

Resumen Loseless Image Compression using List Update Algorithms

Vicente Lermenda Candia

Abstract

Introducción

- ▶ Se debe aprovechar localidad y redundancia de datos.
- ▶ Se quiere minimizar el ratio de compresión.

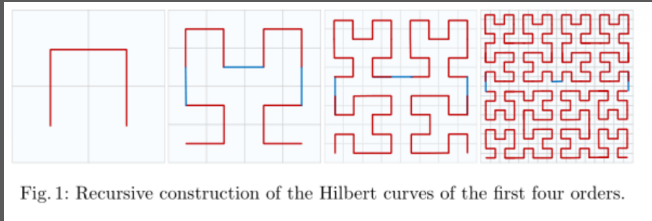
$$\text{Compression Ratio} = \frac{\text{Uncompressed Size}}{\text{Compressed Size}}$$

- ▶ Se debe ser eficiente en términos de recursos computacionales.
- ▶ No obstante, un mayor uso de recursos, en particular tiempo, resulta en mejores ratios de compresión (trade-off).

- ▶ Distintas técnicas de reducción de entropía (prediction-based, wavelet-transform-based, deep learning). No son directamente comparables.
- ▶ Para datos secuenciales (1D), frecuentemente se usan list update algorithms para reducir entropía (Move-To-Front)
- ▶ Finalmente dividimos este esquema en dos etapas:
 1. Linealización con Hilbert space-filling curves ($2D \rightarrow 1D$).
 2. Codificación con list update algorithms (Move-To-Front).

Linealización de una imagen a una secuencia

- ▶ Curva Hilbert orden k cubre una matriz de tamaño $2^k \times 2^k$.



- ▶ La curva de Hilbert se computa eficientemente usando operaciones de bits y lookup tables.

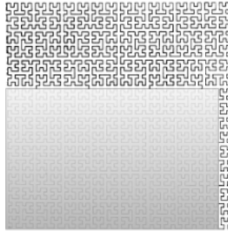


Fig. 2: Linearisation of a non-square image using the Hilbert curve of a larger square whose side-lengths are a power of 2.

- ▶ Se considera "Max-Model" para determinar localidad de datos secuenciales.
- ▶ Dada una secuencia σ , *definimos* :
- ▶ *Ventana* de tamaño w en σ : subsecuencia de w solicitudes consecutivas en σ .
- ▶ Se dice que σ es consistente con alguna función cóncava creciente f , si el número de solicitudes distintas en cualquier ventana de tamaño w , es a lo más $f(w_0)$, $\forall w \in \mathbb{N}$.
- ▶ Extendiendo a imágenes, las ventanas pasan a ser cuadrados de w píxeles.

- **Teorema:** Una imagen tiene localidad Max-Model si y sólo si la secuencia formada por la curva de Hilbert también tiene localidad Max-Model.

List update encoding

- ▶ Frecuentemente usado para reducir la entropía de una secuencia.
- ▶ Guarda items de una secuencia en un linked-list.
- ▶ Accede a la posición i en tiempo i .
- ▶ Al acceder puede acercar el item al frente sin costo adicional (free exchange).
- ▶ Además, el algoritmo puede re-ordenar la lista usando paid exchanges, lo que corresponde a hacer swap entre elementos consecutivos por costo 1.

Move-To-Front

- ▶ Es un list update algorithm.
- ▶ Mueve un item accesado al frente de la lista por free exchange.
- ▶ Se concluye que codificaciones MTF son estrictamente mejores para secuencias formadas por una curva de Hilbert en una imagen.

Iteration	Sequence	List
bananaaa	1	(abcdefghijklmnopqrstuvwxyz)
bananaaa	1,1	(bacdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13	(abcdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13,1	(nabcdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13,1,1	(anabcdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13,1,1,1	(nabcdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13,1,1,1,0	(anabcdefghijklmnopqrstuvwxyz)
bananaaa	1,1,13,1,1,1,0,0	(anabcdefghijklmnopqrstuvwxyz)
Final	1,1,13,1,1,1,0,0	(anabcdefghijklmnopqrstuvwxyz)

- Si hay buena localidad en la imagen, obtendremos una secuencia con valores pequeños, lo cual reduce la entropía.

Experimental results

Image	Column Major	Row Major	Snake	Spiral	Hilbert Curve
Mandrill	0.9656	0.9519	0.9480	0.9537	0.9094
Peppers	0.8904	0.8718	0.8664	0.8735	0.7746
Lena	0.8836	0.8168	0.8086	0.8456	0.7499
Barbara	0.9054	0.8837	0.8791	0.8962	0.8102
Average	0.9112	0.8810	0.8755	0.8922	0.8110

Table 1: Entropy ratio of different linearisation methods for images in the Classic dataset.

Image	Column Major	Row Major	Snake	Spiral	Hilbert Curve
PET1	0.5309	0.5740	0.5686	0.4745	0.4722
PET2	0.5428	0.6024	0.5953	0.5193	0.4901
PET3	0.5294	0.5749	0.5701	0.4892	0.4701
Endoscope1	0.7197	0.6876	0.6803	0.6763	0.6070
Endoscope2	0.6021	0.6016	0.5905	0.5903	0.4454
Eyeground	0.4896	0.4963	0.4921	0.4632	0.3866
Average	0.5690	0.5894	0.5828	0.5354	0.4786

Table 2: Entropy ratio of different linearisation methods for images in the Medical dataset.

Image	Size	Pixels	Column Major	Row Major	Snake	Spiral	Hilbert Curve
Mandrill	3x 512x 512	786,432	0.94	0.93	0.92	0.93	0.89
Peppers	3x 512x 512	786,432	0.86	0.84	0.84	0.84	0.75
Lena	3x 512x 512	786,432	0.86	0.8	0.79	0.82	0.73
Barbara	3x 640x 512	983,040	0.87	0.85	0.85	0.89	0.78
Average	-	-	0.88	0.85	0.85	0.87	0.7875

Table 3: Compression ratio of different linearisation methods for images in the Classic dataset.