for image classification. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (9): 1632–1646. 209

Moosmann, F., Triggs, B., & Jurie, F. (2006) Fast discriminative visual codebooks using randomized clustering forests. In *Advances in Neural Information Processing Systems*, pp. 985–992. 209

Moravec, H. (1983) The Stanford cart and the CMU rover. *Proceedings of the IEEE* **71** (7): 872–884. 352

Mori, G., Ren, X., Efros, A. A., & Malik, J. (2004) Recovering human body configurations: Combining segmentation and recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 326–333. 500

Mundy, J., & Zisserman, A. (1992) *Geometric invariance in computer vision*. MIT Press. 385

Murase, H., & Nayar, S. K. (1995) Visual learning and recognition of 3D objects from appearance. *International Journal of Computer Vision* **14** (1): 5–24. 140

Murino, V., Castellani, U., Etrari, E., & Fusiello, A. (2002) Registration of very time-distant aerial images. In *IEEE International Conference on Image Processing*, pp. 989–992. 420

Murphy, K., Weiss, Y., & Jordan, M. (1999) Loopy belief propagation for approximate inference: an empirical study. In *Uncertainty in Artificial Intelligence*, pp. 467–475. 275

Murphy, K. P. (1998) Switching Kalman filters. Technical report. 567

Mĭcušík, B., & Pajdla, T. (2003) Estimation of omnidirectional camera model from epipolar geometry. In *IEEE Computer Vision & Pattern Recognition*, pp. 485–490. 385

Nadarajah, S., & Kotz, S. (2008) Estimation methods for the multivariate t distribution. *Acta Applicandae Mathematicae: An International Survey Journal on Applying Mathematics and Mathematical Applications* **102** (1): 99–118. 140

Navaratnam, R., Fitzgibbon, A. W., & Cipolla, R. (2007) The joint manifold model for semi-supervised multi-valued regression. In *IEEE International Conference on Computer Vision*, pp. 1–8. 169

Neal, R., & Hinton, G. (1999) A view of the EM algorithm that justifies incremental, sparse and other variants. In *Learning in Graphical Models*, ed. by M. I. Jordan. MIT PRess. 140

Newcombe, R. A., & Davison, A. J. (2010) Live dense reconstruction with a single moving camera. In *IEEE Computer Vision & Pattern Recognition*, pp. 1498–1505. 453

Newcombe, R. A., Lovegrove, S., & Davison, A. J. (2011) DTAM: Dense tracking and mapping in real-time. In *IEEE International Conference on Computer Vision*. 567

Newman, P., & Ho, K. L. (2005) SLAM – loop closing with visually salient features. In *IEEE International Conference on Robotics and Automation*. 568

Nguyen, H. V., & Bai, L. (2010) Cosine similarity metric learning for face verification. In *Asian Conference on Computer Vision*, pp. 709–720. 534

Niebles, J. C., Wang, H., & Li, F.-F. (2008) Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computer Vision* **79** (3): 299–318. 596

Nistér, D. (2004) An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **26** (6): 756–777. 453

Nistér, D., Naroditsky, O., & Bergen, J. R. (2004) Visual odometry. In *IEEE Computer Vision & Pattern Recognition*, pp. 652–659. 567

Nistér, D., & Stewénius, H. (2006) Scalable recognition with a vocabulary tree. In *IEEE Computer Vision & Pattern Recognition*, pp. 2161–2168. 595

Nixon, M., & Aguado, A. S. (2008) *Feature extraction and image processing*. Academic Press, 2nd edition. 19, 352

Nowak, E., & Jurie, F. (2007) Learning visual similarity measures for comparing never seen objects. In *IEEE Computer Vision & Pattern Recognition*. 534

O' Gorman, L., Sammon, M. J., & Seul, M. (2008) *Practical Algorithms for Image Analysis*. Cambridge University Press, 2nd edition. 352

Oh, S. M., Rehg, J. M., Balch, T. R., & Dellaert, F. (2005) Learning and inference in parametric switching linear dynamical systems. In *IEEE International Conference on Computer Vision*, pp. 1161–1168. 567

Ohta, Y., & Kanade, T. (1985) Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **7** (2): 139–154. 269, 274

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **24** (7): 971–987. 352, 534

Oliver, N., Rosario, B., & Pentland, A. (2000) A Bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **22** (8): 831–843. 97, 274

Opelt, A., Pinz, A., & Zisserman, A. (2006) A boundary-fragment-model for object detection. In *European Conference on Computer Vision*, pp. 575–588. 499

Osuna, E., Freund, R., & Girosi, F. (1997) Training support vector machines: an application to face detection. In *IEEE Computer Vision & Pattern Recognition*, pp. 746–751. 210

Özuysal, M., Calonder, M., Lepetit, V., & Fua, P. (2010) Fast keypoint recognition using random ferns. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **32** (3): 448–461. 420

Papoulis, A. (1991) *Probability, Random Variables and Stochastic Processes*. McGraw Hill, 3rd edition. 33

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann. 275

Peel, D., & McLachlan, G. (2000) Robust mixture modelling using the t distribution. *Statistics and Computing* **10** (4): 339–348. 140

Perlibakas, V. (2004) Distance measures for PCA-based face recognition. *Pattern Recognition Letters* **25** (6): 711–724. 533

Petersen, K. B., Pedersen, M. S., Larsen, J., Strimmer, K., Christiansen, L., Hansen, K., He, L., Thibaut, L., Baro, M., Hattinger, S., Sima, V., & The, W. (2006) The matrix cookbook. Technical report. 623

Pham, M., & Chan, T. (2007a) Fast training and selection of Haar features using statistics in boosting-based face detection. In *IEEE International Conference on Computer Vision*. 209, 210

Pham, M., & Chan, T. (2007b) Online learning asymmetric boosted classifiers for object detection. In *IEEE Computer Vision & Pattern Recognition*. 210

Philbin, J., Chum, O., Isard, M., Sivic, J., & Zisserman, A. (2007) Object retrieval with large vocabularies and fast spatial matching. In *IEEE Computer Vision & Pattern Recognition*. 595

Philbin, J., Isard, M., Sivic, J., & Zisserman, A. (2010) Descriptor learning for efficient retrieval. In *European Conference on Computer Vision*, pp. 677–691. 352, 595

Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000) The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **22** (10): 1090–1104. 534

Phung, S., Bouzerdoum, A., & Chai, D. (2005) Skin segmentation using color pixel classification: analysis and comparison. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **27** (1): 147–154. 97

Piccardi, M. (2004) Background subtraction techniques: a review. In *IEEE Int. Conf. Systems, Man and Cybernetics*, pp. 3099–3105. 97

Pinto, N., & Cox, D. (2011) Beyond simple features: A large- scale feature search approach to unconstrained face recognition. In *IEEE International Conference on Automatic Face & Gesture Recognition*. 534

Pollefeys, M. (2002) Visual 3D modeling from images. *On-line tutorial: http://www.cs.unc.edu/marc/tutorial* . 453

Pollefeys, M., Koch, R., & Van Gool, L. J. (1999a) Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision* **32** (1): 7–25. 449

Pollefeys, M., Koch, R., & Van Gool, L. J. (1999b) A simple and efficient rectification method for general motion. In *IEEE International Conference on Computer Vision*, pp. 496–501. 443, 453

Pollefeys, M., & Van Gool, L. J. (2002) Visual modelling: from images to images. *Journal of Visualization and Computer Animation* **13** (4): 199–209. 447, 448

Pollefeys, M., Van Gool, L. J., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J., & Koch, R. (2004) Visual modeling with a hand-held camera. *International Journal of Computer Vision* **59** (3): 207–232. 449, 453

Pons, J.-P., Keriven, R., & Faugeras, O. D. (2007) Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision* **72** (2): 179–193. 454

Poppe, R. (2007) Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding* **108** (1–2): 4–18. 500, 567

Portilla, J., & Simoncelli, E. (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision* **40** (1): 49–70. 317

Pratt, W. H. (2007) *Digital Image Processing*. Wiley Interscience, 3rd edition. 352

Prince, S., Cheok, A. D., Farbiz, F., Williamson, T., Johnson, N., Billinghurst, M., & Kato, H. (2002) 3D live: Real time captured content for mixed reality. In *International Symposium on Mixed and Augmented Reality*, pp. 317–324. 382, 383, 384, 385

Prince, S. J. D., & Aghajanian, J. (2009) Gender classification in uncontrolled settings using additive logistic models. In *IEEE International Conference on Image Processing*, pp. 2557–2560. 201, 210

Prince, S. J. D., Elder, J. H., Warrell, J., & Felisberti, F. M. (2008) Tied factor analysis for face recognition across large pose differences. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (6): 970–984. 522, 524, 533, 534

Prinzie, A., & Van den Poel, D. (2008) Random forests for multiclass classification: random multinomial logit. *Expert Systems with Applications* **35** (3): 1721–1732. 209

Pritch, Y., Kav-Venaki, E., & Peleg, S. (2009) Shift-map image editing. In *IEEE International Conference on Computer Vision*, pp. 151–158. 308, 309, 310, 316

Quan, L., & Lan, Z.-D. (1999) Linear n-point camera pose determination. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **21** (8): 774–780. 385

Rabiner, L. (1989) A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* **77** (2): 257–286. 274

Rae, R., & Ritter, H. (1998) Recognition of human head orientation based on artificial neural networks. *IEEE Transactions on Neural Networks* **9** (2): 257–265. 169

Raguram, R., Frahm, J.-M., & Pollefeys, M. (2008) A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus. In *European Conference on Computer Vision*, pp. 500–513. 420

Ramanan, D., Forsyth, D. A., & Zisserman, A. (2008) Tracking people by learning their appearance. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (1): 65–81. 274

Ranganathan, A. (2009) Semantic scene segmentation using random multinomial logit. In *British Machine Vision Conference*. 210

Ranganathan, A., & Yang, M. (2008) Online sparse matrix Gaussian process regression and vision applications. In *European Conference on Computer Vision*, pp. 468–482. 169

Raphael, C. (2001) Course-to-fine dynamic programming. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **23** (12): 1379–1390. 274

Rasmussen, C. E., & Williams, C. K. I. (2006) *Gaussian Processes for Machine Learning*. MIT Press. 169, 209

Rehg, J., & Kanade, T. (1994) Visual tracking of high DOF articulated structures: an application to human hand tracking. In *European Conference on Computer Vision*, pp. 35–46. 500

Rehg, J. M., & Kanade, T. (1995) Model-based tracking of self-occluding articulated objects. In *IEEE International Conference on Computer Vision*, pp. 612–617. 500

Rehg, J. M., Morris, D. D., & Kanade, T. (2003) Ambiguities in visual tracking of articulated objects using two- and three-dimensional models. *I. J. Robotic Res.* **22** (6): 393–418. 500

Rekimoto, J. (1998) MATRIX: A realtime object identification and registration method for augmented reality. In *Asia Pacific Computer Human Interaction*, pp. 63–69. 420

Ren, X., Berg, A. C., & Malik, J. (2005) Recovering human body configurations using pairwise constraints between parts. In *IEEE International Conference on Computer Vision*, pp. 824–831. 274

Rigoll, G., Kosmala, A., & Eickeler, S. (1998) High performance real-time gesture recognition using hidden Markov

models. In *Int. Workshop on Gesture and Sign language in Human-Computer Interaction*. 274

Rogez, G., Rihan, J., Ramalingam, S., Orrite, C., & Torr, P. (2006) Randomized trees for human pose detection. In *Advances in Neural Information Processing Systems*, pp. 985–992. 209

Romdhani, S., Cong, S., & Psarrou, A. (1999) A multi-view non-linear active shape model using kernel PCA. In *British Machine Vision Conference*. 499

Rosales, R., & Sclaroff, S. (1999) 3D trajectory recovery for tracking multiple objects and trajectory guided recognition of actions. In *IEEE Computer Vision & Pattern Recognition*, pp. 2117–2123. 563, 567

Rosen-Zvi, M., Griffiths, T. L., Steyvers, M., & Smyth, P. (2004) The author-topic model for authors and documents. In *Uncertainty in Artificial Intelligence*, pp. 487–494. 595

Rosenblatt, F. (1958) The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* **65** (6): 386–408. 209

Rosten, E., & Drummond, T. (2006) Machine learning for high-speed corner detection. In *European Conference on Computer Vision*, volume 1, pp. 430–443. 352

Roth, S., & Black, M. J. (2009) Fields of experts. *International Journal of Computer Vision* **82** (2): 205–229. 140, 317

Rother, C., Bordeaux, L., Hamadi, Y., & Blake, A. (2006) Autocollage. *ACM Transactions on Graphics* **25** (3): 847–852. 316

Rother, C., Kohli, P., Feng, W., & Jia, J. (2009) Minimizing sparse higher order energy functions of discrete variables. In *IEEE Computer Vision & Pattern Recognition*, pp. 1382–1389. 317

Rother, C., Kolmogorov, V., & Blake, A. (2004) Grabcut – interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (SIGGRAPH 2004)* **23** (3): 309–314. 316, 317

Rother, C., Kolmogorov, V., Lempitsky, V. S., & Szummer, M. (2007) Optimizing binary MRFs via extended roof duality. In *IEEE Computer Vision & Pattern Recognition*. 316

Rother, C., Kumar, S., Kolmogorov, V., & Blake, A. (2005) Digital tapestry. In *IEEE Computer Vision & Pattern Recognition*, pp. 589–586. 305, 306, 316

Rothwell, C. A., Zisserman, A., Forsyth, D. A., & Mundy, J. L. (1995) Planar object recognition using projective shape representation. *International Journal of Computer Vision* **16** (1): 57–99. 422

Rousseeuw, P. J. (1984) Least median of squares regression. *Journal of the American Statistical Association* **79** (388): 871–880. 420

Rousson, M., & Paragios, N. (2002) Shape priors for level set representations. In *European Conference on Computer Vision*, pp. 78–92. 499

Roweis, S., & Saul, L. (2000) Nonlinear dimensionality reduction by locally linear embedding. *Science* **290** (5500): 2323–2326. 352

Roweis, S. T., & Ghahramani, Z. (1999) A unifying review of linear Gaussian models. *Neural Computation* **11** (2): 305–345. 567

Roweis, S. T., & Ghahramani, Z. (2001) Learning nonlinear dynamical systems using the expectation-maximization algorithm. In *Kalman Filtering and Neural Networks*, ed. by S. Haykin, pp. 175–220. Wiley. 567

Rubin, D., & Thayer, D. (1982) EM algorithms for ML factor analysis. *Psychometrica* **47** (1): 69–76. 140, 499

Rumelhart, D. E., Hinton, G. E., & Williams, R. (1986) Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*, ed. by D. Rumelhart, J. McLelland, & The PDP Research Group, pp. 318–362. 209

Salmen, J., Schlipsing, M., Edelbrunner, J., Hegemann, S., & Lüke, S. (2009) Real-time stereo vision: Making more out of dynamic programming. In *Proceedings of the 13th International Conference on Computer Analysis of Images and Patterns*, CAIP '09, pp. 1096–1103. Springer-Verlag. 274

Salvi, J., Pags, J., & Batlle, J. (2004) Pattern codification strategies in structured light systems. *Pattern Recognition* **37** (4): 827 – 849. Agent Based Computer Vision. 385

Schaffalitzky, F., & Zisserman, A. (2002) Multi-view matching for unordered image sets, or "How do i organize my holiday snaps?". In *European Conference on Computer Vision*, pp. 414–431. 352

Schapire, R., & Singer, Y. (1998) Improved boosting algorithms using confidence-rated predictions. In *Conference on Computational Learning Theory*, pp. 80–91. 209

Scharstein, D., & Szeliski, R. (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* **47** (1): 7–42. 318

Scharstein, D., & Szeliski, R. (2003) High-accuracy depth maps using structured light. In *IEEE Computer Vision & Pattern Recognition*, pp. 194–202. 378, 379, 380, 385

Schlesinger, D., & Flach, B. (2006) Transforming an arbitrary minsum problem into a binary one. Technical Report TUD–FI06–01, Dresden University of Technology. 316

Schmugge, S. J., Jayaram, S., Shin, M., & Tsap, L. (2007) Objective evaluation of approaches to skin detection using roc analysis. *Computer Vision & Image Understanding* **108** (1–2): 41–51. 97

Schneiderman, H., & Kanade, T. (2000) A statistical method for 3D object detection applied to faces and cards. In *IEEE International Conference on Computer Vision*, pp. 746–751. 210

Schölkopf, B., Smola, A. J., & Müller, K.-R. (1997) Kernel principal component analysis. In *International Conf. on Artificial Neural Networks*, pp. 583–588. 352

Schölkopf, B., Smola, A. J., & Müller, K.-R. (1998) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation* **10** (5): 1299–1319. 499

Schüldt, C., Laptev, I., & Caputo, B. (2004) Recognizing human actions: A local SVM approach. In *International Conference on Pattern Recognition*, pp. 32–36. 592, 593, 595

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006) A comparison and evaluation of multi-view stereo reconstruction algorithms. In *IEEE Computer Vision & Pattern Recognition*, pp. 519–528. 454

Seo, H., & Magnenat-Thalmann, N. (2003) An automatic modeling of human bodies from sizing parameters. In *SI3D*, pp. 19–26. 500

Sfikas, G., Nikou, C., & Galatsanos, N. (2007) Robust image segmentation with mixtures of Student's t-distributions. In *IEEE International Conference on Image Processing*, pp. 273–276. 135, 136, 140

Sha'ashua, A., & Ullman, S. (1988) Structural saliency: the detection of globally salient structures using a locally connected network. In *IEEE International Conference on Computer Vision*, pp. 321–327. 274

Shakhnarovich, G., Viola, P. A., & Darrell, T. (2003) Fast pose estimation with parameter-sensitive hashing. In *IEEE International Conference on Computer Vision*, pp. 750–759. 500

Shan, C. (in press) Learning local binary patterns for gender classification on real-world face images. *Pattern Recognition Letters* . 210

Shepherd, B. (1983) An appraisal of a decision tree approach to image classification. In *International Joint Conferences on Artificial Intelligence*, pp. 473–475. 209

Shi, J., & Tomasi, C. (1994) Good features to track. In *IEEE Computer Vision & Pattern Recognition*, pp. 311–326. 453

Shotton, J., Blake, A., & Cipolla, R. (2008a) Multiscale categorical object recognition using contour fragments. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (7): 1270–1281. 499

Shotton, J., Fitzgibbon, A. W., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., & Blake, A. (2011) Real-time human pose recognition in parts from single depth images. In *IEEE Computer Vision & Pattern Recognition*. 207, 209

Shotton, J., Johnson, M., & Cipolla, R. (2008b) Semantic texton forests for image categorization and segmentation. In *IEEE Computer Vision & Pattern Recognition*. 210

Shotton, J., Winn, J., Rother, C., & Criminisi, A. (2009) Textonboost for image understanding: multi-class object recognition and segmentation by jointly modeling texture, layout and context. *International Journal of Computer Vision* **81** (1): 2–23. 203, 204, 205, 209, 210, 316, 335

Shum, H.-Y., & Szeliski, R. (2000) Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision* **36** (2): 101–130. 421

Shumway, R. H., & Stoffer, D. S. (1991) Dynamic linear models with switching. *Journal of the American Statistical Association* **86** (415): 763–769. 567

Shumway, R. H., & Stoffer, D. S. (2007) An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Series Analysis* **3** (4): 253–264. 567

Sidenbladh, H., Black, M. J., & Fleet, D. J. (2000) Stochastic tracking of 3D human figures using 2D image motion. In *European Conference on Computer Vision*, pp. 702–718. 500

Sigal, L., Balan, A. O., & Black, M. J. (2010) HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *International Journal of Computer Vision* **87** (1–2): 4–27. 500

Sigal, L., & Black, M. J. (2006) Measure locally, reason globally: Occlusion-sensitive articulated pose estimation. In *IEEE Computer Vision & Pattern Recognition*, pp. 2041–2048. 274

Sim, R., Elinas, P., Griffin, M., & Little, J. (2005) Vision-based SLAM using the Rao-Blackwellised particle filter. In *IJCAI Workshop on Reasoning with Uncertainty in Robotics*, pp. 9–16. 567

Simon, G., & Berger, M.-O. (2002) Pose estimation for planar structures. *IEEE Computer Graphics and Applications* **22** (6): 46–53. 420

Simon, G., Fitzgibbon, A. W., & Zisserman, A. (2000) Markerless tracking using planar structures in the scene. In *International Symposium on Mixed and Augmented Reality*, pp. 120–128. 420

Sinha, S. N., Mordohai, P., & Pollefeys, M. (2007) Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh. In *IEEE International Conference on Computer Vision*, pp. 1–8. 454

Sinha, S. N., & Pollefeys, M. (2005) Multi-view reconstruction using photo-consistency and exact silhouette constraints: A maximum-flow formulation. In *IEEE International Conference on Computer Vision*, pp. 349–356. 454

Sivic, J., Russell, B. C., Efros, A. A., Zisserman, A., & Freeman, W. T. (2005) Discovering objects and their localization in images. In *IEEE International Conference on Computer Vision*, pp. 370–377. 595

Sivic, J., Russell, B. C., Zisserman, A., Freeman, W. T., & Efros, A. A. (2008) Unsupervised discovery of visual object class hierarchies. In *IEEE Computer Vision & Pattern Recognition*. 595

Sivic, J., & Zisserman, A. (2003) Video google: A text retrieval approach to object matching in videos. In *ICCV*, pp. 1470–1477. 591, 595

Sizintsev, M., Kuthirummal, S., Sawhney, H., Chaudhry, A., Samarasekera, S., & Kumar, R. (2010) GPU accelerated real-time stereo for augmented reality. In *International Symposium 3D Data Processing, Visualization and Transmission*. 444

Sizintsev, M., & Wildes, R. P. (2010) Coarse-to-fine stereo vision with accurate 3D boundaries. *Image and Vision Computing* **28** (3): 352–366. 318

Skrypnyk, I., & Lowe, D. G. (2004) Scene modelling, recognition and tracking with invariant image features. In *International Symposium on Mixed and Augmented Reality*, pp. 110–119. 420

Smith, R., & Cheeseman, P. (1987) On the representation of spatial uncertainty. *International Journal of Robotics Research* **5** (4): 56–68. 567

Smith, R., Self, M., & Cheeseman, P. (1990) Estimating uncertain spatial relationships in robotics. In *Autonomous Robot Vehicles*, ed. by I. J. Cox & G. T. Wilton, pp. 167–193. Springer. 567

Smith, S. M., & Brady, J. M. (1997) SUSAN - a new approach to low level image processing. *International Journal of Computer Vision* **23** (1): 45–78. 352

Snavely, N., Garg, R., Seitz, S. M., & Szeliski, R. (2008) Finding paths through the world's photos. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2008)* **27** (3): 11–21. 450

Snavely, N., Seitz, S. M., & Szeliski, R. (2006) Photo tourism: Exploring photo collections in 3D. In *SIGGRAPH Conference Proceedings*, pp. 835–846. ACM Press. 449

Starck, J., Maki, A., Nobuhura, S., Hilton, A., & Mastuyama, T. (2009) The multiple-camera 3-D production studio. *IEEE Transactions on Circuits and Systems for Video Technology* **19** (6): 856–869. 385

Starner, T., Weaver, J., & Pentland, A. (1998) A wearable computer based amer-

ican sign language recognizer. In *Assistive Technology and Artificial Intelligence*, Lecture Notes in Computer Science, Volume 1458, pp. 84–96. 267, 274

State, A., Hirota, G., Chen, D. T., Garrett, W. F., & Livingston, M. A. (1996) Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In *SIGGRAPH*, pp. 429–438. 420

Stauffer, C., & Grimson, E. (1999) Adaptive background classification using time-based co-occurences. In *IEEE Computer Vision & Pattern Recognition*, pp. 246–252. 97, 140

Stegmann, M. B. (2002) Analysis and segmentation of face images using point annotations and linear subspace techniques. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK–2800 Kgs. Lyngby. 483, 484, 499

Stenger, B., Mendonça, P. R. S., & Cipolla, R. (2001a) Model-based 3D tracking of an articulated hand. In *IEEE Computer Vision & Pattern Recognition*, pp. 310–315. 492, 500

Stenger, B., Mendonça, P. R. S., & Cipolla, R. (2001b) Model-based hand tracking using an unscented Kalman filter. In *British Machine Vision Conference*. 567

Stewénius, H., Engels, C., & Nistér, D. (2006) Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing* **60**: 284–294. 453

Strasdat, H., Montiel, J. M. M., & Davison, A. J. (2010) Real-time monocular SLAM: Why filter? In *IEEE International Conference on Robotics and Automation*, pp. 2657–2664. 567

Sturm, P. F. (2000) Algorithms for plane-based pose estimation. In *IEEE Computer Vision & Pattern Recognition*, pp. 1706–1711. 420

Sturm, P. F., & Maybank, S. J. (1999) On plane-based camera calibration: A general algorithm, singularities, applications. In *IEEE Computer Vision & Pattern Recognition*, pp. 1432–1437. 420

Sturm, P. F., Ramalingam, S., Tardif, J.-P., Gasparini, S., & Barreto, J. (2011) Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends in Computer Graphics and Vision* **6** (1-2): 1–183. 385

Sturm, P. F., & Triggs, B. (1996) A factorization based algorithm for multi-image projective structure and motion. In *European Conference on Computer Vision*, pp. 709–720. 453

Sudderth, E. B., Torralba, A., Freeman, W. T., & Willsky, A. S. (2005) Learning hierarchical models of scenes, objects, and parts. In *IEEE International Conference on Computer Vision*, pp. 1331–1338. 595

Sudderth, E. B., Torralba, A., Freeman, W. T., & Willsky, A. S. (2008) Describing visual scenes using transformed objects and parts. *International Journal of Computer Vision* **77** (1–3): 291–330. 587, 589, 595

Sugimoto, A. (2000) A linear algorithm for computing the homography for conics in correspondence. *Journal of Mathematical Imaging and Vision* **13** (2): 115–130. 420

Sun, J., Zhang, W., Tang, X., & Shum, H. Y. (2006) Background cut. In *European Conference on Computer Vision*, pp. 628–641. 97

Sun, J., Zheng, N., & Shum, H. Y. (2003) Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **25** (7): 787–800. 275, 317

Sutherland, I. E. (1963) Sketchpad: a man-machine graphical communications system. Technical Report Technical Report 296, MIT Lincoln Laboratories. 420

Sutton, C., & McCallum, A. (2011) An introduction to conditional random fields. *Foundations and Trends in Machine Learning* . 316

Szeliski, R. (1996) Video mosaics for virtual environments. *IEEE Computer Graphics and Applications* **16** (2): 22–30. 421

Szeliski, R. (2006) Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision* **2** (1). 421

Szeliski, R. (2010) *Computer vision: algorithms and applications*. Springer. 17, 19, 318, 421

Szeliski, R., & Shum, H.-Y. (1997) Creating full view panoramic image mosaics and environment maps. In *SIGGRAPH*, pp. 251–258. 421

Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., & Rother, C. (2008) A comparative study of energy minimization methods for Markov random fields. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (6): 1068–1080. 317

Taigman, Y., Wolf, L., & Hassner, T. (2009) Multiple one-shots for utilizing class label information. In *British Machine Vision Conference*. 534

Tappen, M. F., & Freeman, W. T. (2003) Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In *IEEE International Conference on Computer Vision*, pp. 900–907. 317

Tarlow, D., Givoni, I., Zemel, R., & Frey, B. (2011) Graph cuts is a max-product algorithm. In *Uncertainty in Artificial Intelligence*. 316

Tenenbaum, J., Silva, V., & Langford, J. (2000) A global geometric framework for nonlinear dimensionality reduction. *Science* **290** (5500): 2319–2315. 352

Tenenbaum, J. B., & Freeman, W. T. (2000) Separating style and content with bilinear models. *Neural Computation* **12** (6): 1247–1283. 527, 534

Terzopolous, D., & Szeliski, R. (1992) Tracking with Kalman snakes. In *Active Vision*, ed. by A. Blake & A. Y. Yuile, pp. 3–29. 567

Thayananthan, A., Navatnam, R., Stenger, B., Torr, P., & Cipolla, R. (2006) Multivariate relevance vector machines for tracking. In *European Conference on Computer Vision*, pp. 124–138. 169

Theobalt, C., Ahmed, N., Lensch, H. P. A., Magnor, M. A., & Seidel, H.-P. (2007) Seeing people in different light-joint shape, motion, and reflectance capture. *IEEE Transactions on Visualization and Computer Graphics* **13** (4): 663–674. 385

Tipping, M., & Bishop, C. M. (1999) Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B* **61** (3): 611–622. 140

Tipping, M. E. (2001) Sparse Bayesian learning and the relevance vector machine. *J. Machine Learning Research* **1**: 211–244. 159, 169, 209, 499

Tomasi, C., & Kanade, T. (1991) Detection and tracking of point features. Technical Report CMU–SC–91–132, Carnegie Mellon. 453

Tomasi, C., & Kanade, T. (1992) Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision* **9** (2): 137–154. 453

Tombari, F., Mattoccia, S., di Stefano, L., & Addimanda, E. (2008) Classification and evaluation of cost aggregation methods for stereo correspondence. In *IEEE Computer Vision & Pattern Recognition*. 318

Torr, P. (1998) Geometric motion segmentation and model selection. *Philosophical Transactions of the Royal Society A* **356** (1740): 1321–1340. 420

Torr, P. H. S., & Criminisi, A. (2004) Dense stereo using pivoted dynamic programming. *Image and vision computing* **22** (10): 795–806. 274

Torr, P. H. S., & Zisserman, A. (2000) MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision & Image Understanding* **78** (1): 138–156. 420

Torralba, A., Murphy, K., & Freeman, W. T. (2007) Sharing visual features for multiclass and multi-view object detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (5): 854–869. 203, 209

Treibitz, T., Schechner, Y. Y., & Singh, H. (2008) Flat refractive geometry. In *IEEE Computer Vision & Pattern Recognition*. 385

Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W. (1999) Bundle adjustment – a modern synthesis. In *Workshop on Vision Algorithms*, pp. 298–372. 453

Tsai, R. (1987) A versatile cameras calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *Journal of Robotics and Automation* **3** (4): 323–344. 385

Turk, M., & Pentland, A. P. (2001) Face recognition using eigenfaces. In *IEEE Computer Vision & Pattern Recognition*, pp. 586–591. 533

Tuytelaars, T., & Mikolajczyk, K. (2007) Local invariant feature detectors: A survey. *Foundations and Trends in Computer Graphics and Vision* **3** (3): 177–280. 352

Urtasun, R., Fleet, D. J., & Fua, P. (2006) 3D people tracking with Gaussian process dynamical models. In *IEEE Computer Vision & Pattern Recognition*, pp. 238–245. 169

Vapnik, V. (1995) *The Nature of Statistical Learning Theory*. Springer Verlag. 209

Varma, M., & Zisserman, A. (2004) Unifying statistical texture classification frameworks. *Image and Vision Computing* **22** (14): 1175–1183. 595

Vasilescu, M. A. O., & Terzopoulos, D. (2002) Multilinear analysis of image ensembles: Tensorfaces. In *European Conference on Computer Vision*, pp. 447–460. 534

Vasilescu, M. A. O., & Terzopoulos, D. (2004) Tensortextures: multilinear image-based rendering. *ACM Transactions on Graphics* **23** (3): 336–342. 530

Vaswani, N., Chowdhury, A. K. R., & Chellappa, R. (2003) Activity recognition using the dynamics of the configuration of interacting objects. In *IEEE Computer Vision & Pattern Recognition*, pp. 633–642. 567

Veksler, O. (2005) Stereo correspondence by dynamic programming on a tree. In *IEEE Computer Vision & Pattern Recognition*, pp. 384–390. 269, 270, 274

Veksler, O. (2008) Star shape prior for graph-cut image segmentation. In *European Conference on Computer Vision*, pp. 454–467. 317

Veksler, O., Boykov, Y., & Mehrani, P. (2010) Superpixels and supervoxels in an energy optimization framework. In *European Conference on Computer Vision*, pp. 211–224. 316

Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003) A survey on pixel-based skin color detection techniques. In *Graphicon*, pp. 85–92. 97

Vicente, S., Kolmogorov, V., & Rother, C. (2008) Graph cut based image segmentation with connectivity priors. In *IEEE Computer Vision & Pattern Recognition*, pp. 1–8. 317

Vincent, E., & Laganiere, R. (2001) Detecting planar homographies in an image pair. In *Image and Signal Processing Analysis*, pp. 182–187. 420

Viola, P., & Jones, M. (2002) Fast and robust classification using asymmetric adaboost and a detector cascade. In *Advances in Neural Information Processing Systems*, pp. 1311–1318. 209, 210

Viola, P. A., & Jones, M. J. (2004) Robust real-time face detection. *International Journal of Computer Vision* **57** (2): 137–154. 202, 203, 209, 210

Viola, P. A., Jones, M. J., & Snow, D. (2005) Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision* **63** (2): 153–161. 203, 204

Vlasic, D., Baran, I., Matusik, W., & Popovic, J. (2008) Articulated mesh animation from multi-view silhouettes. *ACM Transactions on Graphics* **27** (3). 385

Vogiatzis, G., Esteban, C. H., Torr, P. H. S., & Cipolla, R. (2007) Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29** (12): 2241–2246. 316, 450, 451, 454

Vuylsteke, P., & Oosterlinck, A. (1990) Range image acquisition with a single binary-encoded light pattern. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **12** (2): 148–164. 385

Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., & Schmalstieg, D. (2008) Pose tracking from natural features on mobile phones. In *International Symposium on Mixed and Augmented Reality*, pp. 125–134. 420

Wainright, M., Jaakkola, T., & Willsky, A. (2005) MAP estimation via agreement on trees: message passing and linear programming. *IEEE Transactions on Information Theory* **5** (11): 3697–3717. 317

Wang, J. M., Fleet, D. J., & Hertzmann, A. (2007) Multifactor Gaussian process models for style-content separation. In *ICML*, pp. 975–982. 531, 532, 534

Wang, J. M., Fleet, D. J., & Hertzmann, A. (2008) Gaussian process dynamical models for human motion. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **30** (2): 283–298. 567

Wang, X., & Tang, X. (2003) Bayesian face recognition using Gabor features. In *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pp. 70–73. ACM. 533, 534

Wang, X., & Tang, X. (2004a) Dual-space linear discriminant analysis for face recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 564–569. 533

Wang, X., & Tang, X. (2004b) A unified framework for subspace face recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **26** (9): 1222–1228. 533

Wei, L.-Y., & Levoy, M. (2000) Fast texture synthesis using tree-structured vector quantization. In *SIGGRAPH*, pp. 479–488. 317

Weiss, Y., & Freeman, W. (2001) On the optimality of solutions of the max-product belief propagation algorithm in arbitrary graphs. *IEEE Transactions on Information Theory* **47** (2): 723–735. 275, 317

Weiss, Y., Yanover, C., & Meltzer, T. (2011) Linear programming and variants of belief propagation. In *Advances in Markov Random Fields*, ed. by A. Blake, P. Kohli, & C. Rother. MIT Press. 317

Wilbur, R. B., & Kak, A. C. (2006) Purdue RVL-SLLL american sign language database. Technical Report TR–06–12, Purdue University, School of Electrical and Computer Engineering. 244

Williams, C., & Barber, D. (1998) Bayesian classification with Gaussian priors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **20** (2): 1342–1351. 209

Williams, D., & Shah, M. (1992) A fast algorithm for active contours and curvature estimation. *CVGIP: Image understanding* **55** (1): 14–26. 499

Williams, O. M. C., Blake, A., & Cipolla, R. (2005) Sparse Bayesian learning for efficient tracking. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **27** (8): 1292–1304. 167, 169

Williams, O. M. C., Blake, A., & Cipolla, R. (2006) Sparse and semi-supervised visual mapping with the S3P. In *IEEE Computer Vision & Pattern Recognition*, pp. 230–237. 169

Wiskott, L., Fellous, J.-M., Krüger, N., & von der Malsburg, C. (1997) Face recognition by elastic bunch graph matching. In *IEEE International Conference on Image Processing*, pp. 129–132. 533

Wolf, L., Hassner, T., & Taigman, Y. (2009) The one-shot similarity kernel. In *IEEE International Conference on Computer Vision*, pp. 897–902. 534

Woodford, O., Torr, P. H. S., Reid, I., & Fitzgibbon, A. W. (2009) Global stereo reconstruction under second-order smoothness priors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **31** (12): 2115–2128. 316

Wren, C. R., Aazarbayejani, A., Darrell, T., & Pentland, A. P. (1997) Pfinder: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **19** (7): 780–785. 97

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009) Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **31** (2): 210–227. 534

Wu, B., Ai, H., Huang, C., & Lao, S. (2007) Fast rotation invariant multi-view face detection based on real adaboost. In *IEEE Workshop on Face and Gesture Recognition*, pp. 79–84. 210

Wu, C., Agarwal, S., Curless, B., & Seitz, S. (2011) Multicore bundle adjustment. In *IEEE Computer Vision & Pattern Recognition*, pp. 3057–3064. 453

Wu, Y., Lin, J. Y., & Huang, T. S. (2001) Capturing natural hand articulation. In *IEEE International Conference on Computer Vision*, pp. 426–432. 500

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010) Sun database: Large-scale scene recognition from abbey to zoo. In *IEEE Computer Vision & Pattern Recognition*, pp. 3485–3492. 595

Xu, C., & Prince, J. L. (1998) Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing* **7** (3): 359–369. 499

Yang, J., Jiang, Y.-G., Hauptmann, A. G., & Ngo, C.-W. (2007) Evaluating bag-of-visual-words representations in scene classification. In *Multimedia Information Retrieval*, pp. 197–206. 595

Yang, M.-H. (2002) Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods. In *IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 215–220. 533

Yilmaz, A., Javed, O., & Shah, M. (2006) Object tracking: A survey. *ACM Computing Surveys (CSUR)* **38** (4): 1–45. 567

Yin, P., Criminisi, A., Winn, J., & Essa, I. (2007) Tree based classifiers for bilayer video segmentation. In *IEEE Computer Vision & Pattern Recognition*. 209

Yoon, K.-J., & Kweon, I.-S. (2006) Adaptive support-weight approach for correspondence search. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **28** (4): 650–656. 318

Zhang, C., & Zhang, Z. (2010) A survey of recent advances in face detection. Technical Report MSR–TR–2010–66, Microsoft Research, Redmond. 210

Zhang, J., Marszalek, M., Lazebnik, S., & Schmid, C. (2007) Local features and kernels for classification of texture and object categories: A comprehensive study. *International Journal of Computer Vision* **73** (2): 213–238. 595

Zhang, X., & Gao, Y. (2009) Face recognition across pose: A review. *Pattern Recognition* **42** (11): 2876–2896. 533

Zhang, Z. (2000) A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **22** (11): 1330–1334. 420

Zhao, J., & Jiang, Q. (2006) Probabilistic PCA for t distributions. *Neurocomputing* **69** (16–18): 2217–2226. 140

Zhao, W.-Y., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003) Face recognition: A literature survey. *ACM Comput. Surv.* **35** (4): 399–458. 533

Zhou, S., Chellappa, R., & Moghaddam, B. (2004) Visual tracking and recognition using appearance-adaptive models in particle filters. *IEEE Transactions on Image Processing,* **13** (11): 1491–1506. 567

Zhou, S. K., & Chellappa, R. (2004) Probabilistic identity characterization for face recognition. In *IEEE Computer Vision & Pattern Recognition*, pp. 805–812. 533

Zhu, X., Yang, J., & Waibel, A. (2000) Segmenting hands of arbitrary colour. In *IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 446–453. 97

Zitnick, C. L., & Kanade, T. (2000) A cooperative algorithm for stereo matching and occlusion detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **22** (7): 675–684. 318

Index