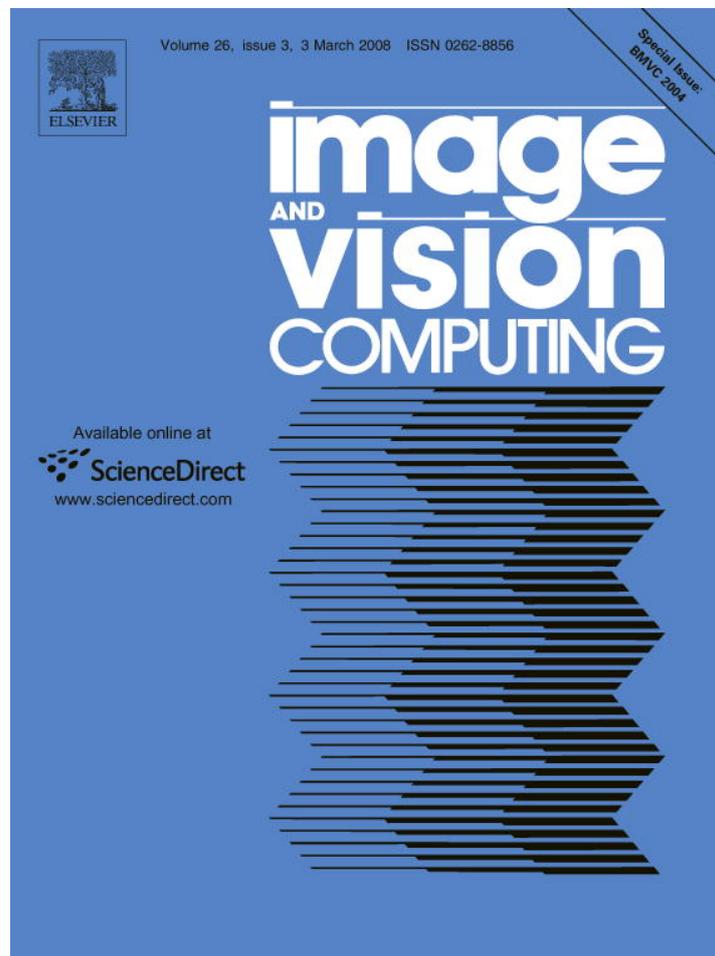


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# Drums, curve descriptors and affine invariant region matching

M. Zuliani \*, L. Bertelli, C.S. Kenney, S. Chandrasekaran, B.S. Manjunath

*Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, USA*

Received 18 February 2005; received in revised form 16 May 2006; accepted 8 December 2006

## Abstract

In this paper we present a new physically motivated curve/region descriptor based on the solution of Helmholtz's equation. The descriptor we propose satisfies the six principles set by MPEG-7: it has a good retrieval accuracy, it is compact, it can be applied in general contests, it has a reasonable computational complexity, it is rotation and scale invariant and provides an hierarchical representation of the curve from coarse to fine. In addition to these properties, the descriptor can be generalized in order to take into account also the intensity content of the image region defined by the curve. The construction of the descriptor can be coupled with a preprocessing step that enables us to describe a curve in an affine invariant fashion. The performance of our approach has been tested in the contest of affine invariant curve and region matching, both within a controlled experimental setup and also using real images. The experiments show that the proposed approach compares favorably to the state of the art curve/region descriptors.

© 2007 Elsevier B.V. All rights reserved.

*Keywords:* Curve descriptors; Curve matching; Helmholtz equation; Affine invariance

## 1. Introduction

The quest for efficient curve and region descriptors has been one of the leading themes in the image analysis community. In general, good descriptors should be invariant under an appropriate class of geometric transformations (like, for example, rotation-scaling-translation or affine), robust in presence of noise, efficient to compute and easy to compare. Zhang et al. [33] classified the curve description approaches into two groups: *contour-based* and *region-based* methods. Each of these groups is further subdivided into two subgroups containing *global* or *structural* approaches. Some of the recently proposed descriptors fall in the contour-based category, as the *curvature scale space* (CSS) descriptor [22] (which has been standardized within the MPEG-7 framework) and the *shape context matrices* [2]. Some others belong to the class of region-based

methods, like the descriptors based on *moments* (geometric [8], Zernike, also standardized within the MPEG-7 framework, and Legendre [29]), on *region frequency representations* (Fourier descriptors [32]), on the *medial axis transform* [23] and on *shock graphs* [27]. Recently Gorelick et al. exploited the properties of the *Poisson equation* to characterize shapes and to derive a set of features that can serve as descriptors [12].

We are interested in a descriptor that can be used in the context of image registration where we will focus on image features defined by Jordan curves (i.e. curves that are closed and do not cross themselves) in order to establish image matches. Such matching should happen in an affine invariant fashion, so that perspective distortions can be handled robustly. The approaches mentioned before either do not always satisfy the MPEG-7 requirements or they are not completely suitable to establish affine invariant matches between curves. The computation of the CSS descriptors is quite demanding, the algorithm converges slowly if the curve is very complex (i.e. the curve presents many points with large curvature) and depends on some empirical parameters that need to be fine tuned. The comparison of

\* Corresponding author. Tel. +1 805 893 5282.

*E-mail addresses:* [zuliani@ece.ucsb.edu](mailto:zuliani@ece.ucsb.edu) (M. Zuliani), [lbartelli@ece.ucsb.edu](mailto:lbartelli@ece.ucsb.edu) (L. Bertelli), [kenney@ece.ucsb.edu](mailto:kenney@ece.ucsb.edu) (C.S. Kenney), [shiv@ece.ucsb.edu](mailto:shiv@ece.ucsb.edu) (S. Chandrasekaran), [manj@ece.ucsb.edu](mailto:manj@ece.ucsb.edu) (B.S. Manjunath).

CSS descriptors is not simple. Shape context matrices provide local curve descriptors that are not very compact (since they consist of the coefficients of a matrix) and their comparison is not very fast. Moment invariants of higher orders do not have a clear physical interpretation and the matching procedure requires a normalization process to compensate for the different dynamic range of the moments of different orders. However recently Zhang et al. [31] experimentally showed that Fourier descriptors and Zernike moment descriptors perform better than the CSS descriptors. Shock graphs are very suitable in scenarios where the similarity between curves is defined in terms of structure, but are not the ideal solution if the notion of equivalence is defined within the class of some specific geometric transformation. Moreover the computational complexity for extracting these descriptors and matching them is quite high. Finally note that, with the exception of moments, it is not straightforward to extend the descriptors listed above in order to represent also the intensity pattern inside the curve. For an extensive quantitative comparison of region descriptors that explicitly take into account the intensity pattern within the region, the interested reader should refer to the survey by Mikolajczyk et al. [21].

The goal of this paper is to develop a curve descriptor that satisfies the six principles set by MPEG-7 and a few other important requirements, such as being rotation-scaling-translation (RST) invariant, having a clear physical interpretation and being capable to take into consideration the intensity content of a closed contour (when the curve identifies an image region). The descriptor we propose is novel in the sense that it combines intimately both the information regarding the shape of a region and its intensity content.

This paper is structured as follows. Section 2 introduces the Helmholtz descriptor, it discusses its analytical properties and presents the numerical scheme used to compute the descriptor. Section 3 will describe a preprocessing step that aims at extracting the *shape* of a curve in order to obtain an affine invariant matching algorithm. In Section 4 we will

show some experimental results and we will evaluate the performance of the descriptors. Finally the conclusions are presented in Section 5.

## 2. The descriptor

In 1966 the mathematician M. Kac published his famous paper entitled “Can One Hear the Shape of a Drum?” [16]. Kac was interested in understanding whether the knowledge of the modes of vibration a drum was sufficient for univocally inferring its geometric structure. The problem posed by Kac can be related to the problem of constructing curve or region descriptors. In fact, if we imagine that the curve we want to label defines the contour of a drum, it is reasonable to think that the spectrum of such curve (in terms of modes of vibration) could be an appealing descriptor, given the fact that it can be easily made RST invariant and has a strong physical characterization. Moreover the intensity inside the image region defined by the curve can be used to model the physical properties of the membrane so that the modes of vibration are related not only to the structure of the boundary but also to the region content. With this in mind, the answer to Kac’s question becomes crucial, i.e. we would like to have the normal modes of vibration of a drum to identify univocally its geometry (so that we can establish a bijection between the space of the Jordan curves modulo a given transformation and the curve descriptors).

The problem posed by Kac remained unsolved until 1992 when the mathematicians C.S. Gordon, D.L. Webb and S. Wolpert proposed a pair of isospectral drums having the same area and perimeter but different contours. In other words “One Cannot Hear the Shape of a Drum” [3,5,11] (see Fig. 1 for some examples of isospectral drums). Even though for our purposes this fact is unfortunate, since it implies that there may exist curves that are not related by an RST transformation and nonetheless have the same spectrum (i.e. possibly the same descriptor), the experiments presented in Section 4 will show how this problem has a limited impact in real life scenarios. Note that an

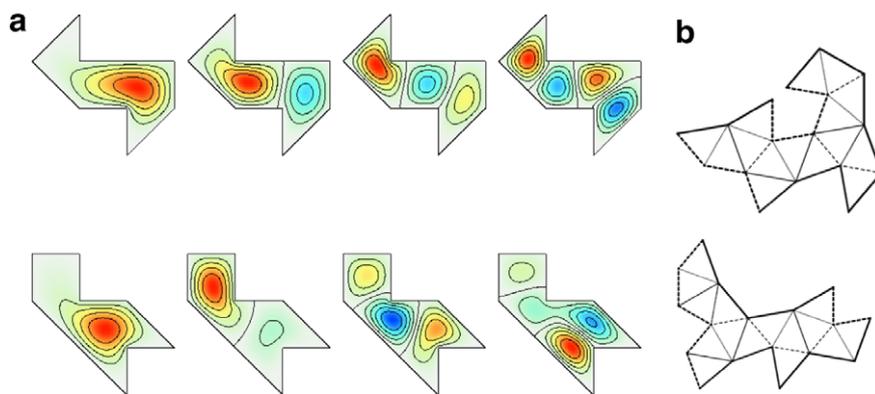


Fig. 1. (a) (courtesy of Dr. T. Driscoll) The first four eigenmodes of an isospectral domain. (b) (courtesy of Dr. Buser, Dr. Conway, Dr. Doyle and Dr. Semmler) Another example of an isospectral domain.

application of the Laplace operator (deeply connected to the spectral properties of a closed region) in the context of image processing and computer vision has also been explored by Saito [26].

In the following subsections we will describe in detail the proposed curve descriptor and the numerical scheme used to compute it.

### 2.1. The Helmholtz equation

Let  $\Gamma$  be a Jordan curve corresponding to the boundary of  $\Omega$ , an open subset of  $\mathbb{R}^2$ . The vibration of the membrane of a drum whose contour is defined by  $\Gamma$  is expressed by the function  $w(\mathbf{x}, t) : \overline{\Omega} \times \mathbb{R} \rightarrow \mathbb{R}$  which solves the wave equation

$$\Delta w(\mathbf{x}, t) - \frac{1}{v(\mathbf{x})^2} \frac{\partial^2 w}{\partial t^2}(\mathbf{x}, t) = 0$$

where  $\Delta$  denotes the Laplacian operator,  $t$  indicates time and  $v(\mathbf{x}) > 0$  indicates the phase velocity of the membrane.<sup>1</sup> This equation can be solved via separation of variables, assuming that  $w$  can be decomposed into a spatial part and into a temporal part according to  $w(\mathbf{x}, t) = u(\mathbf{x})q(t)$ . It can be shown that the spatial part solves the Helmholtz equation, i.e. the elliptic partial differential equation:

$$\Delta u(\mathbf{x}) + \lambda \frac{1}{v(\mathbf{x})^2} u(\mathbf{x}) = 0 \quad (1)$$

where  $\lambda$  is a suitable scalar. The corresponding boundary problem with Dirichlet boundary conditions is:

$$-\Delta u(\mathbf{x}) = \lambda \frac{1}{v(\mathbf{x})^2} u(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \quad (2a)$$

$$u(\mathbf{x}) = 0 \quad \text{for } \mathbf{x} \in \Gamma \quad (2b)$$

### 2.2. The descriptor

Our idea is to use the first  $N_\lambda + 1$  eigenvalues associated to the Helmholtz Eq. (2) to build an RST-invariant descriptor for the curve  $\Gamma$  (in the case where  $v(\mathbf{x}) = v = \text{const}$ ) or for the image patch contained in the region  $\Omega$  (if we set<sup>2</sup>  $v(\mathbf{x})^2 = \frac{1}{I_s(\mathbf{x})}$ , where  $I_s(\mathbf{x})$  denotes the smoothed version of the image intensity at point  $\mathbf{x}$ ). As explained in more detail in Appendix A all the eigenvalues associated to (2) are real and positive and can be sorted in order of increasing value:  $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$  with  $\lambda_k \rightarrow \infty$  as  $k \rightarrow \infty$ . These observations justify the following definition:

**Definition 2.1.** Let  $\Gamma$  be a Jordan curve and let  $\lambda_1, \dots, \lambda_{N_\lambda+1}$  be the first  $N_\lambda + 1$  eigenvalues that solve (2).

<sup>1</sup> For a real membrane the phase velocity is proportional to  $\sqrt{\frac{T}{\sigma(\mathbf{x})}}$ , where  $T$  denotes the membrane tension (expressed in Newtons over meters) and  $\sigma$  the membrane density (expressed in kilograms per square meter), and function of the spatial position  $\mathbf{x}$ .

<sup>2</sup> The physical intuition behind this choice is that the membrane density at  $\mathbf{x}$  is directly proportional to the image intensity at the point  $\mathbf{x}$ .

The correspondent Helmholtz descriptor (HD) is defined as:

$$\mathbf{F}(\Gamma) \stackrel{\text{def}}{=} \left[ \frac{\lambda_1}{\lambda_2} \quad \frac{\lambda_2}{\lambda_3} \quad \dots \quad \frac{\lambda_{N_\lambda}}{\lambda_{N_\lambda+1}} \right]^T \in \mathbb{R}^{N_\lambda} \quad (3)$$

The invariance of the descriptor with respect to an RST transformation can be understood observing that a vibrating membrane will produce the same tones when it is rotated and translated, and that a scaling will only affect their amplitude. This intuition is formalized in the following lemma:

**Lemma 2.2.** Consider the two Jordan curves  $\Gamma_1$  and  $\Gamma_2$  related by an RST transformation:

$$\Gamma_2 = \{\mathbf{x}_2 \in \mathbb{R}^2 : \text{there exists } \mathbf{x}_1 \in \Gamma_1 \text{ such that } \mathbf{x}_2 = sR\mathbf{x}_1 + \mathbf{t}\}$$

where  $s \in \mathbb{R}$  the scaling factor,  $R \in SO(2)$  is a rotation matrix and  $\mathbf{t} \in \mathbb{R}^2$  is a translation vector. Let also  $v_2(\mathbf{x}_2) = v_1(\mathbf{x}_1)$ . Then  $\mathbf{F}(\Gamma_1) = \mathbf{F}(\Gamma_2)$ .

**Proof.** To proof of this lemma follows from the definition of the Laplacian in an orthogonal coordinate system, which is:

$$\Delta = \frac{1}{h_1 h_2} \left[ \frac{\partial}{\partial x_1} \left( \frac{h_2}{h_1} \frac{\partial}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left( \frac{h_1}{h_2} \frac{\partial}{\partial x_2} \right) \right]$$

where  $h_1$  and  $h_2$  are the scale factors of the first fundamental form. It can be easily verified that for a scaling  $s$  and an arbitrary rotation we have  $h_1 = h_2 = s$ . Therefore we can write  $\frac{1}{s^2} \Delta u_2(\mathbf{x}_2) = \Delta u_1(\mathbf{x}_1)$ . Thus the eigenpairs that solve:

$$-\Delta u_1(\mathbf{x}_1) = \lambda \frac{1}{v_1(\mathbf{x}_1)^2} u_1(\mathbf{x}_1) \quad \text{for } \mathbf{x}_1 \in \Omega_1$$

$$u_1(\mathbf{x}_1) = 0 \quad \text{for } \mathbf{x}_1 \in \Gamma_1$$

can be used to construct the solutions for:

$$-\Delta u_2(\mathbf{x}_2) = (s^2 \lambda) \frac{1}{v_2(\mathbf{x}_2)^2} u_2(\mathbf{x}_2) \quad \text{for } \mathbf{x}_2 \in \Omega_2$$

$$u_2(\mathbf{x}_2) = 0 \quad \text{for } \mathbf{x}_2 \in \Gamma_2$$

by scaling the eigenvalues by  $s^2$  and by letting  $u_2(sR\mathbf{x}_1 + \mathbf{t}) = u_1(\mathbf{x}_1)$ . Since the components of the descriptors are ratios of eigenvalues, the scaling factor vanishes and the assertion holds true.  $\square$

### 2.3. Numerical scheme

The second order finite difference scheme we used to solve (2) is a reasonable compromise between accuracy and computational complexity. The step size of the  $N \times N$  discretization mesh is calculated according to

$$h = \frac{\max_{\mathbf{x} \in \Gamma} \|\mathbf{x} - \mathbf{m}(\Omega)\|}{\Delta} \quad (6)$$

where  $m(\Omega)$  is the center of gravity of the region  $\Omega$  (which will be defined formally in Section 3) and  $\Delta$  is a parameter that defines the mesh resolution. The spatial derivatives are approximated by the second order central difference formulae:

$$\frac{\partial^2 u}{\partial x^2}(\mathbf{x}) \approx \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2}$$

$$\frac{\partial^2 u}{\partial y^2}(\mathbf{x}) \approx \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2}$$

which provide the discretized version of (2a):

$$-\frac{u_{p+1,q} + u_{p-1,q} + u_{p,q+1} + u_{p,q-1} - 4u_{p,q}}{h^2} = \lambda \frac{1}{v_{p,q}^2} u_{p,q}$$

where  $0 \leq p \leq N-1$  and  $0 \leq q \leq N-1$  are the indices of the mesh points. Under these assumptions, the solution for (2) is obtained solving a generalized eigenvalue problem:

$$L\mathbf{u} = \lambda V\mathbf{u}$$

where the linear operator  $L$  is given by the sparse symmetric matrix:

$$L = -\frac{1}{h^2} \begin{bmatrix} A & I_N & 0 & \dots & 0 \\ I_N & A & I_N & \dots & 0 \\ 0 & I_N & A & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A \end{bmatrix} \in \mathbb{R}^{N^2 \times N^2}$$

$$A = \begin{bmatrix} -4 & 1 & 0 & \dots & 0 \\ 1 & -4 & 1 & \dots & 0 \\ 0 & 1 & -4 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -4 \end{bmatrix} \in \mathbb{R}^{N \times N}$$

Note that  $I_N$  is the  $N \times N$  identity matrix, the vector  $\mathbf{u} \in \mathbb{R}^{N^2}$  is obtained by row scanning the discretization grid so that:  $\mathbf{u}_{N_{p+q}} = u(x_p, y_q)$  and  $V$  is a diagonal matrix such

that  $V_{N_{p+q}} = \frac{1}{v(x_p, y_q)^2}$ . As we anticipated before,  $v(x_p, y_q)^2$  is chosen to be inversely proportional to the smoothed version of the image  $I_s(x_p, y_q)$ , which is obtained via a convolution with an isotropic Gaussian kernel with standard deviation  $\sigma$  (in our current implementation the standard deviation is set to be equal to 2.5 pixels). This is done to ensure that  $v$  satisfies the required smoothness properties. The size of the problem can be reduced by removing the entries of the vector  $\mathbf{u}$  that correspond to the points outside of the domain  $\Omega$  or to the points on the boundary. This is equivalent to remove the corresponding rows and columns in the matrix  $L$  and  $V$ : after the reduction the matrix  $L$  is no more block tridiagonal (as shown in the sparsity pattern of Fig. 2(b)) but it still is diagonally dominant. In our implementation the eigenvalues/eigenvectors are computed using the Fortran library ARPACK [18] (accessed through Matlab) that takes advantage of the sparse and symmetric structure of  $L$ . To improve the numerical stability of the algorithm we balance the matrices by scaling them, so that  $\|L\|_\infty = \|L\|_\infty = 1$  and we solve the modified sparse eigenvalue problem:

$$(V^{-\frac{1}{2}} L V^{-\frac{1}{2}}) \mathbf{w} = \mu \mathbf{w}$$

where  $\mathbf{u} = V^{-\frac{1}{2}} \mathbf{w}$  and  $\mu$  is the scaled eigenvalue. Fig. 2(b) shows an example of the matrix  $L$  associated to the region outlined by the green boundary in Fig. 2(a). Fig. 2(c) displays the third eigenmode that solves (2). The bumps on the eigenmode surface follow from the fact that the membrane density is proportional to the image intensity. The computation of the HD in the non-uniform case takes on average 1.5 s on a 2.8 GHz Pentium 4 for  $\Delta = 30$ .

#### 2.4. Comparing the descriptors

As mentioned before, it has been theoretically proven that there exist different curves that have the same spectrum. However this event is quite rare (where the notion of “rare” can be formalized more precisely, see [11]) as the experiments presented in Section 4 will confirm.

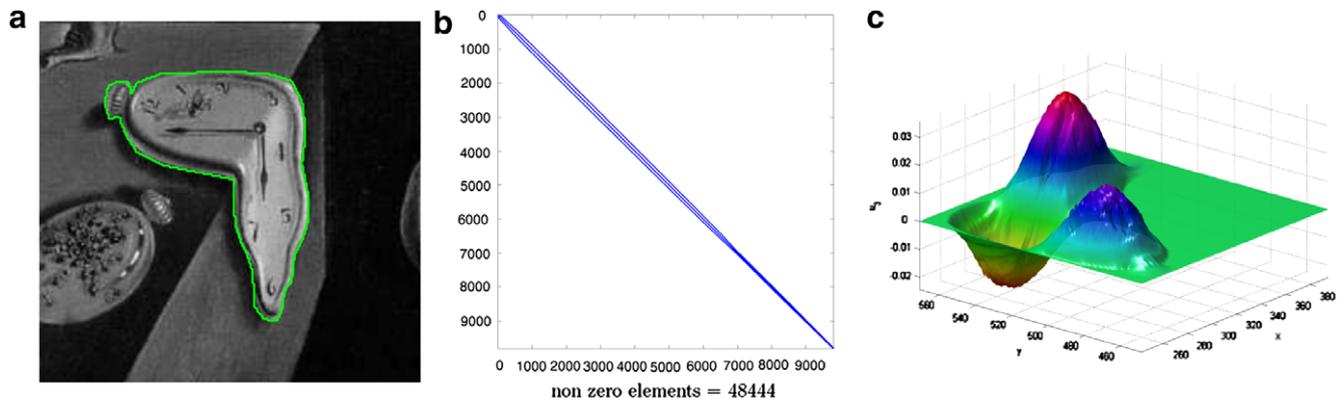


Fig. 2. (b) The sparsity pattern of the matrix  $L$ , i.e. it shows a dot for each non-zero entry of  $L$ . The matrix  $L$  is associated to the image region defined by the green boundary in (a) (cropped from the painting *Persistence of Memory* by Salvador Dali). The size of the matrix  $L$  is about  $10^4 \times 10^4$  but only  $4.8 \times 10^4$  elements are non-zero. (c) The third eigenmode of the Helmholtz Eq. (2).

Because of this, the similarity between the descriptors is defined in terms of the weighted Euclidean:

$$d(\mathbf{F}(\Gamma_1), \mathbf{F}(\Gamma_2)) = \|\mathbf{F}(\Gamma_1) - \mathbf{F}(\Gamma_2)\|_W$$

$$= \sqrt{\sum_{k=1}^N w_k [\mathbf{F}_k(\Gamma_1) - \mathbf{F}_k(\Gamma_2)]^2} \quad (7)$$

where the weights are defined according to  $w_k \stackrel{\text{def}}{=} \exp(\frac{k-1}{N-1} \log \rho)$ . The parameter  $\rho$  defines the ratio of the weight of the last component of the descriptor with respect to the first one. Experimentally we found that  $\rho = 0.75$  is a sensible choice. The rationale behind the introduction of a weighted distance is related to the physical interpretation of the components of the descriptor: the coefficients with larger indices are associated to the fast modes of vibration of the membrane. These modes are more sensitive to perturbations of the shape of the curve and therefore it is reasonable to weight them less when comparing two curve/regions (in [35] some numerical simulations confirmed that the eigenvalues with larger indices are those more affected by the morphological noise). On the other hand the smallest eigenvalues of a matrix are those more affected by the finite precision mathematical operations.

In general the task of studying analytically how the spectrum of a region is affected by the perturbations of the boundary is a complex problem. Even if this problem goes well beyond the scope of this paper, we would like to mention the approaches described in the classical book of Kato ([17], ch. 6, p. 423) and in two recent papers by Noll [25] and by Ngo [24] that attempt to relate quantitatively the perturbations of the domain boundary to the value of the eigenvalues. It is also possible to approach the problem after the Helmholtz equation has been discretized, by considering morphological perturbations that correspond to the removal of rows and columns from  $L$  and  $V$  and evaluating the bounds on the eigenvalues defined by the interlacing theorems thoroughly discussed in [15,20].

### 3. Achieving affine invariance

The descriptors we have introduced in Section 2 are RST-invariant. However very often it is necessary to

match curves or image regions in an affine invariant fashion. As an example, consider planar curves imaged from two different viewpoints using a distant camera, where distant has to be intended with respect to the camera focal length. In this case the perspective distortion can be approximated by an affine transformation (see Fig. 3 for an example). We will describe in detail a procedure that allows to map a curve (or an image region) in a normalized coordinate system where affine-related objects become congruent modulo a geometric rotation (a discussion of related approaches can be found in [1] Chapter 5, [28] and [34]). First we will consider the case where the content of the region is uniform (uniform case) and then we will generalize the results to cases where we take into consideration the intensity content (non-uniform case).

#### 3.1. Uniform case

Let's first introduce the following quantities:

- Let  $V(\Omega) \stackrel{\text{def}}{=} \int_{\Omega} d\mathbf{x}$  be the area of  $\Omega$ , where  $d\mathbf{x}$  is the infinitesimal area element.
- Let  $\mathbf{m}(\Omega) \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} \mathbf{x} d\mathbf{x}$  be the centroid of  $\Omega$ .
- Let  $\Sigma(\Omega) \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} [\mathbf{x} - \mathbf{m}(\Omega)][\mathbf{x} - \mathbf{m}(\Omega)]^T d\mathbf{x}$  be the covariance of  $\Omega$ .

We now have all the ingredients to define the shape of a Jordan curve:

**Definition 3.1.** Let  $\Gamma$  be a Jordan curve. The shape of  $\Gamma$  is a new Jordan curve such that:

$$S(\Gamma) \stackrel{\text{def}}{=} \{s \in \mathbb{R}^2 : s = \Sigma(\Omega)^{-\frac{1}{2}}[\mathbf{x} - \mathbf{m}(\Omega)] \text{ for } \mathbf{x} \in \Gamma\} \quad (8)$$

This definition is important because it allows us to relate affine-transformed curves, as stated in the following theorem:

**Theorem 3.2.** Let  $\Gamma_1$  and  $\Gamma_2$  be two Jordan curves related by an affine transformation:

$$\Gamma_2 = \{\mathbf{x}_2 \in \mathbb{R}^2 : \exists \mathbf{x}_1 \in \Gamma_1 \text{ such that } \mathbf{x}_2 = A\mathbf{x}_1 + \mathbf{b}\}$$

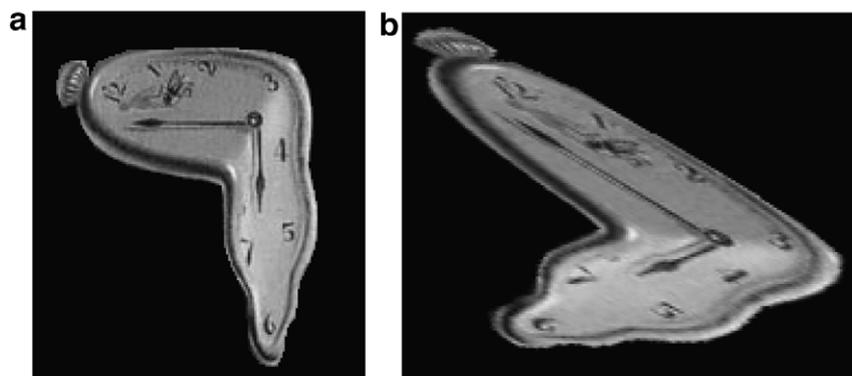


Fig. 3. An example of two image regions related by an affine transformation (cropped from the painting *Persistence of Memory* by Salvador Dalí).

where  $A \in \mathbb{R}^{2 \times 2}$  is a non-singular matrix and  $\mathbf{b} \in \mathbb{R}^2$ . Then the shapes of  $\Gamma_1$  and  $\Gamma_2$  are geometrically congruent via a two-dimensional rotation.

**Proof.** Before beginning with the proof we want to emphasize the fact that all the steps are independent from the dimension  $n$  of the space that hosts the curve. Let  $\Gamma_1 = \partial\Omega_1$  and  $\Gamma_2 = \partial\Omega_2$ . We want to show that the matrix:

$$R \stackrel{\text{def}}{=} \Sigma(\Omega_1)^{\frac{1}{2}} A^T \Sigma(\Omega_2)^{-\frac{1}{2}} \quad (9)$$

establishes the congruence relation between  $S(\Omega_1)$  and  $S(\Omega_2)$ . The first step consists in verifying that (9) is a rotation matrix. To achieve this goal we first prove the following identity:

$$\Sigma(\Omega_2) = A \Sigma(\Omega_1) A^T$$

Since the relation between the area of  $\Omega_1$  and  $\Omega_2$  is:

$$V(\Omega_2) = \int_{\Omega_2} d\mathbf{x}_2 = \int_{\Omega_1} |\det(A)| d\mathbf{x}_1 = |\det(A)| V(\Omega_1)$$

we can write:

$$\begin{aligned} \mathbf{m}(\Omega_2) &= \frac{1}{V(\Omega_2)} \int_{\Omega_2} \mathbf{x}_2 d\mathbf{x}_2 \\ &= \frac{1}{|\det(A)| V(\Omega_1)} \int_{\Omega_1} (A\mathbf{x}_1 + \mathbf{b}) |\det(A)| d\mathbf{x}_1 \\ &= A \frac{1}{V(\Omega_1)} \int_{\Omega_1} \mathbf{x}_1 d\mathbf{x}_1 + \mathbf{b} \frac{1}{V(\Omega_1)} \int_{\Omega_1} d\mathbf{x}_1 \\ &= A\mathbf{m}(\Omega) + \mathbf{b} \end{aligned}$$

and therefore:

$$\begin{aligned} \Sigma(\Omega_2) &= \frac{1}{V(\Omega_2)} \int_{\Omega_2} [\mathbf{x}_2 - \mathbf{m}(\Omega_2)][\mathbf{x}_2 - \mathbf{m}(\Omega_2)]^T d\mathbf{x}_2 \\ &= \frac{1}{|\det(A)| V(\Omega_1)} \int_{\Omega_1} A[\mathbf{x}_1 - \mathbf{m}(\Omega_1)] \\ &\quad \times [\mathbf{x}_1 - \mathbf{m}(\Omega_1)]^T A^T |\det(A)| d\mathbf{x}_1 \\ &= A \Sigma(\Omega_1) A^T \end{aligned}$$

which proves the equality. To show that (9) is indeed a rotation matrix it is enough to verify that:

$$R^T R = \Sigma(\Omega_2)^{-\frac{1}{2}} \underbrace{A \Sigma(\Omega_1)^{\frac{1}{2}} \Sigma(\Omega_1)^{\frac{1}{2}} A^T}_{\Sigma(\Omega_2)} \Sigma(\Omega_2)^{-\frac{1}{2}} = I$$

The proof is concluded observing the following two facts:

- For any  $\mathbf{s}_1 \in S(\Gamma_1)$  there exists  $\mathbf{s}_2 \in S(\Gamma_2)$  such that  $\mathbf{s}_1 = R\mathbf{s}_2$ . To prove this statement note that if  $\mathbf{s}_1 \in S(\Gamma_1)$ , then there exists  $\mathbf{x}_1 \in \Gamma_1$  such that:

$$\mathbf{s}_1 = \Sigma(\Omega_1)^{-\frac{1}{2}} [\mathbf{x}_1 - \mathbf{m}(\Omega_1)]$$

Now let  $\mathbf{x}_2 = A\mathbf{x}_1 + \mathbf{b}$  and  $\mathbf{s}_2 = \Sigma(\Omega_2)^{-\frac{1}{2}} [\mathbf{x}_2 - \mathbf{m}(\Omega_2)] \in S(\Omega_2)$ : we want to show that  $\mathbf{s}_1 = R\mathbf{s}_2$ . This follows immediately from the chain of equalities:

$$\begin{aligned} \mathbf{s}_1 &= \Sigma(\Omega_1)^{-\frac{1}{2}} [\mathbf{x}_1 - \mathbf{m}(\Omega_1)] \\ &= \Sigma(\Omega_1)^{-\frac{1}{2}} A^{-1} [\mathbf{x}_1 - \mathbf{m}(\Omega_2)] \\ &= \underbrace{\Sigma(\Omega_1)^{-\frac{1}{2}} A^{-1} \Sigma(\Omega_2)^{\frac{1}{2}}}_{R^{-T} = R} \Sigma(\Omega_2)^{-\frac{1}{2}} [\mathbf{x}_2 - \mathbf{m}(\Omega_2)] \\ &= R \Sigma(\Omega_2)^{-\frac{1}{2}} [\mathbf{x}_2 - \mathbf{m}(\Omega_2)] = R\mathbf{s}_2 \end{aligned}$$

- For any  $\mathbf{s}_2 \in S(\Gamma_2)$  there exists  $\mathbf{s}_1 \in S(\Gamma_1)$  such that  $\mathbf{s}_2 = R^{-1}\mathbf{s}_1$ . This claim can be proven similarly to what we just did before.

### 3.2. Non-uniform case

Let  $I(\mathbf{x})$  be the intensity value of a single channel image at the location  $\mathbf{x}$ ; we modify the quantities introduced in Section 3.1 as:

- Let  $V(\Omega) \stackrel{\text{def}}{=} \int_{\Omega} I(\mathbf{x}) d(\mathbf{x})$  the weighted area of  $\Omega$ , where  $d\mathbf{x}$  is the infinitesimal area element.
- Let  $\mathbf{m}(\Omega) = \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} I(\mathbf{x}) \mathbf{x} d\mathbf{x}$  be the weighted centroid of  $\Omega$ .
- Let  $\Sigma(\Omega) \stackrel{\text{def}}{=} \frac{1}{V(\Omega)} \int_{\Omega} I(\mathbf{x}) [\mathbf{x} - \mathbf{m}(\Omega)][\mathbf{x} - \mathbf{m}(\Omega)]^T d\mathbf{x}$  be the weighted covariance of  $\Omega$ .

In this case Theorem 3.3 becomes:

**Theorem 3.3.** Let  $\Gamma_1$  and  $\Gamma_2$  be two Jordan curves related by an affine transformation:

$$\Gamma_2 = \{\mathbf{x}_2 \in \mathbb{R}^2 : \exists \mathbf{x}_1 \in \Gamma_1 \text{ such that } \mathbf{x}_2 = A\mathbf{x}_1 + \mathbf{b}\}$$

where  $A \in \mathbb{R}^{2 \times 2}$  is a non-singular matrix and  $\mathbf{b} \in \mathbb{R}^2$ . Moreover suppose that the intensity pattern in  $\Omega_1$  and  $\Omega_2$  is related according to:

$$I_2(\mathbf{x}_2) = I_2(A\mathbf{x}_1 + \mathbf{b}) = I_1(\mathbf{x}_1)$$

Then the shapes of  $\Gamma_1$  and  $\Gamma_2$  are geometrically congruent via a two-dimensional rotation.

**Proof.** The proof follows exactly the same lines of the proof of Theorem 3.2, since it is straightforward to show that:

- $V(\Omega_2) = |\det(A)| V(\Omega_1)$
- $\mathbf{m}(\Omega_2) = A\mathbf{m}(\Omega_1) + \mathbf{b}$
- $\Sigma(\Omega_2) = A \Sigma(\Omega_1) A^T$

also in presence of the weighting factor related to the image intensity.  $\square$

### 3.3. Coupling the normalization procedure with the helmholtz descriptor

If we want to use the descriptors introduced in Section 2 in the context of affine-invariant matching we just have to extract the shape of a curve  $\Gamma$  and then calculate the RST-invariant descriptors of  $S(\Gamma)$  (using or not the content

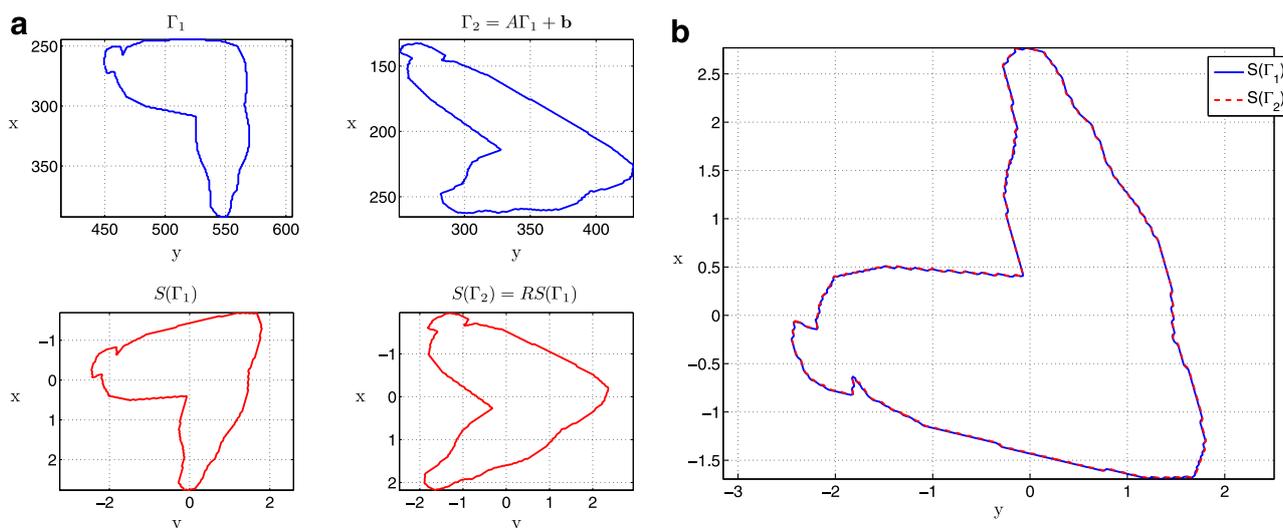


Fig. 4. The four plots on the left show the curve  $\Gamma_1$ , its affine transformation  $\Gamma_2 = A\Gamma_1 + \mathbf{b}$  and the corresponding curve shapes  $S(\Gamma_1)$  and  $S(\Gamma_2)$  in the case where the content of the curve is uniform. The right plot illustrates the congruency between  $S(\Gamma_1)$  and  $S(\Gamma_2)$ . The displayed curves are extracted from Fig. 3(a) and (b).

of the region). This two-step approach can be applied with any RST invariant curve descriptor (such as the Zernike Moment Descriptors) (see Fig. 4 and 5).

#### 4. Experimental results

The experimental results that we will present in this section are divided in two groups. First we will test the performance of the Helmholtz descriptor using a semi-synthetic dataset and then we will use the proposed descriptor to establish matches between images of natural scenes.

##### 4.1. Performance evaluation on a semi-synthetic dataset

The dataset for the experiments described in this section has been extracted from the Amsterdam Library of Object

Images (ALOI, see [10]). We considered 250 frontal images of different objects and for each view we synthetically generated 9 other images by applying an homographic transformation that simulates a change in the position of the camera. Each homography is generated following the procedure that is explained pictorially in Fig. 6(a). The rectangle  $ABCD$ , which represents the boundary of the original image, is transformed into the rectangle  $A'B'C'D'$ , which represents the boundary of the new image. The transformation is parameterized by a single positive scalar  $\alpha$  such that  $A'O = (1 + \alpha)AO$ ,  $B'O = (1 + \alpha)BO$  and  $C'O = (1 - \alpha)CO$ ,  $D'O = (1 - \alpha)DO$ . The images are generated for values of  $\alpha$  uniformly distributed in the interval  $[0.65, 1.35]$ . Further, the image is rotated by a random angle in  $[-\pi, \pi]$ . Fig. 6(b) shows an example of the images generated via this procedure. The objects are segmented using the

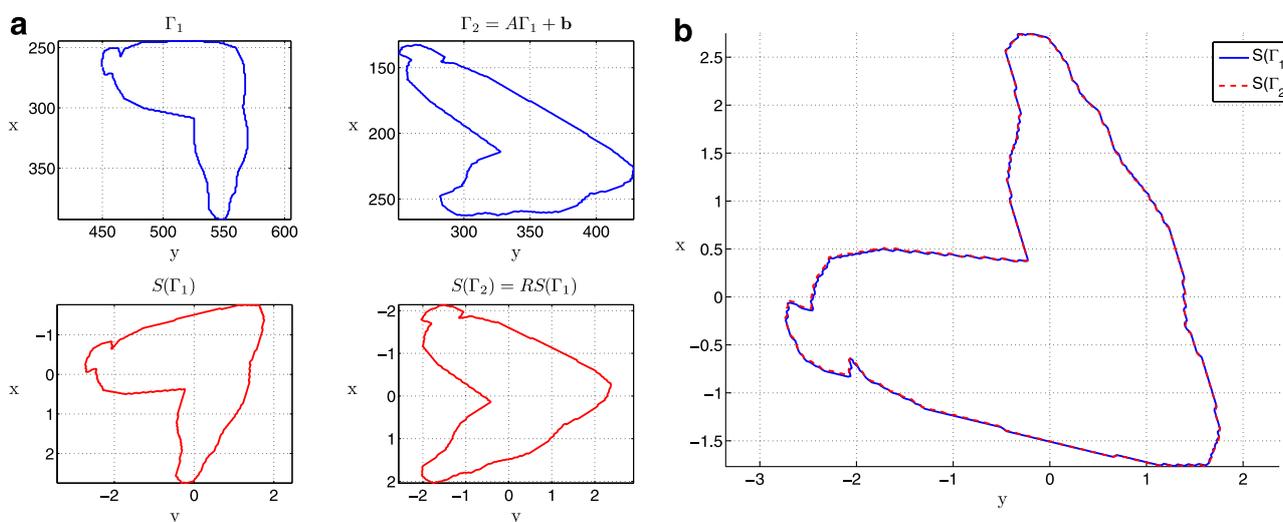


Fig. 5. The four plots on the left show the curve  $\Gamma_1$ , its affine transformation  $\Gamma_2 = A\Gamma_1 + \mathbf{b}$  and the corresponding curve shapes  $S(\Gamma_1)$  and  $S(\Gamma_2)$  in the case where the content of the curve is uniform. The right plot illustrates the congruency between  $S(\Gamma_1)$  and  $S(\Gamma_2)$ . The displayed curves are extracted from Fig. 3(a) and (b).

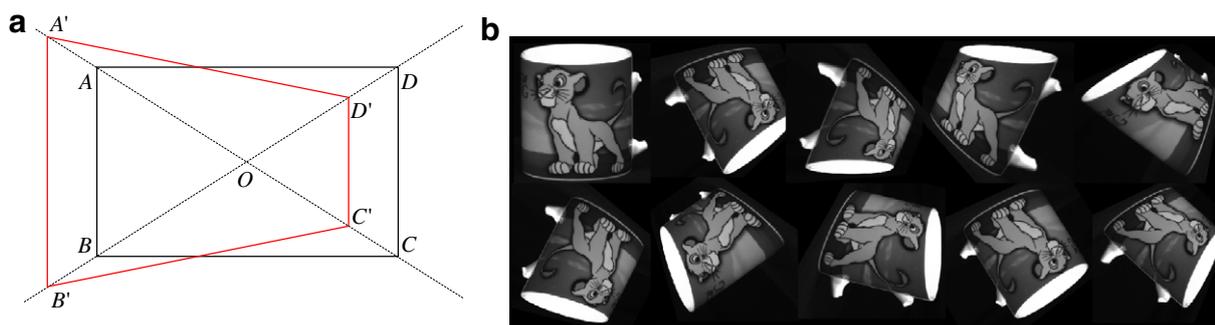


Fig. 6. (a) The random homographies are generated. (b) A set of images synthesized using the homographies generated using the method described in (a) plus an arbitrary rotation.

masking information included in the original ALOI dataset.

In the experiments described in this section, we will compare<sup>3</sup> the performance of the Helmholtz descriptor versus the Zernike Moment Descriptor, which has been shown to perform very well in the context of shape matching and retrieval [31], [33]. Experimental comparisons of the uniform HD versus the Curvature Scale Space Descriptor can be found in [35]. The performance of the descriptors is evaluated using the *precision-recall curve* calculated (over the dataset described previously) as follows. Each curve  $\Gamma$  (or region  $\Omega$ ) is used in turn as the query. Let  $A(\Gamma, N_r)$  denote the set of  $N_r$  retrievals (based on the smallest distances (7) from  $\Gamma$  in the descriptor space) and  $R(\Gamma)$  the set of 10 images in the dataset relevant to  $\Gamma$ . The *precision* is defined as:

$$P(\Gamma, N_r) \stackrel{\text{def}}{=} \frac{|A(\Gamma, N_r) \cap R(\Gamma)|}{N_r}$$

and measures the proportion of items retrieved that are relevant. Similarly, the *recall* is defined as:

$$C(\Gamma, N_r) \stackrel{\text{def}}{=} \frac{|A(\Gamma, N_r) \cap R(\Gamma)|}{10}$$

and measures the proportion of relevant items that are retrieved. Note that the same quantities can be defined in the case where the query is a region  $\Omega$ . The notation  $|\cdot|$  denotes cardinality. The precision recall curve is plotted by averaging precision and recall over all  $\Gamma$ , for different values of  $N_r$ . On the plots, each marker corresponds to a different value for  $N_r$ , ranging from 1 to 20. Moreover dashed red lines refer to the ZMD, whereas continuous blue lines refer to the HD.

Fig. 7(a) compares the performance of the descriptors after the curves/regions have been normalized using the uniform or non-uniform normalization. Both the descriptors have 36 components and are uniformly quantized using 8 bits. For the HD the parameters are  $\Delta = 30$ ,

$\sigma = 2.5$  and  $\rho = 0.75$ . Both for the ZMD and for the HD the performance is better if the regions are normalized using the non-uniform procedure described in Section 3.2. This can be simply explained observing that the dataset contains several objects that have a similar shape but a different and distinctive image content. The non-uniform HD seems to be less affected by the type of normalization used. This can be understood observing that the descriptor combines the information of the shape with the information of the content of the considered region. Fig. 7(b) compares the performance of the uniform HD vs. the non-uniform HD. The parameters of the descriptors are the same as in the previous experiment. The precision recall curves confirm the intuition that the non-uniform Helmholtz descriptor captures the intensity information contained inside the region and that this has a beneficial impact on the overall performance of the approach. Quite surprisingly the performance of the non-uniform HD is essentially equivalent to the performance of the ZMD. We believe that this is due to the fact that the numerical scheme used to solve the Helmholtz equation can be refined and improved. We will elaborate more on this claim at the end of this section. Fig. 8(a) shows the behavior of the non-uniform HD for different resolutions of the discretization mesh parameterized by  $A$  (see Eq. (6)). As before, the remaining parameters for the HD are  $\sigma = 2.5$  and  $\rho = 0.75$  with the descriptors coefficients uniformly quantized using 8 bits. As it was pointed out in [35], the results indicate that the descriptor is reasonably stable for values of  $\Delta \geq 30$ . The experiment illustrated in Fig. 8(b) compares the performance of the ZMD versus the non-uniform HD for different lengths of the descriptor. For both of them the performance fluctuations are rather limited. However we can observe a drop in performance for the HD for  $N_\lambda = 24$ . This might indicate that the components of the Helmholtz descriptor with larger indices bring more information than the corresponding ones for the ZMD. Fig. 9(a) displays the precision recall curves for the ZMD and for the non-uniform HD while varying the number of bits used to quantize the descriptor components. The ZMD presents larger fluctuations than the HD: we hypothesize that this behavior is related to the fact that the coefficients of the Zernike descriptors cover a larger dynamic range than

<sup>3</sup> All the code for the normalization, for the computation of the Helmholtz descriptors and for the computation of the Zernike Moment Descriptors can be downloaded from <http://vision.ece.ucsb.edu/~zuliani/Code/Code.html>.

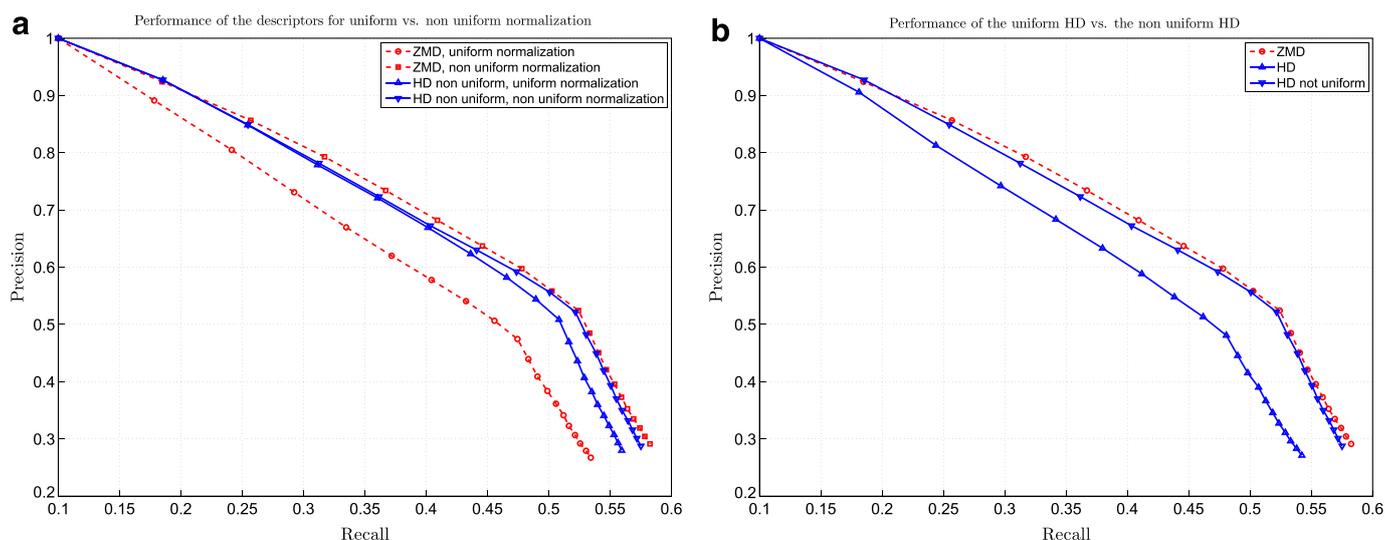


Fig. 7. (a) Compares the performance of the descriptor in the presence of uniform or non-uniform normalization. (b) Compares the performance of the uniform HD vs. the non-uniform HD. In both experiments the descriptors have 36 components and are quantized using 8 bits. For the HD the parameters are  $\Delta = 30$ ,  $\sigma = 2.5$  and  $\rho = 0.75$ .

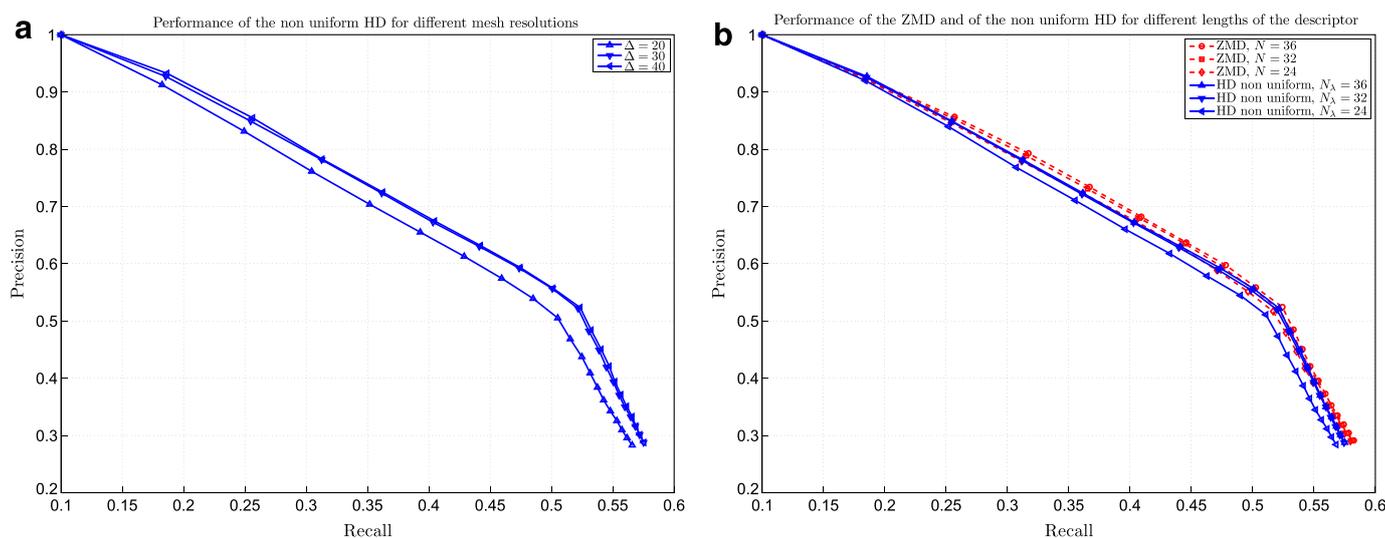


Fig. 8. (a) The behavior of the non-uniform HD for different resolutions of the discretization mesh parameterized by  $\Delta$ . (b) Compares the performance of the ZMD versus the non-uniform HD for different lengths of the descriptor. The parameters for the HD are  $\sigma = 2.5$  and  $\rho = 0.75$ .

the ratios of the eigenvalues of the Laplacian and hence they are more affected by quantization issues.

From these experiments we conclude that the non-uniform Helmholtz descriptor provides a performance (measured in terms of precision recall) that is comparable to the performance obtained by the Zernike Moment Descriptor, which can be considered one of the state of the art descriptors for curve/region description, matching and retrieval [31]. We believe that the numerical method used to compute the solutions of the Helmholtz equation can be greatly improved with an immediate impact on the performances of the HD. This opinion is mainly supported by the observation that the HD provides a joint description of shape and content. The normalization procedure, that can compensate the distortion introduced by an homographic transformation only up to a first order of approximation

is also a critical step in the overall procedure. A visual inspection of the eigenmodes of the normalized regions indicates that perturbations due to a non-satisfactory normalization may produce completely different Helmholtz descriptors. Fig. 9(b) compares the ZMD and the non-uniform HD when the dataset is generated using an affine distortion model for the images (and therefore the normalization procedure carries out its task with no approximations): as expected the results obtained using the non-uniform Helmholtz descriptors are superior to those obtained using the Zernike descriptors. Regarding the problem of isospectrality introduced in Section 2, a manual inspection of a sample set of mismatched curves/regions seems to confirm the intuition that with real imagery the generation of identical spectrums from different curves/regions is an unlikely event. As a concluding remark we

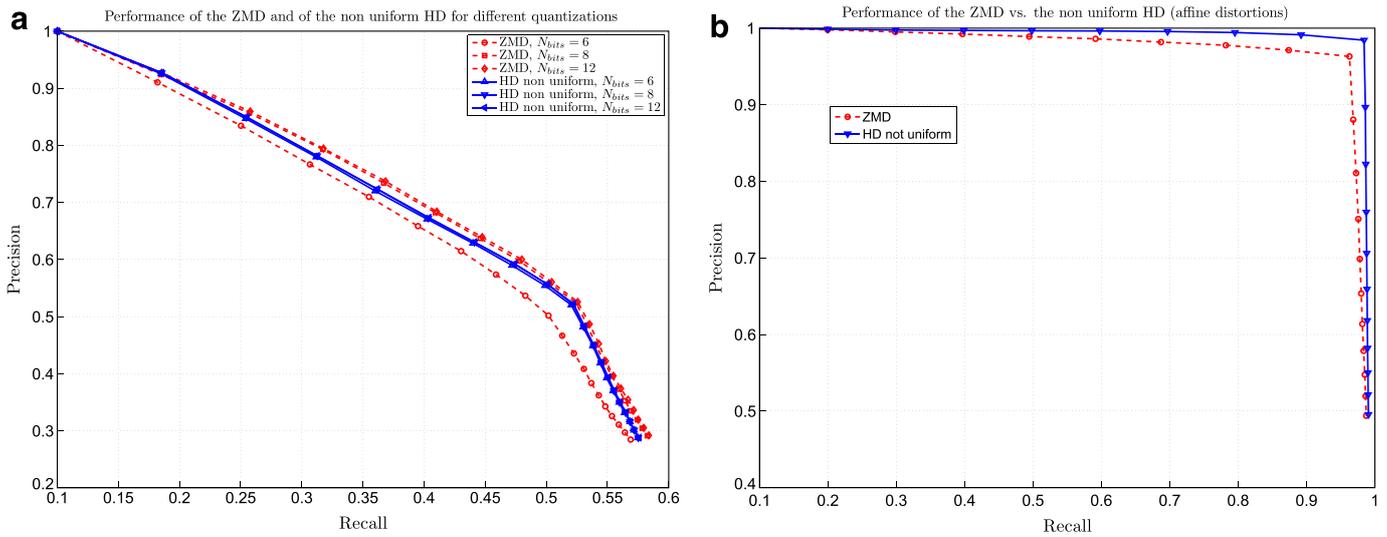


Fig. 9. (a) The precision recall curves for the ZMD and for the non-uniform HD while varying the number of bits used to quantize the descriptor components. The parameters for the HD are  $\sigma = 2.5$  and  $\rho = 0.75$ . (b) Compares the ZMD and the non-uniform HD when the dataset is generated using an affine distortion model for the images.

want to point out that there exists a tradeoff between the distinctiveness and the robustness provided by the descriptors. In the context of image registration, where we would like to establish matches between images, we would like to avoid situations where the distance between the descriptors of different (but visually similar) curves is small enough to produce a mismatch. On the other hand, a certain degree of robustness is needed to compensate for the approximations introduced by the normalization procedure. Both these two aspects will be experimentally explored in greater detail in the next section.

#### 4.2. Performance evaluation on real images

The performance of the descriptors has been tested on a set of pairs of real images, representing outdoor and indoor scenes. Every image pair displays the same scene acquired

under different points of view. The first stage of the processing consists in finding in each image a set of candidate curves for matching. To accomplish this task we used the level set decomposition of the intensity values of the image, which are known to enjoy several important invariance properties [19]. We observed that for our purposes a good strategy is to slice the intensity profile at two levels, one large and one small. This way it is possible to identify dark regions as well as bright regions. Moreover the size of the extracted regions is large enough so that the intensity content is non-uniform. As pointed out in the previous section, this is crucial for the generation of distinctive descriptors in presence of curves that have very similar shapes. After the Helmholtz descriptors are calculated for each curve/region, the matching is performed using the weighted Euclidean distance (7) (with  $\rho = 0.75$ ). In Fig. 10–13 we show the results of the curve matching procedure. Fig. 10(a) and

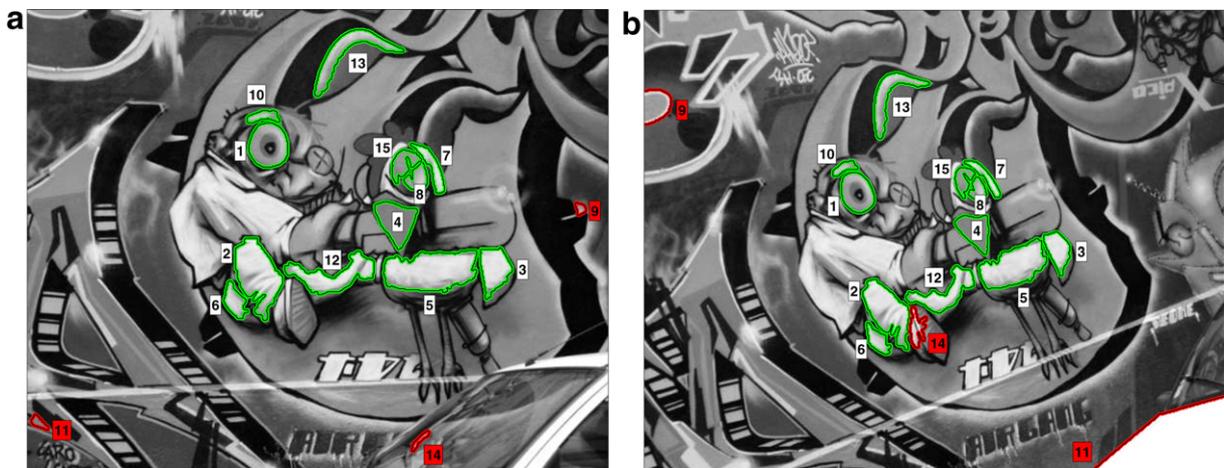


Fig. 10. Results of the matching procedure for the Graffiti scene. In all the examples the HD descriptor is composed by 32 components quantized using 8 bits and  $\Delta = 30$ . The numbers with white background (green region boundaries) identify curve/regions correctly matched, while those with red background (red region boundaries) correspond to mismatches.

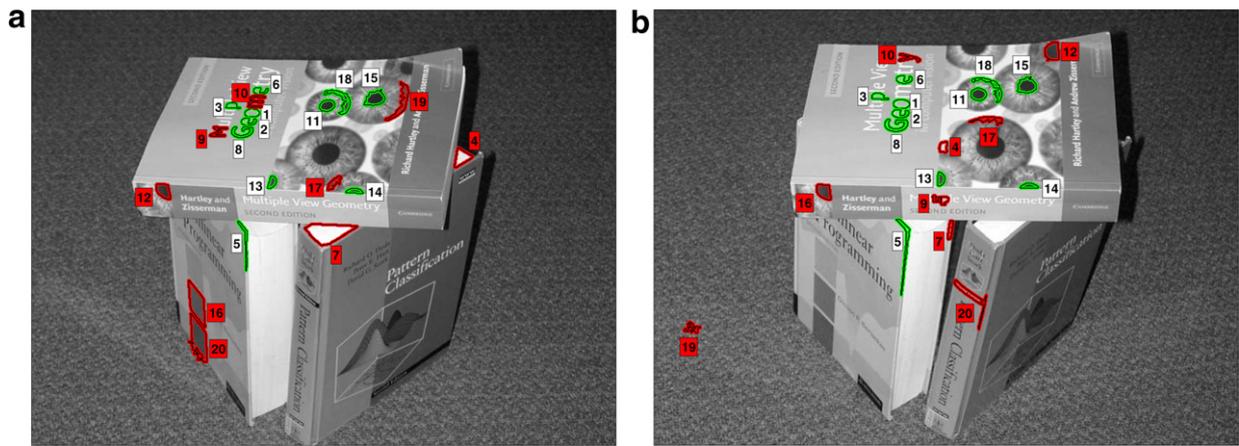


Fig. 11. Results of the matching procedure for the Books scene. See the caption of Fig. 10 for the experimental conditions and the typographical conventions.

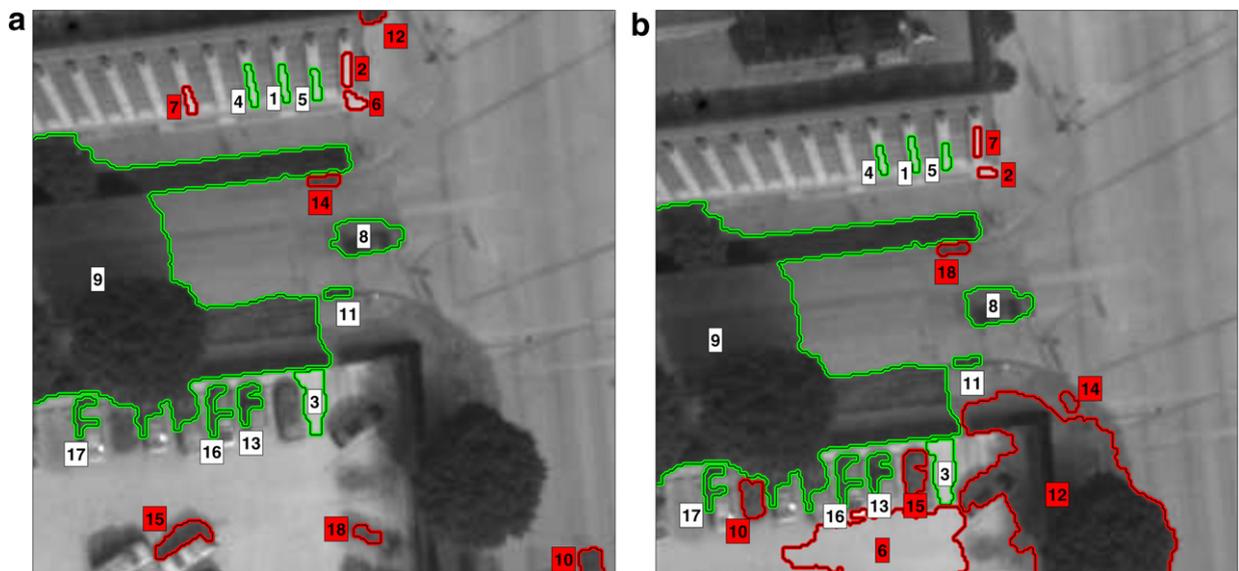


Fig. 12. Results of the matching procedure for the LA street scene. See the caption of Fig. 10 for the experimental conditions and the typographical conventions.

(b) show that the matching using the HD is facilitated if the detected regions are large enough. More specifically, Region 1 has a round shape that is present elsewhere in the image (e.g. the other eye of the puppet or Region 9 in image (b)). However its distinctive intensity content is likely to be captured by the Helmholtz descriptor. Similar considerations can be extended to the Regions 1, 11 and 15 displayed in Fig. 11. In particular Region 1 includes the edges of the inside border of the letter “o”, making such a region distinguishable from the remaining two. In Fig. 12, Regions 1, 4 and 5 have similar shapes but we may still argue that the information contained in the intensity pattern can be relevant to the final matching result. Similar considerations hold for the harbor scene in Fig. 13. As a final remark, we would like to suggest how this approach could be used to bootstrap the estimation of the mappings (such as homographies or fundamental

matrices) that relate the geometry of 3D scenes acquired from points of view separated by a wide base line.

### 5. Conclusions

In this paper we presented a curve/region descriptor that is based on the solution of the Helmholtz equation. This descriptor has a strong physical characterization, since it is related to the modes of vibration of a membrane shaped as the considered region and with a density that is proportional to the region intensity. Together with the descriptor we presented a normalization procedure that is capable of extracting the shape of a curve/region. More precisely, curves (or image regions) are mapped to a normalized coordinate system where affine-related objects become congruent modulo a geometric rotation. The performance of the descriptors has been tested both on a semi-synthetic

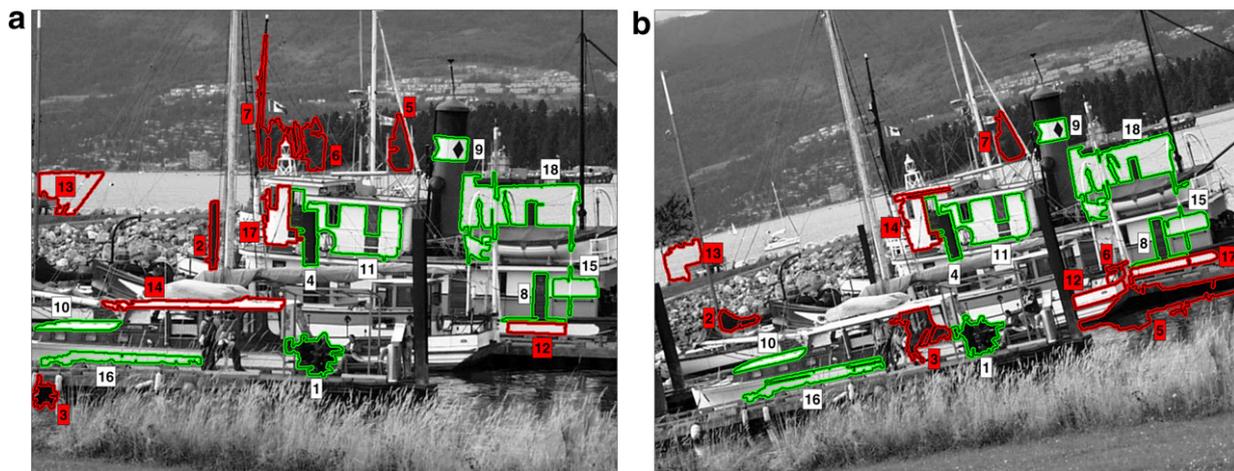


Fig. 13. Results of the matching procedure for the Harbor scene. See the caption of Fig. 10 for the experimental conditions and the typographical conventions.

dataset and on real images and it has been compared with one of the state of the art descriptors, the Zernike moment descriptor. The results of the experiments show that the HD performs well in the context of similarity based curve/region retrieval and curve/region matching. Moreover the normalization procedure proved to be an important tool to compensate for the geometric distortions present in images acquired from different points of view. Both the descriptor and the normalization procedure combine intimately and in an original way the information regarding the shape of the object with the information carried by its visual appearance.

The initial studies that have been presented in this paper open a number of interesting research perspectives. First of all the calculation of the descriptor would greatly benefit from advanced numerical methods [5,9,13,14] that could solve the Helmholtz equation with an higher degree of accuracy and possibly faster (it is known that finite difference schemes may introduce spurious modes and that there exists a dependence between the grid resolution and the largest index of the eigenpair that can be computed). We also believe in the importance of quantitatively characterizing the influence that the perturbations on the boundaries of the curves have on the coefficients of the descriptors or equivalently the sensitivity of the HD with respect to morphological perturbations. Another interesting research perspective consists in understanding the semantics of the descriptors, i.e. how and they relate to the visual properties of the curve/regions [6]. We would also like to emphasize the fact that the theory that supports both the normalization procedure and the calculation of the modes of vibration of a membrane is independent from the dimensionality of the considered objects and could be generalized to deal with regions extracted from three dimensional imagery (such as CAT images). Finally we are interested in region detectors that are able to identify image portions that have a rich intensity content and that present a high degree of repeatability in presence of perspective

distortions. We believe that the curves obtained starting from the level set decomposition of the intensity surface of an image [19] could be a good input for the Helmholtz descriptor.

#### Acknowledgments

The authors thank Dr. Sitaram Bhagavathy for his collaboration during the first stages of this work and Dr. Yosi Keller for some interesting ideas. This work was supported by the Office of Naval Research (Grant #N00014-04-1-0121).

#### Appendix A

For a more thorough treatment of the results presented in this appendix we refer the reader to the texts by Weinberger [30], Carrier et al. [4] and Evans et al. [7].

##### A.1. Some analytical properties of the helmholtz equation

Consider the boundary problem with Dirichlet conditions (2) rewritten here for sake of convenience:

$$-\Delta u(\mathbf{x}) = \lambda \frac{1}{v(\mathbf{x})^2} u(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

$$u(\mathbf{x}) = 0 \quad \text{for } \mathbf{x} \in \Gamma$$

One trivial solution for this problem is  $u(\mathbf{x}) \equiv 0$  everywhere. We are interested in the non-trivial solutions of this problem, which can be proven to exist for a wide class of regions  $\Omega$  for discrete values of the parameter  $\lambda$  and for  $v > 0$ . Such solution are in the form of a countable set of pairs of eigenvalues/eigenfunction, i.e.  $(\lambda_k, u_k)$ . The next lemmas will show that such eigenpairs have the following properties:

- the eigenfunctions are orthogonal,
- the eigenvalues are real and the eigenfunctions can be chosen to be real,
- the eigenvalues are positive.

**Lemma A.1.** Let  $(\lambda_k, u_k)$  and  $(\lambda_l, u_l)$  be two eigenpairs that solve Eq. (2). Then:

$$\langle u_k | u_l \rangle_{\frac{1}{v^2}} \propto \delta_{kl}$$

where  $\delta_{kl}$  is the Kronecker triangle and the notation  $\langle \cdot | \cdot \rangle$  denotes the weighted inner product on  $\Omega$ :  $\langle f | g \rangle_w = \int_{\Omega} w(x) f(x) g(x) dx$ .

**Proof.** Consider the pair of equations:

$$-\Delta u_k = \lambda_k \frac{1}{v^2} u_k \tag{A-2a}$$

$$-\Delta u_l = \lambda_l \frac{1}{v^2} u_l \tag{A-2b}$$

and multiply both sides of (A-2a) by  $u_l$  and of (A-2b) by  $u_k$ . If we subtract both members and we integrate over  $\Omega$  we obtain:

$$\int_{\Omega} (u_l \Delta u_k - u_k \Delta u_l) dx = (\lambda_l - \lambda_k) \int_{\Omega} \frac{1}{v^2} u_k u_l dx. \tag{A-3}$$

The left-hand side can be rewritten using Green's second identity as:

$$\int_{\Omega} (u_l \Delta u_k - u_k \Delta u_l) dx = \int_{\Gamma} \left( u_l \frac{\partial u_k}{\partial \mathbf{n}} - u_k \frac{\partial u_l}{\partial \mathbf{n}} \right) dx$$

where  $\mathbf{n}$  denotes the normal at the boundary. Since  $u_k$  and  $u_l$  are identically equal to zero on the boundary, the left hand side of (A-3) must vanish. Consequently, since  $\lambda_l \neq \lambda_k$ , the proof is concluded observing that the right-hand side of (A-3) yields:

$$\int_{\Omega} \frac{1}{v^2} u_k u_l dx = \langle u_k | u_l \rangle_{\frac{1}{v^2}} = 0 \quad \square$$

**Lemma A.2.** The eigenvalues that satisfy the Eq. (2) are real and the corresponding eigenfunctions can be chosen to be real.

**Proof.** We will proof that the eigenvalues are real by contradiction. Let  $\lambda \in \mathbb{C}$  and let  $u$  be the correspondent eigenfunction (not identically equal to zero) that solves (2). It is straightforward to verify that the complex conjugates of the eigenpairs will also satisfy (2). Hence, letting  $(\lambda_k, u_k) = (\lambda, u)$  and  $(\lambda_l, u_l) = (\lambda^*, u^*)$  and following the same steps of the proof of Lemma A.1 we conclude that:

$$\int_{\Omega} \frac{1}{v^2} u u^* dx = \int_{\Omega} \frac{1}{v^2} |u|^2 dx = 0 \tag{A-4}$$

It follows immediately that (A-4) is satisfied only if  $u \equiv 0$ , which contradicts the hypothesis. Hence  $\lambda \in \mathbb{R}$ . Now suppose there exists a complex eigenfunction corresponding to  $\lambda$ :  $u = v + jw$ . Clearly both  $u$  and  $w$  satisfy (2), hence we can always choose a real eigenfunction.  $\square$

**Lemma A.3.** The eigenvalues that satisfy the Eq. (2) are positive.

**Proof.** If we multiply both members of equation  $\Delta u_k = -\lambda_k \frac{1}{v^2} u_k$  by  $u_k$  and we integrate over the region  $\Omega$  we obtain:

$$\int_{\Omega} u_k \Delta u_k dx = -\lambda_k \int_{\Omega} \frac{1}{v^2} u_k^2 dx$$

Applying Green's first identity to the left hand side we obtain:

$$\int_{\Omega} u_k \Delta u_k dx = \int_{\Gamma} u_k \frac{\partial u_k}{\partial \mathbf{n}} dx - \int_{\Omega} \|\nabla u_k\|^2 dx$$

Since  $u_k$  is identically equal to zero on the boundary we can write:

$$\lambda_k \int_{\Omega} \frac{1}{v^2} u_k^2 dx = \int_{\Omega} \|\nabla u_k\|^2 dx$$

The proof is concluded observing that  $\lambda_k$  can be expressed in terms of the ratio of two positive quantities.  $\square$

We conclude this appendix listing a few other important facts.

- The eigenvalues can be sorted in order of increasing value:  $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$  with  $\lambda_k \rightarrow \infty$  as  $k \rightarrow \infty$ .
- For a given eigenvalue  $\lambda_k$  there is a finite number of linearly independent eigenfunctions (such number is called the multiplicity of  $\lambda_k$ ).
- The first (or principal eigenvalue) has multiplicity 1 and does not change sign over  $\Omega$ .
- The normalized real eigenfunctions  $u_k$  form an orthonormal basis of  $L^2(\Omega)$ , where the normalization is such that  $\int_{\Omega} \frac{1}{v^2} u_k dx = 1$ .
- If the region  $\Omega$  is not bounded it may happen that the set of eigenpairs is no longer discrete.

## References

- [1] K. Åström. Invariancy Methods for Points, Curves and Surfaces in Computational Vision. PhD thesis, Department of mathematics, Lund university, Sweden, 1996.
- [2] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522.
- [3] P. Buser, J. Conway, P. Doyle, K. Semmler, Some planar isospectral domains, International Mathematics Research Notices 9 (1994) 391–400.
- [4] G.F. Carrier, C.E. Pearson, Partial Differential Equations, Theory and Technique, second ed., Academic Press, 1988.
- [5] T.A. Driscoll, Eigenmodes of isospectral drums, SIAM Review 39 (1) (1997) 1–17.
- [6] H. Eidenberger, Statistical analysis of MPEG-7 image descriptions, ACM Multimedia Systems Journal 10 (2) (2004) 84–97.
- [7] L.C. Evans, Partial Differential Equations (Graduate Studies in Mathematics, 19), American Mathematical Society (1998).
- [8] J. Flusser, T. Suk, Pattern recognition by affine moment invariants, Pattern Recognition 26 (1) (1993) 167–174.
- [9] L. Fox, P. Henrici, C. Moler, Approximations and bounds for eigenvalues of elliptic operators, SIAM Journal of Numerical Analysis 4 (1967) 89–102.
- [10] J.M. Geusebroek, G.J. Burghouts, A.W.M. Smeulders, The Amsterdam library of object images, International Journal of Computer Vision 61 (1) (2005) 103–112.
- [11] C. Gordon, D.L. Webb, S. Wolpert, One cannot hear the shape of a drum, Bulletin of the American Mathematical Society 27 (1) (1992) 134–138.

- [12] L. Gorelick, M. Galun, E. Sharon, R. Basri, A. Brandt, Shape representation and classification using the Poisson equation, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, July 2004, pp. 61–67.
- [13] P. Guidotti, J.V. Lambers, Eigenvalue characterization and computation for the laplacian on general domains. Submitted, October 2005.
- [14] K. Höllig, C. Apprich, A. Streit, Introduction to the web-method and its applications, *Advances in Computational Mathematics* 23 (2005) 215–237.
- [15] R. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1999.
- [16] M. Kac, Can one hear the shape of a drum? *American Mathematical Monthly* 73 (2) (1966) 1–23.
- [17] T. Kato, *Perturbation Theory for Linear Operators*, Springer, 1995.
- [18] R.B. Lehoucq, D.C. Sorensen, C. Yang, *ARPACK Users' Guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*, SIAM (1998).
- [19] J.-L. Lisani, L. Moisan, P. Monasse, J.-M. Morel, On the theory of planar shape, *SIAM Journal on Multiscale Modeling and Simulation* 1 (1) (2003) 1–24.
- [20] C.D. Meyer, *Matrix analysis and applied linear algebra*, SIAM (2001).
- [21] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10) (2005) 1615–1630.
- [22] F. Mokhtarian, M. Bober, *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*, Kluwer Academic Publishers., 2003.
- [23] B.S. Morse. *Computation of Object Cores from Grey-level Images*. PhD thesis, University of North Carolina at Chapel Hill, 1994.
- [24] K.V. Ngo, An approach of eigenvalue perturbation theory, *Applied Numerical Analysis and Computational Mathematics* 2 (1) (2005) 108–125.
- [25] A. Noll, Domain perturbations, capacity and shift of eigenvalues, *Journées équations aux dérivées partielles* (1999) 1–10.
- [26] N. Saito, Geometric harmonics as a statistical image processing tool for images defined on irregularly-shaped domains, in: *Proceedings of IEEE Statistical Signal Processing Workshop*, Boreadux, France, July 2005.
- [27] K. Siddiqi, A. Shokoufandeh, S.J. Dickinson, S.W. Zucker, Shock graphs and shape matching, *International Journal of Computer Vision* 35 (1) (1999) 13–32.
- [28] D. Sinclair, A. Blake, Isoperimetric normalization of planar curves, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (8) (1994) 769–777.
- [29] M.R. Teague, Image analysis via the general theory of moments, *Journal of the Optical Society of America* 70 (8) (1979) 920–930.
- [30] H.F. Weinberger, *A First Course in Partial Differential Equations with Complex Variables and Transform Methods*, first ed., Blaisdell Publishing Company, 1965.
- [31] D. Zhang, G. Lu, Evaluation of MPEG-7 shape descriptors against other shape descriptors, *ACM Journal of Multimedia Systems* 9 (1) (2003) 15–30.
- [32] D.S. Zhang, G. Lu. Generic Fourier descriptors for shape-based image retrieval, in: *Proceedings of IEEE International Conference on Multimedia and Expo*, vol. 1, Lausanne, Switzerland, August 2002, pp. 425–428.
- [33] D.S. Zhang, G. Lu, Review of shape representation and description techniques, *Pattern Recognition* 37 (1) (2004) 1–19.
- [34] M. Zuliani, S. Bhagavathy, C.S. Kenney, B.S. Manjunath, Affine-invariant curve matching, in: *IEEE International Conference on Image Processing*, October 2004.
- [35] M. Zuliani, C.S. Kenney, S. Bhagavathy, B.S. Manjunath, Drums and curve descriptors, in: *British Machine Vision Conference*, September 2004.