

Maestría en Ingeniería Matemática

Propuesta de Tesis



Título de la propuesta

Esteganografía basada en modelos esparsos

Identificación del proponente

- Ignacio Ramírez,
- Último título obtenido: PhD (Scientific Computation, University of Minnesota)
- Lugar de trabajo: Instituto de Ing. Eléctrica, Facultad de Ingeniería
- Área de trabajo: Procesamiento Estadístico de Señales
- Información de contacto: tel. 27110974 – cel. 098182068 – mail nacho@fing.edu.uy

Área Temática

Teoría de la Información, Modelado estadístico de señales

Perfil esperado del estudiante

Se busca alguien con buena base matemática en temas como estadística, probabilidades, procesos estocásticos y optimización no lineal, y con interés en aplicaciones a dichas áreas de conocimiento a temas como el procesamiento de señales. Ejemplos de éstos son (sin un orden particular): egresados de las carreras de Ing. Eléctrica o Computación con buena formación en matemática, Licenciados en Estadística y Licenciados en Matemática.

Resumen

La idea de este proyecto es desarrollar un nuevo método para insertar información secreta en imágenes (proceso conocido como *esteganografía*). Un buen algoritmo de esteganografía debe ser capaz de insertar información en una imagen sin modificar de manera perceptible las propiedades estadísticas de lo que sería una *imagen natural típica*, es decir, una foto. Para lograr esto, muchos algoritmos evitan modificar los pixels originales directamente. En su lugar, estos algoritmos primero transforman la imagen a un espacio en donde dichos cambios sean luego repartidos imperceptiblemente en el espacio de origen.

Las transformaciones utilizadas suelen ser las típicas utilizadas en procesamiento de imágenes: wavelets, DCT, Fourier, etc. Lo que este proyecto se propone es realizar algo similar, pero

proyectando los datos a un espacio de dimensión mayor, en una base aprendida específicamente para representar de manera eficiente los datos. A estos modelos, llamados modelos esparsos, los describimos a continuación.

Decimos que una muestra $\mathbf{y} \in \mathbb{R}^m$ admite una representación esparsa si existe un *diccionario* $\mathbf{D} \in \mathbb{R}^{k \times n}$ y un vector de coeficientes $\mathbf{a} \in \mathbb{R}^k$ tales que $\|\mathbf{y} - \mathbf{D}\mathbf{a}\| \leq \epsilon$, $0 \leq \epsilon \ll \|\mathbf{y}\|$, con $\|\mathbf{a}\|_0 \leq p \ll k$. (La pseudo-norma $\|\mathbf{a}\|_0$ mide la cantidad de elementos no nulos en \mathbf{a} .) La estimación de tales modelos suele hacerse mediante la resolución de un problema de mínimos cuadrados penalizado:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{w}} \|\mathbf{y} - \mathbf{D}\mathbf{w}\|_2^2 \quad \text{s.t.} \quad \psi(\mathbf{w}) \leq \lambda, \quad (1)$$

donde $\psi(\cdot)$ es una función $\mathbb{R}^k \rightarrow \mathbb{R}$ que favorece soluciones esparsas. En [1] puede encontrarse una muy buena introducción a este tipo de modelos.

Además del enfoque clásico que implicaría modificar los coeficientes no nulos para acomodar el mensaje secreto, los modelos esparsos abren una nueva posibilidad: la de codificar la información oculta mediante alteraciones de los soportes de las representaciones obtenidas.

Referencias

- [1] A. Bruckstein, D. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, Feb. 2009.