

REGULARIZACIÓN DE SITUACIONES DE AMBIGÜEDAD EN EL REGISTRO DE IMÁGENES MEDIANTE MODELOS DUALES

Juan Morales Sánchez, Jorge Larrey Ruiz, José Luis Sancho Gómez
Departamento de las Tecnologías de la Información y las Comunicaciones
Universidad Politécnica de Cartagena
e-mail: {juan.morales;jorge.larrey;josel.sancho}@upct.es

Abstract—Usually, the most critical step involved in image registration is the image matching. A common methodology when estimating the mapping that geometrically relates two images typically consists of two separate and sequential stages: initial feature matching estimation, and regularization for propagating this matching over all image areas. Parametric models representing fuzzy matching regions are proposed in order to support the initial matching. The classical approach has a main drawback, namely the detection of the common features is ambiguous when there is more than one likely matching. To alleviate this problem, the use of dual models to represent the high similarity regions is also proposed in this paper. An averaged POCS (projection onto convex sets) procedure, combined with regularization based on deformable kernels, is used to solve the multiple-choice dilemma. Implementation of these parametrization and regularization steps is described throughout the paper. The proposed approach is tested on a stereo-pair that presents multiple choices of similar likelihood, with successful results.

I. INTRODUCCIÓN

Sean $I(x, y)$ y $J(x, y)$ dos imágenes, siendo $J(x, y)$ el resultado de la distorsión geométrica de $I(x, y)$ mediante una función espacial biyectiva (continua y diferenciable), $\Psi(x, y) \equiv \{\Psi_x(x, y), \Psi_y(x, y)\}$. En este escenario, el objetivo es encontrar la estimación óptima de este mapeado, Ψ_o , que relaciona geoméricamente las dos imágenes, i.e. $J(x, y) = I(\Psi_x(x, y), \Psi_y(x, y))$. La metodología que proponemos consta de dos fases secuenciales:

- 1) Parametrización, para obtener un conjunto de modelos paramétricos de correspondencia entre $I(x, y)$ y $J(x, y)$, que llamaremos Ω .
- 2) Regularización, para obtener la función de mapeado óptima, Ψ_o , a partir de los modelos paramétricos Ω .

Este artículo está organizado como sigue: en la sección II se presentan las ideas básicas para conseguir una mejor correspondencia de las imágenes mediante el uso de modelos duales; en la sección III se muestra una comparativa de resultados que prueba la validez de la aproximación adoptada, y que ilustra la mejora en la estimación; finalmente, en la sección IV se exponen las conclusiones.

II. MÉTODO PROPUESTO

Asumamos que calculamos un modelo paramétrico $\Omega(\mathbf{x})$ para cada punto $\mathbf{x} \equiv (x, y) \in \Phi$, donde Φ define el dominio de la imagen donde la correspondencia es más probable. Este modelo paramétrico estará presente por tanto en los puntos donde haya evidencia de rasgos comunes entre ambas imágenes, y quedará indeterminado en los pixels restantes.

A. Modelos paramétricos de similitud

Para cada pixel de la imagen $I(x, y)$ se calcula un Mapa de Similitud (MS) de tamaño $(2L+1)$ pixels. Este mapa contendrá la distribución espacial de la similitud entre un bloque de $I(x, y)$ centrado en el pixel $\mathbf{c} \equiv (c_x, c_y)$ y un conjunto de bloques de igual tamaño de $J(x, y)$ (longitud de búsqueda $(2L+1)$). Para calcular el MS se utiliza el coeficiente de similitud $\rho(\mathbf{x})$. Este operador está directamente relacionado con el coeficiente de correlación normalizado de Pearson, $r(\mathbf{x})$, y pseudo-inversamente relacionado con la diferencia cuadrática media normalizada entre dos bloques de imagen, $d(\mathbf{x})$ [1]-[2].

$$\rho(\mathbf{x}) = \begin{cases} \gamma(\mathbf{x}), & \text{si } \gamma(\mathbf{x}) > \lambda \\ 0, & \text{si } \gamma(\mathbf{x}) \leq \lambda \end{cases} \quad \gamma(\mathbf{x}) = \frac{r(\mathbf{x})}{1 + \kappa d(\mathbf{x})} \quad (1)$$

donde κ es una constante positiva que establece la importancia de la correlación frente a la diferencia, y λ es una constante positiva en el intervalo $[0,1]$ que actúa como umbral del coeficiente de similitud. Podemos reemplazar las estructuras de correspondencia bien definidas del MS por un modelo simple que represente las posibles regiones correspondientes. De este modo, la información presente en el MS queda comprimida en un sencillo modelo paramétrico con pocos grados de libertad. Se puede comprobar que en este contexto son suficientes tres modelos paramétricos distintos [2]. En este artículo se introduce el concepto de modelos duales, que consisten en un par de modelos simples, para representar las regiones disjuntas no nulas susceptibles de ser parametrizadas de forma independiente.

Modelo 0 (punto). El MS queda descrito por su centro de masas. Este modelo se aplicará sobre detalles aislados de la imagen. El error cuadrático medio del ajuste a este modelo se

calcula como $E_0 = \sum_{\mathbf{x}} \hat{\rho}(\mathbf{x}) \|\mathbf{v} - \mathbf{x}\|^2$, donde $\hat{\rho}(\mathbf{x})$ es el mapa de similitud normalizado, y \mathbf{v} es el centro de masas calculado.

Modelo 1 (curva). Es una curva cuadrática de coeficientes $\mathbf{a} \equiv \{a, b, c\}$ rotada un ángulo θ . Este modelo aparece fundamentalmente en bordes de objetos. Se proponen los siguientes dos pasos:

- 1) Obtener la orientación principal del MS, θ , por medio de un algoritmo PCA lineal.
- 2) Ajuste por mínimos cuadrados de una función convexa $f(x, \mathbf{a})$ al MS rotado, siendo \mathbf{a} los parámetros que definen esta función. El error de aproximación resulta ser $E_1 = \sum_{\mathbf{x}} \hat{\rho}(\mathbf{x}) (y_\theta - f(x_\theta, \mathbf{a}_0))^2$, donde \mathbf{a}_0 son los parámetros que minimizan E_1 , y x_θ e y_θ son los ejes del plano rotados.

Modelo 2 (sin restricciones). Este modelo no incluye información sobre el MS y por tanto no aplica ninguna restricción sobre el proceso de optimización global.

El proceso de seleccionar el modelo que mejor representa el MS de entre los tres posibles candidatos se basa en la utilización de dos umbrales de error, K_0 y K_1 , que se corresponden con los errores de ajuste E_0 y E_1 respectivamente. Si $E_0 < K_0$, el modelo es un punto; si $E_0 \geq K_0$ y $E_1 < K_1$, el modelo es una curva; en otro caso, el modelo es sin restricciones.

Modelos duales. La detección de modelos duales es la contribución principal de este artículo, dado que pueden aparecer como posibles correspondencias regiones de la imagen no conectadas, debido a la repetitividad típica en el contenido de las imágenes. La obtención de varios modelos en un mismo MS puede lograrse detectando en primer lugar las diferentes áreas de alta similitud que no estén conectadas unas con otras, y calculando a continuación sobre cada región los modelos paramétricos. Las regiones disjuntas que existen en el MS se obtienen mediante procesado morfológico de imagen sobre la imagen binaria del MS. Se aplica un criterio de "vecindad a 8", el cual asume que los pixels vecinos en diagonal están morfológicamente unidos.

B. Fase de regularización

El proceso de forzar las restricciones se efectúa mediante la proyección sobre el modelo. La proyección sobre modelos múltiples no es un tema enteramente resuelto, y va a ser tratado en detalle. Esta proyección necesita ser consistente con ambos modelos, y que no suponga un descarte drástico de ninguno de ellos. Para satisfacer estos requerimientos, se propone una técnica de proyección combinada con una regularización basada en núcleos deformables. El siguiente funcional de error,

$$S(\Psi) = \sum_{\mathbf{x} \in \Phi} \|\mathcal{P}\{\Omega, \Psi\} - \Psi\|^2 + \lambda \sum_{\mathbf{x}} \|\nabla^\sigma \Psi\|^2, \quad (2)$$

contiene un compromiso entre proximidad al modelo paramétrico Ω y suavizado del mapeado estimado Ψ . El objetivo del problema de correspondencia de imágenes es obtener la estimación óptima del mapeado Ψ_0 que minimiza $S(\Psi)$. El funcional (2) consta de dos partes: la primera

representa la compatibilidad del mapeado Ψ con los modelos Ω (siendo $\mathcal{P}\{\Omega, \Psi\}$ el operador de proyección), mientras que el segundo término es la energía de la σ -ésima derivada espacial de la función de mapeado estimada. Una opción para hallar la solución de este problema es usar un sencillo algoritmo de descenso de gradiente [3]-[5], cuya descripción detallada queda fuera del ámbito de este artículo.

C. Proyección sobre modelos simples

La proyección $\mathcal{P}\{\Omega, \mathbf{x}_0\}$ de un punto $\mathbf{x}_0 \equiv (x_0, y_0)$ sobre un modelo paramétrico Ω resulta ser el punto \mathbf{v} perteneciente al modelo que se encuentra más próximo en distancia euclídea a \mathbf{x}_0 . A continuación se muestra el procedimiento para realizar esta operación para cada uno de los tres modelos posibles.

Modelo 0 (punto). La proyección de \mathbf{x}_0 sobre el modelo $\Omega = \{0, \mathbf{v}\}$ resulta ser el punto que define el modelo

$$\mathcal{P}\{\{0, \mathbf{v}\}, \mathbf{x}_0\} = \mathbf{v}. \quad (3)$$

Modelo 1 (curva). La proyección de \mathbf{x}_0 sobre el modelo $\Omega = \{1, \theta, \mathbf{a}\}$ se consigue con tres pasos: rotar \mathbf{x}_0 un ángulo θ , obteniéndose el punto $\mathbf{x}_\theta \equiv (x_\theta, y_\theta)$; proyectar \mathbf{x}_θ sobre la parábola explícita definida por \mathbf{a} ; rotar un ángulo θ la proyección resultante, $\hat{\mathbf{x}}_\theta$, dando como resultado el punto \mathbf{v} . La operación de la proyección puede obtenerse como sigue:

$$\hat{x}_\theta = \arg \min_x \{(x - x_\theta)^2 + (ax^2 + bx + c - y_\theta)^2\}, \quad (4)$$

$$\hat{y}_\theta = a\hat{x}_\theta^2 + b\hat{x}_\theta + c. \quad (5)$$

La solución de (4) es múltiple (puede obtenerse analíticamente con la fórmula de Cardano), y consta de una solución real y dos complejas. Asumiremos que la solución real es el valor de \hat{x}_θ .

Modelo 2 (sin restricciones). Sobre los puntos \mathbf{x} sin restricciones, i.e. $\mathbf{x} \notin \Phi$, no se realiza proyección, y por tanto $\mathbf{v} = \mathbf{x}_0$.

D. Proyección sobre modelos duales

Si consideramos modelos duales en Ω , i.e. $\Omega = \{\Omega_1, \Omega_2\}$, necesitamos un método de proyección de los puntos sobre ellos. Para aplicar las restricciones del modelo dual Ω al punto \mathbf{x}_0 , se propone una idea de proyección balanceada. De este modo, el punto proyectado \mathbf{v} estará en concordancia, tanto como sea posible, con ambos modelos y, por otro lado, no supondrá la elección prematura de uno de los modelos, sino un compromiso entre ellos. La técnica propuesta sitúa el punto \mathbf{x}_0 en una nueva posición dentro de la región del espacio descrita conjuntamente por ambos modelos (espacio de proyección), minimizando la variación de esta posición. La expresión (6) muestra cómo calcular las coordenadas del punto proyectado \mathbf{v} atendiendo a todos estos requerimientos:

$$\mathbf{v} = \mathbf{v}_1 \frac{d_2}{d_1 + d_2} + \mathbf{v}_2 \frac{d_1}{d_1 + d_2}, \quad (6)$$

donde $\mathbf{v}_1 = \mathcal{P}\{\Omega_1, \mathbf{x}_0\}$ y $\mathbf{v}_2 = \mathcal{P}\{\Omega_2, \mathbf{x}_0\}$ son las proyecciones del punto \mathbf{x}_0 sobre cada uno de los modelos considerados en Ω , y d_1 y d_2 son respectivamente las distancias euclídeas de estas proyecciones hasta el punto \mathbf{x}_0 , i.e.

$d_1 = \|\mathbf{v}_1 - \mathbf{x}_0\|$ y $d_2 = \|\mathbf{v}_2 - \mathbf{x}_0\|$. El uso de (6) iterativamente tenderá a no tener efecto sobre el punto proyectado \mathbf{v} . Este hecho no es una desventaja del algoritmo, sino que es el principal objetivo del método de proyección. De este modo, no se realizan decisiones drásticas y la influencia de los pixels vecinos llevarán finalmente el vector proyectado hacia el modelo correcto, siempre dentro de la región descrita por el modelo dual. Este procedimiento iterativo puede ser visto como una técnica de proyección sobre conjuntos convexos (POCS) [6], siempre y cuando $\Omega(\mathbf{x})$ sea una función convexa. A continuación se describe el procedimiento de proyección.

Modelo dual punto-punto. Este modelo aparece en regiones de la imagen con estructuras claras y repetitivas. En imágenes reales, la probabilidad de aparición de estos modelos es sustancial (sobre un 30% de los modelos en la práctica). Ésta es la situación más delicada, porque el proceso de promediado no es posible como tal, y ello causa la inevitable selección de un modelo. El método elige los parámetros del modelo de mínima distancia euclídea al punto. Esta situación se muestra en la Fig. 1. En este modelo, $\Omega = \{\{0, \mathbf{v}_1\}, \{0, \mathbf{v}_2\}\}$, el espacio de proyección tiene la forma de dos puntos aislados, de modo que el punto proyectado puede ser fácilmente obtenido de acuerdo a las distancias euclídeas entre ambos puntos y \mathbf{x}_0 ,

$$\mathbf{v} = \begin{cases} \mathbf{v}_1, & \text{si } \|\mathbf{v}_1 - \mathbf{x}_0\| < \|\mathbf{v}_2 - \mathbf{x}_0\| \\ \mathbf{v}_2, & \text{si } \|\mathbf{v}_1 - \mathbf{x}_0\| \geq \|\mathbf{v}_2 - \mathbf{x}_0\| \end{cases} \quad (7)$$

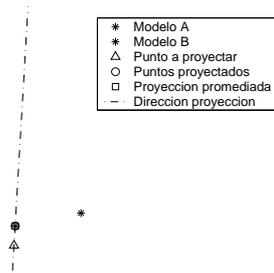


Fig. 1. Modelo punto-punto.

Modelo dual punto-curva. Este modelo aparece típicamente en situaciones similares a las del modelo anterior, pero para estructuras no tan bien definidas. Se corresponde a la situación de menor probabilidad de aparición (20% de los modelos duales). Para obtener la proyección promediada, debe determinarse una dirección de proyección. Ésta es la dirección que une el punto a proyectar \mathbf{x}_0 con el punto que representa al modelo tipo 0 (punto), \mathbf{v}_1 . La proyección este último modelo son las coordenadas que lo definen, mientras que la proyección sobre el modelo tipo 1 (curva) se obtiene de la intersección de este modelo con la dirección de proyección. Una vez se ha calculado la proyección sobre ambos modelos, el vector final proyectado \mathbf{v} queda determinado por medio de un proceso de promediado (expresión (6)). La Fig. 2 ilustra el proceso de proyección. Los parámetros del modelo son $\Omega_1 = \{0, \mathbf{v}_1\}$ y $\Omega_2 = \{1, \theta, \mathbf{a}\}$. Conforme a las ideas previas, la proyección sobre el modelo tipo 0 se obtiene directamente, mientras que

la proyección sobre el modelo tipo 1, \mathbf{v}_2 , se calcula mediante las siguientes expresiones:

$$\phi = \arctan\left(\frac{v_{1,y} - y_0}{v_{1,x} - x_0}\right) \Rightarrow m = \tan(\phi - \theta), \quad (8)$$

$$\begin{bmatrix} x_\theta \\ y_\theta \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad (9)$$

$$\begin{cases} \hat{y} = m(\hat{x} - x_\theta) + y_\theta \\ \hat{y} = a\hat{x}^2 + b\hat{x} + c \end{cases} \Rightarrow \begin{cases} \hat{x} = \frac{-(b-m) \pm \sqrt{(b-m)^2 - 4a(c+m x_\theta - y_\theta)}}{2a} \\ \hat{y} = a\hat{x}^2 + b\hat{x} + c \end{cases} \quad (10)$$

$$\begin{bmatrix} v_{2,x} \\ v_{2,y} \end{bmatrix} = \mathbf{R}^{-1} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}. \quad (11)$$

Si $a = 0$, no es posible aplicar (10) debido a una singularidad. En ese caso, el proceso aún puede llevarse a cabo reemplazando (10) por (12).

$$\begin{cases} \hat{y} = m(\hat{x} - x_\theta) + y_\theta \\ \hat{y} = b\hat{x} + c \end{cases} \Rightarrow \begin{cases} \hat{x} = \frac{-(c+m x_\theta - y_\theta)}{b-m} \\ \hat{y} = b\hat{x} + c \end{cases} \quad (12)$$

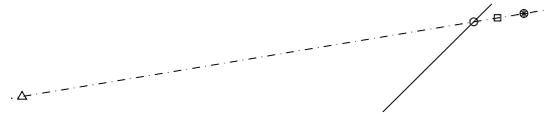
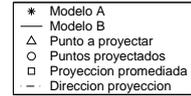


Fig. 2. Modelo punto-curva.

Modelo dual curva-curva. Este es el modelo dual más común (sobre un 50% del total), y resulta ser la situación más ambigua, dado que se le permite mayor libertad a la localización del punto proyectado. La Fig. 3 muestra el procedimiento de promediado de posiciones para este caso. Los parámetros del modelo son $\Omega_1 = \{1, \theta_1, \mathbf{a}_1\}$ y $\Omega_2 = \{1, \theta_2, \mathbf{a}_2\}$. El primer paso para lograr la proyección dual es determinar la dirección normal al ángulo de orientación medio de ambos modelos paramétricos. A partir de este valor y para cada una de las orientaciones, θ_1 y θ_2 , se calculan las pendientes m_1 y m_2 de las rectas rotadas como sigue:

$$m_1 = \cot\left(\frac{\theta_1 - \theta_2}{2}\right), \quad m_2 = -m_1. \quad (13)$$

Se calculan ahora las intersecciones de cada curva, definida por los parámetros del modelo, y la recta correspondiente. De este modo, podemos aplicar las expresiones (8)-(11) una vez por modelo, con los correspondientes valores de los parámetros θ_i , \mathbf{a}_i y m_i . Finalmente, y a partir de las proyecciones calculadas, \mathbf{v}_1 y \mathbf{v}_2 , debemos aplicar la expresión (6) para obtener el punto proyectado \mathbf{v} .

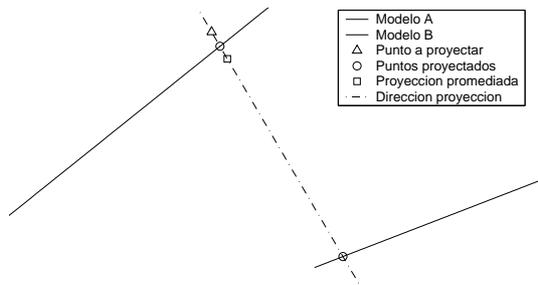


Fig. 3. Modelo curva-curvea.

III. RESULTADOS

La Fig. 4 muestra una comparativa de los resultados de registrar un par de imágenes de tamaño 500×500 . La imagen inicial o de referencia, Fig. 4(a), y la imagen objetivo, Fig. 4(b), son fotografías de un mismo puente sobre una presa, tomadas desde distintas posiciones. De este modo, la deformación entre imágenes procede de diferentes puntos de vista de la cámara. En primer lugar, se realiza una estimación del movimiento con el método propuesto con modelos duales. La Fig. 4(d) muestra los vectores de correspondencia resultantes, que representan el mapeado óptimo, sobre la imagen objetivo. Aplicando este mapeado a la imagen objetivo se consigue el registro de la misma, sin diferencias visibles respecto a Fig. 4(a). Este aspecto se percibe con mayor claridad en la Fig. 4(e), donde se aplican los vectores a una rejilla rectangular. A continuación, repetimos la estimación de la correspondencia sin considerar la posibilidad de múltiples correspondencias, i.e. considerando únicamente modelos simples. En este experimento pueden observarse claramente algunos errores en la correspondencia, Fig. 4(f). De todos modos, cabe destacar que también en el caso de considerar modelos duales la estimación de la deformación resulta pobre, aunque mejor, en algunas regiones de la imagen: en los puntos donde cada imagen presenta objetos o características diferentes, de modo que no es posible una correspondencia perfecta. La Fig. 4(c) muestra las posiciones de los modelos duales que se detectaron en la primera simulación. Esta información no está presente en la segunda simulación, y representa la única diferencia entre ambas, confirmando que los modelos duales son la solución a las situaciones de ambigüedad en la correspondencia.

IV. CONCLUSIONES

En este artículo se ha propuesto el uso de modelos paramétricos duales para tratar con situaciones ambiguas en el registro de imágenes. Estos casos aparecen cuando las características de la imagen inducen a la ambigüedad (múltiples regiones de correspondencia posibles). Los resultados que se han mostrado prueban la eficacia de esta propuesta. La mayoría de los errores en la asignación de los pixels pueden resolverse con esta idea. Este método alivia el problema que presentan algunos algoritmos de correspondencia clásicos, como el típico algoritmo de block-matching. El modelo de parametrización dual está siendo aplicado en la actualidad en el registro de una amplia variedad de imágenes con resultados satisfactorios. En

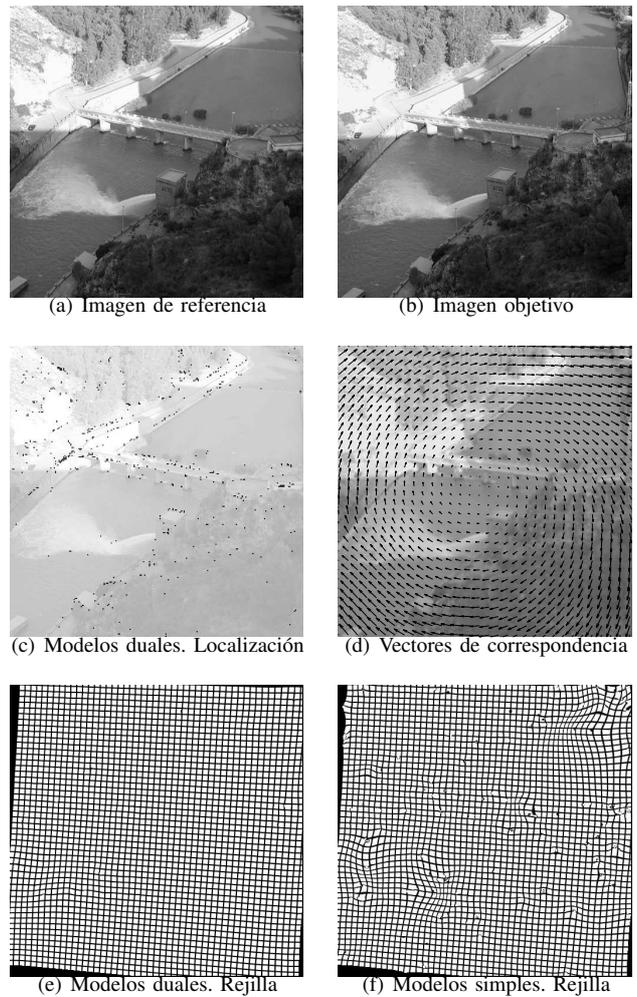


Fig. 4. Resultados de registrar un par de imágenes de muestra.

estos momentos se está estudiando la extensión del método a modelos de mayor complejidad.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología a través del proyecto TIC2002-03033.

REFERENCIAS

- [1] L. Weruaga, J. Morales, L. Núñez, and R. Verdú. Estimating volumetric motion in human thorax with parametric matching constraints. *IEEE Transactions on Medical Imaging*, 22:766–772, June 2003.
- [2] J. Morales, R. Verdú, R. Berenguer, and L. Weruaga. Non-rigid motion estimation based on fuzzy models. In *Proc. IEEE Int. Conf. on DSP'02*, pages 559–562, Santorini, Greece, 2002.
- [3] L. Weruaga, R. Verdú, and J. Morales. Frequency domain formulation of active parametric deformable models. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 26:1568–1578, December 2004.
- [4] Federico Giroso and Tomaso Poggio. Networks and the best approximation property. Technical Report AIM-1164, 1989.
- [5] Federico Giroso, Tomaso Poggio et al. Extensions of a theory of networks for approximation and learning. *Advances in Neural Information Processing Systems*, 3:750–756. Morgan Kaufmann Pub., 1991.
- [6] A. Goshtasby and G.C. Stockman. Point pattern matching using convex hull edges. *IEEE Trans. Systems, Man & Cybernetics*, 15:631–637, 1985.