Tracking of brain tumors using vision and neurosonography

Rubén Machucho Cadena, Sergio de la Cruz Rodríguez
and Eduardo Bayro-Corrochano*

We have developed a method to render brain tumours from endoneurosonography. We propose to track an ultrasound probe in successive endoscopic images without relying on an external optic or magnetic tracking system. The probe is tracked using two different methods: one of them based on a generalised Hough transform and the other one based on particle filters. By estimating the pose of the ultrasound probe in several contiguous images, we use conformal geometric algebra to compute the geometric transformations that yield the 3D position of the tumour, which was segmented in the ultrasound image using morphological operators. We use images from brain phantoms to evaluate the performance of the proposed methods, and our results show that they are robust.

Keywords: medical image processing; endoneurosonography; pattern recognition; generalized hough transform; particle filter

1. Introduction

1.1. Statement of the problem

Recent trends in minimally invasive brain surgery aim at using the joint acquisition of endoscopic and ultrasound (US) images – a technique that has been called endoneurosonography (ENS). Endoscopic images are of great utility for minimally invasive techniques in neurosurgery. Ultrasound images are cheap compared to tomographic and resonance magnetic images (which are very hard to obtain in an intraoperative scenario) and allow surgeons to see beyond the tissues within the brain. Another way would be to extract three-dimensional (3D) information from the combined endoscopic and ultrasound images to help surgeons better locate brain structures (such as tumours). Some work has been done in this direction, mainly in the replacement of the classic ultrasound (2D imaging methodology) with 3D ultrasound equipment (Unsgaard et al. 2006). We have focused our attention on using classic ultrasound techniques and endoscopic images to extract 3D information. We are proposing to track the ultrasound probe in the endoscopic images and then to compute the pose of the ultrasound probe in 3D space without an external method (optical or magnetic).

1.2. Outline of our method

The equipment setup is as follows: the ultrasound probe is introduced through a channel in the endoscope and is seen by the endoscopic camera. Using visual tracking equip-

ment (Polaris), we can calculate the 3D position of the endoscope's tip, and we want to know the pose of the ultrasound probe in order to have the exact location of the US sensor. This is important because the US probe is flexible and rotates around its own axis. It can also move back and forth, and since the channel is wider, there is also random movement around the channel (Figure 1). The US probe is connected to a drive unit for the micro-tip transducer; the transducer is rotated to generate a 360° beam at 10 MHz. By tracking the US probe in the endoscopic image in successive video frames, we can use multiple-view 3D estimation techniques to find the pose of the US probe axis. With this pose and the exact location of the endoscope's tip, we can estimate the 3D coordinates of the US probe tip. This is fundamental since the US image is orthogonal to the US probe's axis (see Figure 2). We know that in one small interval of time x, the ultrasound probe is fixed and the endoscopic camera undergoes a movement equivalent to an inverse motion; i.e., the endoscopic camera is fixed, and the ultrasound probe undergoes a movement. In the Figures 1 and 2, we show the 3D virtual representation of the brain phantom used for the experiments. This model is constructed with magnetic resonance images of the phantom.

1.3. Structure of the paper

This paper is organized as follows. Section 2 describes the techniques used to track the ultrasound probe in the endoscopic images. Two alternative methods are presented: the generalized Hough transform with further refinement

---

*Corresponding author. Email: edb@gdl.cinvestav.mx

Figure 1. ENS equipment setup.

through an analysis of the orientation of the edges (AE) in the image and particle filtering. Section 3 is devoted to ultrasound image processing, and a method to segment brain structures is presented. Section 4 describes how to calculate the 3D pose of the probe using Conformal geometric algebra (CGA) and multiple-view methods. We present the conclusions in Section 5.

2. Endoscopic image processing

2.1. Tracking the ultrasound probe

The goal here is tracking the US probe that is seen in the endoscopic camera images. In order to achieve this, we use two alternative techniques: generalized Hough transform with AE and particle filters. It will be reached by tracking
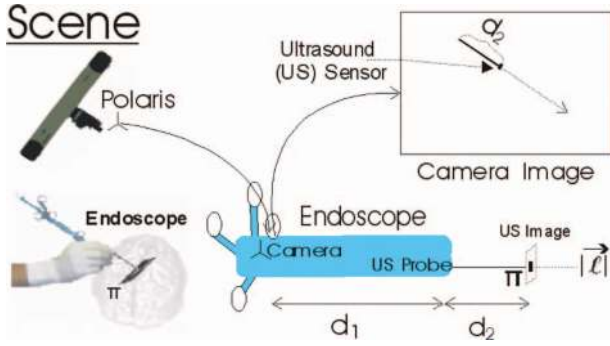
Figure 2. Scene for the virtual representation. The plane Π that is to be calculated contains the US image to be segmented and is orthogonal to the unit vector $\vec{\ell}$.

Figure 3. Generalized Hough transform: definition of R-table components.

the axis line of the US probe throughout the images. Recognizing this line is important because it allows us to compute its 3D coordinates using the projection matrix of the camera. Such a projection is done as follows: every line is back projected to form a plane in the space, which contains the line and the camera's centre. Using two consecutive images, the intersection of its respective planes will yield the line in 3D. To obtain a more accurate result, we are taking only the unit vector of this line (because the calculation of the translation in the projection matrix is up to a scalar factor). We translate the first point (the tip of the endoscope) obtained by the Polaris lecture by a distance d_2 (see Figure 2) along the direction of the unit vector in the direction of the previously obtained line, and this translated point will be the position of the US sensor in 3D, making it possible to obtain the plane Π to make the virtual representation. The Polaris system gives us the linear transformation between itself and the attached tracker to the endoscope. To calculate the linear transformation between the image frame of the endoscopic camera and the tracker, we use our hand-eye calibration method (Bayro Corrochano and Daniilidis 1996). Now we give a brief overview of the generalised Hough transform and the particle filter methods that were independently used to track the axis of the US probe.

2.2. Generalised Hough transform and analysis of edges

2.2.1. Generalised Hough transform

The generalised Hough transform (GHT) is used to find an object in an image, using a template. The matching process is based on a voting system for the possible orientations, positions and also scalings of the template (Vernon 1991). The votes are stored in an accumulator and the location in the accumulator with the greatest amount of votes will be the winner. That is, it will yield the position and orientation of the template in the image. This tracking method is robust to occlusion and illumination changes. We are using it to get a first good approximation of the template in the
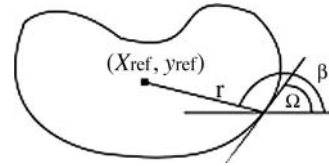
image. Before applying the GHT algorithm, the size of the endoscopic images was reduced (to decrease computational time) with a Gauss pyramid and its edges obtained using a Canny filter.

To start, we need to make an R-table that completely describes the template (see Figure 3). We select a reference point $(x_{\text{ref}}, y_{\text{ref}})$ and draw a line to any point of the boundary. Then we compute the distance and the direction from that point to the reference point, i.e., (r, β). The orientation of the boundary Ω_i will be the index to the (r, β) values. The R-table is built just one time, at the beginning of the matching process, and $(x_{\text{ref}}, y_{\text{ref}})$ is the centroid of the boundary points.

Here we present the Hough transform algorithm. We calculate orientations and positions of the template in the images:

- Represent the shape by building the R-table:
  For all points on the boundary:
  compute orientation Ω (gradient direction $+90^0$),
  compute r and β,
  add an (r, β) entry into the R-table at a location indexed by Ω.
- Quantise the Hough transform space: identify maximum and minimum values of x_{ref}, y_{ref}, ϕ and identify the total number of each one of them.
- Generate an accumulator array $A(x_{\text{ref}}, y_{\text{ref}}, \phi)$; set all values to 0.
- For all edge points (x_i, y_i) in the image:
  compute the orientation Ω (gradient direction $+90^0$),
  compute possible reference points x_{ref}, y_{ref}.
  For each table entry indexed by Ω and for each possible shape orientation ϕ:
  compute $x_{\text{ref}} = x_i + r\cos(\beta + \phi)$,
  $y_{\text{ref}} = y_i + r\sin(\beta + \phi)$,
  increment $A(x_{\text{ref}}, y_{\text{ref}}, \phi)$.
- For all cells in the accumulator array, search for maximum values.
  The coordinates x_{ref}, y_{ref} and ϕ give the position and orientation of the shape in the image.

2.2.2. Analysis of edges

We analysed the orientation of the edges (analysis of edges or AE) in the image in order to increase the accuracy of
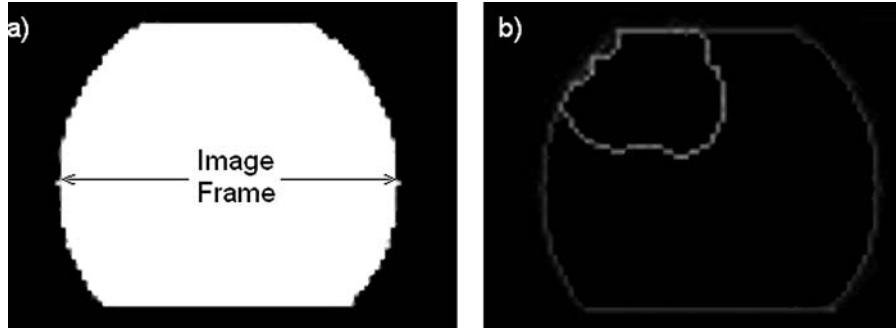
Figure 4. AE. (a) ROI in the endoscopic image. (b) Edges on the image frame.

the GHT tracking algorithm. The AE can be seen as a filter accepting pixels that belong to the US probe and rejecting the others. In order to achieve this, we used a binary mask, which describes the area of interest in the endoscopic images (ROI; see Figure 4a). This mask is computed once in the beginning of the tracking. We checked the orientations of the edges in a neighbourhood and their direct distance to the image frame; the edges that have a position that is physically possible will have a high weight, while the other edges will have a low weight. A neighbourhood of edges has a physically possible position whether it is convex and also has a region support on the image frame (see Figure 4b).

Thus, we have two estimated lines (the first was obtained using GHT and the second was obtained using AE), and these are weighted to obtain a total estimated line (EL):

$$EL = 0.65 \times GHT + 0.35 \times AE, \tag{1}$$

where the factors 0.65 and 0.35 were obtained experimentally.

The results are shown in Figure 5. The figure displays the estimated line of the ultrasound probe and also a horizontal line indicating the end tip.

2.3. Tracking with particle filters

Another approach to performing visual tracking is to use a Bayesian tracker; thus, treating motion tracking as a Bayesian state estimation problem. In order to use a Bayesian framework, one must model the object being tracked as a state vector. Additionally, a method to evaluate how well the predicted states of the state vector fit the observation is needed. The most widely used Bayesian tracker is the Kalman filter. However, Kalman filters require a Gaussian observation probability and a Gaussian posterior probability density. Our observations show that the random movement of the ultrasound probe can hardly be described as Gaussian. This situation made us look for a more general tracker: the particle filter.

2.3.1. The particle filter

Particle filters emerged from the pioneering work of (Isard and Blake 1998). Particle filters were introduced to track objects in visual clutter and can handle multimodal observation probabilities.

Let us assume that x_t represents the state (state vector) of the object at time t and that $X_t = \{x_1, \ldots, x_t\}$ represents its history over time. The vector $Z_t = \{z_1, \ldots, z_t\}$ encloses all the observations z_i up to time t. In our framework, z_t represents an endoscopic image at time t. The particle filter approximates the posterior $p(x_t|Z_t)$ of the probability distribution. The key idea in the particle filtering is to approximate the probability distribution (and consequently the posterior) by a weighted finite set of samples, the particles. Let $S = \{(s_t^{(n)}, \pi_t^{(n)})|n = 1, \ldots, N\}$ be a weighted set of N different samples. Every sample $s^{(i)}$ represents a possible object state, and a weight $\pi^{(i)}$ is associated with it. This weight represents the likelihood for the associated particle to be the true location of the target object. The weights are normalized so that $\sum_{i=1}^{N} \pi^{(n)} = 1$.

By applying Bayes' law, the posterior $p(x_t|Z_t)$ can be expressed recursively as

$$p(x_t|Z_t) = k_t\, p(z_t|x_t)p(x_t|Z_{t-1}). \tag{2}$$

Having the state vector at time $t-1$, the posterior $p(x_{t-1}|Z_{t-1})$ can be obtained by marginalizing over x_{t-1}, making it possible to obtain the distribution $p(x_t|Z_{t-1})$.

$$p(x_t|Z_{t-1}) = \int_{x_{t-1}} p(x_t, x_{t-1}|Z_{t-1})$$
$$= \int_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}|Z_{t-1}), \tag{3}$$

where the chain rule ($p(x_t|x_{t-1})$) is the dynamical model.

To perform the filtering operation, a new set of particles is created by picking up with replacement N particles from the N particles created at time $t-1$. The probability of selecting a particle $s^{(i)}$ is proportional to its normalized
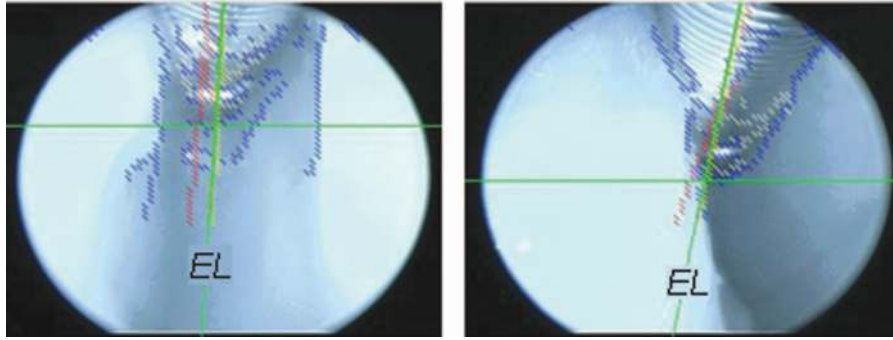
Figure 5. Estimated line obtained using GHT and analysis of edges.

weight $\pi^{(i)}$. This is done as follows: to sort the N particles at time $t - 1$ from greatest to least weight; to create $N * \pi^{(0)}$ particles, to create $N * \pi^{(1)}$ particles and so on until the N new particles have been created. Then the new particles are updated using the evolution model of the system. The new weights for the updated particles are calculated, measuring how well the object position represented by each particle fits the observation z_t at time t. After normalising the weights, the mean state is estimated at each time by $E[S] = \sum_{n=1}^{N} \pi^{(n)} s^{(n)}$.

Subsequent locations of the probe can be represented as a rotation and a translation with respect to the initial line estimate. A state vector can be represented as $s^{(i)} = [d_x^{(i)}, d_y^{(i)}, d_\theta^{(i)}]^\top$, with its components describing this translation and rotation. This model evolves in each stage according to

$$S_t = S_{t-1} + N_t, \qquad (4)$$

with N_t as white Gaussian noise.

To obtain the weight of each particle, the image of the area selected in the first picture is obtained and rotated and translated according to the particle (state vector) components. The transformed image is then compared to the observed image via the Bhattacharyya distance between their colour histograms, as described in Nummiaro et al. (2003).

2.4. Tracking results

The GHT method was applied as explained and the particle filter was applied using up to $N = 300$ particles, but little difference was observed in the results when $N > 100$. Several images and sequences have been tested. Figure 6 shows typical results for the particle filter method. Both tracking methods performed well in practice. However, we prefer the GHT method (Section 2.2) because it is faster and also because the particle filter method is nondeterministic; i.e., under identical initial conditions, the resulting outcome can differ in independent runs because it depends on random estimates. We obtained an accuracy of 94% with the GHT method.

3. Ultrasound image processing

The goal of this stage is to segment interesting structures in the brain images, such as tumours. We are using morphological operators in order to fill small holes that
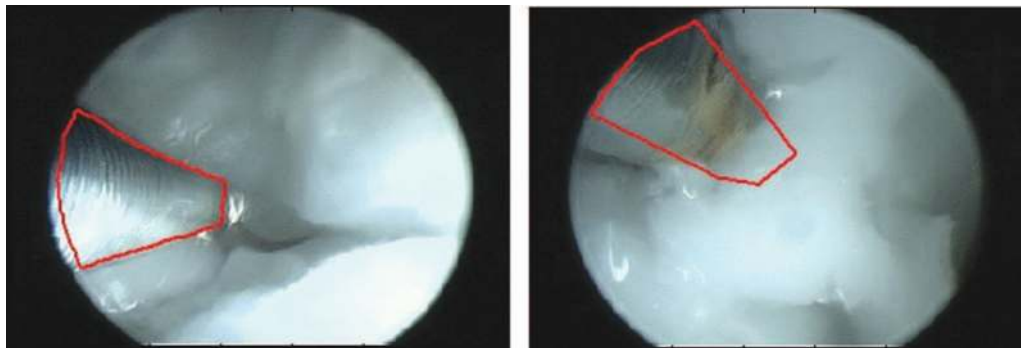

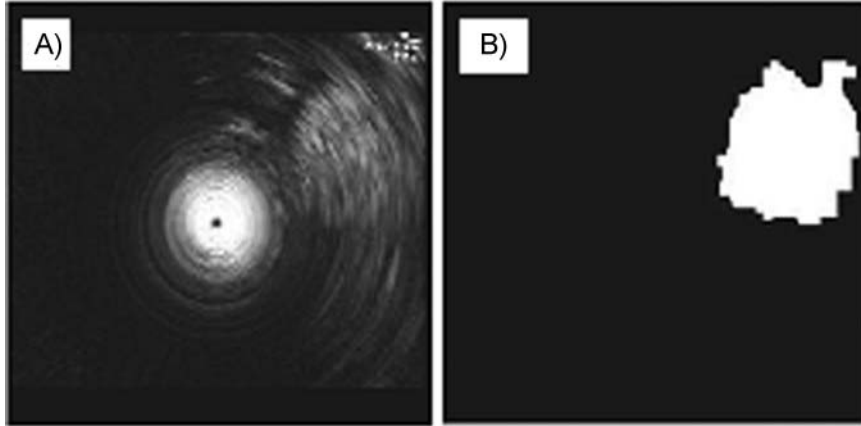Figure 6. The results for the particle filter method.

Figure 7.   A) Original US image. B) Result of segmentation.

appear due to the sub-sampling provided by the Aloka ultrasound system. The closing morphological operator of the image I with the sub-image M (structuring element) is defined as $I \cdot M = (I \oplus M) \ominus M$, where $\oplus, \ominus$ represent the dilation and erosion morphological operators, respectively.

We process the ultrasound images in the following way:

- Do a copy of the original image ([rgb]1,1,1[rgb]1,1,1 [rgb]1,1,1 for not modify it); select a region of interest (ROI), otherwise the ROI will be the complete image. The ROI will be the same for all images.
- If the ROI contains either a section of or the complete central part of the image, we exclude that part of the ROI because it only contains noise.
- Apply a threshold to the grey levels of the ROI to select only the highest levels.
- Apply the closing morphological operator to fill the holes of the ROI.
- Use the chain code to calculate the smallest areas of the ROI and eliminate them.

- Apply a logical AND operation between the ROI and the original image. The result is the segmented tumour, which is to be represented in 3D.

Figure 7, shows an example of the segmented tumour in ultrasound images.

We obtained a processing time of 0.005305 s for the morphological operators method, which is 188 fps. We recommend it for inline implementation, because it is fast and reliable.

4.   Calculating the 3D pose of the tumour

Conformal geometric algebra (CGA) represents geometric entities as points, lines, planes, spheres, pair of points, dilators, etc. in an economical and compact form (Li and Hestenes 2001). It has great potential for applications in medicine, artificial vision and robotics. CGA preserves the Euclidean metric and adds two basis vectors: $\mathbf{e}_+$, $\mathbf{e}_-$ (where $\mathbf{e}_+^2 = 1$ and $\mathbf{e}_-^2 = -1$), which are used to define the point at the origin $\mathbf{e}_0 = \frac{1}{2}(\mathbf{e}_- - \mathbf{e}_+)$ and the point at
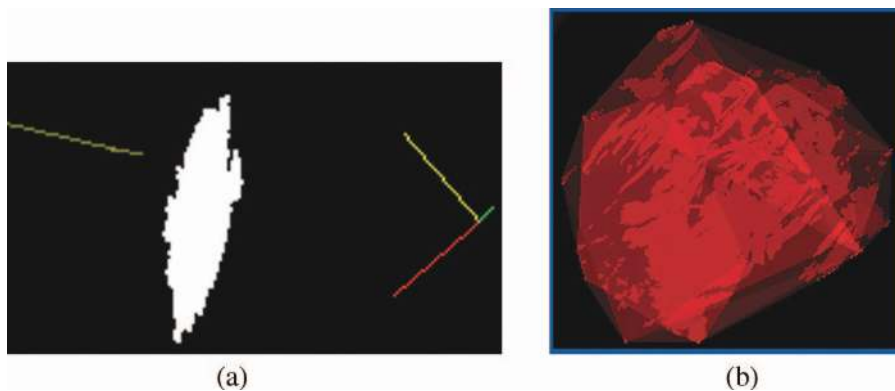

Figure 8.   a) Virtual representation of the segmented US image. b) Applying a convex hull to the results.

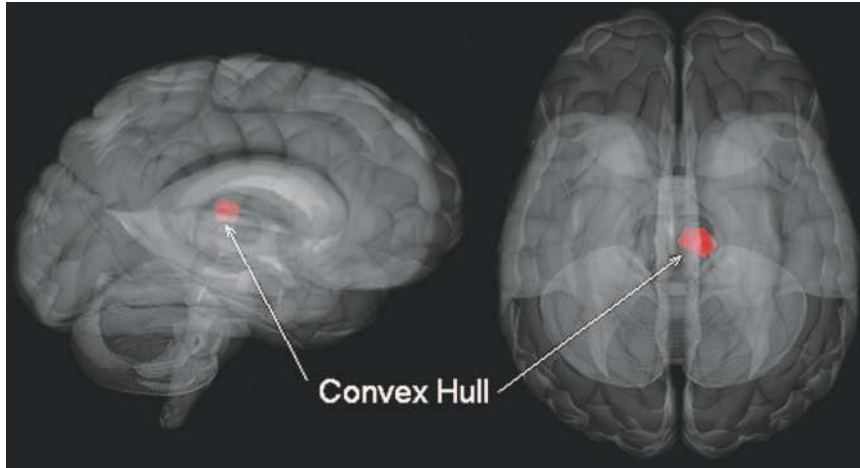Figure 9.   The tumour in a brain phantom model.

infinity $e = e_- + e_+$. The points in CGA are related to the Euclidean space by $\underline{p} = \mathbf{p} + \frac{\mathbf{p}^2}{2}e + e_0$. A sphere in dual form is represented as the *wedge*[1] of four conformal points that lies on sphere $\underline{s}^* = \underline{a} \wedge \underline{b} \wedge \underline{c} \wedge \underline{d}$. Its radius ρ and its centre $\underline{p}$ in $\mathcal{R}^3$ can be obtained using $\rho^2 = \frac{s^2}{(\underline{s}\cdot e)^2}$, $\underline{p} = \frac{s}{-(\underline{s}\cdot e)} + \frac{1}{2}\rho^2 e$. A plane in dual form is defined as a sphere, but the last point is at infinity: $\underline{\pi}^* = \underline{a} \wedge \underline{b} \wedge \underline{c} \wedge e$. A line in dual form is represented as the *wedge* of two points and the infinity point: $\underline{L}^* = \underline{a} \wedge \underline{b} \wedge e$. A line can also be calculated as the intersection of two planes: $\underline{L} = \underline{\pi_1} \wedge \underline{\pi_2}$. This equation is used to calculate the 3D line that represents the ultrasound probe's axis. As we mentioned, we only took the unit vector of this line. To achieve a translation by a distance d_2 from a point $\underline{p_1}$ in the direction of a line and to obtain $\underline{p_2}$: $T = exp(\frac{1}{2}d_2 \underline{L})$, $\underline{p_2} = T\,\underline{p_1}\widetilde{T}$

The last equation is used to find the position of the ultrasound sensor in order to put the segmented ultrasound image in 3D space; $\underline{p_1}$ represents the tip of the endoscope obtained by the Polaris tracking system and d_2 is the distance between two retro-projected points taken from our overlapped template on the found US probe in two consecutive images. One of these two points is the intersection between the horizontal line indicating the end tip and the estimated line of the US probe, as we showed in Figure 5, and the other one is the intersection between the top of our template and the estimated line of the US probe. Figure 8a shows the virtual representation of the segmented US image.

Further results are presented in Figures 8b and 9. The 3D model was built by stacking several estimates and finding the convex hull of their 3D space (Barber et al. 1996).

5.   Conclusions

We addressed the problem of obtaining 3D information from joint ultrasound and endoscopic images obtained with ENS equipment. In order to register both sources, we developed two alternative methods to locate the US probe tip in endoscopic images: one of them using the generalised Hough transform with further refinement through an analysis of the orientation edges in the image and the other one using a particle filter. Some preliminary results are shown. As for the ultrasound image, we presented a method to segment brain structures of interest. Results are shown in 3D space; the 3D information was calculated from the results obtained by the tracking process in endoscopic images. The performance of the proposed approach was demonstrated using several images that were subject to occlusions and changes in illumination and contrast, as we showed in Figures 5 and 6. The results indicate that the proposed approach is robust.

Currently, we are conducting a series of new experiments using an accelerator board in order to estimate contours and to re-create the 3D form of the tumour in real time.

6.   Notes on contributors

Rubén Machucho graduated as an electronic engineer from the Institute Technology of Yucatán in 1997. He gained his Master's degree in 2005, in Electrical Engineering from the Center of Research and Advanced Studies (CINVESTAV) in Guadalajara, México. He is a Ph.D. candidate at the

---

[1] The wedge product, which is also known as the exterior product of multivectors A_r y B_s (of degree r and s, respectively), is defined as $A_r \wedge B_s = \langle A_r B_s \rangle_{r+s}$, where $\langle A_r \rangle_i$ is the blade of degree i taken from the multi-vector of degree r.

same Institute. His research interests include medical image processing, computer vision and computer graphics.

Sergio de la Cruz Rodríguez graduated as a telecommunications and e engineer from Instituto Superior Politécnico José Antonio Echeverría, la Habana, Cuba. He obtained his master's degree in Digital Systems in 2007 at the same Institute. Since 2004 he has been working as full time professor and he is a Ph.D. candidate at the same Institute. His research interests include medical signal and image processing and computer vision.
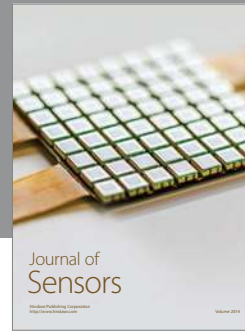
Eduardo Jose Bayro-Corrochano gained his Ph.D. in Cognitive Computer Science in 1993 from the University of Wales at Cardiff. From 1995 to 1999 he was researcher and lecturer at the Institute for Computer Science, Christian Albrechts University, Kiel, Germany and worked on applications of geometric Clifford algebra to cognitive systems. His current research interest focuses on geometric methods for artificial perception and action systems. It includes geometric neural networks, visually guided robotics, colour image processing, Lie bivector algebras for early vision and robot maneuvering. He developed the quaternion wavelet transform for quaternion multi-resolution analysis using the phase concept. He is editor and author of the following books: Geometric Computing for Perception Action Systems, E. Bayro-Corrochano, Springer Verlag, 2001; Geometric Algebra with Applications in Science and Engineering, E. Bayro-Corrochano and G. Sobczyk (Eds.), Birkahauser 2001; Handbook of Computational Geometry for Pattern Recognition, Computer Vision, Neurocomputing and Robotics, E. Bayro-Corrochano, Springer Verlag, 2004. Geometric Computing for Wavelet Transforms, Robot Vision, Learning, Control and action, E. Bayro-Corrochano, Springer Verlag, London, 2009.

References

Barber C, Dobkin D, Huhdanpaa H. 1996. The quickhull algorithm for convex hulls. ACM Trans Math Softw 22(4):469–483.

Bayro Corrochano E, Daniilidis K. 1996. The dual quaternion approach to hand-eye calibration. In: International Conference on Pattern Recognition (ICPR '96), Vol. 1, Vienna (Austria): IEEE Computer Society. pp. 318–322.

Isard M, Blake A. 1998. Condensation – conditional density propagation for visual tracking. Int J Comput Vis. 29(1):5–28.

Li H, Hestenes D, Rockwood A. 2001. Generalized homogeneous coordinates for computational geometry. In: Geometric Computing with Clifford Algebra, G. Sommer (Ed.), Springer-Verlag: London, UK. pp. 27–59.

Nummiaro K, Koller-Meier E, Van Gool L. 2003. An adaptive color-based particle filter. Image Vision Comput. 21(1):99–110.

Unsgaard G, Rygh M, Selbekk T, Müller T, Kolstad F, Lindseth F, Nagelhus Hernes T. 2006. Intra-operative 3D ultrasound in neurosurgery. Acta Neurochir. 148(3):235–253.

Vernon D. 1991. Vernon's machine vision: automated visual inspection and robot vision. Prentice Hall: International (UK) Ltd.