

# Compression

## Projet Mathématiques et Informatique

LABADENS Lucas,  
MARINO Isabelle

Université Paris 7 Diderot

13 Juin 2016

# Sommaire

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

- 1 La compression : définition
- 2 Huffman
- 3 Lempel-Ziv
- 4 Différences entre les algorithmes
- 5 Différences avec les principaux compresseurs
- 6 Conclusion

# La Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Plusieurs types de compression :

# La Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Plusieurs types de compression :

- Compression avec perte de données

# La Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Plusieurs types de compression :

- Compression avec perte de données
- Compression sans perte de données

# La Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Plusieurs types de compression :

- Compression avec perte de données
- Compression sans perte de données
- Exemples : ZIP, GZIP, XZ, BZIP2,...

## Caractéristique :

- Un codage entropique
- repose sur la redondance de caractère
- code un caractère non plus sur un octet mais sur un nombre de bit

## Étape :

- compte des redondances de caractères
- création de l'arbre de compression en fonction des poids des caractères :
  - on crée des feuilles qui sont des poids munie d'un caractère
  - on relie les deux feuilles de poids de plus faible par un nœud qui est un poids muni d'un fils gauche et droit
  - on relie les poids les deux poids les plus faibles par un nouveau jusqu'à ne plus avoir qu'un seul nœud
- récupération des nouveaux codes de chaque caractère
- relecture du fichier pour écrire chaque caractère dans le fichier compresser avec son nouveau code
- on complète le dernière octet par la convention un 1 puis le nombre de 0 nécessaire



# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- récupération de l'arbre de compression

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- récupération de l'arbre de compression
- lecture du fichier compresser bit à bit et parcours de l'arbre :

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- récupération de l'arbre de compression
- lecture du fichier compresser bit à bit et parcours de l'arbre :
  - lecture d'un 0 on se déplace à gauche et sinon à droite

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- récupération de l'arbre de compression
- lecture du fichier compresser bit à bit et parcours de l'arbre :
  - lecture d'un 0 on se déplace à gauche et sinon à droite
  - lorsque l'on tombe sur une feuille on écrit son caractère associé dans le fichier de décompression

## Étape :

- récupération de l'arbre de compression
- lecture du fichier compresser bit à bit et parcours de l'arbre :
  - lecture d'un 0 on se déplace à gauche et sinon à droite
  - lorsque l'on tombe sur une feuille on écrit son caractère associé dans le fichier de décompression
- attention à ne pas lire les bits complétant le dernière octet du fichier

# Exemple

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape 1

Soit un texte ou il est écrit : aaaabcbbbe  
(a,4) (b,3) (c,2) (e,1)

# Exemple

## Compression

La  
compression :  
définition

Huffman

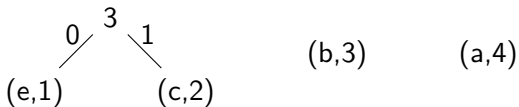
Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape 2



# Exemple

## Compression

La  
compression :  
définition

Huffman

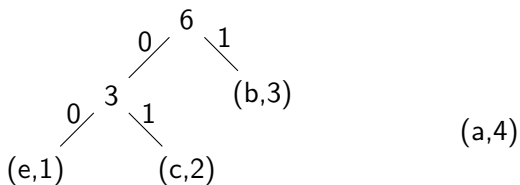
Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape 2





# Exemple

## Compression

La  
compression :  
définition

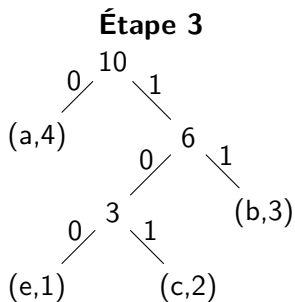
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



# Exemple

## Compression

La  
compression :  
définition

## Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Écriture compresseur

les nouveaux codes sont  $a='0'$   $b='11'$   $c='101'$   $e='100'$

# Exemple : Démonstration

Compression

La  
compression :  
définition

**Huffman**

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Démonstration

# Performances

## Compression

La  
compression :  
définition

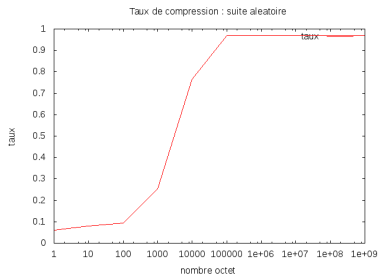
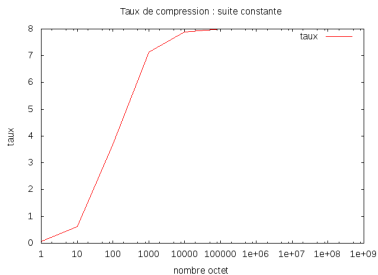
## Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



## Caractéristiques :

## Caractéristiques :

- Codage algorithmique

## Caractéristiques :

- Codage algorithmique
- Bit à bit

## Caractéristiques :

- Codage algorithmique
- Bit à bit
- Code des suites de bits sur le moins de place possible



# Compression

## Compression

La  
compression :  
définition

Huffman

**Lempel-Ziv**

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

**Étape :**

# Compression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création de l'arbre de compression

# Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création de l'arbre de compression
- Étape i : Lecture d'un bit :

# Compression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création de l'arbre de compression
- Étape i : Lecture d'un bit :
  - 0 : déplacement vers la gauche de l'arbre, lecture du bit suivant

# Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création de l'arbre de compression
- Étape i : Lecture d'un bit :
  - 0 : déplacement vers la gauche de l'arbre, lecture du bit suivant
  - 1 : déplacement vers la droite de l'arbre, lecture du bit suivant

# Compression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création de l'arbre de compression
- Étape i : Lecture d'un bit :
  - 0 : déplacement vers la gauche de l'arbre, lecture du bit suivant
  - 1 : déplacement vers la droite de l'arbre, lecture du bit suivant
  - Déplacement impossible : création de nouveau nœud

# Compression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Codage et écriture du nouveau nœud dans le fichier

# Compression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Codage et écriture du nouveau nœud dans le fichier
- Écriture du bit suivant dans le fichier



# Décompression

Compression

La  
compression :  
définition

Huffman

**Lempel-Ziv**

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

**Étape :**

# Décompression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)

# Décompression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé

# Décompression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé
- A l'étape i :

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé
- A l'étape  $i$  :
  - Lecture de  $\lceil \log_2(i) \rceil$  (père)

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé
- A l'étape  $i$  :
  - Lecture de  $\lceil \log_2(i) \rceil$  (père)
  - Lecture du bit suivant (bit)

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé
- A l'étape  $i$  :
  - Lecture de  $\lceil \log_2(i) \rceil$  (père)
  - Lecture du bit suivant (bit)
  - Création de la nouvelle case

# Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Étape :

- Création d'un tableau à 2 entrées (père, bit)
- Lecture bit à bit du fichier compressé
- A l'étape  $i$  :
  - Lecture de  $\lceil \log_2(i) \rceil$  (père)
  - Lecture du bit suivant (bit)
  - Création de la nouvelle case
  - Écriture de la décompression



# Exemple : Compression

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

Soit un fichier contenant "te"  
Suite de bit obtenue :

01110100 01100101

# Exemple : Compression

## Compression

La  
compression :  
définition

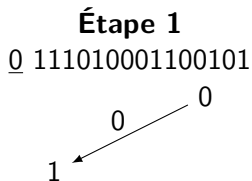
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 0

# Exemple : Compression

## Compression

La  
compression :  
définition

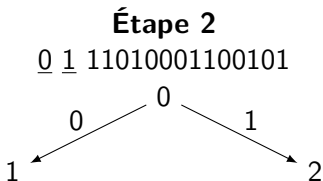
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 01

# Exemple : Compression

Compression

La  
compression :  
définition

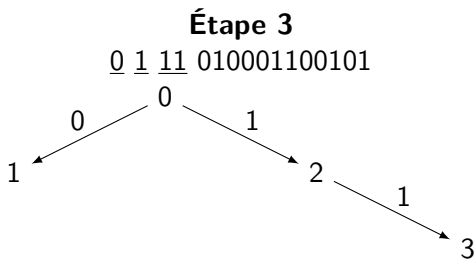
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 101

# Exemple : Compression

Compression

La compression :  
définition

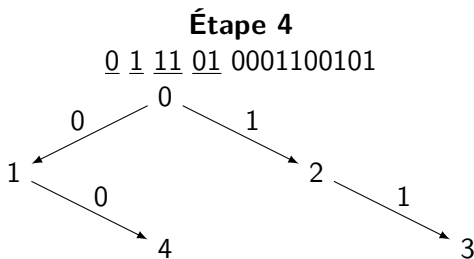
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 011

# Exemple : Compression

Compression

La compression :  
définition

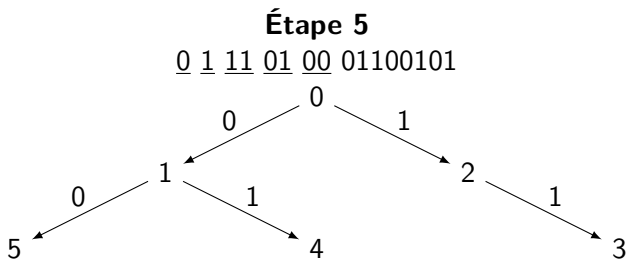
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 0010

# Exemple : Compression

Compression

La compression :  
définition

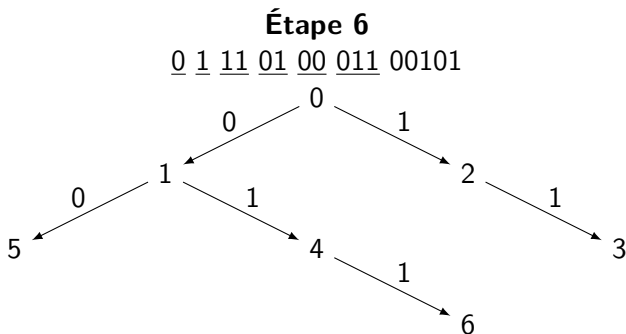
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 1001

# Exemple : Compression

## Compression

La  
compression :  
définition

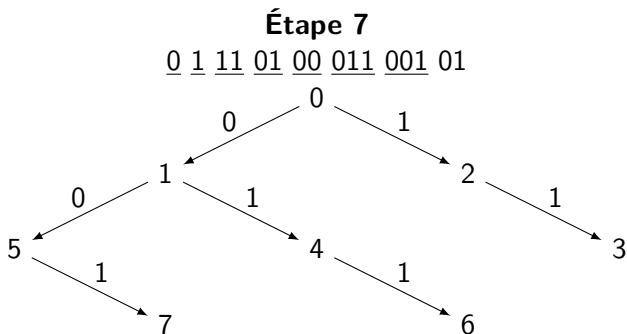
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 1011



# Exemple : Compression

## Compression

La  
compression :  
définition

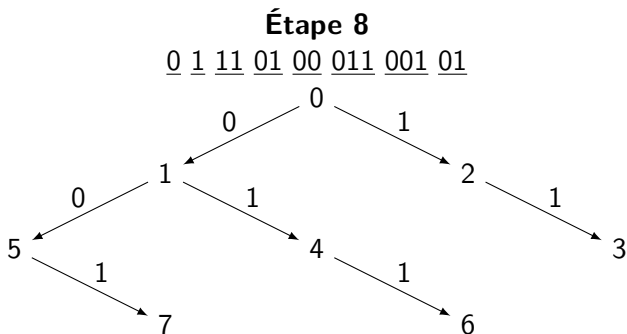
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



Écriture dans le fichier : 1001 et 1000000

# Exemple : Décompression

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

00110101 10010100 11011100 11000000

nœud	0	1	2	3	4	5	6	7	8
père	-1	0	0	2	1	1	4	5	4
bit	-1	0	1	1	1	0	1	1	1

# Exemple : Démonstration

Compression

La  
compression :  
définition

Huffman

**Lempel-Ziv**

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

## Démonstration

# Performances

## Compression

La  
compression :  
définition

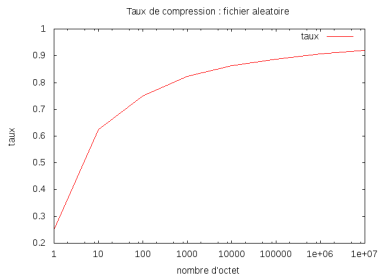
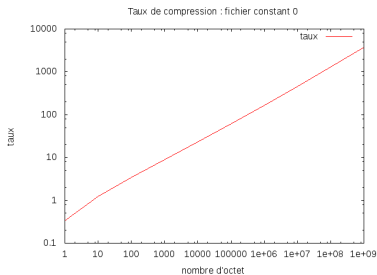
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



# Différences de structures

Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

**Différences  
entre les  
algorithmes**

Différences  
avec les  
principaux  
compresseurs

Conclusion

# Temps d'exécution à la compression

## Compression

La  
compression :  
définition

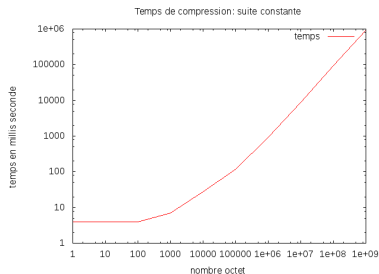
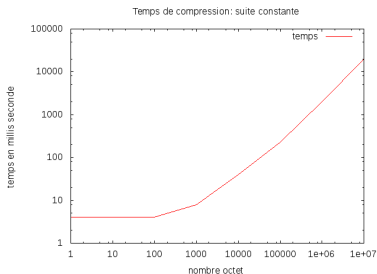
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



# Temps d'exécution à la compression

## Compression

La  
compression :  
définition

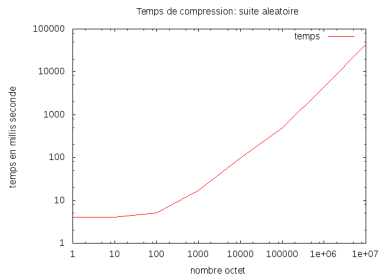
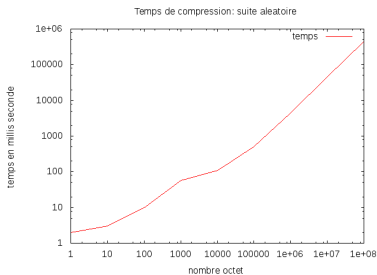
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



# Temps d'exécution à la décompression

## Compression

La  
compression :  
définition

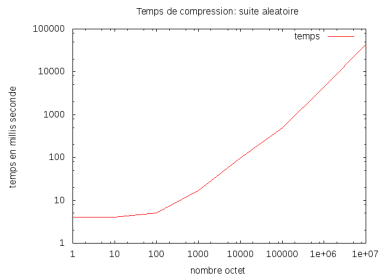
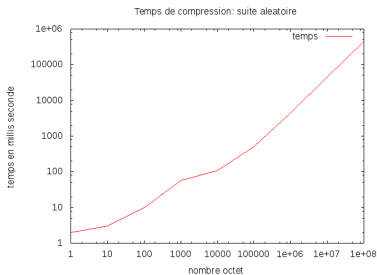
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion





# Temps d'exécution à la décompression

## Compression

La compression :  
définition

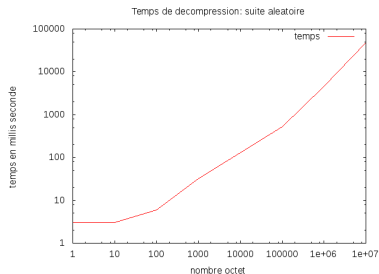
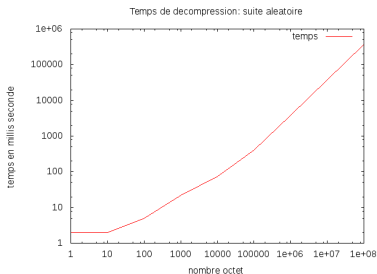
Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion



# Principales différences

## Compression

La compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

	La bible en anglais	Les Fleurs du Mal	Au bonheurs des Dames
taille réelle	883 158	180 199	952 753
zip	3,44	2,51	2,57
Gzip	3,44	2,51	2,644
Bzip2	4,89	3,045	3,719
XZ	4,649	2,826	3,323
Huffman	1,725	1,688	1,747
Lempel-Ziv	1,796	1,310	1,482

# Principales différences

## Compression

La compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion

	fichier PDF	musique MP3	image JPG	vidéo MOV
taille réelle	198 931	8 843 546	3 988 442	54301053
zip	1,246	1,065	1,002	1,0003
Gzip	1,247	1,065	1,002	1,0001
Bzip2	1,212	1,063	1,005	0,9961
XZ	1,261	1,066	1,002	1,0004
Huffman	0,998	1,021	0,999	0,9999
Lempel-Ziv	0,933	0,979	0,929	0,9269

# Conclusion

## Compression

La  
compression :  
définition

Huffman

Lempel-Ziv

Différences  
entre les  
algorithmes

Différences  
avec les  
principaux  
compresseurs

Conclusion