On the Equivalence of Common Approaches to Lighting Insensitive Recognition

Margarita Osadchy
University of Haifa
rita@osadchy.net

David W. Jacobs
University of Maryland
djacobs@cs.umd.edu

Michael Lindenbaum
The Technion
mic@cs.technion.ac.il

Abstract

Lighting variation is commonly handled by methods invariant to additive and multiplicative changes in image intensity. It has been demonstrated that comparing images using the direction of the gradient can produce broader insensitivity to changes in lighting conditions, even for 3D scenes. We analyze two common approaches to image comparison that are invariant, normalized correlation using small correlation windows, and comparison based on a large set of oriented difference of Gaussian filters. We show analytically that these methods calculate a monotonic (cosine) function of the gradient direction difference and hence are equivalent to the direction of gradient method. Our analysis is supported with experiments on both synthetic and real scenes.

1 Introduction

Changes in lighting have a big effect on the way a scene looks. This makes it difficult to match images of an object taken under different lighting conditions. We consider three of the most prominent methods for handling this problem: normalized correlation; comparison of normalized jets of oriented filters such as difference of Gaussians; and comparison of the direction of the gradient. In this paper we show that these methods are equivalent.

These methods are all popular. Normalized correlation is part of the standard toolbox of image processing (eg., [15]). It is widely used to account for lighting variations because it is invariant to changes in the offset (additive changes) and gain (multiplicative changes) of images (eg., [15]). Jets of oriented filters, such as difference of Gaussians (eg., [18, 23, 8, 16]) or Gabor jets (eg., [19]), have become the most ubiquitous representations in computer vision, especially for object recognition and texture analysis. Their use is primarily motivated by advantages unrelated to lighting, but invariance to offset and gain has also been seen as a plus. The direction of the gradient has been proposed for lighting insensitive recognition with increasing frequency ([2, 4, 14, 24, 6, 7, 9]), and has been shown to be invari-

ant or insensitive to changes in lighting direction for many 3D objects ([6]). However, the connections between these methods have not been well understood.

We show that a close relationship exists between normalization and the direction of the gradient. First, we show that in the limit, for a small window, normalized correlation between two image patches computes the cosine of the difference in angle between the direction of gradient in the two patches. Second, we consider normalized correlation of the output of a *jet* of oriented difference of Gaussians. We show that this computes the same thing. We also show that with minor modifications, these two methods compute a similar measure that is invariant to 180 degree changes in gradient direction (ie., its *polarity*). Methods that explicitly compare the direction of gradient may do so slightly differently, for example by computing the sum of square differences in the angle of this direction. However, we confirm experimentally that these subtle differences are not significant.

Currently, a researcher who wishes to perform lighting insensitive matching is faced by a wide range of options. It is not clear what approach might be best, and how this decision might depend on the domain. Determining this experimentally for a particular problem can be quite challenging. By proving that three prominent approaches to achieving lighting insensitivity are equivalent, we simplify this task.

We also feel that our results deepen our understanding of all these methods. In particular, our results show that approaches based on normalized correlation and methods based on jets possess much more powerful insensitivity to lighting variation than is suggested by their invariance to offset and gain. In 3D scenes containing regions of high curvature these methods can also be insensitive to changes in the direction of lighting.

2 Background

We will first briefly review some of the effects of lighting variation on scene appearance. Then we will describe methods for handling them, focusing on the three approaches analyzed in this paper.

We will consider the effects of lighting variation on Lam-

bertian objects. For Lambertian objects, the intensity of a pixel is: $i = l\alpha\hat{n} \cdot \hat{l}$ where i is intensity, l is the magnitude of the light, α is albedo, $\hat{n}$ is the surface normal, and $\hat{l}$ is the direction from the surface to the light. We will assume that light sources are distant and isotropic, and there are no cast shadows, so that light appears at the same intensity and angle relative to all object surfaces.

We first consider the simplest effects of lighting. When light sources change their intensity but do not move, all image intensities are simply scaled. For a planar object and distant light sources, changes in lighting direction also linearly scale the image intensities, because all surface normals have the same angle relative to any light source. For convex objects of uniform albedo, a diffuse light source in which constant intensity light comes from all directions adds a constant to all image intensities. In these simple cases, lighting variation has an additive (*offset*) or multiplicative (*gain*) effect on the image.

Changes in lighting can have a much more complex effect on three-dimensional objects. This is because a change in the direction of lighting changes the angle between lighting and surface normal differently at every scene point, making some points lighter and some dimmer. Chen et al.[6] illustrate this complexity by showing that given any two images, one can construct a Lambertian scene that will produce the first image when illuminated by a point source from one direction, and produce the second image when illuminated by a different point source. So, given one image of a scene, it is impossible to predict anything definite about how it will appear under different lighting.

When a scene contains discontinuities in shape or albedo, edges may be lighting invariant. However, for non-polyhedral objects, changes in illumination can drastically effect the edges produced in an image (eg., [26]). For this reason, edge based methods have been very successful, but only for limited classes of objects, or for matching under controlled lighting conditions.

Another method of coping with illumination is to normalize the intensities of images prior to comparison. The simplest approach is to adjust the offset and gain of the image to standard values (for discussion of another approach, histogram equalization, see, eg., [21, 17]). For example, one can subtract a constant value from the image, or a window in the image, to give it zero mean, and scale the image to give it unit variance. *Normalized cross-correlation* follows this normalization with correlation, and is a standard way to manage the effects of lighting change (eg., [15]. See [5] for an application of this to face recognition).

Many approaches match images using the output of multiscale oriented filters instead of raw pixel intensities (eg., [19, 23, 13, 18, 25]). By using filters that integrate to zero, such as difference of Gaussians or Gabor filters, such approaches become invariant to additive changes in the image.

Normalizing the total magnitude of all filter responses produces invariance to multiplicative changes. This invariance to offset and gain often helps motivate the use of multiscale oriented filters (eg., [29, 30, 8, 3, 16]).

A third approach to handling lighting uses representations of images that are less sensitive to lighting variation. A number of authors have proposed the direction of the image gradient for this purpose ([2, 4, 14, 24, 6, 7, 9]). Sometimes, this has been motivated by the invariance of the direction of gradient to changes in offset and gain (eg., [7, 9]) and indeed to any monotonic change in image intensity ([24]). Chen et al.[6] provide a statistical analysis of the behavior of the direction of gradient, in terms of scene structure. They show that the direction of gradient is invariant to lighting changes at scene discontinuities. Moreover, it is also relatively insensitive to lighting changes for a surface in which the ratio of one principal curvature to the other is high.

There are many possible ways to compare images using the direction of the gradient. Perhaps the simplest is to compute the sum of squares of the differences between two images. The difference between the direction of gradient in two images can range from 0 to π, or, if we compare the images in a manner insensitive to the polarity of the gradient, it ranges from 0 to $\frac{\pi}{2}$. We will call the polarity sensitive comparison 'DIRP', and the polarity invariant comparison 'DIR', for short. Chen et al. also suggest a somewhat superior method, relying on the distribution of direction differences using the statistics of image pairs that come from the same scene, under different lighting conditions.

To illustrate the potential differences between the direction of the gradient and normalization methods, consider the images produced by a uniform albedo, Lambertian cylinder, illuminated by a point light source (see Figure 1). A circular cross-section of the cylinder will produce intensities that look like a cosine function (but are clipped at zero so that they do not become negative). The phase of the cosine varies so that its peak is at the surface normal that points directly towards the light. Even if we normalize two of these intensity patterns from two images, their different phases can cause them to be largely uncorrelated. However, because a cylinder has zero curvature in one direction, the direction of the gradient of its image is always in the direction of maximum curvature. Therefore, for objects such as cylinders, normalization seems to have little ability to compensate for lighting variations, while the direction of the gradient can be completely insensitive to them. We will show, though, that when image normalization is local, this apparent difference between the power of these methods disappears.

Finally, there has been much work on lighting insensitive object recognition that uses cues outside the scope of this paper, such as 3D knowledge of the scene (eg., [27, 11, 1]), color (eg., [28, 12]) and multiple images (eg., [20, 31]).
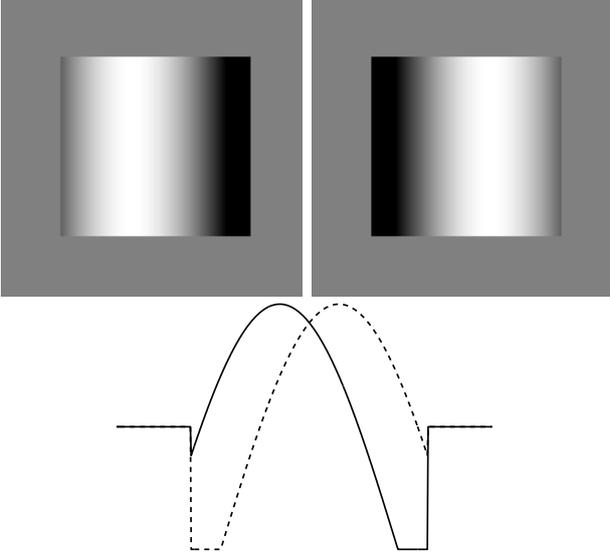
Figure 1. A cylinder illuminated from the left (top left) and from the right (top right). Cross-sections below of the intensities of the two images, after normalization (left, solid line; right, dashed line).

3 Normalized Correlation

We first derive the relationship between normalized correlation and the direction of the gradient.

Proposition: For a small window of an image in which intensities can be approximated as linear, normalized correlation computes the cosine of the difference in the direction of the gradient.

Proof: As a first step in normalized correlation we subtract the mean from each window. Assume I_1 and I_2 are zero mean windows that satisfy the above conditions. Then without loss of generality we can assume that ax is the gradient in I_1, and $bx + cy$ is the gradient in I_2, and θ corresponds to the angle between them.

$$\cos(\theta) = \frac{ab}{a\sqrt{b^2 + c^2}} = \frac{b}{\sqrt{b^2 + c^2}} \qquad (1)$$

The correlation between I_1 and I_2 is defined as[1]:

$$\int_{-1}^{1} \int_{-1}^{1} (abx^2 + acxy) dx dy = \frac{4}{3} ab$$

The normalization factor in I_1 is

$$\int_{-1}^{1} \int_{-1}^{1} a^2 x^2 dx dy = \frac{4}{3} a^2$$

---

[1]The size of the integration interval is chosen arbitrarily, because its magnitude will be cancelled by normalization.

The normalization factor in I_2 is

$$\int_{-1}^{1} \int_{-1}^{1} (b^2 x^2 + c^2 y^2 + 2bcxy) dx dy = \frac{4}{3}(b^2 + c^2)$$

Then the normalized correlation is

$$\frac{4ab}{3} \sqrt{\frac{3}{4a^2}} \sqrt{\frac{3}{4(b^2 + c^2)}} = \frac{b}{\sqrt{b^2 + c^2}}$$

which is the same as eq.1 $\qquad \square$

This demonstrates that with small windows, normalized correlation compares images by summing the cosine of the difference in the direction of the gradient at corresponding image points. This is similar to DIRP since like the square of the difference in angle, the cosine function is also monotonic in this difference, and changes more rapidly as the difference increases (up to a difference of $\frac{\pi}{2}$). We note that normalized correlation is not invariant to the polarity of the gradient direction. However, it is simple to change this, by taking the absolute value of the normalized correlation computed for each window. In this case, we compute the cosine of the difference in angle, allowing this difference to range from 0 to $\frac{\pi}{2}$. Normalized correlation is, of course, an increasing function of the match quality while DIR is a decreasing function. Therefore, in our experiments we use a simple modification: 1 - (absolute value of the normalized correlation), as a distance function. We will denote this comparison method, 'NC'.

Note that with DIR the direction of the gradient is computed at a scale determined by some level of prior smoothing. Similarly, we can control the scale at which normalized correlation compares the direction of the gradient by prior smoothing of the images.

4 Jets of Derivatives of Gaussians

Many authors have remarked that jets of multi-scale, oriented filters provide some insensitivity to lighting variation. A typical explanation for this is given by Wiskott et al.[30]: "Since they are DC-free, they provide robustness against varying brightness in the image. Robustness against varying contrast can be obtained by normalizing the jets." Similar sentiments are echoed by many authors (eg., [23, 8, 16]).

We now analyze a simple version of these approaches using a jet of oriented difference of Gaussian filters at one scale. We show that comparison of the output of these filters effectively computes the cosine of the difference in the direction of gradient of the two images. Or, with a slight variation, it computes a similar quantity that is invariant to polarity. We call this latter comparison method 'DOG'.

Let $D_\theta(\vec{x})$ denote the result at position $\vec{x}$ of convolving an image, I, with a difference of Gaussian filter oriented in the direction θ, where θ indicates the angle rel-

ative to the x axis. Then $D_\theta(\vec{x})$ encodes the magnitude of the gradient of $G \star I$ in the direction θ. A jet is a vector of the output of these filters at $\vec{x}$: $D(\vec{x}) = (D_{\theta_1}(\vec{x}), D_{\theta_2}(\vec{x}), ... D_{\theta_k}(\vec{x}))$. A typical number of oriented filters is eight, eg., $k = 8$. We will also consider a common variation in which the absolute value of the jet is used: $D_a = (|D_{\theta_1}(\vec{x})|, |D_{\theta_2}(\vec{x})|, ... |D_{\theta_k}(\vec{x})|)$. The resulting jets are normalized prior to comparison. The simplest comparison is correlation, in which case we compute $\frac{D_1 \cdot D_2}{\|D_1\|\|D_2\|}$.

Suppose that the direction of the gradient at $\vec{x}$ is along the x axis (this will be without loss of generality when our analysis moves to the continuous domain), and the magnitude of the gradient is M_x. Then: $D_\theta(\vec{x}) = M_x \cos(\theta)$ and we have:

$$D(\vec{x}) = M_x(cos(0), cos(\frac{2\pi}{k}), ... cos(\frac{2(k-1)\pi}{k}))$$

That is, $D(\vec{x})$ is a vector that discretely samples the cosine function, scaled by M_x. If we compute difference of Gaussians at a point, $\vec{y}$, in another image, at which the direction of the gradient is α and its magnitude is M_y we have:

$$D(\vec{y}) = M_y(cos(-\alpha), cos(\frac{2\pi}{k}-\alpha), ... cos(\frac{2(k-1)\pi}{k}-\alpha))$$

DOG compares jets by computing: $\frac{D(\vec{x}) \cdot D(\vec{y})}{\|D(\vec{x})\|\|D(\vec{y})\|}$. To analyze this, it is useful to approximate the discretely sampled cosine with a continuous function. So we take:

$$\|D(\vec{x})\| \approx M_x \sqrt{\int_0^{2\pi} (\cos(\theta))^2 d\theta} = M_x \sqrt{\pi}$$

Similarly, $\|D(\vec{y})\| \approx M_y \sqrt{\pi}$ Therefore:

$$\frac{D(\vec{x}) \cdot D(\vec{y})}{\|D(\vec{x})\|\|D(\vec{y})\|} \approx \frac{1}{\pi} \int_0^{2\pi} cos(\theta) cos(\theta - \alpha) d\theta = cos(\alpha)$$

This is exactly the same comparison measure as normalized correlation.

Next, we consider what happens when we take the absolute values of filter outputs. Since a difference of Gaussian oriented in the direction θ produces a result with the same magnitude as one oriented towards $\theta + \pi$ we only apply filters in a range of directions from 0 to π. We obtain:
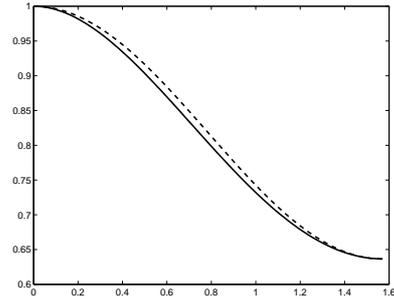
$$\frac{D_a(\vec{x}) \cdot D_a(\vec{y})}{\|D_a(\vec{x})\|\|D_a(\vec{y})\|} \approx \frac{2}{\pi} \int_0^\pi |cos(\theta)||cos(\theta - \alpha)| d\theta$$

We can assume, without loss of generality that $0 \le \alpha \le \frac{\pi}{2}$. We obtain:

$$\frac{2}{\pi} \int_0^\pi |cos(\theta)||cos(\theta - \alpha)| d\theta$$
$$= \frac{2}{\pi}(\frac{\pi}{2} \cos(\alpha) + \frac{\sin(\alpha)}{2} - \alpha \cos(\alpha)$$
$$+ \frac{\cos(\alpha)\sin(2\alpha)}{2} - \frac{\sin \alpha \cos(2\alpha)}{2})$$
$$= \frac{2}{\pi}[(\frac{\pi}{2} - \alpha) \cos(\alpha) + \sin(\alpha)]$$

Function→ Coefficient↓	$\frac{2}{\pi}[(\frac{\pi}{2} - \alpha)\cos(\alpha) + \sin(\alpha)]$	$\frac{2}{11}\cos(2\alpha) + \frac{9}{11}$
1	1.146	1.157
2	-.179	-.180
3	.007	0
4	.020	.022
5	-.002	0
6	-.0003	-.0007
7	0	0

Table 1. Coefficients for representing each function by Legendre polynomials.

Figure 2. $\frac{2}{\pi}[(\frac{\pi}{2} - \alpha)\cos(\alpha) + \sin(\alpha)]$ **(solid line),** $\frac{2}{11}\cos(2\alpha) + \frac{9}{11}$ **(dashed line).**

This first equality is obtained by breaking the integral into intervals of constant sign. The second equality follows from trigonometric identities.

In fact, it turns out that $\frac{2}{\pi}[(\frac{\pi}{2} - \alpha) \cos(\alpha) + \sin(\alpha)] \approx \frac{2}{11} \cos(2\alpha) + \frac{9}{11}$. To show this, we can expand the two functions with Legendre polynomials. The coefficients of the two functions in this orthonormal basis are shown in Table 1. They are almost identical (Figure 2).

Therefore, DOG essentially compares image gradients by taking the cosine of twice the difference in angle. This comparison is insensitive to the polarity of this difference, since it is periodic with a frequency of π. Within this range it is monotonic with the difference in angle, and qualitatively identical to the comparison method in DIR.

Note that these functions are a continuous approximation to the discrete functions actually computed. However, in experiments we have verified that the continuous and discrete versions perform identically, with $k = 8$.

By using a set of eight filters we obtain a highly redundant feature set which, by definition, is not needed to capture the information in the image. However, if we compare these jets by taking inner products, we can see that this is an excellent, discrete approximation to a comparison of a continuous function of the gradient. A non-redundant set of

4
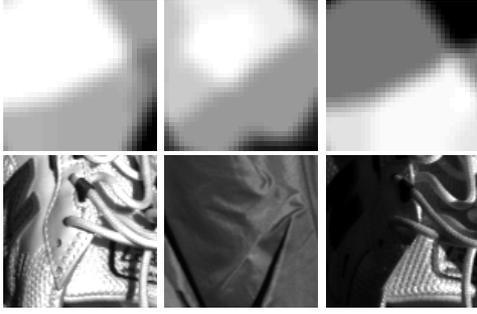
Figure 3. Synthetic objects (top row) and real objects (bottom row). We show two objects with the same lighting (left and center) and a query image (right).

Figure 4. Synthetic objects (top row) and real objects (bottom row). Results (top – real, bottom – synthetic).

two oriented filters would provide a poor approximation to this continuous function.

5 Experiments

While we have analytically shown strong connections between comparison methods using normalized correlation, difference of Gaussian jets, and the direction of gradient there are some subtle differences between the methods. For example, we have analyzed DOG with a continuous approximation to a discrete process. And, while DOG compares the difference in the direction of gradient by taking a cosine, DIR computes the sum of squared differences. Our experiments show that these subtle differences indeed have very little effect on the outcome.

We experiment on both synthetic and real images (Figure 3). Each experiment uses a reference image from each object, taken under identical lighting conditions. We then classify query images of these objects, taken under different lighting, according to the most similar reference image.

We generated synthetic convex objects by starting with a cube, and randomly generating planes that intersected the cube, but not a sphere contained in the cube. With each plane, we carved away a portion of the cube. Then we smoothed the resulting surface. 50 planes produced a polyhedral-like object. Reference images were taken with the light at an angle 45 degrees from vertical. Query images were taken with a light varying by between nine and ninety degrees from the original light, but never varying from frontal by more than 45 degrees. We carved the surface using planes with normals no more than 45 degrees from vertical, so that the resulting image contained no attached shadows.

For real objects, we used the Yale database [6], which contains 20 objects with abrupt changes in albedo and shape. The database consists of 63 images of each object
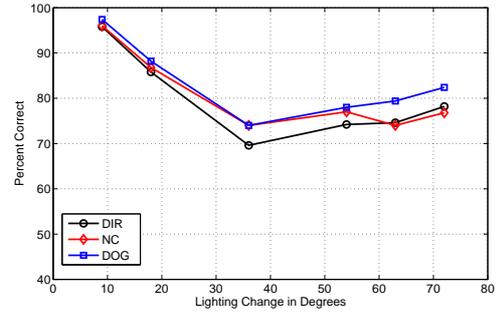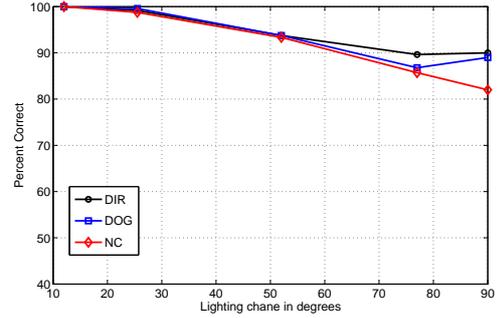
with lighting direction deviating up to 90 degrees from the frontal. A reference image of each object was chosen with frontal illumination. The rest of the images where used as queries.

Figure 4 shows the results of the experiments. The performance of all three methods is almost identical.

6 Conclusions

We have shown that normalized correlation and image comparison based on normalized jets of oriented filters are essentially gradient direction comparison methods. This makes them insensitive to lighting direction for scenes with 3D objects, a result that does not simply follow from their invariance to offset and gain. This accounts for their significant success in image comparison under changing lighting conditions. Understanding this equivalence makes it much simpler for researchers attempting to find the best lighting insensitive image comparison method.

Acknowledgements

References

[1] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **25**(2):218-233, (2003).

[2] M. Bichsel, *Strategies of Robust Object Recognition for the Automatic Identification of Human Faces*, PhD. thesis ETH Zurich, No. 9467, 1991.

[3] C. Bregler and J. Malik, "Learning Appearance Based Models: Mixtures of Second Moment Experts," *Advances in Neural Information Precessing Systems 9*, pp. 845-851, 1997.

[4] R. Brunelli, "Estimation of Pose and Illuminant Direction for Face Processing", MIT AI Memo 1499, 1994.

[5] T. Brunelli and T. Poggio, "Face recognition: Features versus templates", *IEEE Trans. PAMI*, **15**(10):1042-1062, 1993.

[6] H. Chen, P. Belhumeur, D. Jacobs. "In Search of Illumination Invariants", *IEEE Proc. Comp. Vis. and Pattern Recognition*, I:254-261, 2000.

[7] T. Cootes and C. Taylor. "On Representing Edge Structure for Model Matching," *IEEE Proc. Comp. Vis. and Pattern Recognition*, I:1114-1119, 2001.

[8] P. Felzenszwalb and D. Huttenlocher, "Pictorial Structures for Object Recognition," *International Journal of Computer Vision*, (forthcoming).

[9] A. Fitch, A. Kadyrov, W. Christmas, and J. Kittler J., "Orientation Correlation," *British Machine Vision Conference*, **1**:133-142, 2002.

[10] D. Forsyth and J. Ponce, *Computer Vision a Modern Approach*, Prentice-Hall, 2003.

[11] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: Generative Models for Recognition under variable pose and illumination". *IEEE Trans. on PAMI*.23(6):643–660, 2001.

[12] T. Gevers and A.W.M. Smeulders, "Color constant ratio gradients for image segmentation and similarity of textured objects," *CVPR*, pages I:1825, 2001.

[13] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. Anderson, "Overcomplete steerable pyramid filters and rotation invariance", *CVPR*:222-228, 1994.

[14] D. Hond and L. Spacek, "Distinctive descriptions for face processing", *8th British Machine Vision Conference*, Volume 1, pages 320-329, 1997.

[15] B. Jähne, *Digital Image Processing*, Springer-Verlag, Berlin, 1995.

[16] F. Jiao, S. Li, H. Shum, and D. Schuurmans, "Face Alignment Using Statistical Models and Wavelet Features", CVPR '03.

[17] J. Kittler, Y. Li, and J. Matas, "On matching scores for LDA-based face verification", *11th British Machine Vision Conference*, Volume 1, 42-51, 2000.

[18] J. Koenderink and A. van Doorn "Representation of local geometry in the visual system," *Biological Cybernetics*, vol. 55, pp. 367-375, 1987.

[19] M. Lades, J. C. Vortbrggen, J. Buhmann J. Lange, C. von der Malsburg, R. P. Wrtz, and W. Konen. "Distortion Invariant Object Recognition in the Dynamic Link Architecture", *IEEE Transactions on Computers*, 42:300–311, 1993.

[20] S. Narasimhan, V. Ramesh and S. Nayar, "A Class of Photometric Invariants: Separating Material from Shape and Illumination," *International Conference on Computer Vision*, pp. 1387-1394, 2003.

[21] P.J. Phillips and Y. Vardi, "Efficient Illumination Normalization of Facial Images", *Pattern Recognition Letters*, 17: 921-927, 1996.

[22] W. Pratt. *Digital Image Processing (First Edition)*. Wiley, 1978.

[23] R. Rao and D. Ballard. "An active vision architecture based on iconic representations," *Artificial Intelligence*, pp. 461–505, 1995.

[24] S. Ravela and C. Luo, "Appearance-based global similarity retrieval of images," in *Advances in Information Retrieval*, edited by W. Bruce Croft, Kluwer Academic Publisher 2000.

[25] C. Schmid and R. Mohr. "Local Grayvalue Invariants for Image Retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **19**(5), 530-535, 1997.

[26] A. Shashua. "On Photometric Issues in 3D Visual Recognition from a Single 2D Image," *Int. J. of Comp. Vision*, bf 21(1/2), pp. 99-122, 1997.

[27] A. Shashua and T. Riklin-Raviv, "The Quotient Image: Class-Based Re-Rendering And Recognition With Varying Illuminations", *IEEE Trans. on PAMI*, **23**(2), pp.129-139, 2001.

[28] D.A. Slater and G. Healey, "The illumination-invariant recognition of 3d objects using local color invariants," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18**, 2:206-210, 1996.

[29] L. Wiskott, *Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis*, Ph.D. thesis, 1995.

[30] L. Wiskott, J. Fellous, N. Kruger, C. van der Malsburg, "Face Recognition by Elastic Graph Matching," *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, eds. L. Jain et al., CRC Press, Chapter 11, pp. 355-396, 1999.

[31] L. Wolff and J. Fan, "Segmentation of surface curvature with a photometric invariant," *JOSA A*, **11**(11):30903100, 1994.