# Contour Tracking in Complex Environments using Wavelet Basis

Fernando Pérez Nava[1] Antonio Falcón Martel[2]
[1] Dep. de Estadística, Inv. Op. y Computación.
Universidad de La Laguna. Tenerife. Spain
[2] Dep. de Informática y Sistemas,
Universidad de Las Palmas de Gran Canaria.Gran Canaria.Spain
e-mail: [1]fdoperez@ull.es, [2]afalcon@dis.ulpgc.es

**Abstract**

In this paper, a new tracking algorithm for active contours using wavelets is presented. First it is shown how to construct contour descriptions based on a multiscale representation of planar shapes using wavelet basis. Second, in order to model smooth contour transitions between image frames, probabilistic shape priors for modelling contour deformation using wavelets are presented. And finally this new formulation is applied to the problem of tracking a contour in a cluttered environment using stochastic models to predict contour location and appearance in successive image frames. These three components are integrated in the Condensation (Conditional Density Propagation) tracking algorithm which is specially designed to work in cluttered environments. Computational results are given for two real image problems and show that this formulation successfully tracks the objects in the image sequences.

**Keywords:** pattern recognition, computer vision, contour modelling, wavelets, tracking.

## 1 Introduction

Active shape models [1] encompass a variety of forms, principally snakes, deformable templates and dynamic contours. The framework has evolved from the principles of the "snake" which is an elastic model for shapes in motion that can be coupled to image features.

Another theme that is related to the snakes idea has been the representation of geometric prior information which can be incorporated into the tracker by means of a template. Templates have been used effectively in non-dynamic shape-fitting processes Some include statistical learning of shape variations.

Active contours are based on prior geometric models which are defined in terms of a low dimensional parametric representation of shape with B-splines. Prior geometric models are then posed in a probabilistic setting leading to probabilistic models of shape.

In [2] a new probabilistic model of shape was proposed for active contours. It is based on a compact and invariant representation of shape using wavelet basis.

There are a number of salient features in wavelet transforms that make wavelet-domain statistical contour processing attractive:

-Locality: Each wavelet coefficient represents the signal content localized in spatial location and frequency
-Multiresolution: The wavelet transform analyzes the signal at a nested set of scales
-Energy Compaction: The wavelet transform of real-world signals tend to be sparse. A wavelet coefficient is large only if singularities are present within the support of the wavelet basis
-Decorrelation: The wavelet transform of real world signals tend to be approximately decorrelated.

The Locality and Multiresolution properties enable the wavelet transform to efficiently represent sharp contour changes with large coefficients, resulting in the Compaction property. The Compaction and Decorrelation properties simplify the statistical modelling in wavelet domain as compared with a direct spatial domain modelling. Because most of the wavelet coefficients tend to be small, we need to model only a small number of coefficients. This is of particular importance in real time applications.

In this work we will show how to integrate the wavelet contour description and its probabilistic priors in a contour tracker using the Condensation [1] framework.

This paper is divided in five parts: in Section 2 we show a compact and invariant contour coding using the wavelet transform. Then in Section 3 the wavelet based probabilistic shape model is presented and its relation with Besov spaces is established. In Section 4 we introduce the Condensation algorithm. In Section 5 we show several tracking application of this formulation and finally in Section 6 we present the conclusions of this work.

## 2. Wavelet contour description

A wavelet basis uses translations and dilations of a scaling function $\phi$ and a wavelet function $\varphi$. A 1-D function f can be expressed as:

$$f(x)=\sum_{k\in Z} c_{j_0,k}\, 2^{\frac{j_0}{2}}\phi(2^{j_0}x-k)+\sum_{j=j_0}^{\infty}\sum_{k\in Z} d_{j,k}\, 2^{\frac{j}{2}}\varphi(2^{j}x-k) \tag{1}$$

Let then $\mathbf{r}(s)$ = ( x($s$), y(s)) be a discrete parametrized closed planar curve that represents the shape of an object of interest. If the wavelet transform is applied independently to each of the x($s$), y(s) functions, we can describe the planar curve in terms of a decomposition of $\mathbf{r}(s)$:

$$\mathbf{r}(s)=\sum_{k\in Z}\mathbf{c}_{j_0,k}\, 2^{\frac{j_0}{2}}\phi(2^{j_0}s-k)+\sum_{j=j_0}^{\infty}\sum_{k\in Z}\mathbf{d}_{j,k}\, 2^{\frac{j}{2}}\varphi(2^{j}s-k)\quad \mathbf{c}_{j,k}=\begin{pmatrix} c_{j,k;x}\\ c_{j,k;y}\end{pmatrix},\mathbf{d}_{j,k}=\begin{pmatrix} d_{j,k;x}\\ d_{j,k;y}\end{pmatrix} \tag{2}$$

where subindex x and y represent coordinate function pertenence.

### 2.1 Compact coding using the wavelet coefficients

Since one of our goals is to obtain a compact representation of contour it is necessary to develop a contour simplification strategy. In [2] we propose to threshold the $\mathbf{d}_{j,k}$ vector based on its norm. It can be show that this yields an invariant simplification under Euclidean similarities that overperforms other simplification methods based on B-splines or Fourier descriptors.

## 3. Wavelet based Probabilistic Modelling of Curve Deformation

The simplest wavelet transform statistical models [3] are obtained by assuming that the coefficients are independent. Under the independence assumption, modelling reduces to simply specifying the marginal distribution of each wavelet coefficient.

Wavelet coefficients are generally modelled using the generalized gaussian distribution, in this work the usual gaussian distribution will be used as an approximation

For the tractability of the model, all coefficients at each scale are assumed to be independent and identically distributed. That is:

$$\mathbf{d}_{j,k} = \begin{pmatrix} d_{j,k,x} \\ d_{j,k,y} \end{pmatrix} \sim \mathbf{N}_2(\overline{\mathbf{d}}_{j,k}, \sigma^2 \mathbf{I}) \quad (3)$$

where **I** denotes the identity matrix.

Assuming a exponential decay of the variances, the final model is:

$$\mathbf{d}_{j,k} \sim \mathbf{N}_2(\overline{\mathbf{d}}_{j,k}, 2^{-2j\beta} \sigma_D^2 \mathbf{I}) \quad (4)$$

In order to complete the model definition we have to specify the distribution for the coefficient associated with the scaling function $\mathbf{c}_{0,0}$. This coefficient is associated with a rigid traslation of shape and we will assume that it is normally distributed and independent of the non-translation components $\mathbf{d}_{j,k}$.

$$\mathbf{c}_{0,0} = \begin{pmatrix} c_{0,0,x} \\ c_{0,0,y} \end{pmatrix} \sim \mathbf{N}_2(\overline{\mathbf{c}}_{0,0}, \sigma_C^2 \mathbf{I}) \quad (5)$$

We will use this distributions to model smooth changes of shape between frames. A justification for the proposed model comes from the following theorem[3]:

**Theorem**

Let f(x) be decomposed in wavelet coefficients and suppose each coefficient is independently and identically distributed as:

$$d_{j,k} \sim N(0, \sigma_j^2) \text{ with } , \sigma_j = 2^{-j\beta} \sigma_0 \quad (6)$$

with $\beta > 0$ and $\sigma_0 > 0$ then, for $0 < p,q < \infty$, the realizations of the model are almost surely in the Besov Space $B^{\alpha}_{q}(L_p(I))$ if and only if $\beta > \alpha + 1$.

Besov spaces are smoothness spaces: roughly speaking, the parameter $\alpha$ represents the number of well behaved derivatives of f.

With the above assumptions we can then define a prior probabilistic shape model for curve deformation as:

$$p(\mathbf{X}) \propto \exp\left( -\frac{1}{2} (\mathbf{X} - \overline{\mathbf{X}})^T \Sigma^{-1} (\mathbf{X} - \overline{\mathbf{X}}) \right),$$
$$\mathbf{X} = (c_{0,0,x}, d_{0,0,x}, \ldots, d_{j,k,x}, \ldots d_{J-1,2^{J-1}-1,x}, c_{0,0,y}, d_{0,0,y}, \ldots, d_{j,k,y}, \ldots d_{J-1,2^{J-1}-1,y}), \quad (7)$$
$$j = 0..J-1,\ k = 0..2^j - 1$$

where $n=2^J$ is the number of points in the discretized curve, **X** is a vector of $2n$ wavelet coefficients and $\Sigma$ is a diagonal matrix with the above defined variances. Using the theorem we can see that smooth deformations of the curve are prefered with this model.

In the following figure we can see a shape (in discontinous line) with some realizations of the probabilistic model for various values of parameter $\beta$. As expected, when the parameter increases smoother deformations arise. Around the figure we can see in light grey a 99% confidence interval for the points in the curve.

*Fig. 1 Realizations of the probabilistic model. Parameter values are β=0 (no deformation smoothing) for the left image and β=1.6. in the right image. In light grey a 99% confidence interval for the points in the curve is shown.*

A property of the probabilistic model that we will use is the following:

**Theorem**

Let a curve be described as (7) then the mean square displacement along the curve is given by:

$$\overline{\rho}^2 = \frac{\mathrm{Trace}(\Sigma)}{n} \tag{8}$$

Now we present in the next section an application of this probabilistic modelling to the contour tracking problem.

# 4. Contour tracking

This section describes the use of the Condensation algorithm with the wavelet probabilistic formulation. This algorithm is applied to cases where there is a substantial clutter in the background. In this case the probabilistic density of the curve is multimodal and therefore not even approximately gaussian. However change between successive frames can be expected to be smooth so the above formulation can be used.

### 4.1 Dynamic model

An adequate statistical framework for motion tracking must be able to provide a prior for possible motions, in the broad sense of a rigid motion plus a deformation of shape. We will use a second order AR process in shape space to model motion:

$$\mathbf{X}(t_k) - \overline{\mathbf{X}} = \mathbf{A}_2(\mathbf{X}(t_{k-1}) - \overline{\mathbf{X}}) + \mathbf{A}_1(\mathbf{X}(t_{k-2}) - \overline{\mathbf{X}}) + \Sigma^{1/2}\mathbf{w}_k, \quad \mathbf{w}_k \sim \mathbf{N}_{2n}(\mathbf{0}, \mathbf{I}) \tag{9}$$

where $\overline{\mathbf{X}}$ represents a mean shape and $\Sigma$ the noise covariance.

Therefore motion is decomposed as a deterministic drift plus a diffusion process that is assured to be smooth using the above derivations.

### 4.2 Parameter determination

In order to make the model usable it is necessary to show how to determine the parameters $\mathbf{A}_2$, $\mathbf{A}_1$, $\Sigma$. We first decompose motion in several orthogonal linear subspaces $P_i$ determined by their projection matrix $\mathbf{P}_i$. Typically these subspaces spaces are translation, rotation and deformation (euclidean similarities) or translation, affine change and deformation (planar affine motion).

Therefore we can write:

$$\mathbf{X}(t_k) - \overline{\mathbf{X}} = \sum_i \mathbf{P}_i(\mathbf{X}(t_k) - \overline{\mathbf{X}}) \tag{10}$$

and we will model dynamics into each subspace:

$$\mathbf{A}_1 = \sum_i a_i^1 \mathbf{P}_i, \quad \mathbf{A}_2 = \sum_i a_i^2 \mathbf{P}_i, \quad \boldsymbol{\Sigma}^{1/2} = \sum_i b_i\, \mathbf{P}_i, \quad a_i^1, a_i^2, b_i \in \mathrm{R} \tag{11}$$

and use the following theorems:

**Theorem**

Let contour dynamics be given by (9) and (11) and suppose that a steady state distribution exists. Then the distribution is normal with mean $\overline{\mathbf{X}}$ and its covariance matrix $\mathbf{C}_\infty$ verifies:

$$\mathrm{Trace}(\mathbf{C}_\infty)=\Sigma_i\ \mathrm{Trace}(\mathbf{C}_{\infty i}) \tag{12}$$

Where $\mathbf{C}_{\infty i}$ is the covariance of the steady-state into subspace $P_i$

**Theorem**

Let contour dynamics be given by (9) and (11) and suppose that a steady state distribution exists into subspace $P_i$. Then its distribution is normal with mean $\mathbf{P}_i\overline{\mathbf{X}}$ and its covariance matrix $\mathbf{C}_{\infty i}$ verifies:

$$\mathrm{Trace}(\mathbf{C}_{\infty i}) = \frac{(b_i)^2(1-a_i^2)}{(a_i^2+1)(a_i^2-1+a_i^1)(a_i^2-1-a_i^1)}\mathrm{Trace}(\mathbf{P}_i\ \boldsymbol{\Sigma}) \tag{13}$$

**Corollary**

Let contour dynamics be given by (9) and (11) and suppose that a steady state distribution exists into subspace $P_i$. To obtain a mean displacement $\overline{\rho}_i$ we must set $b_i$ to:

$$b_i = \sqrt{n}\,\overline{\rho}_i\sqrt{\left(1-\left(a_i^2\right)^2-\left(a_i^1\right)^2-\frac{2a_i^2\left(a_i^1\right)^2}{1-a_i^2}\right)\frac{1}{\mathrm{Trace}(\mathbf{P}_i\boldsymbol{\Sigma})}} \tag{14}$$

This leads us to determine the parameters associated with the random noise if we can estimate the deterministic motion and the mean square displacement of shape.

In case no steady-state distribution exists we can use the following theorem:

**Theorem**

Let contour dynamics be given by (9) and (11) and suppose that no steady state distribution exists into subspace $P_i$. Then the mean displacement $\overline{\rho}_i(k)$ into subspace $P_i$ at time $t_k$ verifies:

$$\overline{\rho}_i(k) \approx \sqrt{\frac{n}{3}Trace(\mathbf{P}_i\boldsymbol{\Sigma})}\ b_i\ k^{3/2} \tag{15}$$

### 4.3 External observation model

The effect of an external observation $\mathbf{Z}(k)$ is to superimpose a reactive effect on the diffusion in which the density tends to peak in the vicinity of observations so that we can use the posterior probability $p(\mathbf{X}|\mathbf{Z})$ to estimate $\mathbf{X}$ after using evidence $\mathbf{Z}$. We must therefore model the observation density $p(\mathbf{Z}|\mathbf{X})$.

In one dimension observations reduce to a set of scalar positions $\mathbf{z}=(z_1,z_2,...,z_m)$ and the observation density has the form $p(\mathbf{z}|x)$ where $x$ is an scalar position. We will use [1]:

$$p(\mathbf{z} \mid x) \propto 1+\frac{1}{\sqrt{2\pi}\sigma\alpha}\sum_m e^{-\frac{(z_m-x)^2}{2\sigma^2}} \tag{16}$$

where σ represents uncertainty in the position of the $z_i$ and α balances the probability of no one of the $z_i$ to be the corresponding feature in the curve.

In a two dimensional image **Z** is, in principle, the entire set of features in the image, however in order to achieve real time perfomance the two dimensional observation density is evaluated as the product of one-dimensional densities along several curve normals.

### 4.4 The Condensation algorithm

The Condensation algorithm allows us to approximate the posterior density $p(\mathbf{X}|\mathbf{Z})$ using factored sampling it proceeds as follows:

First we will rewrite the AR model (9):

$$\begin{gathered}\tilde{\mathbf{X}}(t_k)-\bar{\tilde{\mathbf{X}}} = \mathbf{A}(\tilde{\mathbf{X}}(t_k)-\bar{\tilde{\mathbf{X}}})+\mathbf{B}\mathbf{w_k}, \\ \tilde{\mathbf{X}}(t_k)=\begin{pmatrix}\mathbf{X}(t_{k-1})\\ \mathbf{X}(t_k)\end{pmatrix}, \bar{\tilde{\mathbf{X}}}=\begin{pmatrix}\bar{\mathbf{X}}\\ \bar{\mathbf{X}}\end{pmatrix}, \mathbf{A}=\begin{pmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{A}_2 & \mathbf{A}_1\end{pmatrix}, \mathbf{B}=\begin{pmatrix}\mathbf{0}\\ \mathbf{B}_0\end{pmatrix}\end{gathered} \tag{17}$$

an we apply the following steps:

**Step 0**

Generate a sample set $S=\{\mathbf{s}_0^{(i)},\pi_0^{(i)},c_0^{(i)}, i=1...N\}$ for time step 0 where:

$\mathbf{s}_0^{(i)}$ is a sample from the prior distribution of the curve

$\pi_0^{(i)}$ is a probability for this curve to be chosen from S (initially $1/N$)

$c_0^{(i)}$ is the cumulative probability distribution for S

Set the time step k=1

**Step 1**

From the previous sample set S construct a new sample set $S_{new}$ as follows:

Set $i$=1

While $i \leq N$ do begin

**Step 1.1 (Selection)**

Select a sample $s_k'^{(i)}$ as follows:

- Generate a random number r uniformly distributed in [0,1]
- Find by binary subdivision the smallest j for which $c_{k-1}^{(j)} \geq$ r
- Set $s_k'^{(i)}= s_{k-1}'^{(j)}$

**Step 1.2 (Prediction)**

Generate a prediction by sampling from:

$$\mathbf{s}_k^{(i)}-\bar{\tilde{\mathbf{X}}} = \mathbf{A}(\mathbf{s'}_k^{(i)}-\bar{\tilde{\mathbf{X}}})+\mathbf{B}\mathbf{w_k} \tag{18}$$

**Step 1.3 (Measurement)**

Measure and weight the new position in terms of the observed features $\mathbf{Z}_k$:

$$\pi_k^{(i)} = p(\mathbf{Z_k} \mid \widetilde{\mathbf{X}}_k = \mathbf{s}_k^{(i)}) \quad (19)$$

Store these values in $S_{new}$. Increment $i$

end while

Normalize weights so that $\sum_i \pi_k^{(i)} = 1$, and evaluate cumulative probabilities as: $c_k^{(0)} = 0, \quad c_k^{(i)} = c_k^{(i-1)} + \pi_k^{(i)}, i = 1...N$

Set S = Snew. Increment $k$. Goto Step1

# 5. Computational results

A set of experiments have been carried out to test the validity of the approach both in an indoor scene and an outdoor scene. The number of wavelet coefficients used has been 16, the wavelet function used is Daubechies LA(8) and the number of elements in set S has been 250.

In the first example (Fig. 2) motion is modelled as traslation plus deformation with parameters:

- Traslation subspace $P_1$: $a_1^1 = 2, \quad a_1^2 = -1, \quad \overline{\rho} = 3$ (No steady-state)
- Deformation subspace $P_2$: $a_2^1 = 0, \quad a_2^2 = 0, \quad \overline{\rho} = 0.5$

and the smoothness parameter for frame to frame deformation has been $\beta$=2.25

To visualize the pdf a set of 20 curves has been sampled from the contour distribution in all frames. As more clutter appears (frame 4) the uncertainty about position grows leading to the curves being more disperse around the hand. In frame 5 we can see how a false matching appears (the distribution becomes multimodal), however the condensation algorithm recovers as the hands goes on moving as can be seen in frame 6.

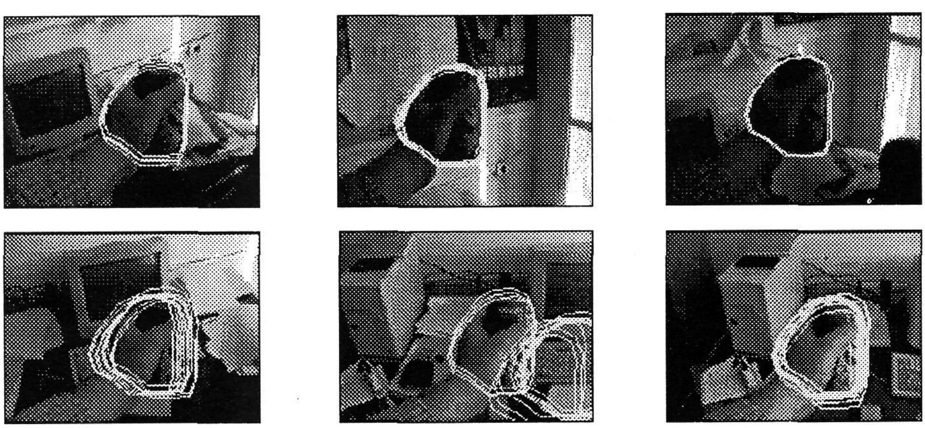

*Fig 2 An indoor scene. Frames are numbered 1-6 from top to bottom and left to right*

In the second case (Fig. 3) the background is cluttered with changes between light and shadows and there are several moving elements interacting with the person being tracked. Parameters in this case have been:

- Traslation subspace $P_1$: $a_1^1 = 2$, $a_1^2 = -1$, $\bar{\rho} = 1$ (No steady-state)
- Deformation subspace $P_2$: $a_2^1 = 0$, $a_2^2 = 0$, $\bar{\rho} = 0.1$

and the smoothness parameter for frame to frame deformation has been $\beta$=2.25

As we can see the head of the girl is in general successfully tracked. In frame 5 the dynamic model fails because the girl suddenly stops leading to the curves being more disperse around the head. However the condensation algorithm recovers as the girl goes on moving as can be seen in frame 6.

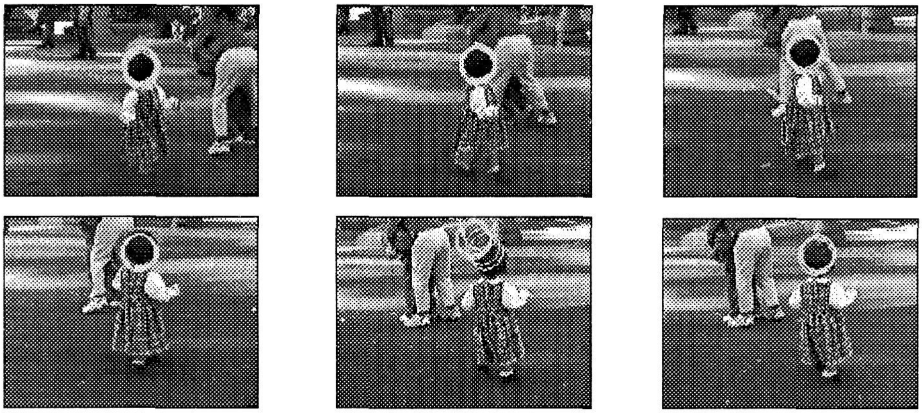

*Fig 3. An outdoor scene. Frames are numbered 1-6 from top to bottom and left to right*

# 6. Conclusions

In this work, a tracking algorithm for active contours using wavelets and the Condensation algorithm has been presented. It is based on a multiscale representation of planar shapes using wavelet basis and a probabilistic prior that favours smooth frame to frame deformations. It is shown how to use this formulation to predict contour location and appearance in successive image frames using stochastic models. These components are integrated in the Condensation tracking algorithm and computational results show that this formulation successfully tracks the objects in the image sequences.

# References


[1] A.Blake A. and M. Isard M. *Active Contours*, Springer, London, 1998.

[2] F. Pérez and A. Falcón, "Planar shape representation based on multiwavelets", *Proc. Eusipco 2000*. Tampere. Finland

[3] H. Choi and R. Baraniuk, "Wavelet Statistical Models and Besov Spaces" , *Proceedings of SPIE Technical Conference on Wavelet Applications in Signal Processing VII* 1999