Image Space I^3 and Eigen Curvature for Illumination Insensitive Face Detection

Christian Bauckhage and John K. Tsotsos

Centre for Vision Research, York University, Toronto, ON, M3J 1P3
http://cs.yorku.ca/LAAV

Abstract. Generally, the performance of present day computer vision systems is still very much affected by varying brightness and light source conditions. Recently, Koenderink suggested that this weakness is due to methodical flaws in low level image processing. As a remedy, he develops a new theory of image modeling. This paper reports on applying his ideas to the problem of illumination insensitive face detection. Experimental results will underline that even a simple and conventional method like principal component analysis can accomplish robust and reliable face detection in the presence of illumination variation if applied to curvature features computed in Koenderink's image space.

1 Motivation and Related Work

In a recent paper on image processing methodology [1], Koenderink fiercely criticized the common practice to understand digital greyscale images as entities embedded in $\mathbb{R}^3$. He observes that if an image was indeed a set of points $(x_i, y_i, z_i)_{i=1...M} \in \mathbb{R}^3$ where the intensity values $z_i = f(x_i, y_i)$ define a surface above the X, Y plane, the geometry of $\mathbb{R}^3$ would allow to rotate this surface about an arbitrary axis. However, such a rotation might cause intensity values to lie in the image coordinate plane and image coordinates to be parallel to the intensity direction. Koenderink argues that a structure that allows for operations leading to physically senseless configurations is not the most reasonable choice for image modeling. As a more appropriate approach to mathematical image modeling he proposes a structure which he calls *image space* $\mathbb{I}^3$. The basic idea is to define $\mathbb{I}^3$ as a fiber bundle that locally looks like $\mathbb{P}^2 \times \mathbb{L}$ where the base manifold $\mathbb{P}^2$ corresponds to the picture plane and the fibers $\mathbb{L}$ are logarithmic scales of the intensity. An analysis of the (differential) geometry of this image space reveals that images in $\mathbb{I}^3$ are (by construction) invariant under different brightness transformations.

In this contribution, we explore the merits this model offers for computer vision. The application domain for our investigation will be illumination insensitive face detection.

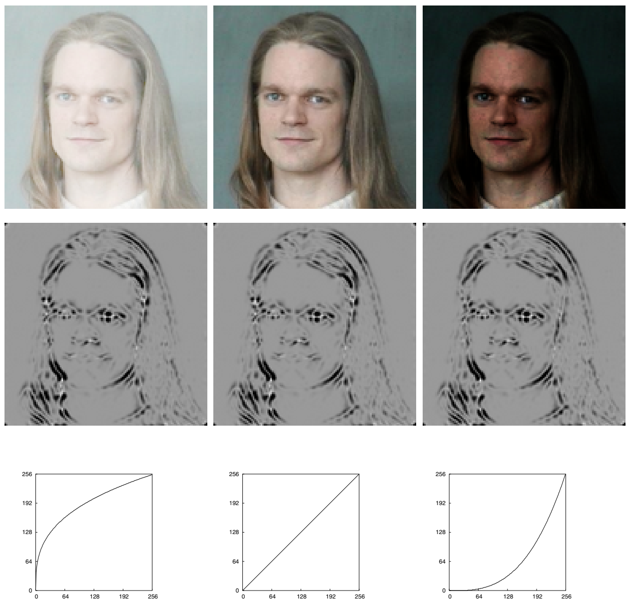
Face detection and recognition are arguably among the most popular topics in computer vision and respective publications are almost innumerable. In fact, the field is so active, it already produced its meta literature (cf. e.g. [2,3,4]). A complete survey of face detection techniques therfore is far beyond the prospects of this report but we shall single out a few contributions which are relevant for our discussion.

Since they were first considered by Sirovich and Kirby [5] and popularized by Turk and Pentland [6], Principal Component Analysis (PCA) based approaches have become

M. Kamel and A. Campilho (Eds.): ICIAR 2005, LNCS 3656, pp. 456–463, 2005.

a widespread tool in face detection. Although there are other subspace techniques like Linear Discriminant Analysis, Independent Component Analysis, or kernelized PCA, simple PCA is still among the most reliable methods [7,8]. However, its performance is known to depend on light source conditions. Recent contributions aiming at illumination invariance hence measure the *gradient similarity* statistics [9] or combine *edge phase congruency* information with local intensity normalization [10]. Others render *eigen-harmonics* to recover a standard illumination [11] or use *eigen light-fields* [12].

Three general trends become apparent from this rough survey: i) Gradient information (i.e. information from the realm of differential geometry) is considered to provide an avenue to illumination invariance. ii) PCA based methods prove to be persistent and are now being applied to more sophisticated data than the mere pixel values of old. iii) There are attempts to embed the abstract concept of *face space* into richer mathematical structures than the usual vector spaces over the field of real numbers.



(a) $\gamma = 1/3$ (b) $\gamma = 1$ (c) $\gamma = 3$

Fig. 1. Gamma transformations of an image and level sets of the corresponding Gaussian curvature K in $\mathbb{I}^3$

In the following, we will bring together all these trends. First, we will survey differential geometry in $\mathbb{I}^3$ and investigate the features it provides for illumination invariant face detection. Then, we will discuss how these feature might be put forth into a PCA based framework. We will address the use of real and complex valued feature vectors and then shall present experimental results. A conclusion will close this contribution.

2 Curvature Features for Face Detection

As a first step towards a more proper mathematical model for image processing, Koenderink examines the most suitable scale for the intensity dimension. Given that the photon count on a CCD chip is Poisson distributed and assuming observations based on different time scales, he shows that a time independent estimation of the Poisson parameter λ leads to a uniform distribution on the log-intensity scale. As a consequence, he proposes to use $Z(x, y) = \log(z(x, y)/z_0)$ for the intensity dimension where z_0 is an arbitrary unit of intensity. Points in $\mathbb{I}^3$ are thus specified by coordinates $\{x, y, Z\}$.

Since his primary concern is a space where the intensity domain and the image plane cannot interfere, Koenderink lists geometric constraints for $\mathbb{I}^3$. The resulting group of possible transformations is the group of direct isotropic similarities. Since these do not affect relations among a set of parallel 3D lines, they do not affect the curvature of surfaces in $\mathbb{I}^3$. Moreover, as the geometry of $\mathbb{I}^3$ comes along with a degenerate metric, Gaussian curvature and Mean curvature of surfaces are given by notably simple expressions. In contrast to the lengthy formulas known from Euclidean geometry they simply correspond to

$$K(x, y) = Z_{xx}Z_{yy} - Z_{xy}^2 \quad \text{and} \quad H(x, y) = \frac{Z_{xx} + Z_{yy}}{2}. \tag{1}$$

where $Z_{xx} = \partial^2 Z/\partial x^2$, $Z_{yy} = \partial^2 Z/\partial y^2$, and $Z_{xy} = \partial^2 Z/\partial x \partial y$.

Using the example of Gaussian curvature K, Fig. 1 underlines that image intensity transformations barely affect surfaces in $\mathbb{I}^3$. In the figure's top row, we see an image whose intensity was subjected to the gamma-transformations that are indicated in the bottom row of the figure. The middle row shows a level set representation of the corresponding Gaussian curvature where bright spots indicate points of high curvature. Obviously, K remains nearly constant across the different transformations.

Using curvature features derived in the fiber bundle $\mathbb{I}^3$ to tackle illumination invariant face detection is thus a tempting idea. In fact, curvature has been used in face recognition before. In [13], a system is presented that combines curvature maps with results from PCA based eye detection in order to improve facial part detection. And [14] reports on using local principal curvature to register 2D face images with 3D models. However, given the ease and success of PCA in face detection and recognition, it is surprising that there are yet no contributions that apply curvature features instead of intensity information to compute eigenfaces.

Extending PCA to curvature features is of course straightforward. Given an input image $I(x, y)$, we can compute the curvature maps $K(x, y)$ and $H(x, y)$. Like in the intensity approach, patches of m pixels can be represented as objects in $\mathbb{R}^m$, i.e. as vectors $\boldsymbol{k}$ and $\boldsymbol{h}$, respectively. Therefore, given a set of n patches representing faces stored in a $m \times n$ data matrix $\boldsymbol{A}$, *eigen curvature faces* can be computed.

In contrast to the intensity based approach, however, curvature based face detection offers a choice of two feature vectors for every pixel. The question is thus if and how to combine Mean and Gaussian curvature? The intuitive approach is to consider the direct sum $\boldsymbol{k} \oplus \boldsymbol{h} \in \mathbb{R}^{2m}$. However, dealing with PCA and a set of n examples where $n \ll m$, we note that even if the data matrix $\boldsymbol{A}$ is a $2m \times n$ matrix, the matrix $\boldsymbol{A}^T\boldsymbol{A}$ used to compute eigenvectors will remain $n \times n$. Hence, doubling the dimension of the data vectors will still result in a maximum of n eigenvectors. Moreover, with $\boldsymbol{A}_K = [\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n]$ and $\boldsymbol{A}_H = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_n]$ denoting the original data matrices and $\boldsymbol{A} = \boldsymbol{A}_{K\oplus H}$ denoting the one resulting from the embedding in $\mathbb{R}^{2m}$, we will have

$$\boldsymbol{A}^T\boldsymbol{A} = \left[\boldsymbol{A}_K^T \boldsymbol{A}_H^T\right] \begin{bmatrix} \boldsymbol{A}_K \\ \boldsymbol{A}_H \end{bmatrix} = \boldsymbol{A}_K^T\boldsymbol{A}_K + \boldsymbol{A}_H^T\boldsymbol{A}_H \tag{2}$$

I.e. up to this point the embedding in $\mathbb{R}^{2m}$ yields a mere additive connection of the information contained in the vectors $\boldsymbol{k}_i$ and $\boldsymbol{h}_i$ where $i \in \{1, \ldots, n\}$. Of course, the characteristic polynomial in the next step of the computation of eigenvectors introduces products and nonlinearity. But nevertheless, it seems worthwhile to consider a more complex entanglement of the two feature spaces.

A more *complex* entanglement is indeed possible if we consider the embedding $\boldsymbol{c} = \boldsymbol{k} + i\boldsymbol{h} \in \mathbb{C}^m$. Since the new data matrix $\boldsymbol{C}$ is composed of complex vectors, the standard approach to computing its eigenvectors requires multiplication with the conjugate transpose $\boldsymbol{C}^\dagger$. And as

$$\left(\boldsymbol{C}^\dagger \boldsymbol{C}\right)_{ij} = \boldsymbol{c}_i^* \boldsymbol{c}_j = \sum_l k_{i_l} k_{j_l} + h_{i_l} h_{j_l} + i(k_{i_l} h_{j_l} - h_{i_l} k_{j_l}) \tag{3}$$

we see that for an embedding in $\mathbb{C}^m$ mixed terms already appear before characteristic polynomials are computed. Note that since $\boldsymbol{C}^\dagger \boldsymbol{C}$ is a Hermitian matrix, its eigenvalues will be real but its eigenvectors will be complex.

3 Experiments

The utility of the above features and feature combinations for illumination insensitive face detection was evaluated by means of different experiments. This section summarizes our findings concerning the following six feature spaces

- I^m, the common m dimensional intensity space
- Q^m, the corresponding intensity space after histogram equalization (often done in the literature to compensate for illumination variation)
- K^m, the Gaussian curvature space resulting from computations in $\mathbb{I}^3$
- H^m, the corresponding Mean curvature space
- $K^m \bigoplus H^m$, the $2m$ dimensional space combining Gaussian- and Mean curvature
- $K^m + iH^m$, the m dimensional complex space combining the curvature features

All experiments reported below were based on the same *small* training set of $n = 34$ images (note that recent approaches based on intensity cues require several thousand training images to cope with lighting variations [3]). The gallery was retrieved from the Internet [15] and shows 27 male and 7 female faces, eleven people are wearing glasses, one person is bearded, 33 subjects are of Caucasian ancestry and one is Asian. All images show frontal views of faces recorded under ambient daylight.

In order to obtain accurate curvature images $K(x, y)$ and $H(x, y)$ all necessary derivations were computed using precise recursive Gaussian filtering [16]. Training data resulted from cropping 80×80 windows (i.e. $m = 6400$) centered at the nose. After vectorization, the data were normalized to unit length and zero mean. Eigenfaces were obtained from a singular value decomposition of the (complex) data matrices. For each of the considered features, the projection of image patches into the corresponding face space was done using the eigenvectors corresponding to the eight largest eigenvalues.

The quality of a face detector was assessed by means of the distance between the pixel where it yielded the highest response and the pixel where it should have occurred.

Fig. 2. Face detection examples under some of the artificial brightness distortions from our first series of experiments. Combined curvature features are most reliable in the case shown here.

To provide ground truth, the locations of the noses were labeled manually for all images in our test sets. Dividing the measured distance by the maximum possible distance to the nose yields the *deviation* $\delta \in [0, 1]$ to which we will refer for the rest of this discussion.

3.1 Semi Synthetic Data

The basis of our first test was formed by a set of 17 face images recorded under the same conditions as the training set. These images show 10 male and 7 female subjects; 6 of them are wearing glasses, 3 are bearded, one is of Asian ancestry the rest are Caucasian. Each test image was subject to 17 different intensity transformations, yielding a set of 289 face images. Examples of some of the distortions can be seen in Fig. 2.

Table 1(a)shows that (combined) curvature features perform better than the intensity based ones. The table summarizes the statistics gathered from the whole test set of 289 images. The first row lists the mean deviations $\mu(\delta)$ resulting from the tested features. The direct sum of curvatures has a one percent lead on the complex combination but face detection using only Mean curvature also performs well. Pure intensity based eigenfaces cannot compete given the kind of distortions in our test set. Eigenfaces from

Table 1. Experimental results: mean (μ) and median (m) deviation in face detection (less is better)

	$\boldsymbol{i}$	$\boldsymbol{q}$	$\boldsymbol{k}$	$\boldsymbol{h}$	$\boldsymbol{k} \oplus \boldsymbol{h}$	$\boldsymbol{k} + i\boldsymbol{h}$
$\mu(\delta)$	0.51	0.25	0.29	0.20	**0.17**	0.18
$\sigma^2(\delta)$	0.14	0.08	0.06	0.04	0.05	0.04
$m(\delta)$	0.52	0.06	0.33	**0.04**	**0.04**	0.05

(a) semi synthetic data

	$\boldsymbol{i}$	$\boldsymbol{q}$	$\boldsymbol{k}$	$\boldsymbol{h}$	$\boldsymbol{k} \oplus \boldsymbol{h}$	$\boldsymbol{k} + i\boldsymbol{h}$
$\mu(\delta)$	0.72	0.29	0.16	0.14	**0.09**	0.13
$\sigma^2(\delta)$	0.07	0.15	0.03	0.03	0.02	0.02
$m(\delta)$	0.83	0.03	0.05	0.03	**0.02**	0.08

(b) real data

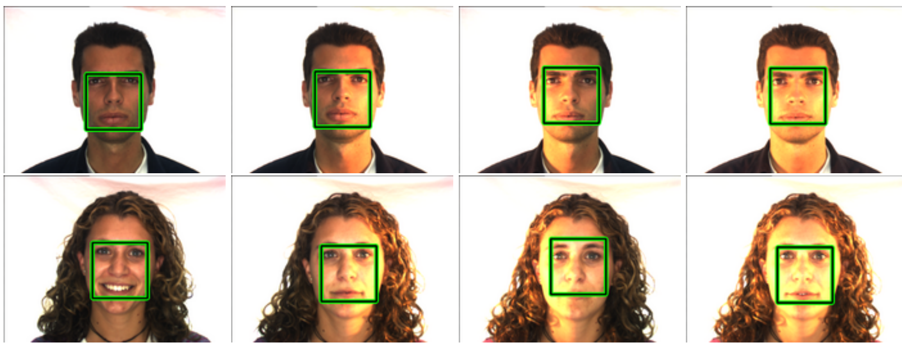


Fig. 3. Exemplary performance of the $\boldsymbol{k} \oplus \boldsymbol{h}$ feature under different lighting conditions

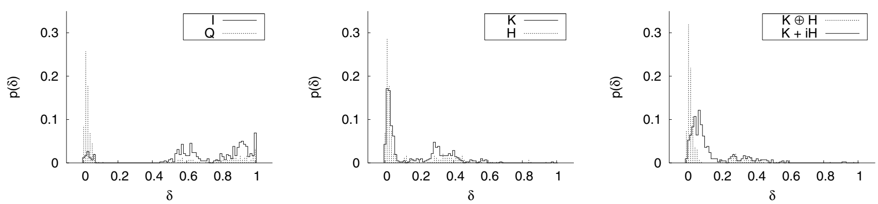


Fig. 4. Distributions of the deviation δ found in experiments with the AR face database

equalized intensity patches yield much better results but are outperformed by three out of four curvature based approaches. As Fig. 2 indicates, the precision in face detection may vary between the considered distortions. Since this behavior is also to be observed across different subjects, the mean might not be the most significant measure to characterize the overall performance of the tested detectors. In the lower most row, Tab. 1(a) thus lists the median deviations $m(\delta)$; they endorse our findings.

3.2 Real Data

In a second series of tests, we experimented with the AR face database [8] whose images show frontal view faces with different facial expressions under different lighting conditions (see Fig. 3). Again, (after scaling the images by 0.45 so that their size com-

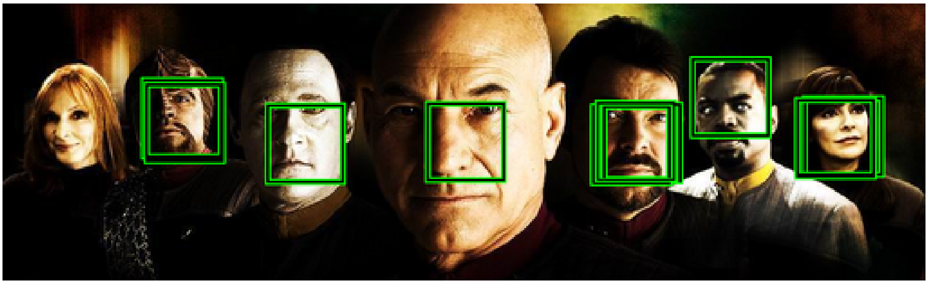


Fig. 5. The largest 10 local maxima in the $\boldsymbol{k} \oplus \boldsymbol{h}$ face response map for an image of extreme lighting conditions. Considering the size of the middle face, the corresponding response admittedly seems arbitrary. The other detection results, however, confirm our systematic experiments.

plied with our training set), the nose was labeled manually and the deviation δ was measured to assess our approach.

Table 1(b) summarizes our findings. Again, face detection by PCA of $\boldsymbol{k} \oplus \boldsymbol{h}$ vectors performs best; its mean deviation from the expected response location and the corresponding variance but also the median $m(\delta)$ are the smallest of all tested features. Figure 3 displays examples that illustrate this performance.

The plots of the distributions of the different deviations in Fig. 4 enable a more detailed analysis of the results. They document the complete breakdown of intensity based face detection and show sharp peaks in the vicinity of $\delta = 0$ for the histogram equalized intensity $\boldsymbol{q}$, the Mean curvature $\boldsymbol{h}$ as well as for the $\boldsymbol{k} \oplus \boldsymbol{h}$ feature. However, while the equalized intensity based approach also yields a notable number of responses for $\delta \in [0.5, 1.0]$, this is not quite as much the case for Mean curvature based face detection and even less so for the direct sum combination of Gaussian and Mean curvature. This observation is also reflected in the true positive and false positive classification rates we measured. Considering a 30 pixels radius around the manually labeled optimal response location acceptable, PCA on vectors $\boldsymbol{q}$ yields a percentage of 67% true positives and 33% false positives on the AR face database. For PCA on $\boldsymbol{h}$ and $\boldsymbol{k} \oplus \boldsymbol{h}$ features, we obtained 75% vs. 25% and 80% vs. 20%, respectively.

3.3 Discussion

Given the fact that all our results were obtained using an approach as simple as PCA trained on a small set of images of faces under ambient daylight, the above findings are remarkable. Curvature maps computed in $\mathbb{I}^3$ definitely provide a promising avenue to illumination insensitive face detection. Not only are these features able to cope with severe artificial illumination distortions but also perform well when applied to images taken under considerably different environmental lighting. Moreover, informal tests on images like the one shown in Fig. 5 revealed that even rather extreme conditions can be dealt with.

Concerning the different types of curvatures features or feature combinations that have been tested, the result is clear. Combining Gaussian and Mean curvature feature vectors in a direct sum yielded the best results in all our experiments. Appearance based face detection using the $\boldsymbol{k} \oplus \boldsymbol{h}$ feature produced useful and reliable results that hardly deviated from manually labeled locations.

4 Conclusion

In this paper, we tested Koenderink's proposal for a new image modeling paradigm. His idea to use a different geometric model than the physically incorrect Euclidean vector space $\mathbb{R}^3$ results in a representation of images that is invariant against several brightness transformations. Moreover, due to its degenerate metric this space comes along with remarkably simple expressions for features like Gaussian or Mean curvature.

Applying this idea to the problem of illumination insensitive face detection shows that curvature features computed in $\mathbb{I}^3$ indeed provide an auspicious but simple solution. Even an off the shelve appearance based method as simple as PCA does not require complex preprocessing or sophisticated tweaking to produce robust and reliable results under a wide range of lighting variations.

References

1. Koenderink, J.J., van Doorn, A.J.: Image Processing Done Right. In: Proc. ECCV. Volume 2350 of LNCS., Springer (2002) 158–172
2. Hjelmås, E., Low, B.K.: Face Detection: A Survey. Comput. Vis. Image Underst. **83** (2001) 236–274
3. Yang, M.H., Kriegman, D., Ahuja, N.: Detecting Faces in Images: A Survey. IEEE Trans. Pattern Anal. Machine Intelli. **24** (2002) 34–58
4. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Comput. Surv. **35** (2003) 399–458
5. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. J. Opt. Soc. Am., A **4** (1987) 519–524
6. Turk, M.A., Pentland, A.P.: Eigenfaces for Recognition. J. Cogn. Neurosci. **3** (1991) 71–86
7. Li, J., Zhou, S., Shekhar, C.: A Comparison of Subspace Analysis for Face Recognition. In: Proc. ICASSP. Volume 3. (2003) 121–124
8. Martínez, A.M., Kak, A.: PCA versus LDA. IEEE Trans. Pattern Anal. and Machine Intelli. **23** (2001) 228–233
9. Chen, H.F., Belhumeur, P.N., Jacobs, D.W.: In Search of Illumination Invariants. In: Proc. CVPR. Volume 1. (2000) 254–261
10. Huang, Y., Lin, S., Li, S.Z., Lu, H., Shum, H.Y.: Face Alignment Under Variable Illumination. In: Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition. (2004) 85–90
11. Qing, L., Gao, S.S.W.: Eigen-Harmonics Faces: Face Recognition under Generic Lighting. In: Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition. (2004) 296–301
12. Gross, R., Baker, S., Matthews, I., Kanade, T.: Face Recognition Across Pose and Illumination. In Jain, A., Li, S., eds.: Handbook of Face Recognition. Springer (2004)
13. Gargesha, M., Panchanathan, S.: A hybrid technique for facial feature point detection. In: Proc. IEEE Southwest Symp. on Image Analysis and Interpretation. (1998) 134–139
14. Tanaka, H.T., Ikeda, M.: Curvature-based face surface recognition using spherical correlation–principal directions for curved object recognition. In: Proc. ICPR. Volume III. (1996) 638–642
15. http://www.techfak.uni-bielefeld.de/ags/ai/members/members.html (retrieved spring 2005)
16. Deriche, R.: Recursively Implementing the Gaussian and Its Derivatives. In: Proc. ICIP. (1992) 263–267