Etude et comparaison de schémas d'analyse Stéganographique d'images numériques

Benoit Roue[1] Patrick Bas[1] Jean-Marc Chassery[1]

[1] Laboratoire des Images et des Signaux de Grenoble

{ Benoit.Roue, Patrick.Bas, Jean-Marc.Chassery}@lis.inpg.fr

Résumé

Ce travail a pour but d'étudier des schémas d'analyse stéganographique. Dans une première partie nous présentons différents algorithmes existants. Dans une deuxième partie nous les évaluons et en décrivons les limites. Enfin nous conclurons sur des améliorations possibles de ces schémas.

Mots clefs

Stéganographie, Stéganalyse, Matrices de coocurrence.

1 Introduction

La stéganographie consiste en l'insertion d'une information dans un support hôte sans que celle-ci ne puisse être décelable. L'intérêt porté à cette discipline s'est accru depuis son apparition dans les documents multimédias (images numériques marquées, logiciels espions...).
La stéganalyse (ou analyse stéganographique) a pour objectif de détecter l'éventuelle présence d'un message inséré et d'en estimer la taille. Plusieurs familles de techniques de stéganographie ont été développées dans la littérature, mais nous nous focalisons ici sur des méthodes basées sur la modification des bits de poids faible (aussi appelés LSB pour *least significant bits*).

2 Schémas d'insertion LSB dans le domaine spatial

Les schémas d'insertion LSB (Figure 1) sont très utilisés en stéganographie d'images numériques car l'insertion d'information ne modifie que les bits de poids faible de l'image (généralement codée sur 256 niveaux de gris) et sont indécelables à l'œil nu. Notre étude se limite à des insertions dans le domaine spatial, mais il existe aussi des schémas d'insertion LSB dans le domaine fréquentiel (*e.g.* insertion dans les coefficient DCT des images JPEG [1]).

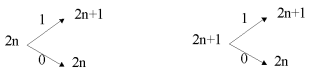


Figure 1 – *Insertion LSB d'un message.*

3 Schémas de stéganalyse des images numériques

3.1 Schéma de Fridrich (Schéma RS)

Cette méthode de stéganalyse, pionnière dans le domaine, consiste tout d'abord à définir un groupe de pixels à valeurs dans $[0...255]$, noté G [2]. Ensuite l'auteur propose de choisir une fonction de discrimination *f* prenant en compte les variations de pixels de G, par exemple *f* telle que (1) :

$$f : \begin{array}{ccc} G & \longmapsto & R \\ x_1 \dots x_n & \longrightarrow & \sum_{i=1}^{n-1} |x_{i+1} - x_i|. \end{array} \quad (1)$$

Enfin, les auteurs définissent les fonctions de permutations des nuances de gris, *F*, telles que (2) :

$$\begin{array}{rcl} F_1 & : & 0 \leftrightarrow 1, 2 \leftrightarrow 3, \dots, 254 \leftrightarrow 255. \\ F_{-1} & : & -1 \leftrightarrow 0, 1 \leftrightarrow 2, \dots, 255 \leftrightarrow 256. \\ F_0 & : & 1 \leftrightarrow 1, 2 \leftrightarrow 2, \dots, 255 \leftrightarrow 255. \end{array} \quad (2)$$

Ces définitions (2) peuvent s'étendre au groupe de pixels G à l'aide d'un masque M de cardinal le cardinal de G, et à valeurs dans $\{-1, 0, 1\}$. L'ensemble des groupes G est alors partitionné en en trois ensembles R, S et U :
$\diamond\ G \in R_M \quad \Leftrightarrow \quad f(F(G)) > f(G)$, G est dit Régulier.
$\diamond\ G \in S_M \quad \Leftrightarrow \quad f(F(G)) < f(G)$, G est dit Singulier.
$\diamond\ G \in U_M \quad \Leftrightarrow \quad f(F(G)) = f(G)$, G est dit Inchangé.

L'analyse se base sur le fait que pour des images naturelles $|R_M| = |R_{-M}|$ et $|S_M| = |S_{-M}|$.Pourtant l'insertion LSB tend à annuler la différence $|R| - |S|$ à mesure que la taille du message augmente. La taille du message inséré est alors obtenue en calculant l'intersection des courbes $|R_M|$ et $|S_M|$. [1]

3.2 Schéma de Dumitrescu

Le schéma de Dumitrescu *et al.* [3] peut être vu comme la formulation mathématique du schéma de Fridrich. Cette méthode est basée sur l'analyse statistique de paires de pixels adjacents d'une image numérique.

[1] $|\ |$ représente le cardinal d'un ensemble.

Définitions. Les auteurs définissent P comme le multi-ensemble de valeurs (en niveaux de gris) de paires de pixels, P est divisé en deux sous-multi-ensembles D_n et C_m :
– D_n est tel que les valeurs des pixels voisins soient $(u, u+n)$ où $(u+n, u)$ où $0 \leq n \leq 2^b - 1$.
– C_m est tel que les valeurs des pixels voisins soient $(u', u'+m)$ où $(u'+m, u')$ où $0 \leq m \leq 2^{b-1} - 1$.
Dans ces deux définitions, b est le nombre de bits utilisés pour coder un pixel. Ainsi les ensembles C_m sont invariants sous insertion LSB.
Finalement D_{2m+1} est fractionné en deux sous-ensembles X_{2m+1} et Y_{2m+1}, définis :
$X_{2m+1} = D_{2m+1} \cap C_{m+1}$ et $Y_{2m+1} = D_{2m+1} \cap C_m$.
Ou plus simplement la paire de pixels dont la composante la plus grande est impaire est dans Y_{2m+1} et celle dont la composante la plus grande est paire est dans X_{2m+1} (cf. Figure 2). Dans une image naturelle, il n'y a aucune

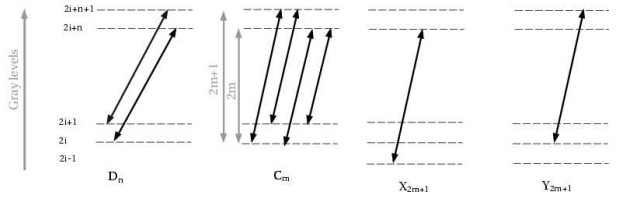


Figure 2 – *Définitions des différents ensembles.*

raison pour que la valeur la plus grande d'un couple de pixels soit paire ou impaire, cela nous amène à l'hypothèse principale du schéma sur laquelle est basée l'analyse stéganographique :

$$E\{|Y_{2m+1}|\} = E\{|X_{2m+1}|\}. \tag{3}$$

Pour résumer, C_m, qui est invariant sous insertion LSB, est divisé en quatre sous-ensembles qui ne le sont pas :
X_{2m-1}, X_{2m}, Y_{2m} et Y_{2m+1}.
L'insertion d'un message cause la variation du nombre d'éléments de ces ensembles et nous permet d'estimer la longueur p de ce message.

Détection de la Stéganographie LSB. Les effets de l'insertion LSB sur les ensembles X_{2m-1}, X_{2m}, Y_{2m} et Y_{2m+1} peuvent être totalement décrits grâces à l'opérateur $\pi \in \{00, 01, 11, 10\}$. Ces modifications sont listées dans le tableau 1. Soient $\rho(\pi, P)$ la probabilité qu'une paire de pixels de P soit modifiée avec π et p la longueur du message inséré. Alors, si le message aléatoirement inseré dans l'image, cela nous amène aux probabilités suivantes :
– $\rho(00, P) = (1 - \frac{p}{2})^2$,
– $\rho(01, P) = \rho(10, P) = \frac{p}{2}(1 - \frac{p}{2})$,
– $\rho(11, P) = (\frac{p}{2})^2$.
Finalement, grâces à l'égalité (3) l'estimation de la longueur du message se fait en résolvant les équations suivantes ((4) et (5)) :

	π			
	00	01	11	10
X_{2m-1}	X_{2m-1}	Y_{2m}	Y_{2m+1}	X_{2m}
X_{2m}	X_{2m}	Y_{2m+1}	Y_{2m}	X_{2m-1}
Y_{2m}	Y_{2m}	X_{2m-1}	X_{2m}	Y_{2m+1}
Y_{2m+1}	Y_{2m+1}	X_{2m}	X_{2m-1}	Y_{2m}

Tableau 1 – *Modifications des différents ensembles sous insertion LSB.*

$$\begin{aligned} &\tfrac{(|C_m| - |C_{m+1}|)p^2}{4} \\ &- \tfrac{(|D_{2m}| - |D_{2m+2}| + 2|Y_{2m+1}| - 2|X_{2m+1}|)p}{2} \\ &+ |Y_{2m+1}| - |X_{2m+1}| = 0, \quad m \geq 1 \end{aligned} \tag{4}$$

et

$$\begin{aligned} &\tfrac{(2|C_0| - |C_1|)p^2}{4} \\ &- \tfrac{(2|D_0| - |D_2| + 2|Y_1| - 2|X_1|)p}{2} \\ &+ |Y_1| - |X_1| = 0, \quad m = 0. \end{aligned} \tag{5}$$

Où p est la proportion de pixels modifiés sous insertion LSB.
De manière à augmenter la précision de l'estimation de la longueur du message, les auteurs proposent d'étendre les équations (4) et (5) avec $0 \leq m \leq 30$.
Dans la suite de ce travail, nous noterons abusivement la somme des cardinaux, $\sum_{m=0}^{30} |X_{2m+1}|$ par $|X_{2m+1}|$ et $\sum_{m=0}^{30} |Y_{2m+1}|$ par $|Y_{2m+1}|$.

3.3 Schéma proposé

Ce nouvel algorithme de détection utilise aussi les statistiques des images naturelles. On définit ainsi 5 ensembles de couples de valeurs de pixels voisins :

$$X \begin{cases} A &= \{(2n+1, 2n+2), (2n+2, 2n+1), n \in P\} \\ B &= \{(2n+1, 2n+3), (2n+3, 2n+1), \\ &\quad (2n+2, 2n), (2n, 2n+2), n \in P\} \\ C &= \{(2n+3, 2n), (2n, 2n+3), n \in P\} \end{cases}$$

$$Y \begin{cases} D &= \{(2n+1, 2n+1), (2n, 2n), n \in P\} \\ E &= \{(2n+1, 2n), (2n, 2n+1), n \in P\} \end{cases}$$

X et Y étant tous deux invariants par permutation LSB. Soit p la probabilité de modification d'un pixel après insertion LSB. En appliquant cette probabilité aux ensembles précédents on obtient[2] ;

$$\begin{cases} |A|' = & (|A| - |B| + |C|)p^2 + (|B| - 2|A|)p + |A| \\ |B|' = & -2(|A| - |B| + |C|)p^2 + 2(|A| - |B| + |C|)p + |B| \\ |C|' = & (|A| - |B| + |C|)p^2 + (|B| - 2|C|)p + |C| \\ |D|' = & 2(|D| - |E|)p^2 - 2(|D| - |E|)p + |D| \\ |E|' = & -2(|D| - |E|)p^2 + 2(|D| - |E|)p + |E| \end{cases}$$

[2]Le symbole ' est utilisé pour désigner un ensemble de l'image après permutation LSB.

Ce qui amène à un système résolvable de 5 équations à 6 inconnues. En effet connaissant A', B', C', D' et E' et en émettant l'hypothèse que $|A| = |E|$. On peut alors déterminer p et en déduire la taille ($\frac{p}{2}$) du message.

4 Résultats et limites des algorithmes présentés

Nous avons testé ces schémas sur la banque d'images Kodak [4] contenant 108 d'images de taille 768×512.

4.1 Protocole de test

Les tests sur les algorithmes ont été effectués de la manière suivante :

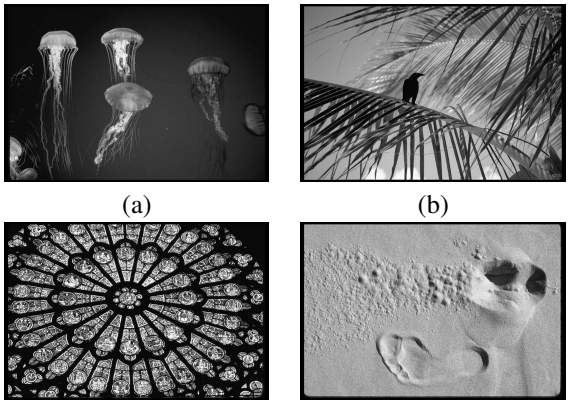
- Le message est insérer aléatoirement avec k clés. En pratique nous avons $k = 10$.
- L'insertion du message est faite sur onze ratios de l'image-test, de telle manière que $0\%, 10\%, 20\%, \ldots, 100\%$ du plan de bit de poids faible de l'image est perturbé.
- Ensuite la longueur du message est estimée pour les trois schémas présentés, sur chaque ratio (m varie de 0 à 30 pour le deuxième).
- Finalement, l' EAM (Erreur Absolue Moyenne) est calculée sur chaque ratio pour chaque clé pour chaque schéma, en calculant la différence entre la longueur réelle la longueur estimée du message.

Cette méthode est utilisée pour chaque image de la banque d'images, et comme résultat final nous avons la moyenne des EAM qui permet de quantifier la performance moyenne de chaque algorithme.
Pour l'algorithme de Dumitrescu nous avons également choisit de calculer la difference relative entre X_{2m+1} et Y_{2m+1} afin d'analyser l'adéquation de la méthode avec l'hypothèse (3) :

$$\epsilon = \frac{2 * ||X_{2m+1}| - |Y_{2m+1}||}{|X_{2m+1}| + |Y_{2m+1}|}$$

4.2 Résultats et limites

Les résultats de ces tests permettent de mettre en évidence une meilleure précision d'estimation concernant le deuxième schéma. En effet une erreur de 7% est obtenue pour la méthode RS, une erreur de 6% pour l'algorithme proposé et seulement une erreur de 2.5% pour le schéma de Dumitrescu.
Cependant une analyse plus en détails des résultats du schéma de Dumitrescu nous a permis de mettre en exergue les points forts mais aussi les limites de cet algorithme. La Figure 3 représente des images qui conduisent à de très bonnes ou de très mauvaises estimations qui sont présentées dans le Tableau 4. Le tableau 3 présente les comportements du schéma Dumitrescu (EAM$\leq 3\%$ pour 70% des images) et indique également la proportion d'images pour lesquelles cet algorithme ne marche pas (pour 2% des images nous avons une erreur d'estimation de 10%).



(a) (b) (c) (d)

Figure 3 – *Quelques images testées : jellies (a), crow (b), notrewindow (c) et sandprints (d).*

Image	EAM Fririch	EAM Dumitrescu	EAM LIS
jellies	6,2	0,5	1,11
crow	5,1	0,8	3,2
notrewindow	26,45	12	12,2
sandprints	17,8	15,2	5,7

Tableau 2 – *EAM des différents schémas.*

EAM	nombre d'images ($\%$)
$\leq 1\%$	13
$\leq 2\%$	42
$\geq 10\%$	2

Tableau 3 – *Résultats du schéma de Dumitrescu sur la banque d'images Kodak.*

Les erreurs d'estimations pour l'algorithme proposé sont plus importantes sauf pour certaines images comme *sandprints*). En effet la sommation sur m utilisée dans l'algorithme de Dumitrescu n'est pas adapté à l'image, alors que dans le schéma proposé elle n'intervient pas.
La méthode RS fournit des estimations fiables sur la plupart des images, cependant certaines erreurs d'estimations, telles que pour l'image *notrewindow* sont très importantes.

4.3 Analyse du schéma de Dumitrescu

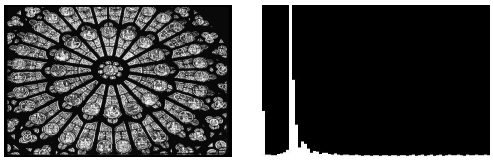
Comme nous l'avons vu, le schéma de Dumitrescu est très fiable sur la plupart des images, mais dans certains cas l'erreur d'estimation devient significative. Notre travail a ensuite conduit à l'identification de ces erreurs qui sont :
- la non-égalité de cardinaux,
- une distribution des statistiques conjointes de l'image non appropriée à l'analyse.

Non-égalité des cardinaux. Dans ce cas-ci, l'hypothèse principale (3) du schéma de Dumitrescu est une hypothèse qui n'est pas vérifiée, car il existe des images naturelles

pour lesquelles elle n'est pas valide.
Comme nous pouvons le voir dans le tableau 4, pour l'image "notredamewindow" la différence relative entre X_{2m+1} et Y_{2m+1} est importante.
L'histogramme de cette image est décrit Figure 4. Sur cette figure nous pouvons remarquer notamment la présence de deux pics singuliers fausse l'égalité des cardinaux (niveaux de gris 9 et 10). Une telle configuration augmente considérablement la valeur de $|Y_1|$ et brise alors l'égalité (3).

Image	EAM (%)	ϵ
jellies	0.54	0.017
crow	0.87	0.004
notredamewindow	12.01	0.22
sandprints	15.18	0.046

Tableau 4 – *Exemples d'erreurs et différences de cardinaux.*



(a) (b)

Figure 4 – *Image notrewindow (a) et son histogramme (b).*

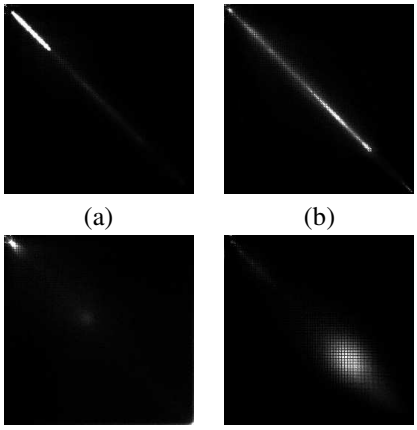
Statistiques conjointes de l'image. Pour plusieurs images le schéma de Dumitrescu n'est pas précis en dépit de la validité de 3), dans ce cas le problème provient du choix empirique des valeurs sur lesquelles m est sommé. Effectivement, si on considère l'image "sandprints", $X_{2m+1} \simeq Y_{2m+1}$, et pourtant l'erreur d'estimation n'est pas négligeable (cf. Tableau 4). Dans la partie 3.2 nous avons vu que le schéma de Dumitrescu est basé sur des statistiques conjointes (statistiques sur des paires de pixels adjacents), nous avons donc étudié ces statistiques pour les images dont l'erreur d'estimation est grande, en utilisant des matrices de coocurrence définies par :

$$MC_t(a,b) = Card\{(s, s+t) \in I^2 \backslash V[s] = a, V[s+t] = b\} \quad (6)$$

Où $MC_t(a,b)$ un élément de la matrice de coocurrence en (a,b), s un pixel de l'image, $V[s]$ sa valeur, I l'image et t la distance entre deux pixels ($t = (1,0)$). Sur la Figure 5 nous pouvons observer les matrices de coocurrence des images présentées Figure 3 qui illustrent la relation entre l'erreur d'estimation et la distribution de la matrice de coocurence :

– Pour les images offrant une faible erreur d'estimation la distribution des maxima des probabilités conjointes est localisée autour de la diagonale, ainsi la sommation pour m de 0 à 30 est adaptée.

– Pour les images offrant une forte une forte erreur d'estimation la distribution de ces maxima n'est pas localisée autour de la diagonale, la sommation pour m de 0 à 30 n'est alors pas adaptée.



(a) (b)

(c) (d)

Figure 5 – *Matrices de coocurrence des Images : jellies (a), crow (b), notredamewindow (c) et sandprints (d).*

5 Conclusions et perspectives

Sur une large majorité des images utilisées les algorithmes ont une précision d'estimation très largement suffisante, et plus spécialement sur celui de Dumitrescu. Cependant quelques images possédant des caractéristiques spécifiques aux probabilités marginales ou conjointes entraînent des erreurs d'estimation importantes. Ces caractéristiques étant connues il parait intéressant de les prendre en compte afin de rendre les schémas plus efficaces. Les perspectives pour la suite de ce travail sont donc :
◇ Etablir une fonction de contraste permettant de classifier plus précisément les images afin de prévoir par quel algorithme la détection sera optimale.
◇ Analyser les images avec un algorithme tenant compte des probabilités (conjointes et marginales) de celles-ci.
◇ Utiliser des techniques de traitement d'image afin de rendre l'estimation plus précise (Filtre passe-bas, manipulation d'histogramme, *etc*).

Références

[1] D.Upham. Jpeg-jsteg, modification of the independent jpeg group's jpeg software for 1-bit steganography in jfif output files. ftp ://ftp.funet.fi/pub/crypt/steganography/, 1992-1997.

[2] J.Fridrich, M.Goljan, and R.Du. Reliable detection of LSB Steganogeraphy in color and grayscale images. In *ACM Workshop on Multimedia and Security*, pages 27–30, 2001.

[3] S.Dumitrescu, X.Wu, and Z.Wang. Detection of LSB steganography via sample pair analysis. In *IEEE transactions on Signal Processing*, pages 1995–2007, 2003.

[4] Kodak database. ftp ://ftp.kodak.com/www/images/pcd/.

[5] H. Farid. Detecting hidden messages using higher-order statistical models. In *International Conference on Image Processing*, NY, 2002.