Karhunen–Loéve Expansion of a Set of Rotated Templates

Matjaž Jogan, Emil Žagar, Aleš Leonardis
Faculty of Computer and Information Science
University of Ljubljana
Tržaška c. 25 1001 Ljubljana, Slovenia
{matjaz.jogan,emil.zagar,alesl}@fri.uni-lj.si

Abstract

In this correspondence we propose a novel method for efficiently calculating the eigenvectors of uniformly rotated images of a set of templates. As we show, the images can be optimally approximated by a linear series of eigenvectors which can be calculated without actually decomposing the sample covariance matrix.

Keywords— **Karhunen–Loéve expansion, eigenvectors, symmetric matrices, Toeplitz matrices**

I. Introduction

There are several applications in computer vision where a template or a set of templates must be learned in a way that enables recognition or matching in every possible orientation of the target. It is clear that when using direct correlation, a large set of templates must be stored and compared with the target. Several researchers have therefore used techniques that compute an optimal approximation of a family of rotated templates [1], [2], [3].

Among the most popular approximation techniques in computer vision is the so-called Principal Component Analysis (PCA), also known as the Karhunen–Loéve expansion, that represents images with features — coefficient vectors — which are the projections of images onto an orthogonal set of eigenvectors.

Eigenvectors are usually calculated using the singular value decomposition of the covariance matrix. Eigenvectors are sorted according to their eigenvalues, which are related to the variance in the set of images that each eigenvector encompasses. By memorizing only a small number of eigenvectors with the largest eigenvalues, we construct an approximation of the learning set of images, which is the optimal linear representation in the least squared error sense.

This work was partly supported by the Ministry of Science and Technology of Republic of Slovenia (Project J2-0414).

In this correspondence we discuss the special properties of the eigenspace of a set of rotated templates. Our work was primarily triggered by a previous contribution of Uenohara and Kanade [4], which explained the tight relationship of the eigenspace of a single rotated template and the Discrete Cosine Transform. In this contribution we further analyze the problem of rotation and prove that for multiple rotated templates there also exists an alternative and faster method of calculation which, in contrast to the case of a single template, is based on complex Discrete Fourier Transform. The organization of the paper is as follows. First we review the related work on the eigenvectors of special matrices, such as those of a Toeplitz or circulant form. In section III we summarize the method of calculating the eigenvectors of a set of rotated versions of a single template. In section IV we show how one can generalize this method for a set of several rotating templates. Finally, in section VI we give a brief overview of the paper.

II. Review of related work

In their correspondence in these Transactions, Uenohara and Kanade [4] describe the relationship between the eigenvectors of a set of uniformly in-plane rotated images of an object and the basis vectors of the DCT. They show that the eigenvectors are completely defined by the fact that the inner product matrix of the image vectors is a symmetric Toeplitz matrix. As they claim, the eigenvectors of the inner product matrix are invariant of the image content and they can be generated much more efficiently by calculating the DCT transforms of the autocorrelation vector. This greatly alleviates the computational expense of the training phase. The authors also mention the relation of their results to those obtained by Perona [2] and Freeman [5] on steerable filters. The main drawback of the method described in [4] is that it can only be applied for templates of a single rotating object. **However some recognition problems require a representation which enables to interpolate the spline of coefficients in order to represent images that**

were not included in the training set. Although this may be not intuitive for sets of templates that are quite different in their appearance, it gets more clear when we look at specific problems, such as recognizing a panoramic view from a training set of panoramic snapshots representing distinct locations in the environment [6], which will be presented in Section V.

As we pointed out in [3], the eigenvectors of a training set consisting of several rotated objects or scenes can not be calculated directly with the use of DCT. However, the structure of the basis functions calculated by SVD resembled that of a combination of harmonics basis. The properties of the eigenvectors of Toeplitz matrices have been thoroughly described before in Sjöström [7] or Gray [8]. In both works it is showed that a circulant matrix (a special case of a Toeplitz matrix) can be diagonalized by the Fourier coefficient matrix, which is a property we will exploit in section IV.

III. Eigenspace representation of a set of rotated versions of a single template

In this section we briefly summarize the procedure for calculating the eigenspace of a set of rotated versions of a single template, as described in [4] and introduce the notation which will be used throughout the paper.

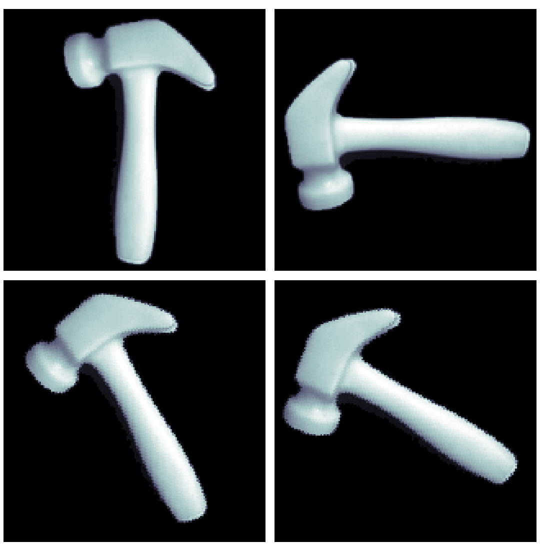

Fig. 1. Template of a toy hammer rotated by 0, (-30), (-60) and (-90) degrees, respectively.
figure

As a set of rotated versions of a template we understand images captured from a single point of view, but under different in-plane rotations. The only constraint is that with the in-plane rotation the information content is preserved. Such is the case when rotating an image of an object on a homogeneous background [4] (Fig. 1), or with panoramic images rotated around the optical axis [3]. To obtain the images of a template in all the possible rotations, we can rotate the original image sequentially by $2\pi/N$. Examples of the first six eigenvectors for the uniformly rotated template of the toy hammer can be seen in Figure 2. **In the digital case, t**his rotation can be described more accurately, if we warp the image to a polar representation as in Fig. 3. In this case, we can rotate the image just by shifting the columns accordingly.

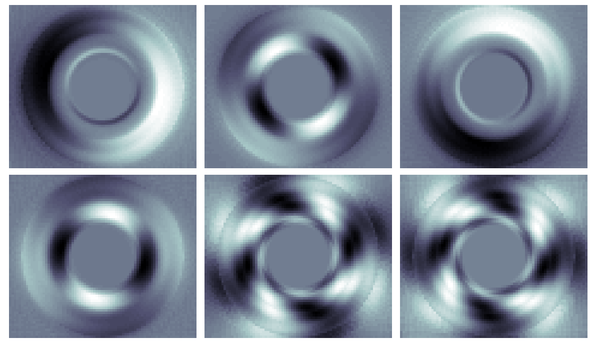

Fig. 2. First 6 eigenvectors of the uniformly rotated template of a toy hammer.
figure

We represent images from the training set as normalized image vectors, from which the mean image is subtracted, in an image matrix $X \in \mathbb{R}^{n\times N}$

$$X = \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{N-1} \end{bmatrix} ,$$

where n is the number of pixels in the image and N is the number of images.

The most straightforward way to solve the eigensystem is to calculate the SVD of the covariance matrix $C \in \mathbb{R}^{n\times n}$ of this normalized vector matrix

$$C = XX^T = \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{N-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0^T \\ \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_{N-1}^T \end{bmatrix} .$$

The eigenvectors $\mathbf{v}_i,\ \ i = 0, \dots, N-1$, form an orthogonal basis. Sorted with respect to descending eigenvalues $\lambda_i,\ \ i = 0, \dots, N-1$, they represent the best linear approximation of the image data. Since the number of pixel elements n in an image is usually high, the computation of the matrix C is a time consuming task of high storage demands. However, it is possible to formulate the equations in such a way that it becomes sufficient to calculate the eigenvectors $\mathbf{v}_i',\ \ i = 0, \dots, N-1$, of the inner product matrix $Q \in \mathbb{R}^{N\times N}$,

Fig. 3. Warped template of a toy hammer rotated 0, (-30), (-60) and (-90) degrees by shifting the columns.

figure

$$Q = X^T X = \begin{bmatrix} \mathbf{x}_0^T \\ \mathbf{x}_1^T \\ \dots \\ \mathbf{x}_{N-1}^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{N-1} \end{bmatrix} \ . \tag{1}$$

Since the eigenvectors $\mathbf{v}'_i$ are the solutions of $X^T X \mathbf{v}'_i = \lambda_i \mathbf{v}'_i$, we can calculate the eigenvectors of XX^T by $XX^T X \mathbf{v}'_i = \lambda_i X \mathbf{v}'_i$ [9]. In this way, we derive the eigenvectors $\mathbf{v}_i$ of the covariance matrix just by projecting the $\mathbf{v}'_i$ on the set of images,

$$\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}} X \mathbf{v}'_i \ .$$

Uenohara and Kanade [4] showed that in the case of an image set consisting of rotated examples of one original image, Q is a **circulant** symmetric Toeplitz matrix and **its eigenvectors $\mathbf{v}'_i$ are therefore not dependent on the contents of the images** [4][1]. Q is of the form

$$Q = \begin{bmatrix} q_0 & q_1 & \dots & q_{N-2} & q_{N-1} \\ q_{N-1} & q_0 & q_1 & \dots & q_{N-2} \\ q_{N-2} & q_{N-1} & q_0 & q_1 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ q_1 & \dots & q_{N-2} & q_{N-1} & q_0 \end{bmatrix} \ .$$

It can be derived from the *shift theorem* [10], that the eigenvectors of a general circulant matrix are the N basis vectors from the Fourier matrix $F = [\mathbf{v}'_0, \mathbf{v}'_1, \dots, \mathbf{v}'_{N-1}]$,[3] where

$$\mathbf{v}'_i = \left[1, \omega^i, \omega^{2i}, \dots, \omega^{(N-1)i}\right]^T, \ \ i = 0, \dots, N-1,$$

and $\omega = e^{-2\pi j/N}$, $j = \sqrt{-1}$. The eigenvalues can be calculated simply by retrieving the magnitude of the DFT of one row of Q,

$$\lambda_i = \sum_{k=0}^{N-1} q_k \, \omega^{i\,k} \ .$$

This interesting property also emphasizes the central point of the Fourier analysis, as it indicates that the Fourier basis diagonalizes every periodic constant coefficient operator, in our case the circular shift operator [11]. In other words, all basis functions of the Fourier transform are eigenvectors of the circular shift operator [10].

Since $q_i = q_{N-i}$, our matrix is circulant symmetric, and therefore we can choose an appropriate set of real–valued orthogonal eigenvectors. As it turns out, the proper basis are the cosine functions from the real and the sine functions from the imaginary part of the Fourier matrix [11].

We can therefore compute the eigensystem of Q just by first computing the autocorrelation vector $[q_0, q_1, q_2, \dots, q_{N-1}]$, and then by calculating the λ_i values, which should be afterwards sorted by decreasing magnitude. The eigenvectors $\mathbf{v}'_i$ corresponding to k largest eigenvalues can then easily be selected from the corresponding basis vectors of the Discrete Cosine Transform (DCT) [4]:

$$v'_{im} = \cos\left[\frac{\pi(2m+1)\,i}{2N}\right]; \quad \begin{array}{rcl} m & = & 0, \dots, N-1 \\ i & = & 0, \dots, k-1 \end{array}$$

Thus, with the help of the DCT, it is possible to compute the basis vectors much more efficiently.

IV. Generalized method for a set of several rotated templates

When dealing with multiple templates, the calculation described in the previous section can not be applied. In this case, we deal with P different templates (images), each of them being rotated N times (Fig. 4). In this case, we cannot directly apply the previous approach to the calculation of eigenvectors of circulant matrices, since the inner product matrix A,

$$A = X^T X = \begin{bmatrix} Q_{00} & Q_{01} & \dots & Q_{0,P-1} \\ Q_{10} & Q_{11} & \dots & Q_{1,P-1} \\ \dots & \dots & \dots & \dots \\ Q_{P-1,0} & Q_{P-1,1} & \dots & Q_{P-1,P-1} \end{bmatrix},$$

is composed of several circulant blocks Q_{jk}, which are, in general, not symmetric. However, as we will show, it is still

[1] **Note that in the case of rotating a digital image, this is just an approximation, unless we use the polar representation of images, as in Fig.3**

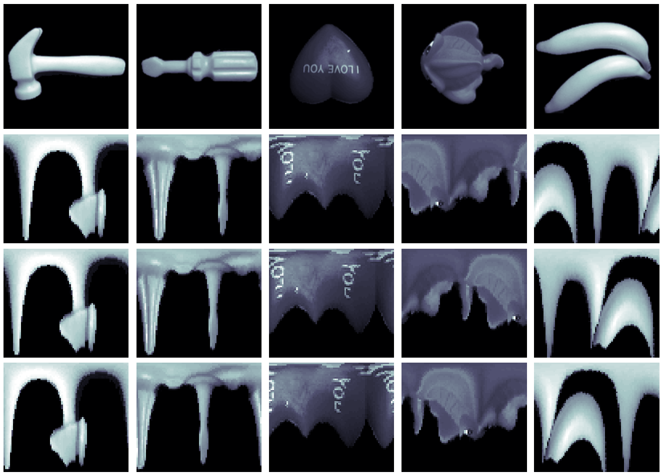

Fig. 4. Templates of five objects and examples of their warped images in three different orientations.
figure

possible to calculate the eigenvectors without performing the SVD decomposition of X.

We have to find a solution of the eigenvalue problem

$$A\mathbf{w}' = \mu\mathbf{w}', \tag{2}$$

where $(\mu, \mathbf{w}')$ is the eigenpair of A. The fact that the matrix blocks Q_{jk} of A are circulant matrices is crucial. As it was already mentioned, every circulant matrix can be diagonalized in the same basis by Fourier matrix F. Consequently all the submatrices Q_{jk} have the same set of eigenvectors $\mathbf{v}'_i$, $i = 0, \ldots, N-1$. Following [12] we shall find the eigenvectors $\mathbf{w}'$ of A among the vectors of the form

$$\mathbf{w}'_i = \left[\alpha_{i0}\mathbf{v}'^T_i, \alpha_{i1}\mathbf{v}'^T_i, \ldots, \alpha_{i,P-1}\mathbf{v}'^T_i\right]^T, \tag{3}$$

where $i = 0, \ldots, N-1$. Equation (2) can be rewritten blockwise as

$$\sum_{k=0}^{P-1} Q_{jk}(\alpha_{ik}\mathbf{v}'_i) = \mu\alpha_{ij}\mathbf{v}'_i, \;\; j = 0, \ldots, P-1 \;\; .$$

Since $\mathbf{v}'_i$ is an eigenvector of every Q_{jk}, the equations simplify to

$$\sum_{k=0}^{P-1} \alpha_{ik}\lambda^i_{jk}\mathbf{v}'_i = \mu\alpha_{ij}\mathbf{v}'_i, \;\; j = 0, \ldots, P-1,$$

where λ^i_{jk} is an eigenvalue of Q_{jk} corresponding to $\mathbf{v}'_i$. This implies a new eigenvalue problem

$$\Lambda^i \boldsymbol{\alpha}_i = \mu\boldsymbol{\alpha}_i \;\; , \tag{4}$$

where

$$\Lambda^i = \begin{bmatrix} \lambda^i_{00} & \lambda^i_{01} & \ldots & \lambda^i_{0,P-1} \\ \lambda^i_{10} & \lambda^i_{11} & \ldots & \lambda^i_{1,P-1} \\ \ldots & \ldots & \ldots & \ldots \\ \lambda^i_{P-1,0} & \lambda^i_{P-1,1} & \ldots & \lambda^i_{P-1,P-1} \end{bmatrix}$$

and

$$\boldsymbol{\alpha}_i = [\alpha_{i0}, \alpha_{i1}, \ldots, \alpha_{i,P-1}]^T \;\; .$$

Since $Q_{jk} = Q^T_{kj}$, it can be proved that Λ^i is Hermitian and we have P linearly independent eigenvectors $\boldsymbol{\alpha}_i$, which provide P linearly independent eigenvectors $\mathbf{w}'_i$ in (3). Since the same procedure can be performed for every $\mathbf{v}'_i$, we can obtain $N \cdot P$ linearly independent eigenvectors of A.

It is therefore possible to solve the eigenproblem using N decompositions of order P. Since P **represents the number of unique templates and is therefore** usually small in comparison to the total number of images $P \cdot N$, this method offers a similar improvement as the method in [4]. **Assuming that the time complexity of decomposing a $n \times n$ matrix is $O(n^3)$, our method accomplishes the task in $N \cdot O(P^3)$ time instead of $O((N \cdot P)^3)$.**

However, by looking at the properties of the circulant matrices one can deduce, that this method works only if we use the complex Fourier basis as the eigenvector set for the circulant matrix. In fact, this set of basis vectors is the only common eigenspace for all the submatrices Q_{ij} from A. Further, as it was shown in Sanchez et al. [13], all the matrices that have the DCT as their eigenvectors are generally full matrices of the form of a Toeplitz matrix combined with a near–Hankel matrix scaled by some constant factors [11]. Since our matrix A does not belong to this class, we cannot use any of the known DCT basis functions to diagonalize it in a real basis effectively as in the case of a single template [4]. Therefore, using a complex basis in the calculation of the final representation results in a complex set of eigenvectors $\mathbf{w}'_i$.

Since both the covariance matrix C and the inner product matrix A are symmetric matrices, they obviously have real eigenvalues and eigenvectors. We can find them by using the following observations. The symmetric structure of the real matrix A implies that the eigenvectors come in conjugated pairs with the same eigenvalue. Therefore, the real and imaginary parts of complex eigenvectors are also real eigenvectors of A. By considering the multiplicity of the eigenvalues, we construct the real set of eigenvectors as follows: in the case of a simple eigenvalue, the corresponding eigenvector is already real, or can be obtained by normalizing the complex one by its nonzero component. In the case of double eigenvalues, the real set of eigenvectors can be generated by selecting just one of the conjugated pair of the complex eigenvectors with the same eigenvalue and taking

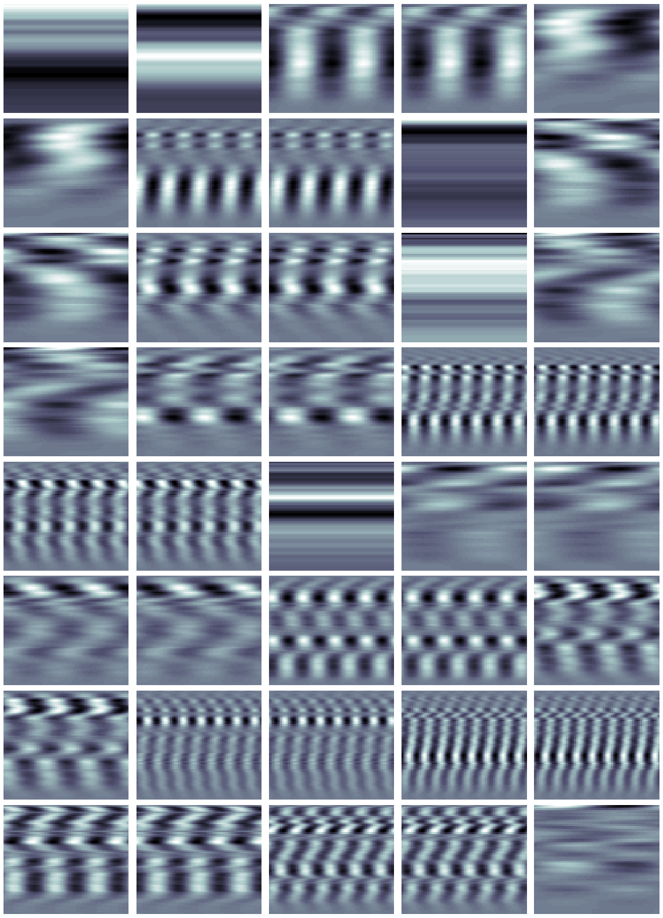

Fig. 5. First forty components of the eigenbasis. Pairs of vectors that differ only in phase form a complex eigenvector.

figure

its real and imaginary parts as a new pair of real eigenvectors with the same eigenvalue. In a general case, the real eigenvectors can be obtained by the orthogonalization of the set of real and imaginary components of the complex eigenvectors. However, in practice, when dealing with images, it is not likely that such a case would occur. So, the proposed method turns out to be very efficient in practical applications. An example of the basis set of eigenvectors can be seen in Figure 5.

However, one can use also the complex basis; in this case we reduce every conjugated pair to a single complex eigenvector. Since pairs of trigonometric functions are more intuitively described in the complex space, calculations with such a basis are even easier than with the real one.

V. A practical demonstration – appearance–based representation of environment using an eigenspace of panoramic images

Our algorithm opens up many possibilities of use in practical applications. As we show here, there exist specific recognition problems that require multiple rotated templates to be encompassed in a single eigenspace which makes possible to interpolate the model in order to represent images that were not explicitly included in the training set. Although this may be not intuitive for sets of templates that are quite different in their appearance, it gets clearer when we look at specific situations. One of the possible applications where our algorithm shows its potential is the task of appearance–based localization, where the appearance of the environment is represented by a training set of panoramic snapshots representing distinct locations in the environment [6]. Examples of four panoramic views can be seen in Fig.7.

It is easy to demonstrate that if the panoramic sensor has a fixed orientation, two images taken at nearby positions tend to be strongly correlated. Since correlation is related to the distance of image projections in the eigenspace, it is obvious that by interpolating that representation one can get a much coarser (although approximate) representation of the appearance of the environment. By virtually rotating all of the panoramic images in order to represent multiple in plane rotations of the robot, the problem to solve is identical to that of multiple rotated templates. One could argue that each location could be represented by its own eigenspace, however an interpolated representation is possible only if all of the positions are represented in an unique eigenspace. We demonstrate this by showing six images (Fig 6), where the first and the last one are from the training set and were taken 50 cm apart, while the other four are reconstructions derived from an interpolated representation at in–between points, representing locations spaced 10 cm apart.

To further clarify the topic, we performed a set of experiments that show the advantage of our method in the case of appearance–based mobile robot localization. As a testing platform, we used a mobile robot equipped with a panoramic camera. The idea was to perform an exploration phase and construct a model of an indoor environment by acquiring panoramic images at 62 measured positions. In our experiment, we chose these positions to be on a 60×60 cm squared grid, and are denoted as squares in Fig 7. The first 56 eigenvectors calculated from these images and their rotated versions can be seen on Fig. 8.

To build a denser representation we interpolated the coefficients of training images on a 5×5 cm grid. The interpolated coefficients for rotated versions of images were also generated. In that way, intermediate positions are also represented in the model although images taken at this po-

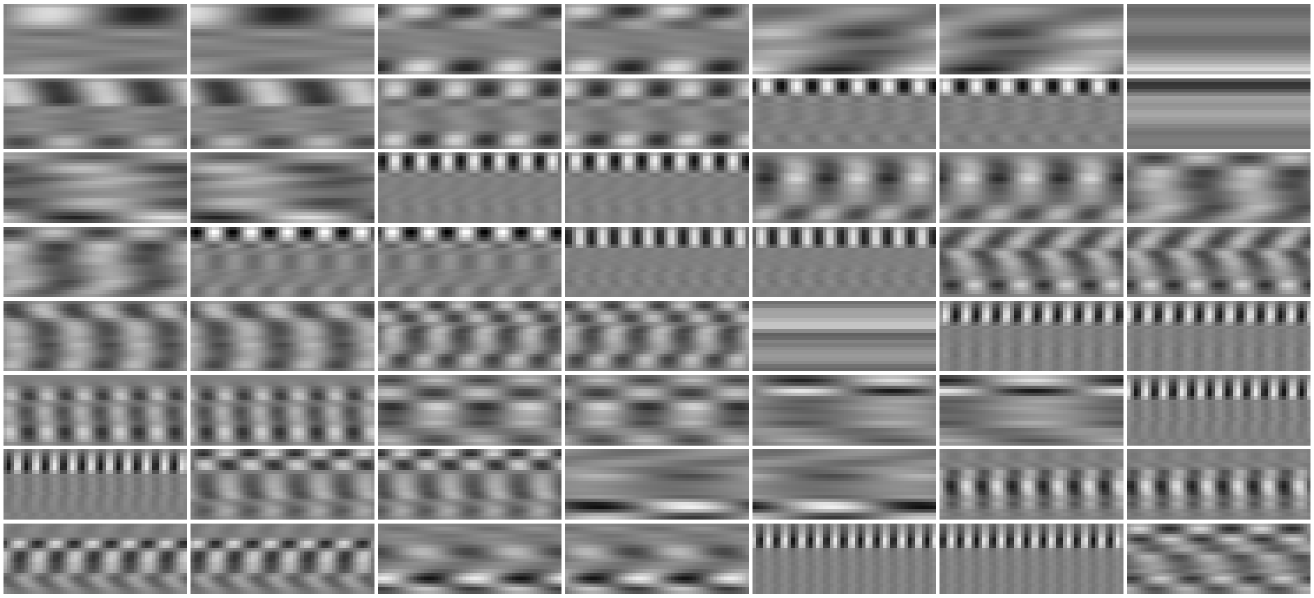

Fig. 8. First 56 eigenvectors constructed from a set of 62 panoramic images, each rotated (shifted) 50 times.
figure

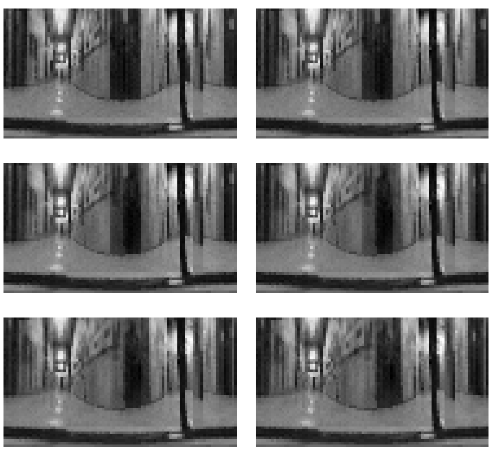

Fig. 6. Six panoramic images, where the intermediate four are virtual images, interpolated from the first and the last image taken from the training set.
figure

sitions were not included in the training set.

VI. Conclusions

We have shown how to compute the K-L expansion of a number of uniformly rotated images arriving from different

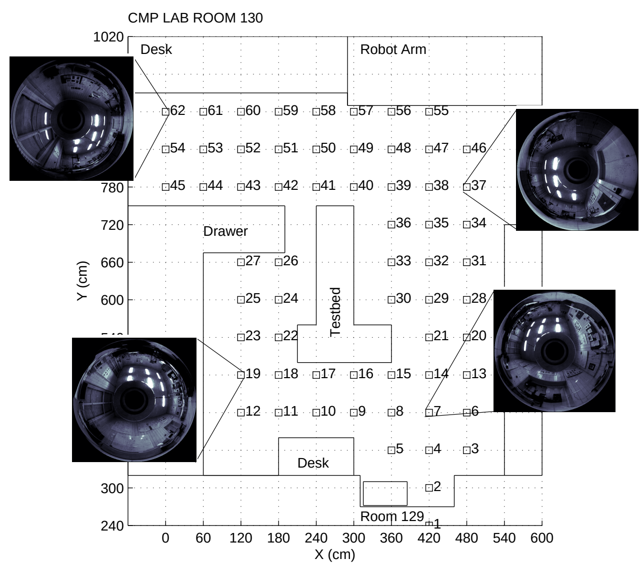

Fig. 7. A map of an indoor environment which was represented by an appearance–based model constructed from a set of images acquired at 62 positions denoted with squares.
figure

objects or scenes. Using our method one can construct an optimal linear basis without performing the Singular Value Decomposition on the whole set of images or their covariance matrix. Instead, by choosing a complex Fourier basis set as eigenvectors for circulant submatrices, we show that the system can be solved by operating on a set of smaller eigenproblems of order P, where P is the number of templates we want to represent. Since P is usually small in comparison with the total number of images ($P \cdot N$), this method offers a similar improvement as the method in [4] in the case of a single image.

Furthermore, the method provides an insight in the calculation of the Karhunen-Loéve expansion for sets of rotated templates; by following the procedure, it can be proven that, also for the set of several rotated templates, the final eigenvectors are composed by locally varying harmonic functions. Once known, these properties can easily be exploited in order to ease the recognition or enable scale invariance.

We demonstrated the practical application of our method by constructing an appearance–based model of environment, which can be used for the task of mobile robot localization. Since there is a need for explicit interpolation of the representation in order to represent in–between positions, the task to solve is identical to that having a set of multiple rotated templates which have to be represented in an unique eigenspace.

References

[1] Shinichi Yoshimura and Takeo Kanade, "Fast template matching based on the normalized correlation by using multiresolution eigenimages," in *Proc. IROS'94*, 1994.

[2] P. Perona, "Deformable kernels for early vision," in *CVPR91*, 1991, pp. 222–227.

[3] Matjaž Jogan and Aleš Leonardis, "Robust localization using eigenspace of spinning-images," in *IEEE Workshop on Omnidirectional Vision*. June 2000, pp. 37–44, IEEE Computer Society.

[4] Michihiro Uenohara and Takeo Kanade, "Optimal approximation of uniformly rotated images: Relationships between Karhunen-Loeve expansion and discrete cosine transform," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 116–119, 1998.

[5] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[6] Matjaž Jogan and Aleš Leonardis, "Robust localization using panoramic view-based recognition," in *15th International Conference on Pattern Recognition*. September 2000, vol. 4, pp. 136–139, IEEE Computer Society.

[7] Eva Lundström, *Singular Value Computations for Toeplitz Matrices and Subspace Tracking*, Ph.D. thesis, Linköpings Universitet, 1996.

[8] Robert M. Gray, "Toeplitz and circulant matrices: A review," Tech. Rep., Stanford University, 2000.

[9] Dana Ballard, *Natural Computation*, The MIT Press, 1996.

[10] Bernd Jähne, Horst Haußecker, and Peter Geißler, Eds., *Handbook of Computer Vision and Applications*, Academic Press, 1999.

[11] Gilbert Strang, "The discrete cosine transform," *SIAM Review*, vol. 41, no. 1, pp. 135–147, 1999.

[12] Sergej Rjasanow, "Effective algorithms with circulant-block matrices," *Linear Algebra Appl.*, vol. 202, pp. 55–69, 1994.

[13] Victoria Sànchez, Pedro García, and Antonio M. Peinado, "Diagonalizing properties of the discrete cosine transform," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2631–2641, November 1995.