PROCEEDINGS OF THE ROYAL SOCIETY A

FirstCite® e-publishing

Proc. R. Soc. A
doi:10.1098/rspa.2008.0212
Published online

A multivariate distribution for sub-images

By S. J. Maybank*

School of Computer Science and Information Systems, Birkbeck College, Malet Street, London WC1E 7HX, UK

A new method for obtaining multivariate distributions for sub-images of natural images is described. The information in each sub-image is summarized by a measurement vector in a measurement space. The dimension of the measurement space is reduced by applying a random projection to the truncated output of the discrete cosine transforms of the sub-images. The measurement space is then reparametrized, such that a Gaussian distribution is a good model for the measurement vectors in the reparametrized space. An Ornstein–Uhlenbeck process, associated with the Gaussian distribution, is used to model the differences between measurement vectors obtained from matching sub-images. The probability of a false alarm and the probability of accepting a correct match are calculated. The accuracy of the resulting statistical model for matching sub-images is tested using images from the Middlebury stereo database with promising results. In particular, if the probability of accepting a correct match is relatively large, then there is good agreement between the calculated and the experimental probabilities of obtaining a unique match that is also a correct match.

Keywords: compressive sensing; discrete cosine transform; image statistics; principal components analysis; random projection; stereo matching

1. Introduction

One of the fundamental tasks in computer vision is to find statistical models for the patterns of pixel values in images. These statistical models have many applications, including object detection (Torralba & Oliva 2003; Charpiat *et al.* 2005), edge detection (Konishi *et al.* 1999), image classification (Srivastava *et al.* 2003; Torralba & Oliva 2003), image denoising (Srivastava *et al.* 2003; Tan & Jiao 2007), super resolution (Tappen *et al.* 2003) and the matching of patterns between images. Statistical models for images can be used to make quantitative predictions about the performance of algorithms. For example, if the distribution of a particular class of patterns is known, then it is possible to estimate the number of such patterns that are likely to be found in an image, and hence to estimate the probability of a false match between patterns or the probability of a false detection of an object associated with a pattern.

It is convenient to restrict the patterns of pixel values to rectangular regions or sub-images within a larger image. Let J be an image of size $n_1 \times n_2$ pixels^2. Then a sub-image w of J is any $m_1 \times m_2$ rectangle of pixel values in J, where $1 \leq m_i \leq n$, $i=1, 2$. In most applications m_1 and m_2 are much smaller than n_1, n_2.

*sjmaybank@dcs.bbk.ac.uk

Received 23 May 2008
Accepted 14 November 2008

Each sub-image corresponds to a point in a measurement space $\mathbb{R}^{m_1 m_2}$ specified by a vector with $m_1 m_2$ components, corresponding to the pixels in w. For a wide range of images, the description of a general sub-image is simplified by making an appropriate choice of basis in $\mathbb{R}^{m_1 m_2}$. The basis associated with the discrete cosine transform (DCT) is a convenient and simple choice (Gonzalez & Woods 2002). The vector $v(w)$ of DCT components of w tends to have many components with relatively small absolute values and only a few components with relatively large absolute values. The few relatively large components can be retained and the remaining components set to zero with only a small loss of information (Field 1987). The components set to zero may be different for different vectors $v(w)$, but this difficulty is overcome using a random projection to a new lower dimensional measurement space $\mathbb{R}^k$. Experiments reported in §3*b* below show that this method of reducing the dimension of the measurement space consistently outperforms principal components analysis (PCA) on a number of natural images.

There is a second reason for using a random projection; the properties of the projected vectors can be summarized by a simple statistical model. The directions of the projected measurement vectors tend to be uniformly distributed on the unit hypersphere in $\mathbb{R}^k$, centred at the origin (Diaconis & Freedman 1984; Dasgupta 1999, 2000). The fit of the directions of the projected vectors to a uniform distribution is improved by reparametrizing $\mathbb{R}^k$. In the reparametrized space, the distribution of the measurement vectors is modelled by the uniform distribution on the unit hypersphere for their directions and an empirical distribution for their lengths. The lengths of the vectors are then scaled such that the resulting vectors are compatible with the Gaussian distribution $\mathcal{N}(0, I(k))$ with expected value 0 and covariance $I(k)$, where $I(k)$ is the $k \times k$ identity matrix. This reparametrization does not lose any information.

The distribution $\mathcal{N}(0, I(k))$ is applied to the matching of sub-images, or more precisely, to the matching of the measurement vectors obtained from sub-images with given dimensions $m_1 \times m_2$. A pair (h_1, h_2) of measurement vectors in $\mathbb{R}^k \times \mathbb{R}^k$ is accepted as a match if h_2 is near to h_1. In order to define 'near' in this context, the difference, $h_2 - h_1$ between h_1 and a correct match, h_2, is modelled by an Ornstein–Uhlenbeck process on $\mathbb{R}^k$ with limiting distribution $\mathcal{N}(0, I(k))$. A false alarm occurs for a given h_1 if h_2 is sampled from $\mathcal{N}(0, I(k))$ and by chance is near to h_1. The distribution for sub-images is used to calculate various probabilities, including the probability of a false alarm and the probability of accepting a correct match. In §6, the distribution for sub-images is tested by using it to predict the performance of a stereo matching algorithm on images from the Middlebury stereo database (Scharstein & Pal 2007), with a good agreement between theory and experiment. To the author's knowledge, this is the first time that a statistical model for sub-images has been used to predict the performance of a stereo matching algorithm.

All the calculations in this paper were carried out using Mathematica (Wolfram 1999).

The remainder of this paper is organized as follows. Related work is described in §2. The reduction in the dimension of the measurement space using a random projection is described in §3. The probability density function for sub-images is obtained in §4 and applied in §5 to the task of finding matching sub-images. The experiments with images from the Middlebury stereo database are reported in §6 and some concluding remarks are made in §7.

2. Related work

The use of a random linear projection to reduce the dimension of a measurement space is one of the main themes of the recent work on compressive sensing. An introduction to compressive sensing is given by Baraniuk (2007) and there is a much more detailed discussion by Donoho (2006). Applications of random projections to dimension reduction for data mining and database search are described by Bingham & Mannila (2001). Dasgupta (1999) showed that the eccentricity of a high-dimensional Gaussian distribution is reduced under a random linear projection and Diaconis & Freedman (1984) showed that for many sets of vectors in high-dimensional spaces, the projections of the vectors to a one- or two-dimensional space are modelled accurately by Gaussian distributions. The application of random projections to face recognition is described by Goel *et al.* (2005) and by Han & Jin (2007). Their experiments show that random projections compare well with PCA. Good accounts of PCA are given by Gonzalez & Woods (2002) and Forsyth & Ponce (2003).

Many authors have shown that there are accurate univariate statistical models for the filter responses obtained from natural images (Field 1987; Huang & Mumford 1999; Huang 2000; Grenander & Srivastava 2001; Srivastava *et al.* 2003). More recently the emphasis has been moved to multivariate statistical models. Grenander & Srivastava (2001) showed how a bivariate distribution can be built-up from marginal distributions. Huang & Mumford (1999) modelled multivariate sets of filter responses using generalized Laplace distributions. Tan & Jiao (2007) used elliptically contoured distributions, among which are included the isotropic distributions, to model wavelet components in natural images, with applications to image denoising and image restoration.

Lee *et al.* (2003) made a detailed analysis of the distribution of high contrast 3×3 sub-images of natural images. These sub-images tend to cluster near to a two-dimensional manifold that corresponds to ideal images of edges. Srivastava *et al.* (2003) reviewed the statistical modelling of natural images and describe statistical models for whole images and for sub-images. They emphasize the non-Gaussian nature of traditional representations of sub-images, as reported by Field (1987), and note that in many cases sub-images cluster near to low-dimensional nonlinear submanifolds of the feature space.

Torralba & Oliva (2003) obtained statistics from the Fourier spectra of entire images. They showed that these statistics are affected by the large-scale properties of scenes in systematic ways, and that as a result images can be assigned to broad categories without the necessity of a detailed segmentation or accurate object detection. A general stochastic model for natural images is described by Mumford & Gidas (2001).

The general view emerging from the above work on image statistics by many authors over the past 20 years is that the filter responses for natural images are remarkably predictable, in that they can be modelled by relatively simple leptokurtic distributions that depend on only a small number of parameters. The purpose of this paper is to extend this work on image statistics by defining a useful multivariate statistical model for sub-images, and then applying it to the matching of sub-images.

The algorithm used in §5 below to test the statistical model for image matching is an example of a dense two-frame stereo algorithm. A comprehensive review of such algorithms is given by Scharstein & Szeliski (2002). The algorithm

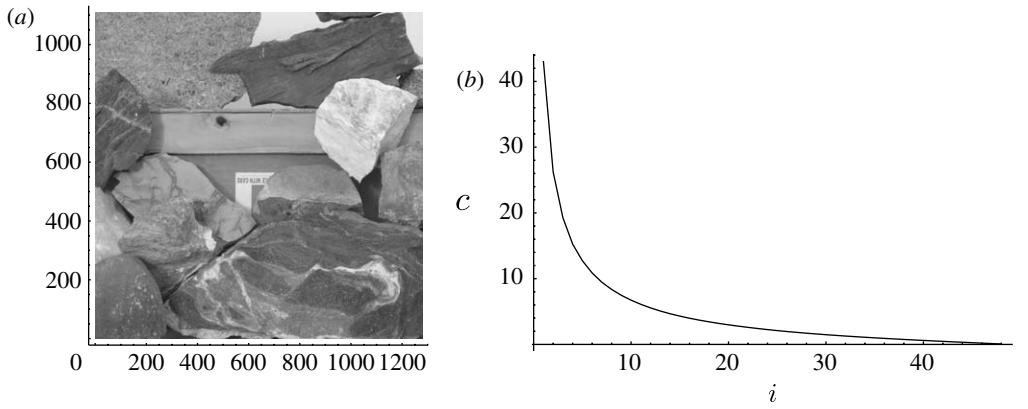

Figure 1. (*a*) Image 1 in the Rocks1 sequence, (*b*) graph of the average values of the ranked absolute values of the DCT components of 7×7 sub-images of image 1.

in §5 has some similarity to the first entry in Scharstein and Szeliski's table 1: traditional sum of squared differences (SSD) method, with aggregation over a square window.

3. Method for reducing the dimension of the measurement space

A new method for reducing the dimension of a measurement space is described in §3*a*, and an experimental comparison of the method with PCA is described in §3*b*. In the experiments, all RGB images are reduced to grey scale images by taking the average of the RGB values at each pixel.

(*a*) *DCT components*

Let w be an $m_1 \times m_2$ sub-image of an image J. It is known that the discrete cosine transform (DCT) of w tends to have only a few components with relative large absolute values (Field 1987). Let the $m_1 \times m_2$ matrix $\mathrm{DCT}(w)$ be the discrete cosine transform of w, indexed such that $\mathrm{DCT}(w)_{11}$ is proportional to the mean grey level in w. Let $v(w)$ be the $m_1 \times m_2 - 1$ dimensional vector obtained by listing the elements of $\mathrm{DCT}(w)$ row by row, except that $\mathrm{DCT}(w)_{11}$ is omitted,

$$v(w) = (\mathrm{DCT}(w)_{12}, \mathrm{DCT}(w)_{13}, \ldots, \mathrm{DCT}(w)_{m_1 m_2})^{\top}.$$

Let $c_i(v)$, $1 \leq i \leq m_1 m_2 - 1$ be the components of $v(w)$ in decreasing order by absolute value. Thus, $c_i(v) = v_{j(i)}$ for some integer $j(i)$ and

$$|\, c_1(v) \,| \geq |\, c_2(v) \,| \geq \cdots \geq |\, c_{m_1 m_2 - 1}(v) \,|.$$

Figure 1*a* shows image number 1 from the Rocks1 sequence of the 2006 Middlebury stereo datasets[1] (Scharstein & Pal 2007). Figure 1*b* shows the corresponding graph of the average value of $|c_i(v(w))|$ as a function of i, where w ranges over a set of 7×7 sub-images of image 1. It is apparent from figure 1*b* that the average value of $|c_i(v)|$ drops rapidly as i increases away from 1, and then

[1] URL http://vision.middlebury.edu/Stereo/data/scenes2006, full size image, illumination 2, exposure 2.

decreases less rapidly for larger values of i. It is known that curves of this type are obtained from a wide range of natural images, and that they are well approximated by power laws (Field 1987; Ruderman 1997).

Figure 1*b* suggests that the information in w can be summarized, to within a small error, by recording the values of the $c_i(v)$ for $1 \leq i \leq j$, where j is a fixed integer less than $m_1 m_2 - 1$. A difficulty arises because a DCT component $v_i(w)$, with i fixed, may be large for one choice of w but small for another choice of w. This difficulty is overcome using random measurements (RM). Let l be the number of DCT components in $v(w) \in \mathbb{R}^{m_1 m_2 - 1}$ that is to be retained. Let $v(l, w)$ be the vector in $\mathbb{R}^{m_1 m_2 - 1}$ obtained by setting to zero the $m_1 m_2 - l - 1$ components in $v(w)$ with the smallest absolute values. The vector $v(l, w)$ thus has at most l non-zero entries. The number of random measurements is made strictly greater than the number of DCT components that are retained. This ensures that with probability 1, the vectors $v(l, w)$ that cannot be recovered exactly from their associated random measurements form a set of measure 0 (Baron *et al.* 2005). This result holds because the set of $v(l, w)$, for varying w, is a union of manifolds of dimension l or less, and each measurement imposes one constraint on those $v(l, w)$ compatible with it.

In this application the number, k, of random measurements is equal to $l+1$. Let Φ be a fixed $k \times (m_1 m_2 - 1)$ matrix with entries Φ_{ij} sampled independently from the Gaussian distribution $\mathcal{N}(0, 1)$ in $\mathbb{R}$ with expected value 0 and variance 1. The components of the projected measurement vector $\Phi v(k-1, w)$ constitute the k random measurements. The fact that $\Phi v(k-1, w)$ uniquely determines $v(k-1, w)$ for almost all $v(k-1, w)$ is sufficient for the applications made below. It is not necessary to undertake the difficult task of recovering $v(k-1, w)$ from $\Phi v(k-1, w)$.

(*b*) *Numerical comparison between PCA and random measurements*

PCA is applied to $n = 5000$ sub-images sampled from image 1 in the Rocks1 sequence. Each sub-image w is of size 7×7. Let C be the covariance of the n vectors $v(w)$, under the assumption that the mean value of the vectors is 0, and let $e(i)$, $1 \leq i \leq 48$, be the eigenvectors of C.

The root mean square error, r.m.s.(PCA, i), of i measurements is defined by

$$\text{r.m.s.(PCA}, i) = \left(\frac{1}{n} \sum_{w} \sum_{j=i+1}^{48} (v(w).e(j))^2 \right)^{1/2}.$$

Similarly, the root mean square error, r.m.s.(RM, i), of i random measurements is defined by

$$\text{r.m.s.(RM}, i) = \left(\frac{1}{n} \sum_{w} \| v(w) - v(i-1, w) \|^2 \right)^{1/2}.$$

It is noted that

$$\text{r.m.s.(PCA}, 0) = \text{r.m.s.(RM}, 1) = \left(\frac{1}{n} \sum_{w} \| v(w) \|^2 \right)^{1/2}.$$

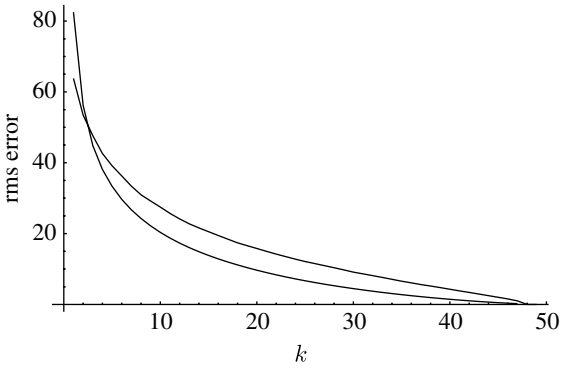

Figure 2. Graphs of $i \mapsto$r.m.s.(PCA, i) and $i \mapsto$r.m.s.(RM, i) for 5000 sub-images are chosen from image 1 in figure 1*a*. The graph of $i \mapsto$r.m.s.(PCA, i) is the upper one at $i=20$.

The calculation of the r.m.s. error using the vectors $v(w)$ is legitimate because the DCT preserves the norms of vectors. The graphs of $i \mapsto$r.m.s.(PCA, i) and $i \mapsto$r.m.s.(RM, i) are shown in figure 2 for the Rocks1 image. The graph of $i \mapsto$r.m.s.(PCA, i) is the upper one at $i=20$. Define $g(\mathrm{PCA}, r)$, $g(\mathrm{RM}, r)$ by

$$g(\mathrm{PCA}, r) = \min\left\{ i, \frac{\text{r.m.s.}(\mathrm{PCA}, i)}{\text{r.m.s.}(\mathrm{PCA}, 0)} < r \right\}, \quad 0 < r < 1,$$

$$g(\mathrm{RM}, r) = \min\left\{ i, \frac{\text{r.m.s.}(\mathrm{RM}, i)}{\text{r.m.s.}(\mathrm{RM}, 1)} < r \right\}, \quad 0 < r < 1.$$

Thus, $g(\mathrm{PCA}, r)$ is the least number of measurements required by PCA to ensure that the normalized r.m.s. error is strictly less than r. Experiments with a range of images show that $g(\mathrm{RM}, r)$ is significantly smaller than $g(\mathrm{PCA}, r)$. The result is particularly noteworthy, because the use of the r.m.s. error favours PCA, due to the fact that PCA minimizes the r.m.s. error over all projections onto a space with a given dimension (Forsyth & Ponce 2003). Some values of $g(\mathrm{PCA}, r)$ and $g(\mathrm{RM}, r)$ are compared in table 1 for three images from the Middlebury stereo database (Scharstein & Szeliski 2002; Hirschmüller & Scharstein 2007), an image from the Groningen natural image data base[2] (van Hateren & van der Schaaf 1998) and the Lena (http://www-ece.rice/∼wakin/images/lena512.bmp) image. The terms Illum2, Exp2, etc. are explained on the Middlebury web pages accessible from http://vision.middlebury.edu/Stereo/data/scenes2006.

The variable r in the second column of table 1 is an upper threshold on the normalized r.m.s. error. The first image, from Rocks1, is figure 1*a* and the four remaining images are shown in figure 3.

4. A multivariate distribution for sub-images

A model is found for the distribution of the measurement vectors obtained from the sub-images of a natural image. The strategy is to reparametrize the measurement space, in order to improve the compatibility of the sampled

[2] The original Groningen image, imk00038.iml, at http://hlab.phys.rug.nl/imlib/l1_200/index.html was converted to the JPEG format for ease of use.

Table 1. Comparison of the number of measurements required by PCA and by RM.

image	r	g(PCA, r)	g(RM, r)
Rocks1, Illum2	0.05	41	31
Exp2	0.1	32	23
view1.png	0.15	25	17
Aloe, Illum3	0.05	38	28
Exp2	0.1	28	20
view1.png	0.15	22	15
Bowling2, Illum3	0.05	40	31
Exp2	0.1	32	22
view1.png	0.15	25	16
Groningen	0.05	47	38
imk0003.iml	0.1	44	31
	0.15	40	26
Lena	0.05	41	30
	0.1	31	21
	0.15	23	15

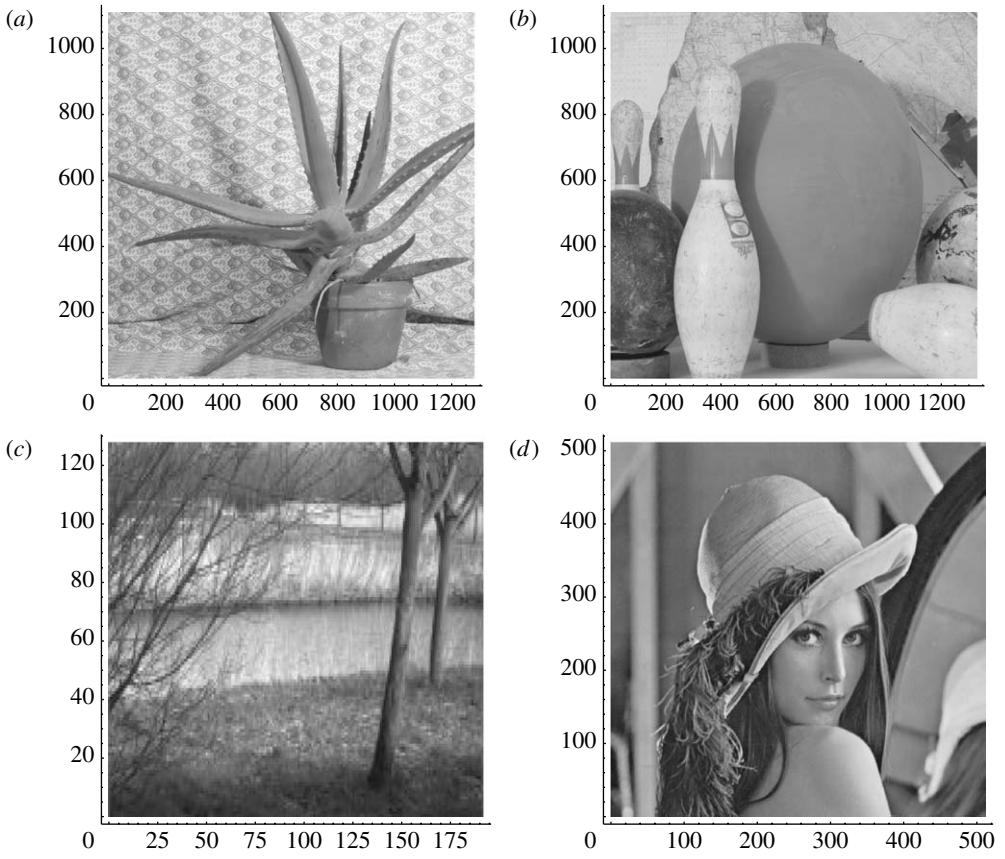

Figure 3. Four of the images used in table 1, (*a*) Aloe, (*b*) Bowling2, (*c*) Groningen, (*d*) Lena.

measurement vectors with an isotropic distribution, and then scale the lengths of the sampled vectors in order to make them compatible with the Gaussian distribution $\mathcal{N}(0, I(k))$. If a second natural image is given, similar to the first, for example, the two images might be a stereo pair, then it is assumed that the measurement vectors from the second image are modelled by $\mathcal{N}(0, I(k))$, after reparametrization and scaling using the functions learnt from the first image.

Statistical tests to verify the compatibility of the reparametrized and scaled measurement vectors with $\mathcal{N}(0, I(k))$ are reported in §4*b*, and some remarks on statistical modelling are made in §4*c*.

(*a*) *Isotropic statistical models*

Let J be an image, let m_1, m_2 be given strictly positive integers and let $\boldsymbol{\Phi}$ be a $k\times(m_1m_2-1)$ matrix with entries $\boldsymbol{\Phi}_{ij}$ sampled independently from $\mathcal{N}(0, 1)$. Let

$$V = \{v(k-1, w), \quad w \text{ is an } m_1 \times m_2 \text{ sub-image in } J\}, \tag{4.1}$$

where $v(k-1,w)$ is as defined in §3*a*. Let $|V|$ be the number of elements in V, define the matrix C by

$$C = \frac{1}{|V|}\sum_{v\in V}(\boldsymbol{\Phi}v)\otimes(\boldsymbol{\Phi}v),$$

and define the set $\tilde{V}$ by

$$\tilde{V} = \{C^{-1/2}\boldsymbol{\Phi}v, v \in V\}. \tag{4.2}$$

The subtraction of the mean value from the vectors in V or in $\{\boldsymbol{\Phi}v, v\in V\}$ prior to calculating C does not improve the results. Most of the vectors in V and $\{\boldsymbol{\Phi}v, v\in V\}$ are clustered about the origin, but the mean is strongly affected by a few vectors far from the origin. If the mean is subtracted out, then the directions of the resulting vectors are less isotropic.

The results of Diaconis & Freedman (1984) and Dasgupta (1999) suggested that an isotropic Gaussian distribution is a possible model for $\tilde{V}$. With this in mind, a function $v\mapsto f(v)$ is sought, such that the directions of the vectors $\{f(v), v \in \tilde{V}\}$ are compatible with the uniform distribution on S^{k-1}. The uniform distribution on S^{k-1} has expected value 0 and covariance $k^{-1}I(k)$. Let v be a non-zero vector in $\mathbb{R}^k$ and consider a perturbation of the form

$$v \mapsto \|v\|\|v + Mv\|^{-1}(v + Mv), \tag{4.3}$$

where M is a symmetric $k\times k$ matrix such that the Frobenius norm, $\|M\|$, is small. The asymmetric part of M is set to zero because to first order its effect on $v+Mv$ is a rotation, which does not change the fit of a set of vectors to an isotropic distribution. In detail,

$$\|v + Mv\|^2 = \|v\|^2 + 2v^\top Mv + O(\|M\|^2),$$

and the antisymmetric part of M makes no contribution to $\|v+Mv\|$, to first order, as claimed. Let $\hat{v}=\|v\|^{-1}v$. The covariance $C(M)$ of the directions of the perturbed vectors in $\tilde{V}$, omitting any zero vector in $\tilde{V}$, is

$$\begin{aligned} C(M) &= \frac{1}{|\tilde{V}|}\sum_{v\in\tilde{V}}\|v + Mv\|^{-2}(v + Mv)\otimes(v + Mv), \\ &= \frac{1}{|\tilde{V}|}\sum_{v\in\tilde{V}}\hat{v}\otimes\hat{v} + M\hat{v}\otimes\hat{v} + \hat{v}\otimes\hat{v}M - 2(\hat{v}^\top M\hat{v})\hat{v}\otimes\hat{v} + O(\|M\|^2). \end{aligned} \tag{4.4}$$

The matrix M is estimated by omitting the $O(\|M\|^2)$ term from the right-hand side of (4.4), setting the remaining terms equal to $k^{-1}I(k)$, and then solving for M. The matrices $tI(k)$ for $t \in \mathbb{R}$ are excluded from the solution space because the term in (4.4) linear in M vanishes if $M = tI$.

The perturbation (4.3) is iterated. In practice, four iterations are sufficient to ensure that the covariance of the directions of the resulting vectors is close to $k^{-1}I(k)$. Let the four $k \times k$ matrices obtained as successive values of M be M_i, $1 \leq i \leq 4$, define the matrices A_i, $1 \leq i \leq 4$, by $A_i = I(k) + M_i$, and set $A = A_4A_3A_2A_1$. Then $v \mapsto f(v)$ is defined by $f(0) = 0$ and

$$f(v) = \|v\| \|A\|^{-1} Av, \quad v \neq 0. \tag{4.5}$$

An isotropic probability density function in $\mathbb{R}^k$ is determined by the density for the lengths of vectors (Samorodnitsky & Taqqu 1994; Tan & Jiao 2007). Thus, a given isotropic density can be converted to any other isotropic density by an appropriate scaling of the lengths of vectors. The vectors in $f(\tilde{V})$ are scaled such that the resulting vectors are modelled by the distribution $\mathcal{N}(0, I(k))$. The scaling function is estimated using the empirical distribution of the lengths of the vectors in $f(\tilde{V})$ or equivalently the lengths of the vectors in $\tilde{V}$.

Define the integer valued function $r \mapsto n(r)$ by

$$n(r) = |\{v, v \in \tilde{V} \text{ and } \|v\| \leq r\}|, \quad r \geq 0.$$

If x is sampled from $\mathcal{N}(0, I(k))$, then

$$\begin{aligned} P(\|x\| \leq r) &= \frac{k}{(k/2)!2^{k/2}} \int_0^r s^{k-1} \exp\left(-\frac{1}{2}s^2\right) \mathrm{d}s, \\ &= 1 - \frac{\Gamma(k/2, r^2/2)}{\Gamma(k/2)}, \end{aligned}$$

where $(a, z) \mapsto \Gamma(a, z)$ is the incomplete gamma function (Abramowitz & Stegun 1965; Wolfram 1999). The required scaling $r \mapsto \rho(r)$, $r \geq 0$, is obtained by numerical solution of the equation

$$1 - \frac{\Gamma(k/2, \rho(r)^2/2)}{\Gamma(k/2)} = \frac{n(r)}{|\tilde{V}|}, \quad r \geq 0.$$

The scaling of the lengths of the vectors in $f(\tilde{V})$ is

$$v \mapsto \rho(\|v\|)\|v\|^{-1}v, \quad v \neq 0.$$

The resulting function ϕ from V to vectors in $\mathbb{R}^k$ compatible with $\mathcal{N}(0, I(k))$ is

$$\phi(0) = 0 \quad \text{and} \quad \phi(v) = \frac{\rho(\|C^{-1/2}\Phi v\|)AC^{-1/2}\Phi v}{\|AC^{-1/2}\Phi v\|}, \quad v \in V, \quad v \neq 0. \tag{4.6}$$

(*b*) *Statistical tests*

Let $H = \{\phi(v), v \in V\}$, where V is defined by (4.1) and ϕ is defined by (4.6). The distribution of the vectors in H is tested for compatibility with $\mathcal{N}(0, I(k))$. There are three tests. In the first test, the vectors in H are separated by a random

Table 2. Empirical standard deviations $\hat{\sigma}(Y)$ and $\hat{\sigma}(Z)$ for the same five images as in table 1.

image	$\hat{\sigma}(Y)$	$\hat{\sigma}(Z)$
Rocks1	0.0055	0.0048
Aloe	0.0106	0.0081
Bowling2	0.0157	0.0114
Groningen	0.0593	0.0441
Lena	0.0147	0.0109

hyperplane through the origin. If H is modelled by $\mathcal{N}(0, I(k))$, then the fraction of vectors in a given half space should be close to 0.5, with a small standard deviation. The second test is similar, except that the hyperplane is chosen at a given non-zero distance from the origin. In the third test, the vectors in H are projected onto a random one-dimensional subspace. If H is modelled by $\mathcal{N}(0, I(k))$, then the empirical standard deviation of the projected vectors should be close to 1.

Let u be a random vector uniformly distributed on S^{k-1}, let $H_1(u)$ be the set of vectors $h \in H$ for which $u^\top h > 0$ and let $Y(u) = |H_1(u)|/|H|$. The hypersphere S^{k-1} was sampled 1000 times and the empirical standard deviation, $\hat{\sigma}(Y)$, was calculated. If $\mathcal{N}(0, I(k))$ is a good model for the elements of H, then $\hat{\sigma}(Y)$ should be near to zero with a high probability.

Let $u \in S^{k-1}$ and define $A(u, d)$ for $d \in \mathbb{R}$ by $A(u, d) = \{h, h \in \mathbb{R} \text{ and } u^\top h \leq d\}$. Let $Z(u) = |A(u, d) \cap H| / |H|$, for $d = 0.8$. The empirical standard deviation $\hat{\sigma}(Z)$ was found for 1000 values of u sampled independently and uniformly from S^{k-1}. If $\mathcal{N}(0, I(k))$ is a good model for the elements of H, then $\hat{\sigma}(Z)$ should be near to zero.

The values of $\hat{\sigma}(Y)$ and $\hat{\sigma}(Z)$ are shown in table 2 for the same five images as in table 1. In each case, $k = 12$, the sub-images are of size 7×7 and $|H| = 9000$.

Typical results for a set H consisting of 9000 samples drawn randomly from $\mathcal{N}(0, I(k))$ are $\hat{\sigma}(Y) = 0.00578$, $\hat{\sigma}(Z) = 0.00449$. The values of $\hat{\sigma}(Y)$, $\hat{\sigma}(Z)$ obtained from the five images are comparable with those obtained using samples from $\mathcal{N}(0, I(k))$.

In the final test, the projections of H are examined. If $u \in S^{k-1}$ and the elements of H are sampled from $\mathcal{N}(0, I(k))$, then the elements of $H(u) = \{u^\top h, h \in H\}$ are sampled from $\mathcal{N}(0, 1)$. Conversely, it is a consequence of the Cramér–Wold theorem that if X is a random variable defined on $\mathbb{R}^k$ such that $u^\top X \sim \mathcal{N}(0, 1)$ for all $u \in S^{k-1}$, then $X \sim \mathcal{N}(0, I(k))$ (Moran 1984). As in the case of table 1, $k = 12$, $m_1 = m_2 = 7$ and $|H| = 9000$. The normalized histogram of $H(u)$ for the Rocks1 image in figure 1*a* is shown in figure 4, superposed on the density function for $\mathcal{N}(0, 1)$. The vector u was sampled from the uniform distribution on S^k. There is a good agreement between the histogram and the density.

The standard deviation of $H(u)$ is estimated under the assumption that the elements of $H(u)$ are sampled from a zero mean Gaussian distribution. Let σ be a candidate value for the standard deviation. The real line, $\mathbb{R}$, is divided into 10 intervals such that each interval is assigned a probability of 1/10 by the distribution $\mathcal{N}(0, 1)$. Let $H_i(u)$ be the number of elements of $H(u)$ in the ith interval, and let $p_i(u) = |H_i(u)| / |H(u)|$, $1 \leq i \leq 10$. An application of Bayes rule yields

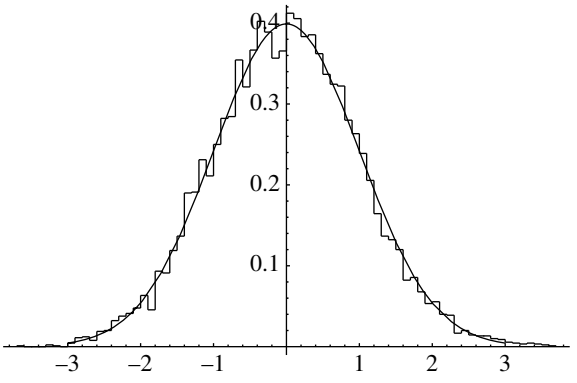

Figure 4. Histogram of $H(u)$ superposed on the density for $\mathcal{N}(0, 1)$.

Table 3. Mean and standard deviations of 100 values of $\hat{\sigma}(u)$ for each of five images.

image	mean	s.d.
Rocks1	0.9988	0.0086
Aloe	1.0015	0.0138
Bowling2	1.9996	0.0085
Groningen	1.0005	0.0076
Lena	1.0002	0.0128

$$P(\sigma \,|\, \{p_1(u), \ldots, p_{10}(u)\}) = \frac{P(\{p_1(u), \ldots, p_{10}(u)\} \,|\, \sigma) P(\sigma)}{P(\{p_1(u), \ldots, p_{10}(u)\})}. \tag{4.7}$$

Let q_i be the probability mass assigned by $N(0, \sigma^2)$ to the ith interval, $1 \leq i \leq 10$. The probability $P(\{p_1(u), \ldots, p_{10}(u)\} \,|\, \sigma)$ is obtained using the multinomial distribution,

$$P(\{p_1(u), \ldots p_{10}(u)\} \,|\, \sigma) = \begin{pmatrix} & |\, H(u) \,| & \\ |\, H_1(u) \,| & \ldots & |\, H_{10}(u) \,| \end{pmatrix} \prod_{i=1}^{10} q_i^{|\, H_i(u) \,|},$$

and $P(\sigma)$ is obtained from the scale invariant prior, $P(\sigma) = \sigma^{-1}$ (Jaynes 2003). An estimate $\hat{\sigma}(u)$ of σ is obtained by maximizing the numerator on the right-hand side of (4.7),

$$\hat{\sigma}(u) = \operatorname{argmax}_{\sigma \mapsto} \frac{1}{\sigma} \begin{pmatrix} & |\, H(u) \,| & \\ |\, H_1(u) \,| & \ldots & |\, H_{10}(u) \,| \end{pmatrix} \prod_{i=1}^{10} q_i^{|\, H_i(u) \,|}.$$

Values of $\hat{\sigma}(u)$ were obtained for 100 samples u from S^{k-1}. The mean and the standard deviations of the 100 values of $\hat{\sigma}(u)$ are shown in table 3 for the five images listed in table 2.

Typical results for a set H consisting of 9000 samples from $\mathcal{N}(0, I(k))$ are mean 0.9987 and standard deviation 0.0097, both of which are comparable with the appropriate entries in table 2.

(c) *Discussion*

The fact that the empirical covariance of H is equal to $I(k)$ and the fact that the results of the tests in §4*b* are compatible with the hypothesis that H is modelled by $\mathcal{N}(0, I(k))$ do not prove that the vectors in H are sampled from $\mathcal{N}(0, I(k))$. There is always the possibility that H inherits from the original image structures that are not modelled by $\mathcal{N}(0, I(k))$. However, the use of random measurements and the reduction of the dimension of the measurement space from $m_1 m_2 - 1$ to k both tend to improve the compatibility of H with $\mathcal{N}(0, I(k))$, and the results on stereo matching, reported in §6 below, confirm that the model $\mathcal{N}(0, I(k))$ does contain a significant amount of useful information about H.

5. Theory of matching sub-images

The distribution for sub-images obtained in §4*a* is applied to the matching of sub-images. An Ornstein–Uhlenbeck process (Karatzas & Shreve 1988) is used to model the difference between a given measurement vector and a matching measurement vector. Formulae for the probability of a single false alarm and the probability of a single correct match with no false alarm are obtained.

(a) *Statistical model for the differences between sub-images*

Let J_1, J_2 be a pair of images and consider the task of matching a sub-image sampled from J_1 with a sub-image sampled from J_2. Let the random variables W_1, W_2 model the sub-images of J_1, J_2, respectively, and let the random variable $D(W_1)$ model the difference between W_1 and the matching random sub-image in J_2. It follows that $W_2 = W_1 + D(W_1)$.

It is assumed that W_1 and W_2 have the same distribution. This assumption places a strong constraint on $D(W_1)$. The correct choice of distribution for $D(W_1)$ becomes apparent on considering the random vectors $G_i = \phi(v(k-1, W_i))$, $i = 1, 2$ in $\mathbb{R}^k$, where ϕ is defined by (4.6). Let $\tilde{D}(G_1)$ be the perturbation such that

$$G_2 = G_1 + \tilde{D}(G_1), \tag{5.1}$$

and suppose that $G_1 \sim \mathcal{N}(0, I(k))$. It is necessary to define $\tilde{D}(G_1)$ such that $G_2 \sim \mathcal{N}(0, I(k))$. The random vectors G_1, G_2 are regarded as nearby states in a stochastic process on $\mathbb{R}^k$, and $\tilde{D}(G_1)$ is regarded as an increment in the process. If the process has $\mathcal{N}(0, I(k))$ as a limiting distribution, then this distribution is preserved under the increment $\tilde{D}(G_1)$, giving $G_1 + \tilde{D}(G_1) \sim \mathcal{N}(0, I(k))$, as required. The simplest option is to use the Ornstein–Uhlenbeck process $s \mapsto X_s$, $0 \leq s$, in $\mathbb{R}^k$ (Karatzas & Shreve 1988), as described by the Itô stochastic differential equation

$$\mathrm{d}X_s = -X_s \, \mathrm{d}s + \sqrt{2}\mathrm{d}B_s, \quad 0 \leq s.$$

The process B is a standard Brownian motion on $\mathbb{R}^k$. If X_0 is given, then the expected value m_s and covariance C_s of X_s are

$$m_s = \exp(-s)X_0, \qquad 0 \le s,$$
$$C_s = (1-\exp(-2s))I(k), \qquad 0 \le s.$$

The diffusion model for image matching preserves the symmetry between the two images J_1, J_2, in that the random vectors G_1, G_2 are required to have the same distribution, $\mathcal{N}(0, I(k))$. The symmetry is broken when a single sub-image, w_1, is chosen in J_1 and a matching sub-image is sought in J_2. The diffusion model biases the matching in favour of those sub-images that appear frequently in J_2. This bias is more extreme if the vector, h_1, in $\mathbb{R}^k$ obtained from w_1 is in a region of $\mathbb{R}^k$ in which the probability density function for $\mathcal{N}(0, I(k))$ takes very small values.

The two random vectors G_1, G_2 are regarded as states of a single Ornstein–Uhlenbeck process X. The pair (h_1, h_2) is defined to be a correct match, or equivalently a realization of (G_1, G_2), if $X_0 = h_1$ and $X_t = h_2$ where t is a fixed time, to be determined. Let $x \mapsto f_t(h_1, x)$ be the probability density function for X_t, given that $X_0 = h_1$. The density $f_t(h_1, x)$ is Gaussian with expected value $m_t(h_1) = \exp(-t)h_1$ and covariance $C_t = (1-\exp(-2t))I(k)$. Let $b(\delta)$ be defined for $0 \le \delta < 1$ such that

$$\int_{\|x - m_t(h_1)\| \le b(\delta)} f_t(h_1, x)\mathrm{d}x = \delta. \tag{5.2}$$

A pair (h_1, h_2) is accepted as a match if $\|h_2 - m_1(h_1)\| \le b(\delta)$. The quantity δ is the probability that (h_1, h_2) is accepted, given that (h_1, h_2) is a correct match.

It remains to estimate t. A training set (h_{i1}, h_{i2}), $1 \le i \le m$, of correctly matched pairs of vectors is required. It is assumed that h_{i2} is sampled from the density $x \mapsto f_t(h_{i1}, x)$ and t is estimated numerically using maximum likelihood,

$$t = \operatorname{argmax} s \mapsto \frac{1}{m}\sum_{i=1}^{m} \ln(f_s(h_{i1}, h_{i2})). \tag{5.3}$$

The above matching algorithm is related to the sum of squared differences (SSD) algorithm, as listed by Scharstein & Szeliski (2002) in their table 1. The difference is that the above algorithm does not calculate the SSD using the original sub-images but instead calculates a version of the SSD in a measurement space, $\mathbb{R}^k$, in which the measurement vectors have the distribution $\mathcal{N}(0, I(k))$.

It is noted that other diffusion models for sub-images are possible. For example, the feature space $\mathbb{R}^k$ could be parametrized such that the probability density function for sub-images is the uniform density on the unit ball. The natural diffusion for sub-images is then a Brownian motion in the unit ball with reflection at the boundary. Unfortunately, the numerical calculations associated with this diffusion are extremely complicated.

(*b*) *Probabilities of false alarm and acceptance*

Let $E(h_1, h_2)$ be the event that (h_1, h_2) is a correct match and let $A(h_1, h_2)$ be the event that (h_1, h_2) is accepted as a match. It follows from (5.2) and the criterion for accepting a match that

$$P(A(h_1, h_2) \,|\, E(h_1, h_2), h_1) = \delta. \tag{5.4}$$

It is assumed that there are c candidate matches to h_1 and that the correct match is included among them. It follows that:

$$P(E(h_1, h_2) \,|\, h_1) = c^{-1}.$$

The probability $P(A(h_1, h_2)|E(h_1, h_2),h_1)$ is affected by the assumption that the correct match is included in the c candidate matches, but this effect is neglected under the assumption that c is large.

If (h_1, h_2) is not a correct match, then it is assumed that h_2 is sampled from $\mathcal{N}(0, I(k))$, independently of h_1. A false alarm occurs if (h_1, h_2) is accepted, even though it is not a correct match. Let $\bar{E}(h_1, h_2)$ be the complement of the event $E(h_1, h_2)$. The probability of a false alarm is given by

$$P(A(h_1, h_2) \,|\, \bar{E}(h_1, h_2), h_1) = \frac{1}{(2\pi)^{k/2}} \int_{\|x - m_t(h_1)\| \le b(\delta)} \exp\left(-\frac{1}{2}\|x\|^2\right) \mathrm{d}x.$$

Let $g(h_1, \delta) = P(A(h_1, h_2) \,|\, \bar{E}(h_1, h_2), h_1)$. A short calculation yields

$$\begin{aligned} P(A(h_1, h_2) \,|\, h_1) &= P(A(h_1, h_2) \,|\, E(h_1, h_2), h_1) P(E(h_1, h_2), h_1) \\ &\quad + P(A(h_1, h_2) \,|\, \bar{E}(h_1, h_2), h_1) P(\bar{E}(h_1, h_2), h_1), \\ &= c^{-1}\delta + c^{-1}(c-1) g(h_1, \delta). \end{aligned}$$

Let $N(h_1)$ be the event that no match to h_1 is accepted. It follows that:

$$P(N(h_1) \,|\, h_1) = (1-\delta)(1-g(h_1, \delta))^{c-1}.$$

Suppose that vectors $H = \{h_{i1}, 1 \le i \le n\}$ are obtained from sub-images of J_1, and each vector h_{i1} has c candidate matches obtained from sub-images of J_2, of which one is a correct match to h_{i1}. The average probability, $P(N(H)|H)$, that no match is accepted is

$$P(N(H) \,|\, H) = \frac{1}{n}(1-\delta) \sum_{i=1}^{n} (1 - g(h_{i1}, \delta))^{c-1}. \tag{5.5}$$

Let $F(h_1)$ be the event that the correct match to h_1 is not accepted and that there is a single false alarm. It follows that:

$$P(F(h_1) \,|\, h_1) = (c-1)(1-\delta) g(h_1, \delta)(1-g(h_1, \delta))^{c-2}$$

and

$$P(F(H) \,|\, H) = \frac{1}{n}(c-1)(1-\delta) \sum_{i=1}^{n} g(h_{i1}, \delta)(1-g(h_{i1}, \delta))^{c-2}. \tag{5.6}$$

Let $T(h_1)$ be the event that the correct match to h_1 is accepted and that there are no false alarms. It follows that:

$$P(T(h_1) \,|\, h_1) = \delta(1-g(h_1, \delta))^{c-1}$$

and

$$P(T(H) \,|\, H) = \frac{\delta}{n} \sum_{i=1}^{n} (1 - g(h_{i1}, \delta))^{c-1}. \tag{5.7}$$

In §6, the calculated probabilities $P(N(H)|H)$, $P(F(H)|H)$, $P(T(H)|H)$ are compared with their empirical estimates.

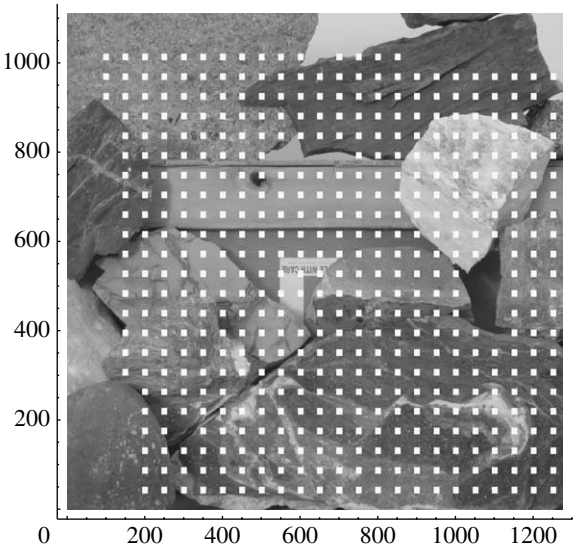

Figure 5. The 500 sub-images for stereo matching with $m_1 = m_2 = 15$.

6. Experiments on stereo matching

The theory of matching described in §5 was tested experimentally on stereo images from the Middlebury stereo database. The aim of the test was to show that it is possible to predict the performance of the matching algorithm.

(*a*) *Experiments*

All RGB colour images were reduced to grey scale images by taking the average of the RGB values at each pixel. The Rocks1 image sequence contains seven images numbered from 0 to 6, and rectified such that if (x_1, y_1) and (x_2, y_2) are matching points in different images then $x_1 = x_2$. The disparities $y_2 - y_1$ for the matches between image 1 and 5 are available, rounded to integer values, in the file disp1.pgm. Image 1 is shown above in figure 1*a*, with the y-axis horizontal.

Suitable values for k, m_1, m_2 were chosen and the function ϕ in (4.6) was calculated using 9000 $m_1 \times m_2$ sub-images sampled from image 1 at the vertices of a square grid. Stereo matching was carried out between image 1 and 5 using 500 sub-images sampled from image 1, with the centres of the sub-images at the vertices of another square grid. The 500 sub-images are shown in figure 5 for $m_1 = m_2 = 15$.

Let $D(x_1, y_1)$ be the disparity, rounded to an integer, at (x_1, y_1) in image 1. If the disparity at (x_1, y_1) is not defined, for example, due to occlusion, then $D(x_1, y_1) = 0$. A match $(x_1, y_1) \leftrightarrow (x_5, y_5)$ was regarded as correct if

$$| y_1 - D(x_1, y_1) - y_5 | \leq 2. \tag{6.1}$$

The bound of 2 in (6.1) is reasonable, first because y_1, y_2 and $D(x_1, y_1)$ are rounded to integers and second because the correlations between nearby sub-images lead to additional matches in the neighbourhood of a correct match. A consequence of (6.1) is that correct matches to (x_1, y_1) occur in groups of five.

Let $D_{\max}$, $D_{\min}$ be, respectively, the maximum and minimum of the non-zero entries of D. The number of candidate matches to (x_1, y_1) is $D_{\max} - D_{\min} + 1$. The number c of candidate matches, as defined in §5*b*, is defined pessimistically by $c = D_{\max} - D_{\min} - 3$. This definition of c is obtained by subtracting four from $D_{\max} - D_{\min} + 1$, to take account of the fact that five of the matches are grouped together and considered to be a single correct match. The value of c is adjusted for the relatively small number of sub-images that are near to the boundaries of the images. The parameter t was estimated using (5.3) applied to 1000 sub-images sampled from image 1 and their matching sub-images in image 5.

(*b*) *Results*

The results of the experiments are shown in tables 4–8. The same random projection matrix $\boldsymbol{\Phi}$ was used throughout. The variable δ in the first column of each table is the probability (5.4) that a correct match is accepted. The average probabilities $P(N(H)|H)$, $P(F(H)|H)$, $P(T(H)|H)$ are as defined by (5.5)–(5.7), respectively, and $\hat{P}(N)$, $\hat{P}(F)$, $\hat{P}(T)$ are the empirical estimates of these average probabilities.

The empirical estimates were calculated as follows. Any sub-image with 11 or more false alarms was discarded on the grounds that the large number of false alarms was incompatible with the statistical model for matching described in §5*a*. To give a numerical example, in the case of the Rocks1 image, if $k = 12$ and $\delta = 0.95$, then (5.2) yields $b(\delta) = 2.6328$, giving a probability of a false alarm less than 0.1379, and an exceedingly small probability of 11 or more false alarms. If, nevertheless, 11 false alarms are observed, then this is a sign that the model has failed for the particular sub-image under consideration. Let n be the number of the remaining sub-images and let $n(F)$ be the number of sub-images for which no match was found. Then $\hat{P}(N)$ is defined by $\hat{P}(N) = n(F)/n$, and $\hat{P}(F)$, $\hat{P}(T)$ are defined analogously. The values of n are given in the second column of each table.

The final column in tables 4–8 gives the value $\sigma = (p(1-p)/n)^{1/2}$ of the standard deviation in the estimate of the probability $p = P(T(H)|H)$ by the mean value of n independent Bernoulli random variables, each of which takes the value 1 with probability p. The prediction of $\hat{P}(T)$ by $P(T(H)|H)$ is the best possible if $\hat{P}(T) - P(T(H)\,|\,H) = O(\sigma)$.

The probability $\hat{P}(F)$ is similar to the percentage $B_{\overline{\mathcal{O}}}$ of unoccluded badly matched pixels defined by Scharstein & Szeliski (2002) in their §5.1. The values of $P(\hat{F})$ in tables 4–8 are similar to the values of $B_{\overline{\mathcal{O}}}$, in spite of the fact that the values of c are much larger than the maximum disparity of 15 given by Scharstein and Szeliski in their table 2.

The agreement between $P(T(H)|H)$ and $\hat{P}(T)$ is good in tables 4–6 for the larger values of δ. In contrast to $P(T(H)|H)$ and $\hat{P}(T)$, the differences between $P(N(H)|H)$ and $\hat{P}(N)$ and between $P(F(H)|H)$ and $\hat{P}(F)$ in tables 4–6 are small at low values of δ, and tend to increase as δ increases. The increase in the value of k for table 5 led to increases in $P(T(H)|H)$ and $\hat{P}(T)$, when compared with table 4, but to a reduction in the accuracy with which $P(T(H)|H)$ predicted $\hat{P}(T)$.

There is less agreement between $P(T(H)|H)$ and $\hat{P}(T)$ in table 7 for the Bowling2 sequence. The values of n are low, showing that a large number of sub-images have 11 or more false alarms. In addition, $P(T(H)|H)$ and $\hat{P}(T)$ are very low. The Bowling2 image, as shown in figure 3*b*, is dominated by the uniform surface of the bowling ball.

Table 4. Rocks1 sequence, $m_1=7$, $m_2=7$, $k=12$, $t=0.2$, $c=114$.

δ	n	$\hat{P}(N)$	$P(N(H)\|H)$	$\hat{P}(F)$	$P(F(H)\|H)$	$\hat{P}(T)$	$P(T(H)\|H)$	σ
0.8	473	0.152	0.128	0.046	0.047	0.437	0.515	0.022
0.85	460	0.132	0.085	0.050	0.038	0.439	0.483	0.023
0.9	452	0.110	0.046	0.037	0.027	0.376	0.414	0.023
0.95	428	0.084	0.015	0.030	0.013	0.282	0.294	0.022

Table 5. Rocks1 sequence, $m_1=7$, $m_2=7$, $k=24$, $t=0.3$, $c=114$.

δ	n	$\hat{P}(N)$	$P(N(H)\|H)$	$\hat{P}(F)$	$P(F(H)\|H)$	$\hat{P}(T)$	$P(T(H)\|H)$	σ
0.8	490	0.183	0.170	0.026	0.025	0.575	0.682	0.021
0.85	487	0.151	0.119	0.030	0.024	0.554	0.674	0.021
0.9	478	0.129	0.069	0.025	0.021	0.525	0.629	0.022
0.95	452	0.099	0.026	0.017	0.013	0.482	0.509	0.022

Table 6. Aloe sequence, $m_1=7$, $m_2=7$, $k=12$, $t=0.2$, $c=165$.

δ	n	$\hat{P}(N)$	$P(N(H)\|H)$	$\hat{P}(F)$	$P(F(H)\|H)$	$\hat{P}(T)$	$P(T(H)\|H)$	σ
0.8	443	0.099	0.111	0.045	0.051	0.395	0.446	0.023
0.85	429	0.093	0.071	0.046	0.039	0.368	0.403	0.023
0.9	411	0.070	0.037	0.051	0.025	0.321	0.338	0.023
0.95	389	0.056	0.012	0.041	0.010	0.251	0.230	0.021

Table 7. Bowling2 sequence, $m_1=7$, $m_2=7$, $k=12$, $t=0.3375$, $c=159$.

δ	n	$\hat{P}(N)$	$P(N(H)\|H)$	$\hat{P}(F)$	$P(F(H)\|H)$	$\hat{P}(T)$	$P(T(H)\|H)$	σ
0.8	272	0.077	0.010	0.040	0.018	0.154	0.043	0.012
0.85	240	0.058	0.004	0.037	0.009	0.133	0.027	0.010
0.9	204	0.053	0.001	0.034	0.003	0.078	0.014	0.008
0.95	145	0.013	0.000	0.041	0.000	0.068	0.005	0.005

Table 8. Bowling2 sequence, $m_1=15$, $m_2=15$, $k=56$, $t=0.5$, $c=159$.

δ	n	$\hat{P}(N)$	$P(N(H)\|H)$	$\hat{P}(F)$	$P(F(H)\|H)$	$\hat{P}(T)$	$P(T(H)\|H)$	σ
0.8	471	0.144	0.158	0.050	0.034	0.458	0.634	0.022
0.85	456	0.135	0.105	0.039	0.034	0.427	0.596	0.022
0.9	438	0.111	0.055	0.038	0.028	0.406	0.498	0.023
0.95	403	0.084	0.016	0.032	0.014	0.362	0.307	0.022

This suggests that the results in table 7 arise because the perturbations $\tilde{D}(G_1)$ in (5.1) are comparable in size with the differences between the pairs of 7×7 sub-images sampled from the area covered by the bowling ball. Table 8 shows that the larger values of n and larger values of $P(T(H)|H)$ are obtained if the sub-images are increased in size to 15×15, and k is increased proportionately to 56. (Note that for this large value of k, the covariance of the directions of the vectors in $\tilde{V}$ (see (4.2)) is near to $k^{-1}I(k)$ and thus the matrix A in (4.5) can be set equal to $I(k)$.) The agreement between $P(T(H)|H)$ and $\hat{P}(H)$ is less in table 8 than in tables 4–6.

7. Conclusion

The sub-images of an image have been mapped to a measurement space $\mathbb{R}^k$ in which they are modelled by a Gaussian distribution $\mathcal{N}(0, I(k))$. The differences between the measurement vectors of matching sub-images are modelled using an Ornstein–Uhlenbeck process which has $\mathcal{N}(0, I(k))$ as a limiting distribution. The resulting statistical model for image matching is tested by using it to calculate certain probabilities that measure the performance of a stereo matching algorithm. The calculated probabilities are compared with those obtained experimentally by applying the algorithm to images from the Middlebury stereo database. The performance of the algorithm is successfully predicted. In particular, if the probability of accepting a correct match is relatively large, then there is good agreement between the calculated and the experimental probabilities of obtaining a unique match which is also a correct match.

The probability density function for sub-images and the associated Ornstein–Uhlenbeck process has many potential applications, including image registration, object detection and stereo matching. In the case of stereo matching it may be possible to predict the amount of information contributed by each part of a matching algorithm, and in this way find out which parts of the algorithm are optimal and which parts could be improved.

References

Abramowitz, M. & Stegun, I. A. (eds) 1965. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. New York, NY: Dover.

Baraniuk, R. G. 2007 Compressive sensing. *IEEE Signal Process Mag.* **24**, 118–121. (doi:10.1109/MSP.2007.4286571)

Baron, D., Wakin, M. B., Duarte, M. F., Sarvotham, S. & Baraniuk, R. G. Submitted. Distributed compressed sensing. Preprint. See http://www.dsp.ece.rice.edu.

Bingham, E. & Mannila, H. 2001 Random projection in dimensionality reduction: applications to image and text data. In *Proc. 7th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* (*KDD-2001*), *San Francisco, CA, USA*, pp. 245–250.

Charpiat, G., Faugeras, O. & Keriven, R. 2005 Image statistics based on diffeomorphic matching. In *Proc. Int. Conf. on Computer Vision, ICCV 2005*, vol. 1, pp. 852–857.

Dasgupta, S. 1999 Learning mixtures of Gaussians. In *Fortieth Annual IEEE Symp. on Foundations of Computer Science* (*FOCS*), pp. 634–644.

Dasgupta, S. 2000 Experiments with random projections. In *Proc. 16th Conf. on Uncertainty in Artificial Intelligence*, pp. 143–151.

Diaconis, P. & Freedman, D. 1984 Asymptotics of graphical projection pursuit. *Ann. Stat.* **12**, 793–815. (doi:10.1214/aos/1176346703)

Donoho, D. L. 2006 Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306. (doi:10.1109/TIT.2006.871582)

Field, D. 1987 Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. Ser. A* **4**, 2379–2394. (doi:10.1364/JOSAA.4.002379)

Forsyth, D. A. & Ponce, J. 2003 *Computer vision, a modern approach*. Upper Saddle River, NJ: Prentice Hall.

Goel, N., Bebis, G. & Nefian, A. 2005 Face recognition experiments with random projection. In *Proc. SPIE 5779 Defence and Security Symposium, Orlando, Florida*, pp. 426–437.

Gonzalez, R. C. & Woods, R. E. 2002 *Digital image processing*, 2nd edn. Upper Saddle River, NJ: Prentice Hall.

Grenander, U. & Srivastava, A. 2001 Probability models for clutter in natural images. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 424–429. (doi:10.1109/34.917579)

Han, P. Y. J. & Jin, A. T. B. 2007 Random projection with robust linear discriminant analysis model in face recognition. In *Proc. Computer Graphics Imaging and Visualization, CGIV 2007*, pp. 11–15.

Hirschmüller, H. & Scharstein, D. 2007 Evaluation of cost functions for stereo matching. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8. *CVPR 2007, Minneapolis, MN.*

Huang, J. 2000 Statistics of natural images. PhD thesis, Division of Applied Mathematics, Brown University.

Huang, J. & Mumford, D. 1999 Statistics of natural images and models. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* vol. 1, pp. 541–547. *CVPR 1999, Fort Collins, CO.*

Jaynes, E. T. 2003 *Probability theory: the logic of science*. Cambridge, UK: Cambridge University Press.

Karatzas, I. & Shreve, S. E. 1988 *Brownian motion and stochastic calculus Graduate texts in mathematics*, vol. 113. New York, NY: Springer.

Konishi, S., Yuille, A. & Coughlan, J. 1999 Fundamental bounds on edge detection: an information theoretic evaluation of different edge cues. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* pp. 573–579. *CVPR 1999, Fort Collins, CO.*

Lee, A. B., Pederson, K. S. & Mumford, D. 2003 The nonlinear statistics of high-contrast patches in natural images. *Int. J. Comput. Vis.* **54**, 83–103. (doi:10.1023/A:1023705401078)

Moran, P. A. P. 1984 *An introduction to probability*. Oxford, UK: Oxford Science Publications, Oxford University Press.

Mumford, D. & Gidas, B. 2001 Stochastic models for generic images. *Q. Appl. Math.* **LIV**, 85–111.

Ruderman, D. L. 1997 Origins of scaling in natural images. *Vision Res.* **37**, 3385–3398. (doi:10.1016/S0042-6989(97)00008-4)

Samorodnitsky, G. & Taqqu, M. S. 1994 *Stable non-Gaussian random processes: stochastic processes with infinite variance*. New York, NY: Chapman and Hall.

Scharstein, D. & Pal, C. 2007 Learning conditional random fields for stereo. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2007, pp. 1–8. CVPR, Minneapolis, MN.*

Scharstein, D. & Szeliski, R. 2002 A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* **47**, 7–42. (doi:10.1023/A:1014573219977)

Srivastava, A., Lee, A. B., Simoncelli, E. P. & Zhu, S.-C. 2003 On advances in statistical modelling of natural images. *J. Math. Imag. Vis.* **18**, 17–33. (doi:10.1023/A:1021889010444)

Tan, S. & Jiao, L. 2007 Multivariate statistical models for image denoising in the wavelet domain. *Int. J. Comput. Vis.* **75**, 209–230. (doi:10.1007/s11263-006-0019-7)

Tappen, M. F. R., Russell, B. C. & Freeman, W. T. 2003 Exploiting the sparse derivative prior for super-resolution and image demosaicing. In *Third IEEE Workshop on Statistical and Computational Theories of Vision.*

Torralba, A. & Oliva, A. 2003 Statistics of natural image categories. *Netw. Comput. Neural Syst.* **14**, 391–412. (doi:10.1088/0954-898X/14/3/302)

van Hateren, J. H. & van der Schaaf, A. 1998 Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. B* **265**, 359–366. (doi:10.1098/rspb.1998.0303)

Wolfram, S. 1999 *The Mathematica book*, 4th edn. Cambridge, UK: Cambridge University Press.