3D Pose Estimation using Coupled Snakes

Benjamin Hohnhäuser and Günter Hommel
Technische Universität Berlin, Fakultät IV
Institut für Technische Informatik
und Mikroelektronik (TIME)
Einsteinufer 17, D-10587 Berlin, Germany
bh@cs.tu-berlin.de

ABSTRACT

An extension to the classical concept of active contour models is proposed. Besides the introduction of a new technique to couple several contours and treat them as one, the coupling of active contours beyond image domains is presented. The coupling is realized by a force, that controls the mutual attraction or repulsion of each active contour, depending on its definition. The coupling across the borders of images acquired by different sources provides a higher integration of information for each extracted image feature implying an appropriate registration of the images.

Keywords Active Contour; Snake; Segmentation; Sensor Fusion; Coupling Force; Coupling Energy; Distance Function

1. INTRODUCTION

Active contour models, also known as *snakes*, are elastic curves defined within the image domain that move under influence of internal forces coming from within the curve itself and external forces computed from the image data. They deform towards image features satisfying certain smoothness constraints.

Since their introduction by Kass et al. [KWT87], active contour models have been subject to extensive research activities as well as to applications in various fields of computer vision. The classical approach based on deforming an initial contour towards object boundaries was extended and improved in many ways.

Several improvements addressed the problems of initialisation and poor convergence to object boundaries. In [Coh91] Cohen introduced an inflation force that pushes the snake towards the edge like a balloon which is inflated. The initial curve needed no longer to be close to the solution in order to converge. A different approach using two interlinked snakes (*dual snakes*) was proposed by Gunn et al. in [GN97]. The advantage over classical snakes is the higher robustness against local minima and a lower sensitivity to initialisation and parameters. In their paper [CC97] Cham et al. present the *Stereo Coupled Active Contours*. They describe a technique for tracking objects in the two images of a stereo vision system by snakes which are coupled. The *dual snakes* and the *Stereo Coupled Active Contours* can be considered as a special case or subset of the coupled snakes concept proposed in this paper.

In [XP97] Xu et al. presented the new external force called the *Gradient Vector Flow (GVF)*, that in addition to the enhancement of robustness against a weak initial position also improves the convergence to boundary concavities.

The work by McInerney et al. [MT95] overcame the topological limitations of the original snake concept. In their work [GGO00] Giraldi et al. combined the approach of topological changes with the benefits of the *dual snakes* idea.

The main fields of active contour applications nowadays is mainly the segmentation of medical images as well as areal images and object tracking. Some examples for the application of active contours to medical images from different sources are shown in [SD96][LKS00][CHTH93] and [HCG00]. The second major field of the application of snakes is the segmentation of areal images acquired by airborne and space-borne sensory. Like in medical image segmentation the task is to find areas or edges which have no regular shape in the presence of noise (e.g. *speckle* in

Journal of WSCG, Vol.12, No.1-3, ISSN 1213-6972
WSCG'04, February 2-6, 2003, Plzen, Czech Republic.

SAR[1] images)[GR01][Hor99]. In the area of tracking moving objects there is the recent work of Chen et al. using a Kalman Filter[CRH02] and work of MacCormick et al. that describes an approach for tracking multiple objects with a probabilistic exclusion principle in [MB00].

For a more detailed and broader review on active contours we refer to [JZDJ98].

2. BACKGROUND

The classical formulation [KWT87] of an active contour defines it as a two dimensional, planar curve $\mathbf{v}(s) = (x(s), y(s)), s \in [0,1]$ that minimises its energy by deformation and movement within an image. The energy results from forces that are derived from the contour itself and from the image. The words *force* and *energy* are used synonymously. The total energy E_{snake} of an active contour $\mathbf{v}$ is given by

$$E_{snake} = \int_0^1 (\alpha(s) \left|\frac{d\mathbf{v}}{ds}\right|^2 + \beta(s) \left|\frac{d^2\mathbf{v}}{ds^2}\right|^2 + E_{ext}(\mathbf{v}) ds \quad (1)$$

where α and β are weighting parameters which control the snake's tension and rigidity, respectively. The external energy results from the forces that originate from the image and any other source, like user induced or constraint forces.

3. APPROACH

In times of more complex and also faster sensor systems often more than only intensity data of a scene are acquired (e.g. infra-red or dense range data). In most automatic or semi-automatic recognition schemes the segmentation is performed *separately* for all image data provided by the different sources. Thus the segmentation processes do not mutually affect each other and the integration of collected image features has to be performed at a higher level instead.

A more suitable approach is to use as much information as possible about a feature in the segmentation process. Assuming the images of the different sources are correctly registered before, it is preferable to perform the segmentation in all images at the same time, incorporating all information available. So each extracted feature is now in some way more exactly aggregating data from different image sources. In order to do this a segmentation technique must be used that works upon a single image domain. This seems to be a promising approach to extend the concept of active contours.

3.1 Extending the Classical Snake Formulation

We extended the original formulation in two aspects. First of all we consider a set of n classical active contours $\mathbf{v_1} \cdots \mathbf{v_n}$ as an entity that we call coupled active contour $\mathbf{C}$. The classical snake is a special case of our definition of the coupled snake. A coupled active contour is defined as:

$$\mathbf{C} = \{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_n}\} \quad (2)$$

consisting of n classical active contours. Furthermore we introduce a new internal energy E_{coup}. It is defined by the coupling force between the contours within $\mathbf{C}$. The function $d_{coup}(\mathbf{v_i}(s), \mathbf{v_j}(s))$ calculates the coupling force between the snakes $\mathbf{v_i}$ and $\mathbf{v_j}$ at s. The coupling energy E_{coup} for the whole coupled active contour $\mathbf{C}$ is then defined as follows

$$E_{coup} = \int_0^1 D_{coup}(\mathbf{C}, s) ds \quad (3)$$

where

$$D_{coup}(\mathbf{C}, s) = \sum_{i=1}^{n} \sum_{j=1}^{n} \xi_{ij} d_{coup}(\mathbf{v_i}(s), \mathbf{v_j}(s)) \quad (4)$$

where ξ_{ij} is a weighting factor that defines how much the position of $\mathbf{v_j}$ relative to $\mathbf{v_i}$ influences the energy minimising process of $\mathbf{v_i}$. Since we extended the definition of active contours we consider E_{coup} as resulting from an internal force as part of the internal energy term. The central point is now the definition of the coupling force, that interlinks all active contours in $\mathbf{C}$. For each pair of contours it is possible to define its own force function which even must not necessarily be symmetric at all. Thus it is possible to define within the snake concept itself *volcanoes* and *springs* attached to certain points or regions in an image like mentioned in [SA98]. But more important for vision tasks is the possibility to use certain features detected by a snake like an anchor, that guides other snakes to their assumed destination similar to [SK01].

To reduce the computational complexity we rely on some restrictions of the definition of the coupling force respectively the function d_{coup} in our implementation. This is no general limitation to the concept but only to our application. First, each classical contour in $\mathbf{C}$ has the same number of control-points. Second, each control-point of contour i is only interlinked to control-points of the other contours with the same index. These restrictions are useful with focus on real-time applications but of course not necessary if enough computational power is available.

It follows a short illustration of the concept by some standard setups. With $n = 1$ (only one active contour in $\mathbf{C}$) we have the classical active contour. Rising the number of classical contours in $\mathbf{C}$ up to two and defining the function d_{coup} in a symmetric manner as the Euclidian distance (or another meaningful metric or a even more complex distance measure) leads to a behaviour like described in [GN97].

Since this force is symmetric there is no need to distinguish between $\mathbf{v_i}$ and $\mathbf{v_j}$. The next example shows an unsymmetric definition for d. We want $\mathbf{v_1}$ to push

[1] Synthetic Aperture Radar

away $\mathbf{v_2}$ while $\mathbf{v_2}$ has no effect on $\mathbf{v_1}$. We define d_{coup} as follows:

$$d_{coup}(\mathbf{v_i}(s),\mathbf{v_j}(s)) = \begin{cases} 0 & \text{if } i=1, j=2 \\ \frac{1}{d_{max}} & \text{if } i=2, j=1 \end{cases} \quad (5)$$

with

$$d_{max} = max((x_i(s)-x_j(s)),(y_i(s)-y_j(s))) \quad (6)$$

The more classical contours are contained in **C** the more complex the definition of d_{coup} can become.

3.2 Coupling Beyond One Image Domain

Due to the extension of the classical active contour the coupled active contour concept enables us to let all single active contours in **C** interact within one image domain. Further it is also possible to let each single active contour in **C** move exclusively in its own image, like proposed in [CC97].

For that reason the formulation for classical snakes has to be extended by the indication of the image in which it operates. Thus, the above definition of $\mathbf{v_i}$ turns into

$$\mathbf{v}(s) = (x(s), y(s), \iota), s \in [0,1], \iota \in I \quad (7)$$

where I is the set of all images from different image sources of the scene.

The main point is that after the images have been acquired by different sensor systems and a proper registration is performed, each segmentation result obtained by a coupled active contour will represent a higher amount of information for each extracted feature. This intrinsic integration of more data into one feature during its detection is a great advantage over the conventional separation of detection and integration, because it directly influences the segmentation process to be more accurate and results in higher accuracy.

4. EXPERIMENTAL RESULTS

In our application a special sensor system is used (laser-range-camera [SFR99] by ASTRIUM - *European Space Systems Company*) which provides *pixel-synchronous* dense range images as well as intensity images as shown in Figure 1 and Figure 2. If the different image sources provide different image formats (e.g. width and height) adequate trimming or interpolation has to be deployed.

The main advantage of this sensor system is to directly get metric range data which enables us to determine the spatial position and orientation of recognised objects for manipulation. On the other hand the main drawback is the poor quality of the range data that includes a very high amount of noise (see Figure 2). Concerning the noise we have to perform some filtering before we can segment the images. For this scene it appeared reasonable to use a median filter with a 15×15 operator window. The drawback is that we lose a high amount of range information so that the spatial relation of the extracted features is inconsistent in many cases.

Figure 1: Intensity image of the scene

In our system we especially investigate active contours which are interlinked beyond image domains (intensity and range). For that purpose we concentrate on a coupled contour that consists of only two interlinked classical contours.

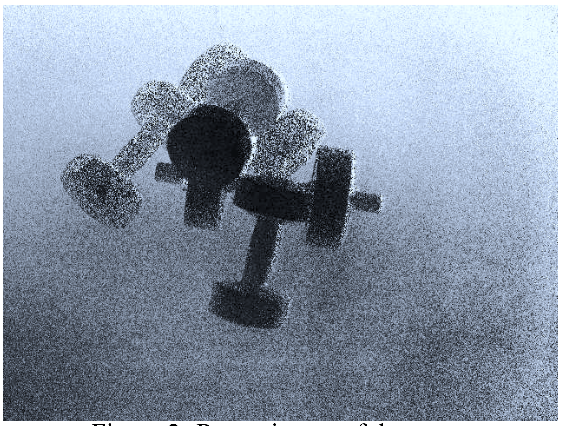

Figure 2: Range image of the scene

For the initial position of the active contours which is a crucial point we use a semi-automatic method. After edge detection and simple automatic evaluation (edge length and strength) for both images (range and intensity) is performed the system proposes some candidates to the user. Since this choosing algorithm is fairly conservative, regarding the total number of results, some of the well suited candidates are unfortunately discarded. These candidates can be accepted or rejected. For all accepted candidates a segmentation process by coupled active contours is triggered. Due to the already accomplished edge detection the initial position of the two snakes in **C** is pretty good.

On average it takes about ten energy optimising iterations to reach the final configuration for a coupled active contour with about 30 control points each and a 5×5 search window. One iteration takes below one millisecond on a PC with a 1200 MHz Pentium II processor.

Figure 3 shows segmentation results. Because of the median-filtering the range data is not very accurate anymore, so the estimation of the spatial position of some features is limited. For that kind of filtered images it appeared to be necessary for the snake to consist of about 50 control points. For the elliptical surface of the barbell in the foreground the assumption is met. The estimated radius is about two millimetres inaccurate. The calculated plane where the arc lies in has a deviation of about 5°. The position and orientation of all other features is more or less incorrect and not suitable to form a consistent hypothesis of the object's pose.

5. SUMMERY, CONCLUSIONS, AND FUTURE WORK

The above described extensions on active contours gain the benefits of a more compact formulation of basic concepts like *springs* and *volcanoes*, which can be formulated as part of the coupled active contour. Furthermore it is now possible to integrate information from different image sources into one feature to increase the accuracy. Compared to an approach where only the calculation of the image energy for a single snake is performed in several images the coupled active contours provide an individual set of parameters for each snake in each image. In addition it is possible to formulate some of the earlier extensions in the area of active contours with coupled active contours. Particularly, the classical snake is a special case of a coupled active contour.

One major drawback is if there are several or at least only two different image sources in at least one stage of the whole process some kind of image registration/sensor fusion has to be performed. Another disadvantage is the missing ability of coupled active contours in changing their topology.

There are a lot of desirable improvements. First of all a more accurate sensor system is inevitable. First attempts into that direction have already been performed. Further, the up to now semi-automatic stage for selecting assumingly promising candidates for the segmentation must be improved and extended to an automatic stage, where no user guidance is needed. This can be achieved by analysing and interpreting the range data close to a potentially relevant feature detected in the preceeding edge detection procedure. Since the presented approach is a very open and basic improvement, extensions into a variety of directions are possible.

ACKNOWLEDGEMENTS

This work has been performed in cooperation with DaimlerChrysler AG, Research and Technology, Intelligent Systems Group, Berlin. Special thanks to Dr. Andreas Stopp and Dr. Frieder Lohnert for their helpful comments and annotations.

REFERENCES

[CC97] Tat-Jen Cham and Roberto Cipolla. Stereo coupled active contours. In *IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition 1997 (CVPR'97)*, pages 1094–1099, 1997.

[CHTH93] Timothy F. Cootes, A. Hill, Christopher J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. In *IPMI*, pages 33–47, 1993.

[Coh91] Laurent D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing*, 53(2):211–218, Mar 1991.

[CRH02] Yunqiang Chen, Yong Rui, and Thomas Huang. Parametric contour tracking using unscented kalman filter. In *Proceedings of IEEE ICIP 2002*, volume III, pages 613–616, September 2002.

[GGO00] Gilson Giraldi, Luiz Goncalves, and Antonio Oliveira. Dual topologically adaptable snakes. In *Proceedings of the Fifth Joint Conference on Information Sciences (JCIS'2000) - Third International Conference on Computer Vision, Pattern Recognition, and Image Processing*, pages 103–106, 2000.

[GN97] Steve R. Gunn and Mark S. Nixon. A robust snake implemantation; a dual active contour. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(1):63–68, Jan 1997.

[GR01] O. Germain and Ph. Réfrégier. Edge location in sar images: performance of the likelihood ratio filter and accuracy improvement with an active contour approach. *IEEE Trans. on Image Processing*, 10(1):72–78, January 2001.

[HCG00] G. Hamarneh, A. Chodorowski, and T. Gustavsson. Active contour models: Application to oral lesion detection in color images. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 4, pages 2458 –2463, 2000.

[Hor99] M.S. Horritt. A statistical active contour model for sar image segmentation. *Image and Vision Computing*, 17(3/4):213–224, March 1999.

[JZDJ98] Anil K. Jain, Yu Zhong, and Marie-Pierre Dubuisson-Jolly. Deformable template models: A review. *Signal Processing*, 71(2):109–129, December 1998.

[KWT87] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *Inernational Journal of Computer Vision*, pages 321–331, 1987.

[LKS00] S. Loncaric, D. Kovacevic, and E. Sorantin. Semi-automatic active contour approach to segmentation of computed tomography volumes, 2000. url: http://citeseer.nj.nec.com /loncaric00semiautomatic.html.

[MB00] John MacCormick and Andrew Blake. A probabilistic exclusion principle for tracking multiple objects. *International Journal of Computer Vision*, 39(1):57–71, 2000.

[MT95] T. McInerney and D. Terzopoulos. Topologically adaptable snakes. In *Proc. of the ICCV95*, pages 840–845, 1995. url: http://citeseer.nj.nec.com /mcinerney95topologically.html.

[SA98] J. Schnabel and S. Arridge. Snake implementation in the uk iue project, 1998.

[SD96] L.H. Staib and J.S. Duncan. Model-based deformable surface finding for medical images. *IEEE Trans. on Medical Imaging*, 15(5):720–731, October 1996.

[SFR99] W. Schroeder, E. Forgber, and G. Roeh. Laser range camera application. Technical report, 1999.

[SK01] Thitiwan Srinark and Chandra Kambhamettu. A framework for multiple snakes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’01)*, volume 2, pages 202–209, 2001.

[XP97] Chenyang Xu and Jerry L. Prince. Gradient vector flow, a new external force for snakes. In *IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition 1997 (CVPR’97)*, pages 66–71, 1997.

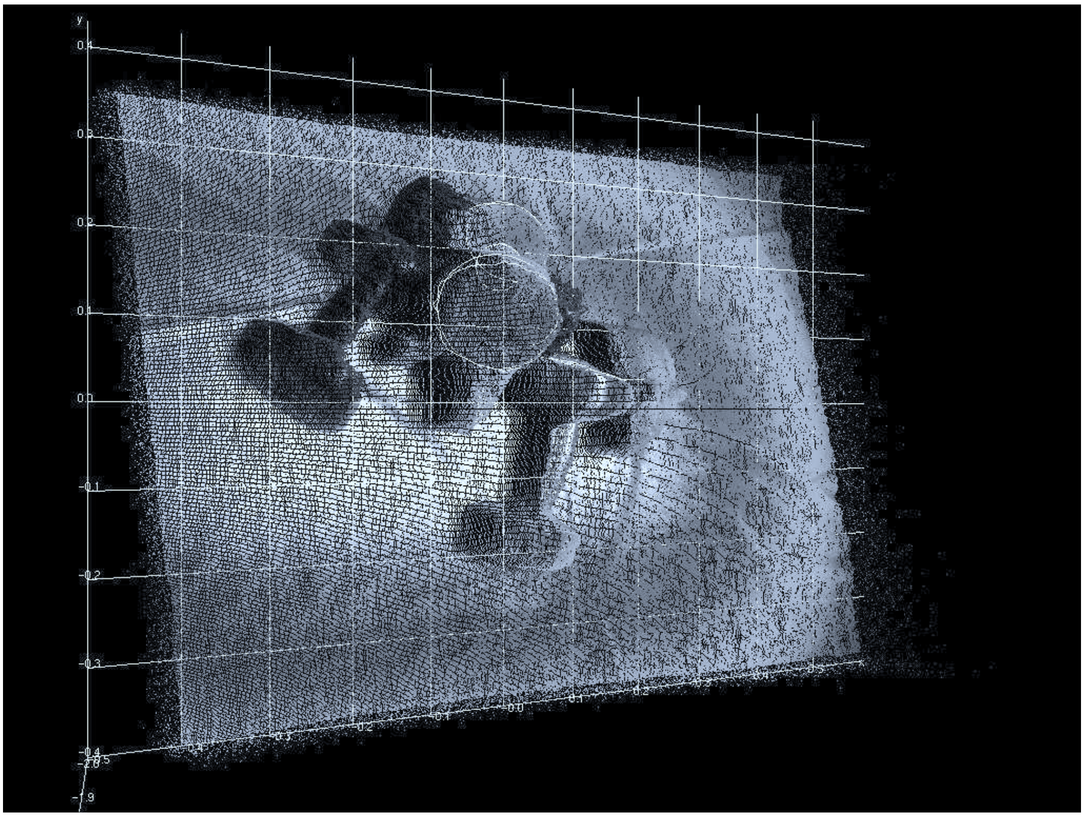

Figure 3: Segmentation results