

# Aplicaciones de Teoría de la Información en Procesamiento de Imágenes (ATIPi 2018)

## Proyecto de Evaluación

### Consideraciones generales

- El trabajo consiste en la implementación de algoritmos de compresión y descompresión de imágenes en software.
- Los programas deberán cumplir con los requisitos siguientes:
  - El lenguaje de programación debe ser C++ o C.
  - Los experimentos solicitados en el trabajo deben ser compilables, ejecutables, y reproducibles por los docentes en plataformas de computación genéricas (PC con Windows o Linux, Mac, etc.), sin requerir paquetes especiales. La entrega debe incluir instrucciones claras de cómo compilar y ejecutar los programas (incluyendo archivos makefile, proyectos de MS Visual Studio, etc., apropiados a la plataforma).
  - Las imágenes de entrada y salida serán en formato PGM (8bpp, escala gris, modo P5). Ver descripción del formato en [http://en.wikipedia.org/wiki/Netpbm\\_format](http://en.wikipedia.org/wiki/Netpbm_format).
  - Los experimentos requeridos se harán sobre un conjunto de imágenes estándar que estarán disponibles en la página de web del curso.
  - Las decisiones de diseño e implementación que influyen sobre el rendimiento de los programas son importantes y serán evaluadas. Los programas deberán permitir la realización de los experimentos requeridos, sobre las imágenes provistas, en tiempo razonable (medido en segundos).
- En la letra del proyecto se especifica la funcionalidad mínima requerida. Decisiones de diseño que no estén definidas explícitamente quedan a criterio del programador (p.ej. los detalles del encabezamiento del archivo comprimido, opciones del programa, etc.). Mejoras en funcionalidad, si son interesantes y efectivas, recibirán puntaje extra.
- La entrega del proyecto consistirá en
  1. Un paquete comprimido (formato zip o tgz) con los archivos fuente y ejecutables de los programas, incluyendo instrucciones precisas para su compilación y ejecución.
  2. Un informe en formato pdf conteniendo: descripción de los programas (incluyendo decisiones de diseño y detalles que no se desprendan obviamente de la letra del proyecto), resultados de los experimentos y su discusión, y respuestas a las preguntas formuladas.

- *Está prohibido* bajar soluciones de cualquier tipo, totales o parciales, de internet, o copiarlas de ejercicios similares en ediciones anteriores del curso.

Tenga en cuenta que transgresiones a estas prohibiciones son relativamente fáciles de identificar—existen herramientas sofisticadas de dominio público que permiten detectar con gran confiabilidad la copia, con o sin alteraciones, de trabajos existentes, sea de internet o de otras fuentes. De constatarse dichas transgresiones, serán severamente sancionadas.

## Proyecto D2: Compresión de imágenes en colores

Se implementarán compresores y descompresores sin pérdidas para imágenes de tono gris (8 bpp), y para imágenes en colores (24 pbb). El algoritmo de compresión para imágenes de tono gris consistirá de los elementos que se detallan a continuación.

### Imágenes de tono gris

- Se utiliza el mismo predictor fijo MED utilizado en JPEG-LS. El predictor no tiene parte adaptativa.
- Se codifican errores de predicción, utilizando códigos de Golomb de orden  $m = 2^k$ ,  $k = 0, 1, 2, \dots$ , y utilizando la función

$$M(e) = \begin{cases} -2e + 1 & e < 0, \\ 2e & e \geq 0, \end{cases} \quad (1)$$

para trasladar errores de predicción  $e = 0, -1, 1, -2, 2, \dots$  al rango no negativo  $M(e) = 0, 1, 2, 3, 4, \dots$ . No se utiliza la función alternativa  $M'(e) = M(-1 - e)$  mencionada en el curso.

- Se determina el parámetro  $k$  del código de Golomb basado en dos estadísticas:  $A$  (suma de valores absolutos de errores) y  $n$  (número de muestras). *No hay modelo de contextos global*. Las estadísticas se toman del contexto local de la muestra, según el procedimiento detallado abajo, donde se supone que las muestras están numeradas en orden “raster” lineal  $x_1, x_2, x_3, \dots, x_{w \times h}$ , siendo  $w$  el número de columnas y  $h$  el número de filas de la imagen.
1. Se fija un tamaño de contexto  $N \geq 1$ , y una plantilla causal tomando las  $N$  muestras ya procesadas más cercanas, en distancia euclidiana, a la muestra a procesar,  $x_i$  (con algún criterio razonable en caso de empates en distancia). Para ello, se recomienda precalcular una tabla de pares  $(\Delta w, \Delta h)$  de desplazamientos relativos a la muestra modelada, en orden creciente de magnitud. Por ejemplo, para  $N = 5$ , y la plantilla causal de abajo,

	$c$	$b$	$d$
$e$	$a$	$x$	

esa tabla podría ser

$$\{\{-1, 0\}, \{0, -1\}, \{-1, -1\}, \{1, -1\}, \{-2, 0\}\}$$

que corresponde al orden  $a, b, c, d, e$ .

2. Inicialización:  $i = 1$ .

3. Procesamiento de la muestra  $x_i$ :

- a) Se calcula el valor predicho  $\hat{x}_i$ .
- b) Se define  $n$  como el número de píxeles del contexto de  $x_i$  que están dentro del área válida de la imagen. (O sea que para la gran mayoría de la muestras,  $n = N$ ;  $n$  puede ser menor que  $N$  para muestras cercanas a los bordes de la imagen).
- c) Si  $n = 0$  (solo se da con  $i = 1$ ), se toma  $k = 3$   
Si  $n > 0$ , sean  $x_{j_1}, x_{j_2}, \dots, x_{j_n}$  las  $n$  muestras en el contexto de  $x_i$ . Se calcula

$$A = \sum_{\ell=1}^n |x_{j_\ell} - \hat{x}_{j_\ell}|,$$

(esta es la suma de magnitudes de errores de predicción para símbolos en el contexto). Se determina  $k$  a partir de  $A$  y  $n$  utilizando el mismo procedimiento que en los contextos regulares de JPEG-LS (donde  $n$  y  $A$  se obtienen de manera distinta, pero son también un número de muestras, y una suma de valores absolutos de  $n$  errores de predicción).

(Note que en una implementación eficiente, conviene almacenar algunas filas de valores predichos y filas de la imagen, o tal vez filas de errores de predicción, para evitar repetir el cálculo de predicción muchas veces para cada muestra.)

- d) Se codifica  $x_i - \hat{x}_i$  con el mapeo (1) y el código de Golomb de orden  $m = 2^k$ .
- e) Si  $i \leq w \times h$ , volver al paso 3. En caso contrario, terminar.

## Imágenes en colores

El compresor para imágenes en colores se basará en el de imágenes de tono gris. Se estudiarán dos estrategias para el manejo de planos de color:

- A. Se comprimen los planos de color  $R, G, B$  por separado, como si cada uno fuera una imagen de tono gris independiente. El archivo comprimido consiste de 3 secciones independientes, una para cada plano. El descompresor descomprime cada plano y entrelaza las cadenas  $R, G, B$  en formato PPM.

B. Antes de comprimir los planos de color, se efectúa la transformación

$$\begin{aligned}R' &\leftarrow R - G \text{ mód } 256, \\G' &\leftarrow G, \\B' &\leftarrow B - G \text{ mód } 256.\end{aligned}$$

Se comprimen los planos  $R', G', B'$  como en A. El descompresor recupera los planos  $R', G', B'$ , y efectúa la transformación inversa para recuperar los planos  $R, G, B$ .

### Tareas:

1. Implementar el compresor y el descompresor correspondiente, dejando el valor  $N$  y, en el caso de imágenes en colores, la elección de opción A o B, como parámetros del programa de compresión. Los valores de estos parámetros se incluirán al principio del archivo comprimido, para uso del descompresor.
2. Comprimir y descomprimir cada una de las imágenes en el conjunto de prueba (8 bpp y 24 bpp), probando los valores  $N = 2, 4, 8, 16, 64, 256$ . En el caso de las imágenes en colores, probar también las opciones A y B. Comprobar en cada caso que la compresión y descompresión funcionan correctamente (sin pérdidas), y tabular las tasas de compresión (en bits/pixel) para cada una de las combinaciones de imagen, valor de  $N$ , y opción A/B que se hayan probado.
3. Explicar la función del parámetro  $N$ .
4. Explicar la función de la transformada  $(R, G, B) \rightarrow (R', G', B')$ .
5. Recomendar un valor por defecto para el parámetro  $N$ .
6. Comparar las tasas de compresión obtenidas con el parámetro recomendado en el inciso anterior con las obtenidas con JPEG-LS.
7. Explicar los resultados de la comparación, especialmente en casos en que se registren diferencias grandes en la tasa de compresión.