

# Whitening for Photometric Comparison of Smooth Surfaces under Varying Illumination

Margarita Osadchy<sup>1</sup>, Michael Lindenbaum<sup>2</sup>, and David Jacobs<sup>3</sup>

<sup>1</sup> NEC Labs America, Princeton, NJ

<sup>2</sup> Dept. of Computer Science, The Technion, Haifa, Israel

<sup>3</sup> Dept. of Computer Science, The University of Maryland, College Park, Maryland

**Abstract.** We consider the problem of image comparison in order to match smooth surfaces under varying illumination. In a smooth surface nearby surface normals are highly correlated. We model such surfaces as Gaussian processes and derive the resulting statistical characterization of the corresponding images. Supported by this model, we treat the difference between two images, associated with the same surface and different lighting, as colored Gaussian noise, and use the whitening tool from signal detection theory to construct a measure of difference between such images. This also improves comparisons by accentuating the differences between images of different surfaces. At the same time, we prove that no linear filter, including ours, can produce lighting insensitive image comparisons. While our Gaussian assumption is a simplification, the resulting measure functions well for both synthetic and real smooth objects. Thus we improve upon methods for matching images of smooth objects, while providing insight into the performance of such methods. Much prior work has focused on image comparison methods appropriate for highly curved surfaces. We combine our method with one of these, and demonstrate high performance on rough and smooth objects.

## 1 Introduction

Comparing images is a fundamental part of computer vision systems that perform recognition, alignment and tracking. Many approaches have tackled the critical problem of accounting for lighting variations [6, 11, 13, 2, 3] when making comparisons. These methods work well on rough objects containing discontinuities or places of rapid change in albedo or shape. However, comparing images of smooth surfaces with no edges or texture under varying illumination remains a challenging problem. This problem is important since most real surfaces contain rough and smooth regions. Handling smooth regions is important for improved recognition or dense registration or tracking of such objects. In this paper we propose a new measure for image comparison of smooth surfaces, and demonstrate its value on the problem of object identification under fixed pose but varying lighting.

There are three things that seems to be very important in constructing a representation for image comparison. First, finding a representation that captures similarities between images of the same object (eg., through quasi-invariance). Second, also capturing dissimilarity between images of different objects. Third, choosing an optimal measure for comparing the resulting representations. Most previous methods have focused on the first problem, by choosing representations of images that are invariant, or quasi-invariant to lighting. Edges are a classic example. [3] discuss the quasi-invariance to lighting changes of operators that use derivatives. Gabor jets are also widely used for image comparison, in part because they are also considered to be insensitive to lighting changes (eg., [13]). [6] point out that the direction of the gradient is relatively insensitive to lighting changes. However, it is well-known that quasi-invariance to lighting changes is difficult to achieve for smooth objects.<sup>4</sup> Hence we will not focus on invariant representations, but tackle the other two problems: increasing dissimilarity between images of different objects while constructing an optimal comparison measure.

The primary problem presented by smooth objects is that nearby albedos and surface normals are highly correlated, which causes correlations in nearby intensities in their images. Consequently, comparisons that treat neighboring pixels as independent, such as sum-of-squared-differences (SSD) are not statistically valid. Moreover, correlations between image pixels improve the chances that images of two different objects will match well, since if they are similar at one point, they are likely to be similar at many. We approach this problem by constructing a statistical model of the dependencies between neighboring portions of smooth shapes. We then use this to model the effect that lighting changes have on the appearance of a smooth object. We can then design operators to decorrelate the pixels in images of these objects.

We use whitening to lessen dependencies in the difference between two images. Signal detection theory tells us that this is the optimal approach when the

---

<sup>4</sup> This is made explicit in the analysis of [6], which shows that gradient direction is truly invariant to lighting direction for surfaces with discontinuities, and varies more rapidly with smoother objects.

difference between images of the same object consists of colored (non-independent) Gaussian noise [17]. We show that for a simple model of smooth surfaces, this is a good characterization.

Whitening has often been used for decorrelation of images in image processing tasks such as watermarking [8, 7], image restoration [18, 4, 5], and texture feature extraction [9, 14]. Many methods have used some differential operators or the Laplacian [16] to approximate the whitening filter, though [14] used a 2D causal linear prediction model to derive whitening filters.

Whitening decorrelates image intensities, but it does not make them insensitive to lighting variation. In fact, we prove that no linear filter can produce an image representation that is more insensitive to lighting variation than the original image. One consequence of this is to prove that non-linear lighting insensitive methods for rough surfaces, such as the direction of gradient, are more lighting insensitive than any possible linear filter.

To summarize, whitening, like any linear filtering, does not make images of the same object more similar. However, it helps to increase dissimilarity between images of different objects and allows us to use SSD as the optimal measure for comparison. These make whitening a superior comparison method for smooth surfaces, which we confirm in our experiments on synthetic and real data. We combine whitening with the direction of gradient to produce a comparison method that performs very well on both smooth and rough objects.

We first presented the idea of using whitening for image comparison in our recent workshop paper [1]. The discussion of invariance in Section 3 is completely new, as is the motivation for using whitening described in Section 2.4 and the combined recognition method and experiments described in Section 4.3.

## 2 The Whitening Approach

As mentioned above, discrimination between smooth objects is difficult due to the high correlation between nearby pixels in their images. One consequence of this is that pixel by pixel comparisons such as SSD are not optimal. In this section we show how to derive linear filters that remove correlations between neighboring pixels. These *whitened* images can then be optimally compared using SSD. We take a statistical approach, regarding the difference image,  $I_d = I_1 - I_2$  as a random variable ( $I_1$  and  $I_2$  denote two images of the same surface). We analyze this considering a Lambertian surface illuminated by distant point sources. Neglecting shadows, we can model the images as:  $I_1 = \rho \hat{N} s_1$  and  $I_2 = \rho \hat{N} s_2$ , where  $\hat{N}$  are surface normals,  $\rho$  is albedo, and  $s_1, s_2$  are light sources in two images. Then

$$I_d = \rho \hat{N} s_1 - \rho \hat{N} s_2 = \rho \hat{N} (s_1 - s_2) \quad (1)$$

Dependencies that exist between nearby surface normals of an object lead to dependencies in  $I_d$ , which we treat by modeling  $I_d$  as *colored* Gaussian noise. (Colored Gaussian noise captures noise with dependencies, whereas white noise is independent.) While this model is not strictly true, it is a valuable approximation

that opens the way to using a whitening filter, which is a standard tool in signal detection, to reduce dependency in the difference image.

## 2.1 Whitening in Signal Processing

First we describe whitening. Let  $\mathbf{n}$  represent a set of pixels in the difference image, as a vector. Assume that  $\mathbf{n}$  is Gaussian colored noise. This implies that it is fully characterized by its first and second order statistics. In particular, the whitening filter may be designed using the covariance matrix. Let  $C = E[\mathbf{nn}^T]$  be the covariance matrix characterizing the distribution of  $\mathbf{n}$  ( $E$  denotes expected value). Let  $W$  be a matrix composed of the scaled eigenvectors of  $C$ ,  $\frac{1}{\sqrt{\lambda_i}}\mathbf{e}_i$  as rows. Then, the components of  $\mathbf{y} = W\mathbf{n}$  are independent, as implied from their Gaussianity and their covariance:

$$E[\mathbf{yy}^T] = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$$

That is, the multiplication by the matrix  $W$  “whitens” the vector  $\mathbf{n}$ .

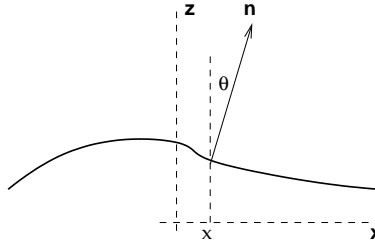
## 2.2 A Model for Natural Images - Rough Plane Covariance

To whiten a surface’s images, we must understand their covariance structure. Consider a surface characterized by normal vectors that make small random perturbations about a common direction (without loss of generality the  $z$  axis). We refer to such a surface as *roughly planar* and assume that locally a smooth surface behaves like a roughly planar surface. This is a generalization of the common facet model [10]. Considering the simplified, 1D, variant, the “surface” is described by a function  $z = f(x)$ . The normals at every point  $x$  are random (but not independent!) and each of them is specified by a single parameter  $\theta$ , which is its angle relative to the  $z$  axis (Figure 1). Quantitatively we characterize the function  $\theta(x)$  as a wide sense (w.s.) stationary Gaussian random process [15]. That is, we assume that the expected value at every point is constant  $\mu_\theta = 0$ , that the variance  $C_\theta(x, x) = \sigma_\theta^2$  is constant as well, and that the auto-correlation  $C_\theta(x_1, x_2) = r(x_1, x_2)\sigma_\theta^2 = r(|x_1 - x_2|)\sigma_\theta^2$  depends only on the distance between two points.  $r(|x_1 - x_2|)$  is a correlation coefficient. We also assume that the surface is Lambertian, and that its albedo  $\rho$ , is constant, at least locally.

**Proposition 1:** Under the above assumptions and for a distant light source, illuminating the surface at angle  $\phi$  (relative to the  $z$  axis), the reflected light function  $I(x)$  is a random w.s. stationary process. Its expected value, variance and auto-correlation are:

$$\begin{aligned} E[I(x)] &= \rho \cos \phi e^{-\sigma_\theta^2/2} \\ \sigma_I^2 &= \frac{1}{2} \rho^2 (\sin^2 \phi (1 - e^{-2\sigma_\theta^2}) + \cos^2 \phi (1 - e^{-\sigma_\theta^2})^2) \\ C_I(x_1, x_2) &= \frac{1}{2} \rho^2 (\sin^2 \phi e^{-\sigma_\theta^2} (e^{r\sigma_\theta^2} - e^{-r\sigma_\theta^2}) + \cos^2 \phi (e^{-\sigma_\theta^2} (e^{r\sigma_\theta^2} + e^{-r\sigma_\theta^2}) - 2e^{-\sigma_\theta^2})) \end{aligned} \quad (2)$$

where  $x_1, x_2$  are the two points for which the correlation coefficient of the tangent direction is  $r = r(|x_1 - x_2|)$ .



**Fig. 1.** The roughly planar (random) surface is specified (in 2D approximation) by the angle  $\theta(x)$  that the normal makes with the  $z$  direction.

*Proof.* (For details see [1]) The reflected light function  $I(x)$  is a random process. Let  $x_1, x_2$  be two points for which the correlation coefficient of the tangent direction is  $r = r(\|x_1, x_2\|)$ . Then, their autocorrelation is

$$\begin{aligned}
 C_I(x_1, x_2) &= E[(I(x_1) - E[I(x)])(I(x_2) - E[I(x)])] \\
 &= \rho^2 E[(\sin\phi\sin\theta_1 + \cos\phi\cos\theta_1 - \cos\phi E[\cos\theta_1]) \cdot \\
 &\quad (\sin\phi\sin\theta_2 + \cos\phi\cos\theta_2 - \cos\phi E[\cos\theta_2])] \\
 &= \rho^2 (\sin^2\phi E[\sin\theta_1\sin\theta_2] + \cos^2\phi E[\cos\theta_1\cos\theta_2] - \cos^2\phi E[\cos\theta]^2) \\
 &= \frac{1}{2}\rho^2 (\sin^2\phi e^{-\sigma_\theta^2} (e^{r\sigma_\theta^2} - e^{-r\sigma_\theta^2}) + \cos^2\phi (e^{-\sigma_\theta^2} (e^{r\sigma_\theta^2} + e^{-r\sigma_\theta^2}) - 2e^{-\sigma_\theta^2}))
 \end{aligned}$$

Note that all  $\sin\theta_i\cos\theta_j$  terms vanish due to symmetry. The rest of the derivation requires us to change variables, to the sum and difference of  $\theta_1$  and  $\theta_2$ , which are independent. Simple trigonometric expressions and the Gaussian integral  $\int_{-\infty}^{\infty} \cos x e^{-x^2/2a^2} dx = \sqrt{2\pi}|a|e^{-a^2/2}$  are used as well.

For rougher surfaces (larger  $\sigma_\theta^2$ ) correlation decreases while for the (impossible) white surface (independent normals,  $r = 0$ ), the image is white as well.

The covariance in eq. 2 is non-stationary and it varies with  $\phi$ . It can be shown however, that of the two additive terms in the covariance expression the first is dominant, provided the surface is smooth ( $\sigma_\theta$  is small) and that the illumination angle  $\phi$  is not very small. This readily implies that:

**Covariance characterization for rough Lambertian plane:** the second order statistical behavior of a rough Lambertian, planar surface, illuminated by a single source, is characterized by an autocorrelation function which, for nearly every illumination, is approximately invariant of the illumination direction up to a multiplicative factor.

See [1] for experimental validation of this result for real objects.

### 2.3 Whitening using AR models

Designing a whitening filter by estimating the covariance is problematic as the covariance (and the mean) are nonstationary. Fortunately, fitting a parametric

Autoregressive (AR) model, allows us to get the whitening filter directly without explicitly estimating covariance [12].

A sequence  $x(n)$  is called an AR process of order  $p$  if it can be generated as the output of the recursive causal linear system

$$x(n) = \sum_{k=1}^p a(k)x(n-k) + \varepsilon(n), \forall n \quad (3)$$

where  $\varepsilon(n)$  is white noise, and the sum  $\bar{x}(n) = \sum_{k=1}^p a(k)x(n-k)$ , is the best linear mean squared (MS) predictor of  $x(n)$  based on the previous  $p$  samples. Given a random sequence (with possible dependencies), an AR model can be fitted using SVD to estimate the overdetermined parameters  $a(k)$  which minimize the empirical MS prediction error  $\sum_n (x(n) - \bar{x}(n))^2$ . For Gaussian signals the prediction error sequence:  $\varepsilon(n) = x(n) - \bar{x}(n)$  is white, implying that the filter  $W = (1, -a_1, \dots, -a_p)$  is a whitening filter for  $x(n)$ . We have adopted a 2D “causal” model described in [12], where a gray level  $x(n)$  is predicted from the previous gray levels in a  $p \times p$  neighborhood in column by column scan. Using a non-causal neighborhood leads to a lower SSD, but the prediction error sequence is not white [12].

Note that scaling all the grey levels by the same factor would give a correlation function which is the same up to a multiplicative constant. This is essentially what happens when the angle between the average normal and the illumination direction changes. Fortunately, this does not change either the AR coefficients, or the resulting whitening filter, implying that it can be space invariant.

The whitening filter depends on the image statistics. Intuitively, for smoother images the correlation is larger and decorrelating it requires a wider filter. For images which are not so smooth the decorrelation is done over a small range, and the filter looks very much like the Laplacian, which is also known to have some whitening effect. Therefore, for rougher images, we do not expect to perform better than an alternative procedure using the Laplacian. As we shall see later, for smooth objects the performance difference is significant.

## 2.4 Whitening Images from Different Objects

Signal detection theory tells us that whitening is useful for image comparison because whitened images from the same object can be optimally compared using SSD. Whitening has another advantage, it makes images from different objects more distinctive.

To see this, let  $S$  denote a 3D surface. We will take two pictures of  $S$  in a fixed pose with two different point sources of light,  $s_1$  and  $s_2$ .  $s_1, s_2$  are each  $3 \times 1$  vectors that encode lighting direction and magnitude.  $p_{i,j}$  denotes a patch of the surface corresponding to an image pixel. We approximate  $p_{i,j}$  as a planar patch, with surface normal  $\hat{N}_{i,j}$ , and albedo  $\rho_{i,j}$ . It will be convenient to denote the scaled surface normal  $\rho_{i,j}\hat{N}_{i,j}$  by  $N_{i,j}$ . We denote the image pixels corresponding to  $p_{i,j}$  by  $I_{1,i,j}, I_{2,i,j}$  in the two images. So we may write, for example,  $I_{1,i,j} = N_{i,j}^T s_1$ , since we ignore the effects of shadows.

Let  $L$  denote a whitening filter, represented discretely as a matrix with elements  $L_{k,l}$ . Without loss of generality we suppose  $L$  is square and  $-n \leq k, l \leq n$ . If we apply this filter to the image  $I_1$  we denote the output as  $\mathcal{I}_1$ . So:

$$\mathcal{I}_{1,i,j} = \sum_{k=-n}^n \sum_{l=-n}^n L_{k,l} I_{1,i+k,j+l}$$

We can define a new surface,  $\mathcal{S}$ , such that its scaled surface normals are:

$$\mathcal{N}_{i,j} = \sum_{k=-n}^n \sum_{l=-n}^n L_{k,l} N_{i+k,j+l}$$

Intuitively,  $\mathcal{S}$  can be thought of as the surface filtered by  $L$ . According to our model, while the original normals are highly correlated, the whitened normals will be white noise, with randomized directions and scales. As high-dimensional, white noise, different whitened surfaces will also be uncorrelated with each other, with high probability. This is analogous to taking a smooth, white surface and splattering it with gray paint. Smooth surfaces are easily confused with each other, while highly textured ones are not. Of course, whitening does not add differences to signals, it makes explicit the differences that are already there.

More formally, communication theory tells us that discriminating between correlated models is difficult. Specifically, for two unit energy signals  $z_1(x), z_2(x)$ , the correlation coefficient is  $\rho_{ij} = \int z_i(x) z_j(x) dx$ . For best performance, the correlation coefficient between any pair of models should be as low as possible. For two signals the lowest correlation is  $-1$ , and choosing  $z_2(x) = -z_1(x)$  is optimal. When the number of signals is large, such correlations between all signals are not possible, and the best we can get is  $\rho \approx 0$  [17].

Whitening treats the signals and the noise equally and therefore leaves the signal to noise ratio (SNR) the same. However the whitened signals become uncorrelated and therefore with the same SNR we get better performance. The correlation between the original images associated with different objects is high initially and is almost zero afterwards, so the improvement is significant.

### 3 Invariance and Linear Filtering

While most prior work has focused on finding lighting insensitive image comparisons, we have not argued that whitening is lighting insensitive. We now prove a result that casts doubt on the ability of any linear filter to produce lighting insensitive representations.

**Theorem 1.** *Suppose that the lighting directions  $s_1$  and  $s_2$  are drawn from a uniform distribution, and that we neglect the effects of shadows in images. Then  $I_{1,i,j}/I_{2,i,j}$  and  $\mathcal{I}_{1,i,j}/\mathcal{I}_{2,i,j}$  are identically distributed. That is, the distribution of the ratio of intensities between one image of an object and another are unaffected by filtering with an arbitrary linear filter. In this sense, no linear filter can produce a lighting insensitive representation.*

*Proof.* This follows immediately once we consider that linearly filtering the images is equivalent to filtering the surface, as described above. Let  $\mathcal{N}$  denote the filtered normals, as above, but now for an arbitrary linear filter. Let  $\hat{\mathcal{N}}_{i,j}$  denote a unit vector in the direction of  $\mathcal{N}_{i,j}$ .

$$\frac{I_{1,i,j}}{I_{2,i,j}} = \frac{\hat{N}_{i,j} s_1^T}{\hat{N}_{i,j} s_2^T} \quad \frac{\mathcal{I}_{1,i,j}}{\mathcal{I}_{2,i,j}} = \frac{\hat{\mathcal{N}}_{i,j} s_1^T}{\hat{\mathcal{N}}_{i,j} s_2^T}$$

Since  $s_1$  and  $s_2$  are uniformly distributed it is clear from symmetry that these two fractions are identically distributed, because  $\hat{N}$  and  $\hat{\mathcal{N}}$  are identical up to a rotation. In sum, we have created a filtered surface that is affected by lighting changes exactly as the original surface.

It is possible to extend this result to handle the case of attached shadows by restricting the distribution of light sources to appear in a hemisphere above the surface normal. However we omit details of this for lack of space.

Our result may seem to contradict the well known fact that derivative operators are invariant to constant changes in illumination, but this is not so. The effect of constant illumination on a surface patch is equivalent to that of a point source in the direction of  $\hat{N}$ , since this is the average of all the illumination that strikes the surface. Consequently, the unfiltered surface normal is sensitive to constant illumination, since it points in this direction, while the derivative of the surface normal is insensitive to constant lighting, since it is orthogonal to  $\hat{N}$ . The unfiltered and filtered surfaces are equally sensitive to lighting changes from random directions, but the unfiltered surface is most sensitive to lighting changes from above. However, while lighting itself contains significant constant (DC) components, it is not clear that lighting *changes* tend to have large constant components.

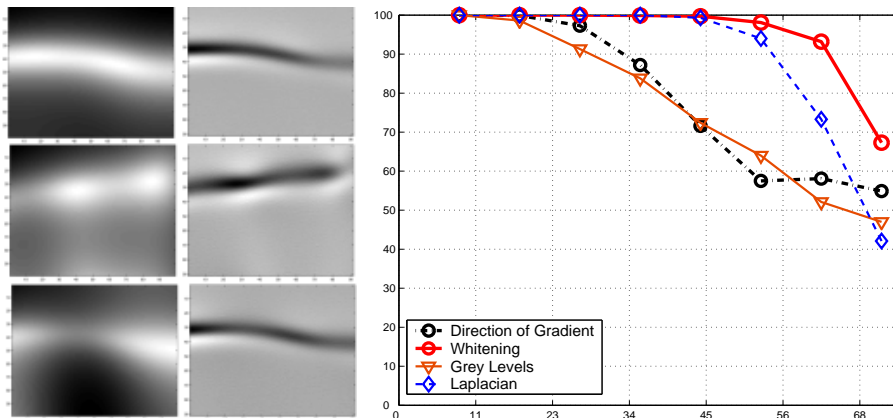
Our results suggest that we should not seek lighting insensitive representations with linear filters, although as we have shown, linear filters that decorrelate images can improve image comparison. Another possible response to this result is to seek lighting insensitivity in non-linear operations, such as edge detection. In fact, one implication of our result is that the direction of the gradient is more insensitive to lighting than any possible linear operator, since [6] shows that the direction of the gradient is less sensitive to lighting than the image itself.

## 4 Experiments

We tested our ideas by applying them to object recognition. A set of objects is represented in a library containing one image for every object. Let  $I_{M_1}, I_{M_2}, \dots$  be reference images in the library. Let  $I_Q$  be the query image of one of the objects from this set, taken with the same pose, but different illumination. The task is to decide which of the objects is the one in the query image.

Since the reference image  $I_{M_j}$  was taken with a different illumination intensity than the test image, every scaled version of it is a valid model as well.





**Fig. 2.** The top and the center images in the left column correspond to different surfaces and one illumination. The bottom image is created from the same scene used for the top image, but with a different illumination. The center column shows the whitened images and illustrates that whitening reveals hidden differences. The plot on the right shows recognition performance of the tested methods on the synthetic images. The success rate is plotted against the average angle between the illumination source and the average surface normal.

Minimizing the SSD over all scaled versions is equivalent to taking the SSD between the normalized whitened images. This normalization also compensates for the fact that some objects are rougher than others, which makes the difference between two differently illuminated images of them larger.

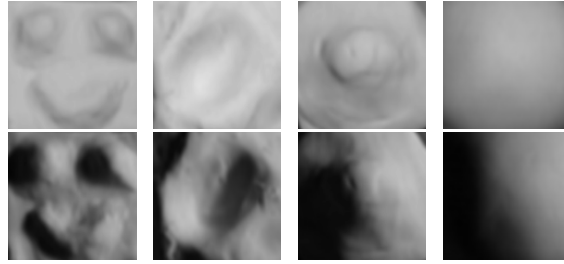
Therefore we perform the following steps: 1) For every reference image,  $I_{M_j}$ , use the whitening operator  $W$ , to calculate the normalized  $L_2$  norm  $E_j = \left\| \frac{W(I_{M_j})}{\|W(I_{M_j})\|} - \frac{W(I_Q)}{\|W(I_Q)\|} \right\|$ . 2) Choose the model associated with the smallest whitened error norm,  $E_j$ .

We tested the whitening approach on smooth textureless surfaces. We also integrated whitening with a comparison method designed for rough surfaces, and showed that this combined method could work on rough and smooth surfaces.

#### 4.1 Synthetic images

The first set of experiments was done using synthetic images. Every scene was created as a sum of random harmonic functions, with fixed amplitudes but random directions and phases. This provides an ensemble of images with similar statistical properties. These were rendered as Lambertian surfaces with point sources.

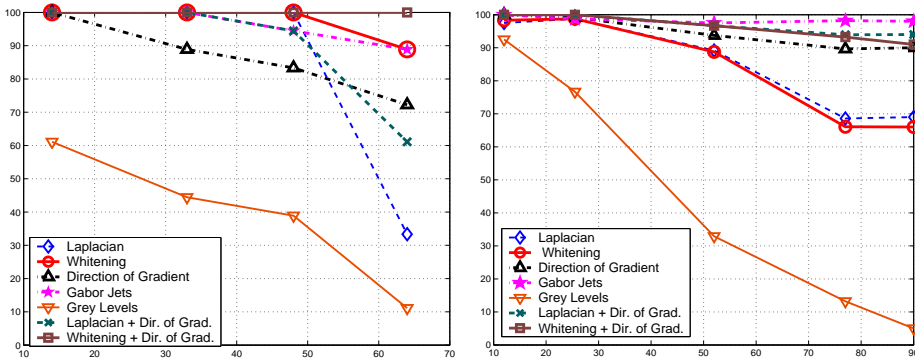
We trained a whitening filter using 1000-5000 images with a fixed illumination, deviating 67.5 degrees from the  $z$  direction. The training set was independent of the test set. A test was done as follows: two random scenes were illuminated by the same nearly vertical illumination to create two references



**Fig. 3.** Samples from the smooth real objects data set; top – frontal illumination, bottom – side illumination.

images  $I_r, I'_r$ . The test image  $I_t$  was synthesized from the first scene, with a different illumination, making an angle  $\phi$  with the  $z$  axis (see Figure 2).

For comparison we also tested other algorithms using the SSD of the gray level image, a Laplacian filtered image, and the direction of the gradient<sup>5</sup>. See Figure 2 for the results.



**Fig. 4.** Recognition performance of the tested methods on real smooth objects on the left and rough objects (Yale database) on the right. The success rate is plotted against the average angle (in degrees) between the illumination source and the average surface normal.

We came to several conclusions. First, whitening was the most successful method. Second, whitening worked best with a large filter, but it also worked substantially better than other methods even with a  $7 \times 7$  filter, except for extreme illumination angles. In particular whitening was always better than the Laplacian, even when a  $3 \times 3$  filter was used, implying that both large distance correlations and causality are important.

<sup>5</sup> We did not test Gabor Jets on the synthetic images, but later experiments show they are not especially effective on smooth surfaces.

## 4.2 Real Smooth Objects

Next, we describe experiments with real, smooth objects that produce images with substantial shadows (Figure 3). We created eighteen objects from clay and illuminated them by a single light source moving along a half circle, so that its distance from the object was roughly fixed. We used a camera placed vertically above the object, and took 14 images of every object with different lighting directions at angles in the range  $[-70, 70]$  degrees to the vertical axis. One image of each object, associated with a nearly vertical illumination, were chosen as the reference images.

The whitening filter was trained on the difference images between reference images and corresponding images associated with the same object and six other illuminations. Only twelve images associated with 2 objects (out of 18) were used. We learned the whitening filter, as a 2D causal filter with 25 coefficients inside  $7 \times 7$  windows. All images of the 18 objects except the reference images were used as query images (234 images). We divided the query images into four groups according to their angular lighting direction:  $10^\circ - 25^\circ$ ,  $26^\circ - 40^\circ$ ,  $41^\circ - 55^\circ$ , and  $56^\circ - 70^\circ$ .

The plot in Figure 4 shows our results. Whitening again performed better than the other methods. We also observed that for a few of the roughest objects, the Laplacian, whitening and gradient angle performed equally well. For smoother objects, however, whitening worked considerably better. The Laplacian couldn't whiten the smooth surfaces, because its size was insufficient to handle the high correlations between the grey levels of the smooth surfaces.

## 4.3 The Combined Method

To handle objects that may be rough or smooth, we propose that whitening be combined with a measure that is geared towards handling rough objects, such as the direction of gradient. We have done a proof-of-concept implementation of a simple combined method. Direction of gradient is naturally normalized to the  $[0, \pi]$  range. Whitening, however, requires normalization prior to combining. Let  $s_1, s_2, \dots, s_n$  denote the distances between the query image and  $n$  reference images after whitening. We normalize them to the  $[0, 1]$  range by dividing all the distances by  $\max |s_i|$ . Different areas in the image can have different roughness levels. We compensate for this effect by choosing the normalization factor adaptively in  $10 \times 10$  pixel areas instead of the whole image. Our experiments showed that adaptive normalization yields better results. We have also scaled the direction of gradient output to the  $[0, 1]$  range. We have tested the combined method on both smooth (Figure 4 left) and rough data (Figure 4 right) sets. As a smooth set we took the clay objects described in the previous section. As a rough set we took the Yale database [6], which contains 20 objects with abrupt changes in albedo and shape. The database consists of 63 images of each object with lighting direction deviating up to 90 degrees from the frontal. Our experiments showed that the combination of whitening and direction of gradient (CWD) was better than either whitening or direction of gradient alone on both data sets;

and CWD had the best (and perfect) performance on the smooth set. On the rough data the combined method performed very well, but not as well as Gabor Jets. In future work we plan to continue this approach and try to find a more clever combining technique that will integrate whitening with some variation of Gabor Jets. We also tested a combination of Laplacian and direction of gradient. This combination performed less well than CWD on smooth data and similar to CWD on the rough data. The Laplacian has some whitening effect, which explains its good performance on smooth data. On the other hand, decorrelation in the rough objects occurs over a small range, and the whitening filter looks very much like the Laplacian explaining the results on the rough set.

## 5 Conclusions

In this work we have proposed a measure for image comparison of smooth surfaces under varying illumination. The measure was motivated by a simple statistical model of smooth surfaces. This model showed that the error between two images associated with the same object under different lighting may be modelled as colored noise. We adapted well-known techniques of whitening to perform matching of images corrupted by such noise.

We found that whitening was more effective than other representations for comparing images of smooth surfaces taken under varying illumination conditions. Previous methods have commonly used the Laplacian or the magnitude of gradient, as whitening approximations. This seems to be adequate for rough images but leads to inferior results for smoother ones.

We believe that recognition (or image comparison in general) should use all the image information. Many current methods neglect photometric information and thus cannot handle smooth objects. Our preliminary results showed that a proper combining method, using both the information in edges and the in smooth patches, would yield superior results, especially in hard tasks.

## References

- 1.
2. P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class-specific linear projection. *PAMI*, 19(7):711–720, July 1997.
3. R. Brunelli and T. Poggio. Face recognition: Features versus templates. *PAMI*, 15(10):1042–1062, 1993.
4. B. Bundschuh. A linear predictor as a regularization function in adaptive image restoration and reconstruction. In *5th Int. Conf. on Computer Analysis of Images and Patterns*, 1993.
5. H. Bundschuh, B. Schulz and D. Schneider. Adaptive least squares image restoration using whitening filters of short length. In *Second HST Image Restoration Workshop*, 1993.
6. H.F. Chen, P.N. Belhumeur, and D.W. Jacobs. In search of illumination invariants. In *CVPR00*, pages I: 254–261, 2000.

7. M. L. Cox, I. J. Miller and Bloom J. A. *Digital Watermarking*. Morgan Kaufmann, 2002.
8. T. Depovere, G. Kalker and J.P. Linnartz. Improved watermark detection using filtering before correlation. In *IEEE Int. Conf. on Image Processing*, pages I: 430–434, 1998.
9. O.D. Faugeras and W.K. Pratt. Decorrelation methods of texture feature extraction. *PAMI*, 2(4):323–332, July 1980.
10. R.M. Haralick and L.G. Shapiro. Computer and robot vision. In *Addison-Wesley*, 1992.
11. D.W. Jacobs, P.N. Belhumeur, and R. Basri. Comparing images under variable illumination. In *CVPR98*, pages 610–617, 1998.
12. A.K. Jain. Fundamentals of digital image processing. In *Prentice Hall*, 1989.
13. M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *TC*, 42(3):300–311, March 1993.
14. Z. Lin and Y. Attikiouzel. Two-dimensional linear prediction model-based decorrelation method. *PAMI*, 11(6):661–665, June 1989.
15. A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 3rd edition, 1991.
16. W.K. Pratt. *Digital Image Processing (First Edition)*. Wiley, 1978.
17. H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I*. Wiley, New-York, 1965.
18. L.P. Yaroslavsky. *Digital Picture Processing. An Introduction*. Springer Verlag, Berlin, Heidelberg, 1985.