

# Karhunen-Loève Transformation for Optimal Color Feature Generation

*Mehmet Celenk and Ivan Chang*  
*School of Electrical Engineering & Computer Science*  
*Ohio University, Stocker Center*  
*Athens, Ohio, USA*

## Abstract

In many applications, it is desirable to reduce the number of color and texture features to a sufficient minimum. Computational complexity and storage cost are some of the obvious reasons for this requirement. A related reason is that although two features may carry good discriminatory information when treated separately, there is little gain if they are combined together in a feature vector, because of a possible high mutual correlation between them. Thus, complexity increases without much gain. The major task of this paper is to address the problem of selecting the most important features for a given textured color image so as to maintain the optimal number of color and texture characteristics.

The basic approach explored in this paper is based on the discrete Karhunen-Loève Transform (KLT). The reason behind the selection of KLT-based texture and color features is that an appropriately chosen transform can exploit and remove information redundancies, which usually exist in a wide range of color and texture scenes obtained by measuring devices. If the transform chosen is subjected to appropriate constraints, maximum information can be preserved in the output with reduced dimension, and without nuances that do not exist in the original input samples. This is crucial if any subsequent feature-analyzing function is to be able to produce meaningful results. Another reason is that KLT seems closely related to early visual pathway of primates.

## 1. Introduction

### 1.1. Karhunen-Loève Transform

*Karhunen-Loève Transform*<sup>\*</sup> is, in a sense, a linear filter whose response is optimized with respect to some performance measure in measurement space. Consider a set of physical measurements  $\{\xi_i\}$ , where  $\xi_i$  is an instance of

<sup>\*</sup>Note that this technique has been given different names: *empirical orthogonal functions*, *principal component analysis*, *singular value decomposition*, *least squares methods*, *factor analysis*, and *matched filtering*.

measurement vector. Quantities, which do not depend on the choice of origin in measurement space, remain invariant under translation. For convenience, we choose the reference such that  $\langle \xi \rangle = 0$ , where  $\langle \dots \rangle$  denotes the ensemble average over the set  $\{\xi_i\}$ . The variance of data set  $\{\xi_i\}$  along a unit vector  $\mathbf{x}$  is given by

$$\sigma_{\mathbf{x}}^2 = \langle (\xi^T \mathbf{x})^2 \rangle \quad (1)$$

Equation (1) can be written as

$$\sigma_{\mathbf{x}}^2 = \mathbf{x}^T \mathbf{C} \mathbf{x} \quad (2)$$

where  $\mathbf{C} = \langle \xi \xi^T \rangle$ .  $\mathbf{C}$  is a symmetric and positive semi-definite matrix<sup>†</sup>. All eigenvalues of  $\mathbf{C}$  are real and non-negative, and its eigenvectors can be taken as orthogonal. Let  $\{\mathbf{e}^\alpha\}$  be the orthonormal basis formed by the eigenvectors of  $\mathbf{C}$  such that  $\mathbf{e}^\alpha$  corresponds to eigenvalue  $\lambda^\alpha$ . Suppose we take the eigenvalues to be in decreasing order

$$\lambda^1 \geq \lambda^2 \geq \dots \geq \lambda^N \quad (3)$$

where  $N$  denotes dimension of the raw data set. The variance  $\sigma_{\mathbf{x}}^2$  can be decomposed along the eigenvectors

$$\sigma_{\mathbf{x}}^2 = \sum_{\alpha} \lambda^\alpha x_\alpha^2 \quad (4)$$

where  $x_\alpha$  is the component of  $\mathbf{x}$  along  $\mathbf{e}^\alpha$ . By the chosen order in equation (3), we have

$$\sigma_{\mathbf{x}}^2 = \sum_{\alpha} \lambda^\alpha x_\alpha^2 \leq \lambda^1 \sum_{\alpha} x_\alpha^2 = \lambda^1 \quad (5)$$

for all unit vectors  $\mathbf{x}$ . The unconstrained maximum of  $\sigma_{\mathbf{x}}^2$  on the  $N$ -dimensional unit sphere is equal to the maximum eigenvalue ( $\lambda^1$ ) and occurs when we choose  $x_1 = \pm 1$  and 0 for all other components (*i.e.*,  $\mathbf{x} = \mathbf{e}^1$ ). If we constrain  $\mathbf{x}$  to lie in the subspace perpendicular to  $\mathbf{e}^1$ , then  $x_1 = 0$  and

$$\sigma_{\mathbf{x}}^2 = \sum_{\alpha > 1} \lambda^\alpha x_\alpha^2 \leq \lambda^2 \sum_{\alpha > 1} x_\alpha^2 = \lambda^2 \quad (6)$$

<sup>†</sup>Proof: For any vector  $\mathbf{x}$ ,  $\mathbf{x}^T \mathbf{C} \mathbf{x} = \mathbf{x}^T \langle \xi \xi^T \rangle \mathbf{x} = \langle \mathbf{x}^T \xi \xi^T \mathbf{x} \rangle = \langle (\xi^T \mathbf{x})^2 \rangle \geq 0$ .

The maximum of  $\sigma_{\mathbf{x}}^2$  is  $\lambda^2$  and occurs when  $x_2 = \pm 1$  and 0 for all other components (*i.e.*,  $\mathbf{x} = \mathbf{e}^2$ ). By following this induction process, we obtain a decomposition of the data space using linear filters<sup>‡</sup> that form an orthogonal basis set and that account for as much as possible the variance of the data set.

## 1.2. Early Visual Pathway and Karhunen-Loève Transform

Primate visual information processing occurs in multiple stages of increasing abstraction starting from low-level signal processing to various high-level cognitive behaviors. In this section, we present a simple description on the early stage of primate visual pathway and explain its close connection to KLT.

Visual image processing begins in a region called *retina* located at the back of the eye. The retina has an orderly layered anatomical arrangement with the foremost layer lined with photoreceptors sensitive to light. The human retina contains two types of photoreceptors - *rods* and *cones*. Unlike most neurons, rods and cones do not fire action potentials [3]. Instead, their responses are the graded changes in membrane potentials. Color vision is mediated by three types of cones each differentiated by a visual pigment that is more sensitive to a different part of the visible spectrum. It has been found experimentally that individual cones contain only one type of the three pigments. One pigment type is primarily sensitive to long wavelengths in the visible spectrum and makes a strong contribution to the perception of red. Another is selective for middle wavelengths and makes a strong contribution to the perception of green. The third is responsive to short wavelengths and makes a strong contribution to the perception of blue.

The middle layer contains interneurons which relay the photoreceptor outputs to the output neurons in the third layer. The output neurons converge upon the optic nerve which enters the brain for higher order processing stages. The neurons in the middle and third layer are composed of cells with a center-surround receptive field which forms an antagonistic pair. This antagonistic arrangement gives the cell the performance of responding strongly to contrasting signals (*i.e.*, zero-crossings).

In color vision, the antagonistic center-surround receptive field is responsible for the color-opponent neural channels: red-green, yellow-blue, and white-black. Similar channels are observed when we apply KLT to a few images from the *VisTex* Database [4], as indicated by the signs of the eigenvector components in Fig. 1-4<sup>§</sup>. In Fig. 1-4, the

original images from *VisTex* are shown in the top-left position, while the KLT-based orthogonal channels are shown in the top-right, bottom-left, and bottom-right positions in order of decreasing eigenvalues. The KLT-based channels in Fig. 1-4 have been normalized to 8-bit grayscale in order to display the spatial variation in each channel clearly, as suggested by [8]. Fig. 5-8 show how the KLT-based channels in Fig. 1-4 would have been like in the original RGB channels. The eigenvectors and eigenvalues of the KLT-based channels in Fig. 1-4 are computed as follows: (a) Fig. 1: [ 0.5090 0.5308 0.6776 ] (3328.2), [ -0.7666 -0.0784 0.6373 ] (399.6), [ 0.3914 -0.8438 0.3671 ] (15.7); (b) Fig. 2: [ 0.6196 0.5594 0.5506 ] (6401.5), [ -0.6878 0.0489 0.7243 ] (55.5), [ 0.3783 -0.8274 0.4151 ] (4.6); (c) Fig. 3: [ 0.7418 0.5835 0.3306 ] (4670.1), [ -0.4670 0.0956 0.8791 ] (98.0), [ 0.4814 -0.8065 0.3434 ] (4.6); (d) Fig. 4: [ 0.7753 0.5319 0.3405 ] (773.9), [ -0.5596 0.3287 0.7608 ] (76.5), [ 0.2928 -0.7804 0.5525 ] (7.1).

The eigenvectors corresponding to the largest eigenvalues have responses with the same sign. Their convolutions with the original images produce grayscale images which preserve most of the original information, as can readily be seen from Fig. 1-4. The rest of the eigenvectors exhibit some sort of antagonistic pair(s) (at least one) so that the early primate visual pathway may very well perform some transformation that resembles KLT.

Linsker [5] proposed a self-organization model in the primate visual pathway using a version of *Hebbian* learning in a *layered* network. Mathematically, he showed the connection between Hebbian learning in a layered architecture, KLT, and the maximum Shannon information. A summary of his work can be found in [6].

## 2. Spatial Texture Features

As discussed in the previous section, primate visual perception is a complex process that involves multiple stages. We have presented a simple biological description of early visual pathway and showed its close connection to KLT. The output from the retina will undergo further processing stages in specialized regions of the brain to compute more sophisticated features. The most ubiquitous of these are perhaps spatial and temporal textures. The ability to recognize these features underlies many important cognitive behaviors such as object recognition and motion detection. How do these abilities arise from the simple quasi-linear response of the retina and the subsequent non-linear transformations in the rest of the visual pathway? This is an area which the authors are still investigating and results will be reported in the future.

For our present purpose, we employ Haralick's method of co-occurrence [1] which bears certain psychophysical significance. In the co-occurrence method, the relative fre-

<sup>‡</sup>The eigenvector components correspond to impulse response coefficients of a typical  $N$ -th order moving average model whose output is simply the convolution of the response coefficients with the input data.

<sup>§</sup>The original images are all of size  $128 \times 128$  with 8 bits/R,G,B but have been reduced in size fourfold for illustration purposes.

quencies of gray-level pairs of pixels at certain relative displacements are computed and stored in a matrix of *co-occurrence*  $\mathbf{P}$ . For  $N_g$  gray-levels in the image, the co-occurrence matrix will be of size  $N_g \times N_g$ . If  $N_g$  is large compared to the image size, the number of pixel pairs contributing to each element  $p_{ij}$  in  $\mathbf{P}$  will be statistically insignificant as well as computationally expensive. On the other hand, if  $N_g$  is small compared to image size, most of the texture information will be averaged out. Ohanian and Dubes [2] reported that  $N_g = 8$  was an appropriate choice for images of size  $32 \times 32$ . As suggested by other researches [7], the combination of the nearest neighbor pairs at orientations  $0, \frac{\pi}{4}, \frac{\pi}{2},$  and  $\frac{3\pi}{4}$  are used in computing the texture features. Haralick [1] suggested 14 texture features based on the co-occurrence matrix. The four most widely used features are angular second moment (asm), contrast (con), correlation (cor), and entropy (ent) as given by the following definitions

$$asm = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{ij}^2 \quad (7)$$

$$con = \sum_{n=0}^{N_g-1} n^2 \sum_{|i-j|=n} p_{ij} \quad (8)$$

$$cor = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{ijp_{ij} - i_j p_i^{(x)} p_j^{(y)}}{\sigma_x \sigma_y} \quad (9)$$

$$ent = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{ij} \log p_{ij} \quad (10)$$

where  $\sigma_x, \sigma_y$  are the standard deviations corresponding to the distributions

$$p_i^{(x)} = \sum_{j=0}^{N_g-1} p_{ij} \quad (11)$$

$$p_j^{(y)} = \sum_{i=0}^{N_g-1} p_{ij} \quad (12)$$

### 3. Numerical Results

We label the four original images as shown in Fig. 1-4 as 001-004 respectively. The subscript  $g$  corresponds to the same image contaminated by 64% Gaussian white noise. Tables 1 through 5 summarize the numerical values of the mean, asm, con, cor, and ent features based on the normalized KLT-based channels of the largest eigenvalues of Fig. 1-4, respectively.

Table 1: Mean Color Feature

$\mu$	001	001 <sub>g</sub>	002	002 <sub>g</sub>	003	003 <sub>g</sub>	004	004 <sub>g</sub>
R	117	119	124	124	132	129	115	115
G	148	147	110	112	106	106	102	103
B	185	180	101	104	75	77	95	96

Table 2: Angular Second Moment Feature ( $\times 10^{-2}$ )

asm	001	001 <sub>g</sub>	002	002 <sub>g</sub>	003	003 <sub>g</sub>	004	004 <sub>g</sub>
1, 0	29.3	4.04	4.42	2.92	6.09	3.09	7.35	3.93
1, $\frac{\pi}{4}$	27.8	4.02	4.04	2.85	6.61	3.14	5.08	3.94
1, $\frac{\pi}{2}$	28.6	4.06	5.43	3.01	6.30	3.12	6.18	3.92
1, $\frac{3\pi}{4}$	27.5	4.00	4.24	2.84	4.91	2.98	5.01	3.93

Table 3: Contrast Feature ( $\times 10^{-2}$ )

con	001	001 <sub>g</sub>	002	002 <sub>g</sub>	003	003 <sub>g</sub>	004	004 <sub>g</sub>
1, 0	49.5	261	206	351	215	348	86.4	337
1, $\frac{\pi}{4}$	82.4	267	248	384	152	320	211	354
1, $\frac{\pi}{2}$	54.8	261	140	316	179	330	165	350
1, $\frac{3\pi}{4}$	82.3	271	244	384	363	424	217	354

Table 4: Correlation Feature ( $\times 10^{-2}$ )

cor	001	001 <sub>g</sub>	002	002 <sub>g</sub>	003	003 <sub>g</sub>	004	004 <sub>g</sub>
1, 0	89.4	36.1	62.5	38.6	61.5	34.1	80.8	13.2
1, $\frac{\pi}{4}$	82.3	34.7	54.8	32.9	72.7	39.2	53.1	8.7
1, $\frac{\pi}{2}$	88.2	36.0	74.5	44.7	69.6	37.3	63.4	9.7
1, $\frac{3\pi}{4}$	82.3	33.7	55.6	32.8	34.8	19.6	51.9	8.5

Table 5: Entropy Feature

ent	001	001 <sub>g</sub>	002	002 <sub>g</sub>	003	003 <sub>g</sub>	004	004 <sub>g</sub>
1, 0	2.83	4.98	4.97	5.40	4.61	5.32	4.32	5.01
1, $\frac{\pi}{4}$	3.03	4.99	5.08	5.43	4.55	5.29	4.82	5.02
1, $\frac{\pi}{2}$	2.91	4.98	4.74	5.35	4.54	5.30	4.61	5.02
1, $\frac{3\pi}{4}$	3.05	4.99	5.05	5.43	4.80	5.37	4.84	5.02

## 4. Discussion

In this paper, we have presented a KLT-based optimal color feature generation method using Haralick's co-occurrence method. In our simulation of four images, it appears that there is a close connection between the early stages of the primate visual pathway and the KLT method. A 3-layered computational model seems to be an appropriate approximation to the primate visual pathway. Development of this model remains one of the primary focuses of our future research work. Use of the extracted color features for texture and shape recognition requires further investigation.

## 5. Acknowledgement

The second author would like to thank Dr. Rolf Hankel for useful discussions.

## References

- [1] R. Haralick, Statistical and Structural Approaches to Texture, *Proc. IEEE*, (1979).
- [2] P. P. Ohanian and R. C. Dubes, *Pattern Recognition*, **25**, 8 (1992).
- [3] E. R. Kandel *et. al.*, *Principles of Neural Science*, Appleton & Lange, CT, 1991, pg. 400-417, 467-480.
- [4] Vision and Modeling Group, *VisTex Database*, MIT Media Lab, MA, 1995.
- [5] R. Linsker, From Basic Network Principles to Neural Architecture, *Proc. Natl. Acad. Sci., USA*, (1986).
- [6] R. Linsker, *Computer*, (March 1988).
- [7] J. Strand and T. Taxt, *Pattern Recognition*, **27**, 10 (1994).
- [8] R. Hankel<sup>¶</sup>, *Private Communication*.

## 6. Biography

Mehmet Celenk received his BS and MS degrees in Electrical Engineering from the Istanbul Technical University, Istanbul, Turkey, in 1974 and 1976, and his Ph.D. degree in Electrical Engineering and Computer Science from the Stevens Institute of Technology, Hoboken, NJ, in 1983. He served in the Turkish Army as a second lieutenant from 1984 to 1985, where he received the distinguished service award for his R&D contributions to the Signal School in Ankara, Turkey. He joined the Ohio University in 1985, where he is presently an Associate Professor of Electrical Engineering and Computer Science. He received the Russ Research Award in 1988 and the Academic Challenge Faculty Fellowship from 1988 to 1992 from the Russ College of Engineering and Technology of Ohio University. He published numerous articles and reviewed for various professional organizations including IEEE, SPIE, IS&T, IEE, and NSF. His research interests include computer vision and image processing, distributed computing and multimedia communications in computer networks, and reconfigurable digital systems. He is a member of IEEE, IEEE Computer & Communications Societies, ACM, SPIE, ASEE, IS&T, and Eta Kappa Nu.

---

<sup>¶</sup>Present address: Institut für Theoretische Neurophysik, Universität Bremen, Germany.



Figure 1. Image GroundWaterCity.0002.ppm from VisTex.



Figure 5. KLT-based channels of Fig. 1 in RGB-space

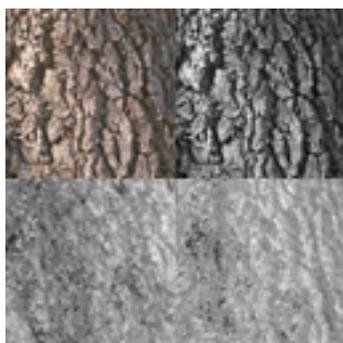


Figure 2. Image Bark.0009.ppm from VisTex.

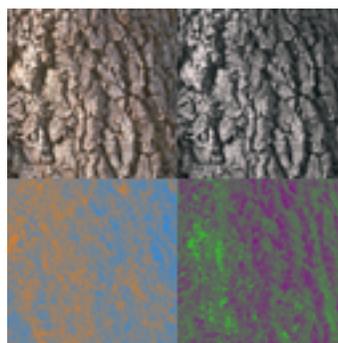


Figure 6. KLT-based channels of Fig. 2 in RGB-space

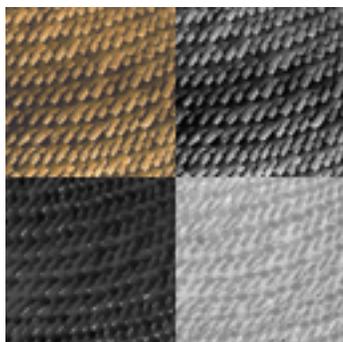


Figure 3. Image Fabric.0003.ppm from VisTex.

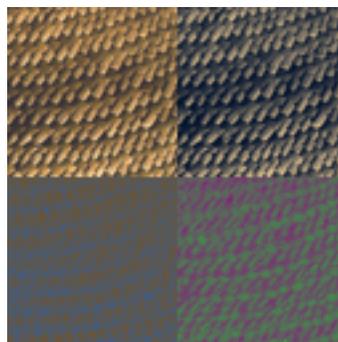


Figure 7. KLT-based channels of Fig. 3 in RGB-space

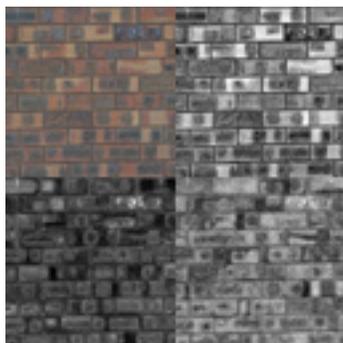


Figure 4. Image Brick.0001.ppm from VisTex.

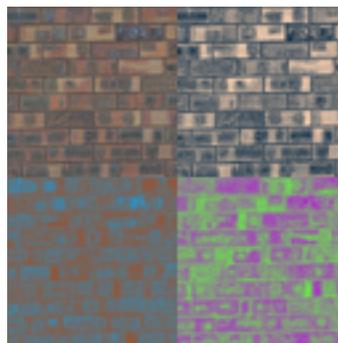


Figure 8. KLT-based channels of Fig. 4 in RGB-space