



Universidad  
Rafael Landívar  
Tradición Jesuita en Guatemala

# Compresión basadas en Diccionarios RLE

## Algoritmo de compresión RLE

Run-Length Encoding

Archivo sin comprimir



EMEZETA.COM



Archivo comprimido

Facultad de ingeniería

# Tipos compresores sin pérdida

Los métodos de compresión sin pérdida pueden clasificarse en dos tipos: **basados en diccionario** y **estadísticos**. La principal distinción entre ambos es que en el caso de los métodos estadísticos, la codificación de un símbolo se basa en el contexto en el que este ocurre, mientras que los métodos basados en diccionario agrupan símbolos creando un tipo de contexto implícito.

# Tipos compresores sin pérdida

En general, un compresor estadístico logra mejores razones de compresión que un compresor basado en diccionario pero la complejidad computacional y requerimientos de memoria para los compresores estadísticos son mucho mayores que en su contraparte

# Métodos basados en diccionarios

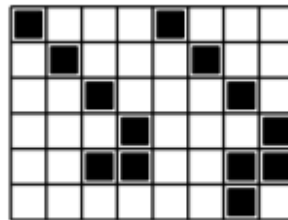
Los métodos basados en diccionario usan el principio de remplazar cadenas de datos con codewords que identifican a esa cadena dentro de un diccionario.

El diccionario contiene una lista de subcadenas y códigos asociados a cada una de ellas. El diccionario que contiene las cadenas de símbolos puede ser estático o adaptable (dinámico). Los métodos basados en diccionario, a diferencia de los métodos estadísticos, usan códigos de longitud fija y no requieren de las estadísticas de los símbolos para realizar la compresión.

# Compresión RLE

- El método de compresión **RLE** (*Run Length Encoding*) es utilizado por muchos formatos de imagen para su rápida transferencia o almacenamiento. Se basa en la repetición de elementos consecutivos.
- La compresión de píxeles es un método que permite almacenar píxeles de forma óptima: para una imagen monocromática hay, por definición, sólo dos colores. Por lo tanto, se puede codificar un punto de la imagen en un solo bit para ganar espacio en la memoria.

0 1 0 1



# Compresión RLE

- El principio fundamental consiste en codificar un primer elemento al dar el número de repeticiones de un valor y después el valor que va a repetirse. Por lo tanto, según este principio, la cadena “AAAAHHHHHHHHHHHHHHH” cuando está comprimida da como resultado “5A14H”. La ganancia de compresión es  $(19-5) / 19$ , es decir, aproximadamente 73,7%.
- Por otro lado, para la cadena “CORRECTLY”, donde hay poca repetición de caracteres, el resultado de la compresión es “1C1O2R1E1C1T1L1Y”. Por lo tanto, la compresión genera un costo muy elevado y una ganancia de compresión negativa de  $(9-16) / 9$ , es decir, ¡-78%!

# Compresión RLE

- La compresión RLE tienen diferentes tipos de funcionamientos, cada uno de ellos patentados. Pero todos se basan en la misma repetición de símbolos para comprimir la información

## Algoritmo de compresión RLE

Run-Length Encoding

Archivo sin comprimir



EMEZETA.COM



Archivo comprimido

◊ **Ejercicio 1.3** (sol. en pág. 1051): ¿Cuál sería el archivo comprimido para el siguiente bitmap de  $6 \times 8$ ?

