Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Level-Set Segmentation with Contour based Object Representation

Daniel Weiler, Julian Eggert

2009

Preprint:

This is an accepted article published in Proceedings of the 2009 International Joint Conference on Neural Networks (IJCNN). The final authenticated version is available online at: https://doi.org/[DOI not available]

Weiler, D., Röhrbein, F., Eggert, J. (2009). Level-Set Segmentation with Contour based Object Representation, *Proceedings of International Joint Conference on Neural Networks* (IJCNN), 3327-3334.

Honda Research Institute Europe GmbH
http://www.honda-ri.de/
Carl-Legien-Strasse 30
63073 Offenbach am Main
Germany

Level-Set Segmentation with Contour based Object Representation

Daniel Weiler, Florian Roehrbein, Julian Eggert

***Abstract*— In this paper we present an approach for contour based object representation. To this end we use a curvature signal gained by a level-set segmentation method. The advantage of that curvature signal is that it generates no computational overhead as it is a byproduct of standard level-set segmentation methods. Different methods for the description of the segmented objects, so called object descriptors are presented. The object descriptors are all invariant against translation, rotation and scale of the object. Furthermore we show a sparse and memory efficient representation of the descriptors for a series of objects. Finally an approach for classification of unknown objects based on "memorized" objects is proposed.**

I. Introduction

In the field of image segmentation, two major approaches can be distinguished: *multi region segmentation* and *figure-background segregation*. While the former tries to group similar (by their image features) and related (by their spatial properties like location, etc.) pixels of an image into separate regions, the latter attempts to find a few salient regions of an image considering them as a foreground "figure", labeling all the reminder without any further differentiation as background. In this paper we address the problem of figure-background segregation based on level-set methods, with a special focus on a contour based representation of the segmented objects.

The segmentation occurs by means of level-set methods [1], which separate all image pixels into two disjoint regions by favoring homogeneous image properties for pixels within the same region and dissimilar image properties for pixels belonging to different regions. The level-set formalism describes the region properties using an energy functional that implicitly contains the region description and that has to be minimized. The formulation of the energy functional dates back to e.g. Mumford and Shah [2] and to Zhu and Yuille [3]. Later on, the functionals were reformulated and minimized using the level-set framework e.g. by [4] and [5]. In recent years level-set methods became a powerful tool for image segmentation. State-of-the-art level-set methods are able to work on arbitrary feature maps [6]. These feature maps may incorporate the three color components of an image but might be extended by any other characteristic property of a region (e.g. texture and motion). Most level-set methods assume the feature maps to be independent and commonly utilize a feature vector composed of three color and three texture components to perform the segmentation [7]. In [8] an approach was introduced that yields competitive results by employing only the three color components but considering these image features as located in a common, multi-dimensional feature space comparable to the probabilistic color distributions modeled by means of Gaussian Mixture Models in state-of-the-art figure-background segregation algorithms [9], [10], [11].

Among all segmentation algorithms from computer vision, level-set methods provide perhaps the closest link with the biologically motivated, connectionist models as represented e.g. by [12]. Similar to neural models, level-set methods work on a grid of nodes located in image/retinotopic space, interpreting the grid as having local connectivity, and using local rules for the propagation of activity in the grid. Time is included explicitly into the model by a formulation of the dynamics of the nodes activity. Furthermore, the external influence from other sources (larger network effects, feedback from other areas, inclusion of prior knowledge) can be readily integrated on a node-per-node basis, which makes level-sets appealing for the integration into biologically motivated system frameworks.

In addition to color and texture, shape is a fundamental feature of an object. In recent years a large number of shape descriptors were introduced. In [13] a comprehensive shape review is given. Generally, two types of shape representation can be distinguished: contour-based and region-based methods. While the former exploit only shape boundary features (e.g. curvature and perimeter), the latter consider all pixels of an object (e.g. area, compactness and geometric moments). In this paper we address the former contour-based methods.

In [14] the *population coding of shape in area V4* of macaque monkeys is studied. Based on the tuning of many macaque V4 neurons for the curvature of an objects contour, it is shown that populations of these neurons represent entire shapes. To this end, the responses of 109 neurons in area V4 were evaluated. The neurons were stimulated by 49 basic shapes. The population responses were arranged on a two-dimensional grid: angular postion (abscissa) times boundary curvature (ordinate). Based on this population code, the curvature signal, i.e. the shape of the original object, that was presented to the neurons could be retrieved. The "curvature" activity maps shown in [14] are similar to the results we show in this paper in Sect. IV-B in Fig. 5. Based on and motivated by this similarity of level-set segmentation curvature signals to the primate visual cortex we evaluate in this paper curvature based object descriptors and sparse and memory efficient representation of these descriptors.

The paper is organized as follows. In Sect. II we describe the level-set method applied for image segmentation and the approaches to extract a curvature signal from the level-set

Daniel Weiler is with Darmstadt University of Technology, Institute of Automatic Control, Control Theory and Robotics Lab, Darmstadt D-64283, Germany (phone: +49-6151-16-6066; fax: +49-6151-16-2507; email: Daniel.Weiler@rtr.tu-darmstadt.de).

Florian Roehrbein and Julian Eggert are with Honda Research Institute (Europe) GmbH, Offenbach D-63073, Germany (phone: +49-69-89011-750; fax: +49-69-89011-749; email: Julian.Eggert@honda-ri.de).

978-1-4244-3553-1/09/$25.00 ©2009 IEEE

function. Section III introduces the proposed method for translation, scale and rotation invariant object description based on a curvature signal. The results of the proposed algorithms are presented in Sect. IV. A short discussion finalizes the paper.

II. Level-Set Segmentation

A. Standard Level-Set based Region Segmentation

Level-set methods are front propagation methods. Starting with an initial contour, the figure-background segregation task is solved by iteratively moving the contour according to the solution of a partial differential equation (PDE). The PDE is often originated from the minimization of an energy functional. The solution to the PDE constitutes an initial value problem that is solved by a gradient descent. Depending on the initialization of the problem (i.e. on the initial contour) the gradient descent will, in cases of reliable initialization, succeed in finding the global minimum of the energy functional or, in cases with unreliable initialization, fail in doing so and be stuck in a local minimum. Famous representatives of energy functionals for image segmentation problems are those by Mumford and Shah [2] and by Zhu and Yuille [3]. While the former work in its original version on gray value images (i.e. on scalar data), utilize the mean gray value of a region as a simple region descriptor and were only later extended to vector valued data [6] (e.g. color images), the latter use more advanced probabilistic region descriptors that are based on the distributions of each feature channel inside and outside the contour. In many cases it is sufficient to model these distributions by uni-modal Gaussian distributions. In some rare cases the distributions are approximated in a multi-modal way [5] e.g. by Gaussian Mixture Models or Nonparametric Parzen Density Estimates [15]. Within a region the models of all features together add up to the region descriptor.

Similar to state-of-the-art figure-background segregation algorithms [9], [10], [11], level-set methods use a smoothness term to control the granularity of the segmentation. A common way is to penalize the length of the contour, that can be formulated in the energy functional by simply adding the length of the contour to the energy that is to be minimized. In doing so, few large objects are favored over many small objects as well as smooth object boundaries over ragged object boundaries.

Compared to "active contours" (snakes) [16], that also constitute front propagation methods and explicitly represent a contour by supporting points, level-set methods represent contours implicitly by a level-set function that is defined over the complete image plane. The contour is defined as an iso-level in the level-set function, i.e. the contour is the set of all locations, where the level-set function has a specific value. This value is commonly chosen to be zero, thus the inside and outside regions can easily be determined by the Heaviside function $H(x)$ [1].

[1] $H(x) = 1$ for $X > 0$ and $H(x) = 0$ for $X \leq 0$.

The proposed object description framework is based on a standard two-region level-set method for image segmentation [5], [8], [17]. In a level-set framework, a level-set function $\phi \in \Omega \mapsto \mathbb{R}$ is used to divide the image plane Ω into two disjoint regions, Ω_1 and Ω_2, where $\phi(x) > 0$ if $x \in \Omega_1$ and $\phi(x) < 0$ if $x \in \Omega_2$. Here we adopt the convention that Ω_1 indicates the background and Ω_2 the segmented object. A functional of the level-set function ϕ can be formulated that incorporates the following constraints:

- Segmentation constraint: the data within each region Ω_i should be as similar as possible to the corresponding region descriptor ρ_i.
- Smoothness constraint: the length of the contour separating the regions Ω_i should be as short as possible.

This leads to the expression[2]

$$E(\phi) = \nu \int_{\Omega} |\nabla H(\phi)| dx - \sum_{i=1}^{2} \int_{\Omega} \chi_i(\phi) \log p_i \, dx \qquad (1)$$

with the Heaviside function $H(\phi)$ and $\chi_1 = H(\phi)$ and $\chi_2 = 1 - H(\phi)$. That is, the χ_i's act as region masks, since $\chi_i = 1$ for $x \in \Omega_i$ and 0 otherwise. The first term acts as a smoothness term, that favors few large regions as well as smooth regions boundaries, whereas the second term contains assignment probabilities $p_1(x)$ and $p_2(x)$ that a pixel at position x belongs to the inner and outer regions Ω_1 and Ω_2, respectively, favoring a unique region assignment.

Minimization of this functional [1] with respect to the level-set function ϕ using gradient descent leads to

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \log \frac{p_1}{p_2} \right] \qquad (2)$$

with the smeared-out delta-function $\delta(\phi)$ as a numerically approximation to the derivation of $H(\phi)$ e.g.:

$$\delta(\phi) = \frac{1}{\pi} \cdot \frac{\tau}{\tau^2 + \phi^2} \ . \qquad (3)$$

A region descriptor $\rho_i(\mathbf{f})$ that depends on the image feature vector $\mathbf{f}$ serves to describe the characteristic properties of the outer vs. the inner regions. The assignment probabilities $p_i(x)$ for each image position are calculated based on an image feature vector via $p_i(x) := \rho_i(\mathbf{f}(x))$. The parameters of the region descriptor $\rho_i(\mathbf{f})$ are gained in a separate step using the measured feature vectors $\mathbf{f}(x)$ at all positions $x \in \Omega_i$ of a region i.

B. Signed-Distance Functions and Curvature Signals

In general, level-set methods evaluate exclusively the sign of the level-set function to determine an object and its surroundings. The exact value of the level-set function is not considered by most approaches. Signed-distance functions [18] are a common means of regulating the value of the level-set function, as they enforce the absolute value of the gradient of the level-set function to be one.

$$|\nabla \phi| = 1 \ . \qquad (4)$$

[2] Remark that ϕ, χ_i and p_i are functions over the image position x.

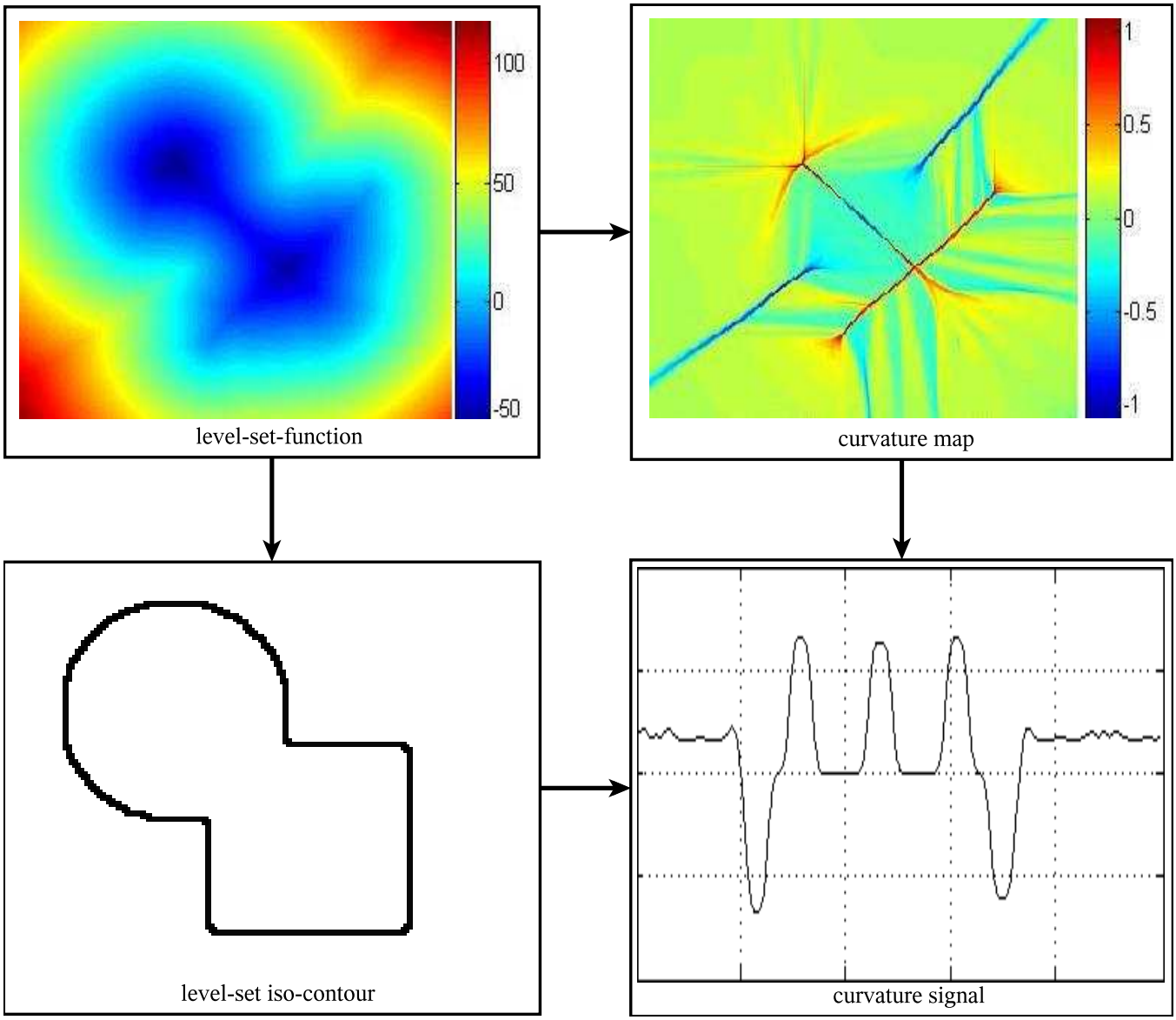

Fig. 1. The way from a level-set function to the curvature signal. Starting in the upper left. From the level-set-function (upper left) the curvature κ (upper right) of each position of the level-set function and the level-set iso-contour $\partial\Omega$ (lower left) can be derived. Tracing the level-set iso-contour $\partial\Omega$ and reading the curvature values κ at the position of the contour yields the curvature signal $C(\psi)$ (lower right).

Thus many level-set algorithms requiring the computation of the gradient become simpler and furthermore, the value of the level-set function corresponds to the distance from the contour. A common signed distance regularization equation evaluates the difference of the absolute value of the gradient to the requested value one:

$$\frac{\partial \phi}{\partial t} = -S(\phi_0)\left(|\nabla \phi| - 1\right) \tag{5}$$

with $S(\phi_0)$ as a smeared out signum function e.g.:

$$S(\phi_0) = \frac{\phi_0}{\sqrt{\phi_0^2 + 1}} \quad . \tag{6}$$

More advanced methods for the level-set reinitialization can be found in e.g. [19], [20], [21], [22], [23]

The curvature κ of the boundary $\partial\Omega$, where $\phi(x) = 0$ if $x \in \partial\Omega$ between the regions Ω_1 and Ω_2 is defined as the divergence of the normal

$$N = \frac{\nabla \phi}{|\nabla \phi|} \tag{7}$$

yielding

$$\kappa = \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \quad . \tag{8}$$

The curvature κ features positive values for convex regions of the boundary, zero for straight regions of the boundary and negative values for concave regions, respectively.

Starting at an arbitrary boundary point, walking (counter clock wise) along the boundary and collecting the curvature values κ yields a curvature signal $\kappa(\psi)$ (see Fig. 1). The

curvature signal exhibits a period of 2π. One period of $\kappa(\psi)$ with $\psi \in [0; 2\pi[$ will be denoted as the curvature signal $C(\psi)$ of a segmented level-set object Ω_2.

Remark: One property of level-set methods for image segmentation are segmented regions Ω_2 that are not connected. In our case we assume that the segmented level-set object Ω_2 features only one continuous curvature and thus does not consist of two or more unconnected regions. This is without loss of generality as in the case of two or more unconnected regions, the proposed methods can be applied to each region separately.

III. Contour-based, Scale and Rotation Invariant Object Representation

A. Curvature Function in Spatial Domain

The curvature signal $C(\psi)$ of a segmented level-set object Ω_2 can already be used as an object descriptor. Such a descriptor would be translation invariant but not scale or rotation invariant. To achieve a scale invariant representation, the length L of the boundary $\partial\Omega$ is normalized to 2π, which corresponds to the perimeter of the unit circle. The length L of the boundary is defined as follows:

$$L(\partial\Omega) = \int_{\Omega} \partial\Omega \; dx \;\; . \tag{9}$$

The scale invariant normalization of $C(\psi)$ requires a scale factor

$$s = \frac{2\pi}{L(\partial\Omega)} \tag{10}$$

which yields the final scale normalized curvature signal

$$\tilde{C}(\psi) = \frac{1}{s} \cdot C(\psi) = \frac{L(\partial\Omega)}{2\pi} \cdot C(\psi) \;\; . \tag{11}$$

To achieve also a rotation normalization, i.e. a normalization with respect to the starting point on the contour and a rotation of the object, the scale normalized curvature signal $\tilde{C}(\psi)$ has to be aligned to the phase φ_1 of its fundamental frequency, which yields a final translation, scale and rotation invariant curvature signal

$$\hat{C}(\psi) = \tilde{C}(\psi + \varphi_1) \;\; . \tag{12}$$

The phase φ_1 of the fundamental frequency of the curvature signal is defined by Fourier series [24] as follows:

$$\varphi_1 = \angle \left(\int_0^{2\pi} \hat{C}_{\phi}(\psi) \mathrm{e}^{-\mathrm{i}\psi} \; d\psi \right) \;\; . \tag{13}$$

Remarks: 1) For curvature signals that exhibit a period of $n \cdot 2\pi$ with $n \in \mathbb{N}$ and $n > 1$ the fundamental frequency vanishes, i.e. its magnitude becomes zero and thus no phase can by detected. In this case the phase of the first overtone, whose magnitude is not zero should be used for rotation normalization. 2) For circles, whose curvature signals are straight lines, no phases can be detected at all. To be able to handle also that case, a pre-test for circles has to be done, which might e.g. be done by evaluating

$$\neg\mathrm{circle} = \int_0^{2\pi} \left(\tilde{C}(\psi) - 1 \right)^2 \; d\psi \tag{14}$$

which equals zero for scale normalized circles.

Implementing the above developed contour-based, scale and rotation invariant object representation, the curvature signals for a single object vary depending on its scale and rotation due to numerical quantization errors. For this purpose we decided to collect normalized curvature signals $\hat{C}^i(\psi)$ of the same objects but with different scales and rotations. To combine all these curvature signals to one common representative signal for the object – the object descriptor $\bar{C}(\psi)$ – we propose three different methods.

The first methods uses the mean curvature signal as defined by

$$\bar{C}^M(\psi) = \frac{\sum_i \hat{C}^i(\psi)}{\sum_i 1} \tag{15}$$

as an object descriptor $\bar{C}(\psi)$.

The second method exploits a Gaussian approximation of the distribution (over i) of the collected curvature signals $\hat{C}^i(\psi)$ as an object descriptor $\bar{C}(\psi)$, defined as follows:

$$\bar{C}^G(\psi) = \mathcal{N}\left(\mu(\psi), \sigma^2(\psi)\right) \tag{16}$$

where $\mu(\psi)$ equals $\bar{C}^M(\psi)$ and $\sigma^2(\psi)$ is defined as follows:

$$\sigma^2(\psi) = \frac{\sum_i \left(\hat{C}^i(\psi) - \bar{C}^M(\psi) \right)^2}{\sum_i 1} \;\; . \tag{17}$$

The third method represents the object descriptor $\bar{C}(\psi)$ extensively in a grid based way – a histogram that reflects a non-parametric representation of the distribution of the collected curvature signals $\hat{C}^i(\psi)$, defined as follows:

$$\bar{C}^H(\psi) = H(\psi, m) = (h_m(\psi)) \;\; . \tag{18}$$

The histogram $H(\psi, m) = (h_m(\psi))$ is defined as:

$$h_m(\psi) = \frac{\sum_i \left(H(\hat{C}^i(\psi) - b_m) - H(\hat{C}^i(\psi) - b_{m+1}) \right)}{\sum_i 1} \tag{19}$$

with bins indexed by m and borders of the histogram bins defined by b_m. This object descriptor (the histogram $H(\psi, m) = (h_m(\psi))$) is similar to the neural activity maps shown in [14]. Here the authors show neural responses in a two-dimensional grid. One axis corresponds to the angular position ψ and the other axis to the curvature value κ.

B. Curvature Function in Frequency Domain

So far we looked at the objects level-set curvature signal $C(\psi)$ only in the spatial domain. In the context of periodical signals Fourier analysis constitutes a means of evaluating these signals. A periodical signal $C(\psi)$ with the period $T = 2\pi$ and the angular frequency $\omega = \frac{2\pi}{T} = \frac{2\pi}{2\pi} = 1$ can be written as [24]:

$$C(\psi) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{\mathrm{i}n\psi} \tag{20}$$

with the complex Fourier coefficients

$$c(n) = \frac{1}{2\pi} \int_0^{2\pi} C(\psi) \mathrm{e}^{-\mathrm{i}n\psi} \, d\psi \ . \tag{21}$$

To achieve scale invariance the curvature signal needs to be scaled analog to Eq. 11 in spatial domain. This yields the scale invariant complex Fourier coefficients

$$\tilde{c}(n) = \frac{1}{s} \cdot c(n) \tag{22}$$

with the scale factor s as defined in Eq. 10.

To achieve rotation invariance two methods are possible. In the first method, the phase φ_1 of the fundamental frequency has to be determined as in Eq. 13. All complex Fourier coefficients can be rotation normalized by the following equation:

$$\hat{c}^F(n) = \mathrm{e}^{\mathrm{i}n\varphi_1} \cdot \tilde{c}(n) \ . \tag{23}$$

The second method we propose for rotation invariant object representation completely ignores the phases of the Fourier coefficients and uses only their magnitude

$$\hat{c}^{|F|}(n) = |\tilde{c}(n)| \ . \tag{24}$$

Equation 24 neglects some information of the curvature signal. In our experience the reduced information is still able to act as an object descriptor. In particular when comparing a "new" object descriptor to descriptors given in a "memory" (see Ch. III-C for a detailed discussion). Furthermore, it is not necessary to use all Fourier coefficients $\hat{c}(n)$ with $n \in \mathbb{Z}$. Also the first N Fourier coefficients $\hat{c}(n)$ with $|n| \leq N$ are sufficient to represent the objects contour. Disregarding the higher Fourier coefficients is equivalent to a low-pass filtering of the curvature signal, thus very similar objects, that differ only in high details on their surface, are represented by the same object descriptors. This observation would also occur due to quantization errors, when leaving the mathematical formulation and going to real implementations. Furthermore this constitutes a well received generalization property of the proposed approach.

Following the above discussion, Eq. 23 and Eq. 24 with $|n| \leq N$ leads to a discrete object descriptor. Together with a series of curvature signals in frequency domain $\hat{c}^i(n)$, acquired at different scales and rotations, again three different methods of obtaining object descriptors $\bar{c}$ are proposed. Analog to those in ch. III-A an object descriptor $\bar{c}^M(n)$ based on the mean values (cp. Eq. 15), $\bar{c}^G(n) = \mathcal{N}(\mu(n), \sigma^2(n))$ a Gaussian (i.e. parametric) distribution (cp. Eq. 16) and $\bar{c}^H(n, m)$ a non-parametric, histogram-based distribution (cp. Eq. 18) are proposed.

C. Object classification

Until now the objects are segmented and represented by an object descriptor, derived from a curvature signal of the object's contour. For the object descriptor three different methods of creating a scale and rotation invariant representation were presented. One in the spatial domain (Eq. 12: $\hat{C}(\psi)$) and two in the frequency domain, with and without considering the phase information (Eq. 23: $\hat{c}^F(n)$ and Eq. 24: $\hat{c}^{|F|}(n)$, respectively). Independent from those we have presented three methods for collecting the information of a series of object descriptors $\hat{C}^i(\psi)$ or $\hat{c}^i(n)$ for the same objects. These methods are based one the mean values (Eq. 15), a Gaussian distribution (Eq. 16) and a histogram that approximates the distribution (Eq. 18). As every method of object descriptor creation can be combined with every method of object descriptor collection we arrive at nine different methods for obtaining an object descriptor ($\bar{C}(\psi)$ and $\bar{c}(n)$ in spatial and frequency domain, respectively) of one object, as depicted in Fig. 2.

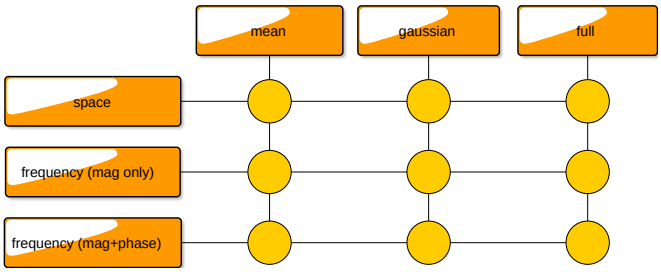

Fig. 2. Three different methods of representing the curvature signal (left column) and three different methods of collecting a serious of curvature signals for the same object and creating an object descriptor (top row).

In this section we will develop a method to use a number of known ("memorized") object descriptors to recognize an unknown object based on its object descriptor derived from its curvature signal. For this purpose all nine above proposed methods may be used to build an object descriptor database ("memory") $\bar{C}_k(\psi)$ and $\bar{c}_k(n)$ of known objects k. The object descriptor of e.g. an unknown object $\hat{c}^{|F|}(n)$ can be compared to all known object descriptors $\bar{c}_k^{|F|}(n)$ by e.g. computing the corresponding Euclidean distances

$$d_k = ||\bar{c}_k^{|F|}(n) - \hat{c}^{|F|}(n)|| = \sum_n \left((\bar{c}_k^{|F|}(n) - \hat{c}^{|F|}(n))^2 \right) \ . \tag{25}$$

In this way the most probable object is found by determining the index k_{min} of the minimal distance

$$k_{\min} = \arg\min(d_k) \ . \tag{26}$$

For a given number of objects and when all objects' descriptors are known ("memorized"), the length of the object descriptors, i.e. the number of components in the vectors $\bar{c}_k(n)$ can be reduced. Even with such a sparse object

descriptor representation, unknown objects can sufficiently well be assigned to a known object from the database ("memory"). In the following we will present two methods to reduce the number of descriptor dimensions.

The first method works with object descriptors $\bar{c}_k^M(n)$ for each known object k that are based on the absolute value of the Fourier coefficients that are represented only by the mean values of the collected curvature signals

$$\bar{c}_k^M(n) = \frac{\sum\limits_i \hat{c}_{i,k}^{|F|}(n)}{\sum\limits_i 1} \tag{27}$$

with $\hat{c}_{i,k}^{|F|}(n)$ as defined in Eq. 24. The number of required components can be reduced by applying standard PCA (principle component analysis) [25]. PCA constitutes a linear transformation, that transforms the original $(N+1)$-dimensional ($n \in \{0, 1, ..., N\}$) objects descriptor $\bar{c}_k(n)$ to a lower, Q-dimensional ($Q < N+1$) object descriptor

$$\bar{c}_k(q) = T \cdot \bar{c}_k(n) \tag{28}$$

with the transformation matrix T. The transformation matrix T is determined by the PCA in a way that investigates the direction of the greatest variance of the object descriptors' components. The directions of great variance are preserved, while the directions with low variance are refused. Every object descriptor of an unknown object $\hat{c}(n)$ has now to be transformed in the low dimensional space via Eq. 28. In this low-dimensional space, the corresponding known object descriptor can be found analog to Eq. 25 and Eq. 26. In this way only the transformation matrix T and the k known low-dimensional object descriptors have to be memorized.

The second method is also based on the absolute values of the Fourier coefficients. In contrast to the first method the distribution $\bar{c}_k(n)$ of the descriptors component n of the known object k are assumed to be known. The distribution might be approximated by a Gaussian distribution $\bar{c}_k(n) = \mathcal{N}\left(\mu_k(n), \sigma_k^2(n)\right)$ (cp. Eq. 16) or represented discrete by a histogram $\bar{c}_k(n) = H_k(n, m)$ (cp. Eq. 18). In the following we will represent the distribution of the n's components of the k's known object by

$$D^{k,n}\left(\hat{c}^{|F|}(n)\right) = D^{k,n}(c) = P^n(c|k) \tag{29}$$

which equals the conditional probability $P^n(c|k)$ of the n's descriptors component for the object descriptor value c, given the known object k. For a not assigned object descriptor $c = \hat{c}^{|F|}(n)$ the classification probability $P(k|c)$ for each known object k, given the object descriptor c can be computed as follows:

$$P_k = P(k|c) = \prod_n P^n(k|c) \tag{30}$$

The object will than be assign to the known object $k_{\max}$ with the highest probability P_k:

$$k_{\max} = \arg\max(P_k) \tag{31}$$

with

$$P^n(k|c) = \frac{P^n(c|k) \cdot P^n(k)}{P^n(c)} \tag{32}$$

according to Bayes' theorem [26]. Without preferring any object

$$P^n(k) = \frac{1}{K} \tag{33}$$

where K denotes the total number of known objects and

$$\begin{aligned} P^n(c) &= \sum_k P^n(c \cap k) \\ &= \sum_k P^n(c|k) \cdot P^n(k) \\ &= \frac{1}{K} \sum_k P^n(c|k) \end{aligned} \tag{34}$$

that finally reads as

$$P^n(k|c) = \frac{P^n(c|k)}{\sum\limits_k P^n(c|k)} \tag{35}$$

To reduce the number of components in the object descriptor vectors the probability $\mathbf{P}^n$ for a correct classification of all known objects k, based only on one single component can be computed for all components n. Choosing only the Q components with the highest probabilities $\mathbf{P}^n$ to represent the objects and ignoring the reminding components yields a new low-dimensional object descriptor. The probability $\mathbf{P}^n$ accumulates the correct classification of all known objects k and thus is defined as the product of all probabilities $P^{n,k}$ of the correct classification of each object k

$$\mathbf{P}^n = \prod_k P^{n,k} \tag{36}$$

with the correct classification probability $P^{n,k}$ for the known object k and the object descriptors component n defined as follows:

$$P^{n,k} = \frac{P_{\text{HIT}}^{n,k} + P_{\text{CR}}^{n,k}}{P_{\text{HIT}}^{n,k} + P_{\text{FA}}^{n,k} + P_{\text{MISS}}^{n,k} + P_{\text{CR}}^{n,k}} \ . \tag{37}$$

Defining the sets

$$S_{n,k} = \{c \mid \arg\max_{\tilde{k}}(P^n(\tilde{k}|c)) = k\} \tag{38}$$

that contain all values c that yield a decision for the known object k leads to the *hit*-probability

$$P_{\text{HIT}}^{n,k} = P^n(S_{n,k}|k) = \int\limits_{S_{n,k}} P^n(c|k)\ dc \tag{39}$$

the *false alarm*-probability

$$P_{\text{FA}}^{n,k} = P^n(S_{n,k}|\neg k) = \sum_{\tilde{k} \neq k} \int\limits_{S_{n,\tilde{k}}} P^n(c|\tilde{k})\ dc \tag{40}$$

the *miss*-probability

$$P_{\text{MISS}}^{n,k} = P^n(\neg S_{n,k}|k) = \int\limits_{\neg S_{n,k}} P^n(c|k)\ dc \tag{41}$$

and the *correct rejection*-probability

$$P_{\mathrm{CR}}^{n,k} = P^n(\neg S_{n,k}|\neg k) = \sum_{\tilde{k}\neq k} \int\limits_{\neg S_{n,\tilde{k}}} P^n(c|\tilde{k})\ dc\ . \quad (42)$$

Both presented methods have in common that the number of components in the low-dimensional object descriptor space has to be predefined and is not determined by the algorithms themselves. Please see Sect. IV for reasonable values.

IV. Main Results

A. Simple database with six basic shapes

We used a simple database for basic tests of the proposed method. Figure 3 depicts the six basic shapes used for the results presented in this section.

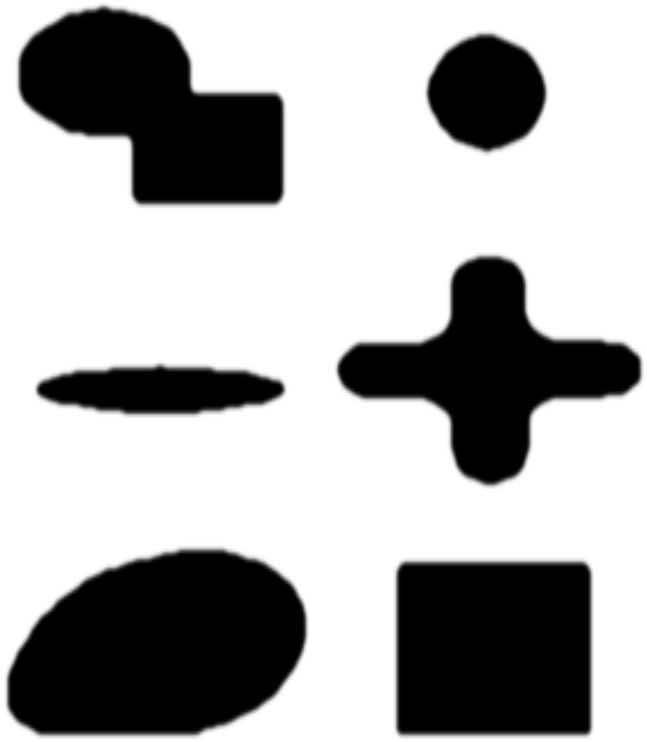

Fig. 3. Database with six basic shapes. These basic shapes were used to produce the results shown in this section

These basic shapes where chosen to demonstrate the principles of the approach and to show easily comprehensible examples. The proposed methods are not restricted to these basic shapes, but when using more complex shapes, the curvature signals and object descriptors also get more complex and thus the examples would become less descriptive.

B. Curvature Function in Spatial Domain

Figure 4 depicts a spatial domain object descriptor for the first basic shape (Fig. 3, upper left). The object descriptor is based on a Gaussian distribution $\bar{C}^G(\psi) = \mathcal{N}\left(\mu(\psi), \sigma^2(\psi)\right)$ (cp. Eq. 16). The horizontal axis represents the angular position ψ and the vertical axis the mean curvature $\kappa = \mu$ and its standard deviation σ. The curvature signal starts in the middle of the great circle that corresponds to a section of almost constant curvature. This section is followed by a small section of great concavity, i.e. negative values. Thereafter are three peaks of positive curvature located, i.e. convex parts of the contour. Theses peaks correspond to the three corners in the lower right part of the object. Finally a section with constant curvature can be observed, that belongs to the great circle again.

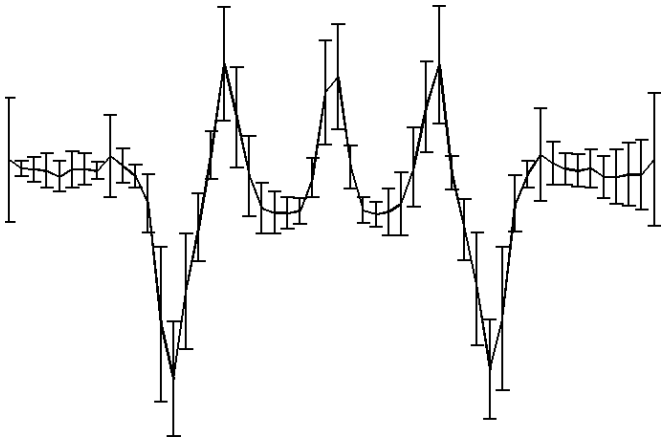

Fig. 4. Spatial domain object descriptor for the first basic shape (Fig. 3, upper left). The object descriptor is represented by mean values and the corresponding standard deviation, the parameters of a Gaussian distribution. The horizontal axis represents the angular position ψ and the vertical axis the mean curvature $\kappa = \mu$ and its standard deviation σ.

A histogram based object descriptor $\bar{C}^H(\psi)$ (cp. Eq. 18) is depicted in Fig. 5. The horizontal axis represents the angular position ψ and the vertical axis the distribution of the curvature κ. Dark areas of the image denote high values

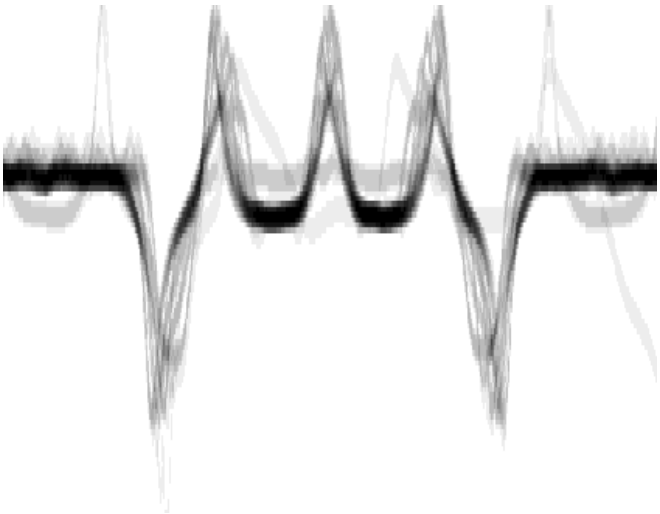

Fig. 5. Histogram based object descriptor $\bar{C}^H(\psi)$ for the first basic shape (Fig. 3, upper left). The object descriptor is represented by discrete distributions of the curvature κ. The horizontal axis represents the angular position ψ and the vertical axis the curvature κ. Dark areas of the image denote high values (probabilities) and bright areas denote low values (probabilities).

(probabilities) and bright areas denote low values (probabilities). This object descriptor is similar to the population responses of neurons in area V4, shown in [14].

C. Curvature Function in Frequency Domain and Object Classification

Figure 6 depicts all six frequency domain object descriptors $\bar{c}^G(n) = \mathcal{N}(\mu(n), \sigma^2(n))$ based on Gaussian (i.e. parametric) distributions (cp. Eq. 16) with $N = 15$ for the basic shapes shown in Fig. 3. The horizontal axis represents the component $n \in \{0, 1, \ldots, 15\}$ of the object descriptor. The colors denote the different objects k.

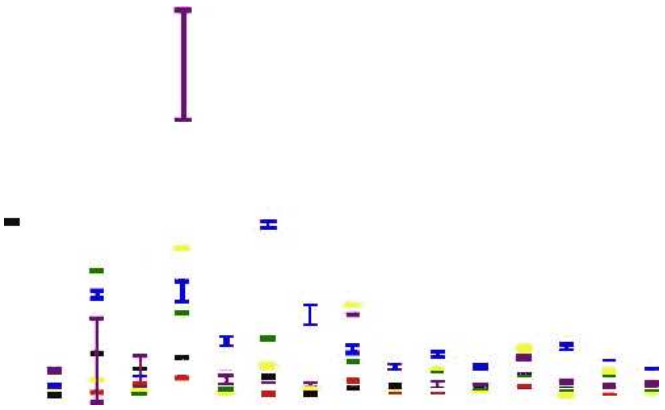

Fig. 6. Frequency domain object descriptors for all six objects with basic shapes from Fig. 3. The different colors denote the objects. Shown are the mean values and standard deviations of each object descriptors component. The component $n = 4$ shows the highest probability to classify all objects correct.

In the column $n = 4$ in Fig. 6 all six colors, i.e. all six objects, can clearly be distinguished from each other. The components either have only little variance or, in case of large variance (the one on the top), have a large distance to their nearest neighbors. Thus the probability $\mathbf{P}^n$ for a correct classification of all objects (see Eq. 36) is maximal for the column $n = 4$. Sparse object descriptors for the database of basic shapes from Fig. 3 would work quite well, using only the fourth component of all non-sparse object descriptors.

V. Conclusion

In this paper we gave a brief introduction to level-set methods for image segmentation. We showed the regularization by signed-distance functions and how to create a curvature signal from a given level-set function. Three different methods for the translation, scale and rotation invariant representation of curvature signals were shown. Furthermore we proposed three methods for creating object descriptors from a series of curvature signals for the same object. Finally we introduced an approach for object classification based on the proposed object descriptors. This object classification approach included two methods for dimension reduction of the used object descriptors.

Future work will include the evaluation of the integration of the proposed approaches in a larger object detection and classification framework.

References

[1] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Cmpt. Phys.*, vol. 79, pp. 12–49, 1988.

[2] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math*, vol. 42, pp. 577–685, 1989.

[3] S. C. Zhu and A. L. Yuille, "Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation," *PAMI*, vol. 18, no. 9, pp. 884–900, 1996.

[4] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[5] J. Kim, J. W. Fisher, A. J. Yezzi, M. Çetin, and A. S. Willsky, "Nonparametric methods for image segmentation using information theory and curve evolution," in *International Conference on Image Processing, Rochester, New York*, vol. 3, Sept. 2002, pp. 797–800.

[6] M. Rousson and R. Deriche, "A variational framework for active and adaptative segmentation of vector valued images," *IEEE Workshop on Motion and Video Computing, Orlando, Florida*, Dec. 2002.

[7] T. Brox, M. Rousson, R. Deriche, and J. Weickert, "Unsupervised segmentation incorporating colour, texture, and motion," *Computer Analysis of Images and Patterns*, vol. 2756, pp. 353–360, 2003.

[8] D. Weiler and J. Eggert, "Multi-dimensional histogram-based image segmentation," in *International Conference on Neural Information Processing (ICONIP)*, Kitakyushu, Japan, 2007.

[9] Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on Computer Vision*, vol. 1, pp. 105–112, 2001.

[10] Y.-Y. Chuang, B. Curless, D. Salesin, and R. Szeliski, "A bayesian approach to digital matting." in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 2001, pp. 264–271.

[11] C. Rother, V. Kolmogorov, and A. Blake, ""GrabCut": Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.

[12] Grossberg, Stephen, Hong, and Simon, "A neural model of surface perception: Lightness, anchoring, and filling-in," *Spatial Vision*, vol. 19, no. 2-4, pp. 263–321, 2006.

[13] S. Loncaric, "A survey of shape analysis techniques," *Pattern Recognition*, vol. 31, pp. 983–1001, 1998.

[14] A. Pasupathy and C. E. Connor, "Population coding of shape in area v4," *Nat Neurosci*, vol. 5, no. 12, pp. 1332–1338, December 2002.

[15] E. Parzen, "On the estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.

[16] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal for Computer Vision*, vol. 1, no. 4, pp. 321–331, Jan. 1988.

[17] T. Chan, B. Sandberg, and L. Vese, "Active contours without edges for vector-valued images," *J. Visual Communication Image Representation*, vol. 11, no. 2, pp. 130–141, 2000.

[18] *Level Set Methods and Dynamic Implicit Surfaces*. Springer, Berlin, 2002, ch. 2 & 7.

[19] R. P. Fedkiw, T. Aslam, B. Merriman, and S. Osher, "A non-oscillatory eulerian approach to interfaces in multimaterial flows (the ghost fluid method)," *J. Comput. Phys.*, vol. 152, no. 2, pp. 457–492, 1999.

[20] D. Peng, B. Merriman, S. Osher, H. Zhao, and M. Kang, "A pde-based fast local level set method," *J. Comput. Phys.*, vol. 155, no. 2, pp. 410–438, 1999.

[21] G. Russo and P. Smereka, "A remark on computing distance functions," *J. Comput. Phys.*, vol. 163, no. 1, pp. 51–67, September 2000.

[22] M. Sussman and E. Fatemi, "An efficient, interface-preserving level set redistancing algorithm and its application to interfacial incompressible fluid flow," *SIAM J. Sci. Comput.*, vol. 20, no. 4, pp. 1165–1191, 1999.

[23] M. Sussman, P. Smereka, and S. Osher, "A level set approach for computing solutions to incompressible two-phase flow," *Journal of Computational Physics*, vol. 114, pp. 146–159, sep 1994.

[24] *The Analytical Theory of Heat*. Dover Publications, 1878.

[25] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.

[26] *Pattern Classification*. John Wiley & Sons, 2001, ch. 2.