Technologic Papers

Improving Images Steganography

מאת אפיק קסטיאל (cp77fk4r)

פתיחה

"סטגנוגרפיה היא האמנות והמדע של הסתרת מסרים, באופן שאף אחד זולת המקבל לא יוכל לראותם או לדעת על קיומם. בניגוד לקריפטוגרפיה שבה קיום המידע עצמו אינו מוסתר, אלא רק תוכנו. המילה סטגנוגרפיה מקורה בלטינית, סטגנו פירושה מכוסה או חבוי . זוהי אמנות עתיקה למדי, כבר לפני הספירה הנוצרית יש תיעוד על שימוש בסטגנוגרפיה, כמו כיסוי מסר שנכתב על לוח עץ בעזרת שעווה. בדרך כלל מסר סטגנוגרפי נראה על פניו כמשהו תמים אחר, כגון תמונה, קטע עיתונות, רשימת קניות או כל דבר אחר שאינו מעורר חשד, המשמש ככיסוי למסר האמיתי."
(מתוך ויקידפיה העברית)

במאמר זה אני מתכוון להציג מספר הבטים או "שיטות" בהן הסטגנוגראפים משתמשים בכדי ליעל את האלגוריתמים שלהם. נתחיל בהצגת מספר הנחות בסיס ונגדיר מספר ביטויים חשובים , לאחר מכן נציג שיטות שונות ליעול האלגוריתמים.

מונחים כללים

כאשר אנו מדברים על סטגנוגרפיה ויזואלית, המושג "רעש" אינו מתייחס לרעשים ווקאלים אלא מצביע על סטיה בצבעי התמונה המקורית. הרעיון העומד מאחורי מדידת רעש הוא לקבוע את כמות הרעש שיכול להופיע בתמונה מבלי שעין הצופה הממוצע תוכל לקבוע בוודאות כי אכן בוצעו שינויים בתמונה. מספר מונחי יסוד הקיימים בנושא יוכלו לעזור לנו להגדיר מה הקריטריונים שעל-פיהם נעבוד:

- **SNR**- [Signal-to-Noise Ratio] - **היחס בין עוצמת האות לעוצמת הרעש הכולל בהעברת אות ממקור מסוים ליעד**. הרעש עשוי להיות רעש נלווה לאות המקורי (למשל רעש שוט), רעש שקיים בתווך שבו מועבר השידור ,או רעש בגלאי.
- **PSNR**- [Peak Signal-to-Noise Ratio] - **היחס המרבי האפשרי בין הספק האות לבין הספק הרעש** שמשפיע על אמינות האות המוצג.
- **MSE**- [Mean Squared Error] – הוא **ממוצע השגיאות (השינוי היחסי) בריבוע**. (מסתכלים על ממוצע בריבוע בגלל שככל שהממוצע יגדל כך נוכל לעקוב אחריו בקלות יותר)

מתחילים

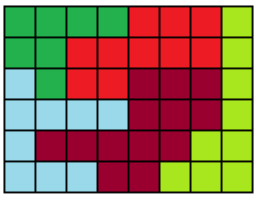
אנחנו מחפשים עקרונות ונקודות בהם אנחנו יכולים להשתמש בכדי להשפיע על איכות ה-Embed (הטמעה) של כל אלגוריתם סטגנוגראפי שהוא. נתחיל בהצגת משפט חשוב:

- אלגוריתם סטגנוגראפי "קלאסי" ישאף לחלק את המידע אותו הוא מעוניין להטמיע (Payload) למספר רב של חלקים ככל הניתן, ביחס לגודל האיזור בו הוא מחביא את המידע (Container), בכדי לשנות את התוצר כמה שפחות (שאיפה ל-MSE כמה שיותר קטן).

זה נשמע הגיוני- נניח ואנחנו רוצים להחביא מרשם סודי להעשרת אורניום בתוך תמונה, כך שנוכל להעביר אותה בשדה-התעופה מבלי שיגלו אותנו. ההנחה הראשונה אומרת שככל ש"נפזר" את המרשם ביותר מקומות בתמונה כך הצופה ירגיש בשינוי כמה שפחות- כל שנפזר את ה-Payload כך חלקיו יהיו קטנים יותר וכך ירגישו בהם פחות.

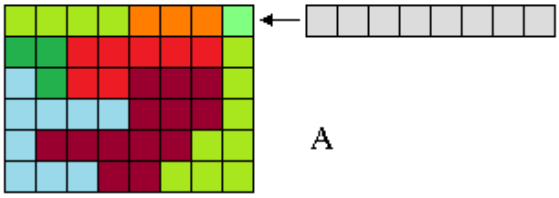
דוגמא

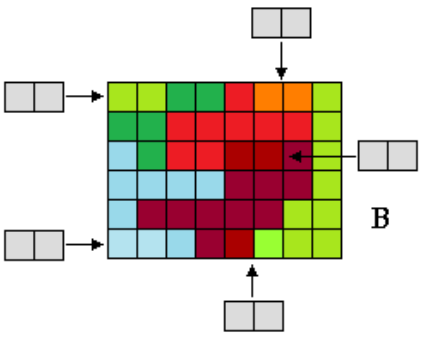
נניח שיש לנו את התמונה הבאה: (48 בתים).



ואנחנו מעוניינים להחביא בה את המידע הבא: (8 בתים)

שימו לב להבדל בתוצאה (פחות או יותר- MSE) בין שתי האפשרויות שלנו:





- A - הכנסנו את ה-Payload בחתיכה אחת, פשוט כמו שהוא.
- B - חתכנו את ה-Payload לארבעה חתיכות בנות 2 בתים כל אחת.

הדוגמאות עצמן אינן מדגימות במדויק את שינוי הגוונים שהיו אמורים להתרחש אלא גוונים "קרובים" פחות או יותר בכדי להמחיש את הרעיון.

הכוונה ב- "ירגישו בהם כמה שפחות" היא שאם מישהו באמת יקח את שתי התמונות, יציב אותן אחד ליד השניה וינסה למצוא הבדלים הוא ישים לב לכמה שפחות שינויים , זאת בדיוק ההגדרה של אלגוריתם סטגנוגרפי מוצלח לעומת אלגוריתם סטגנוגרפי פחות מוצלח.

אם על ידי שימוש ביחס של ה-Container וה-Payload הצלחנו לגלות, את מספר החלקים המירבי בו נוכל לחלק את ה-Payload, מה עוד נשאר לשפר באלגוריתם שלנו?

כמובן שמבחינת שיפור ה- MSE, בעזרת מציאת כמות החלוקה האפקטיבית ביותר אין יותר מה לשפר , יחס זה יחס , אבל אנחנו יכולים ל שפר נתון אחר- נוכל לשפר את האלגוריתם שלנו ע "י מציאת המיקום האפקטיבי ביותר להחבאת ה-Payload!

כאן נשאלת השאלה- האם המיקום של כל חתיכה משנה? ואם כן, איפה הכי כדאי למקם אותן?

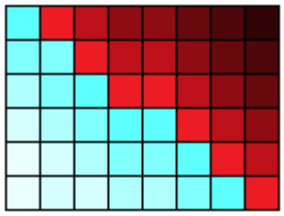
בכדי לענות על שאלות אלה אנחנו צריכים לזכור דבר אחד - **מדובר בהשוואה שלא מתבצעת בעזרת מחשב, אלא בעזרת העין האנושית**. אתם שואלים מה הקשר? הרעיון הוא לנצל חולשות בעין האנושית (או יותר נכון במח האנושי שמפרש את מה שהעין שלנו רואה) ולפיכך למקם את החתיכות שלנו.

מספר עובדות חשובות על העין האנושית

- **לעין האנושית קשה יותר לזהות סטיות בין גוונים כהים מאשר סטיות בין גוונים בהירים.** משמעות הדבר שאותו PSNR בגוונים כהים יתן לנו MSE נמוך הרבה יותר מבגוונים בהירים.
- **לעין האנושית קשה יותר לזהות סטיות בשולי התמונה מאשר סטיות במרכזה , ניתן להבין מכאן** שאותו PSNR בשולי התמונה יתן לנו MSE נמוך מבמרכזה.
- **לעין האנושית קשה יותר לזהות סטיות כשמדובר במצב של חפיפה בין ניגוד גוונים.** לפיכך, אם נמצא אזור בתמונה שיש לו SNR גבוה יחסית לשאר התמונה, נוכל ליצור שם PSNR שיתן לנו MSE נמוך יחסית לPSNR שממוקם באיזור בעל SNR נמוך ביחס לשאר התמונה.

מיפוי ביחס לגוונים:

נבחן את העובדה הראשונה, לפיה לעין האנושית קשה יותר לזהות סטיות בין גוונים כהים. נראה איך זה מתקיים. נניח וקיימת לנו התמונה הבאה:



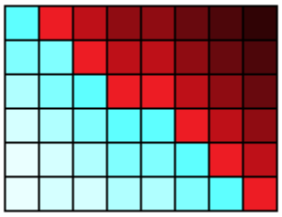
אנחנו צריכים לבצע בדיקת גוון (Shade Analysis), לבדוק מה ממוצע הגוונים בתמונה ולמפות את התמונה לאיזורים ולכל איזור לתת ניקוד ביחד לממוצע. נניח ולאחר שליפת הממוצע מניתוח הגוונים קיבלנו שמרכז התמונה הוא הממוצע. המפוי שלנו יראה כך:

0	0	1	2	2	3	4	5
-1	0	0	1	1	2	3	4
-2	-1	0	0	0	1	2	3
-3	-2	-1	0	0	0	1	2
-4	-3	-2	-1	-1	0	0	1
-5	-4	-3	-2	-2	-1	0	0

בדוגמא זו נוכל לראות שככל שהגוון כהה יותר , כך הרלוונטיות שלו בלהכיל PSNR שייתן לנו MSE-קטן יחסית- גדלה.

העובדה השנייה מבהירה כי לעין האנושית קשה יותר לזהות סטיות בשולי התמונה מאשר במרכזה. לעין האנושית (או שוב, למוח האנושי) **קשה יותר לקלוט סטיות בגוונים המקוריים ככל שהם מרוחקים מהמרכז**. למה? בגלל "הפוקוס" האוטומטי שהעין מבצעת. אם תחזיקו עיפרון מולכם ותסתכלו רק עליו , העפרון יראה בבירור והאובייקטים בצדדיו (לא מאחוריו) יהיו חסרי פרטים, העין שלנו רגישה פחות ופחות לפרטים במסגרות התמונה.

לכן, כמעט תמיד נשאף ליצור את הסטיות שלנו רחוק ככל שניתן מהמרכז . בואו נראה איך זה בא לידי ביטוי, אם קיימת לנו התמונה הבאה:



אנחנו (בהתייחסות רק לעובדה השניה) ננקד אותה באופן הבא:

3	2	2	2	2	2	2	3
3	2	1	1	1	1	2	3
3	2	1	0	0	1	2	3
3	2	1	0	0	1	2	3
3	2	1	1	1	1	2	3
3	2	2	2	2	2	2	3

כמובן שבמציאות אנחנו לא מתחשבים רק במיקום הבתים, אך לכל דבר הזמן שלו. אם נממש אלגוריתם שיבצע מיפוי רלוונטיות על פי שתי עובדות אלו, התוצאה תראה כך:

0	0	1	2	2	3	4	5
-1	0	0	1	1	2	3	4
-2	-1	0	0	0	1	2	3
-3	-2	-1	0	0	0	1	2
-4	-3	-2	-1	-1	0	0	1
-5	-4	-3	-2	-2	-1	0	0

+

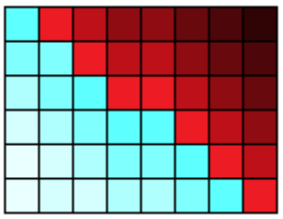
3	2	2	2	2	2	2	3
3	2	1	1	1	1	2	3
3	2	1	0	0	1	2	3
3	2	1	0	0	1	2	3
3	2	1	1	1	1	2	3
3	2	2	2	2	2	2	3

=

3	2	1	3	4	3	5	8
2	2	2	2	2	3	5	7
1	1	1	0	0	2	4	6
0	0	0	0	0	1	3	5
-1	-1	-1	0	0	1	2	4
-2	-2	-1	0	0	1	2	3

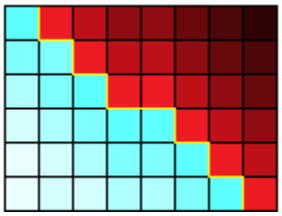
שימו לב לתוצאה כאשר מתחשבים בשתי העובדות ולא רק באחת מהן , ננסה להבין ולשלב גם את העובדה השלישית שלנו.

לפי העובדה השלישית, לעין האנושית קשה יותר לזהות סטיות כשמדובר במצב של חפיפה בין ניגוד גוונים. עובדה זו מעט מסובכת יותר היות ומדובר בהבנה של מצב ניגודי בין גוונים:



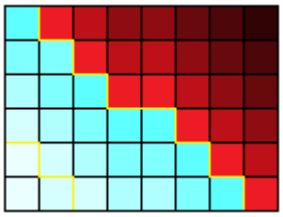
(כאן הזיהוי עוד פשוט יחסית- תחשבו שמדובר בפיקסלים הרבה הרבה יותר קטנים)

העין האנושית כמובן תמשך ישירות למרכז התמונה היות ושם מעבר הגוונים הוא החד ביותר:



הגבול הצהוב מסמל את הניגוד המירבי ביותר בתמונה , או בשפה מקצועית יותר - כאן ה-SNR הוא הגבוה ביותר. העובדה השלישית מבהירה כי אם ניצור PSNR- סטייה עם גוון לא קשור - דווקא איפה שיש SNR גבוה, העין האנושית פחות תשים לב לסטייה.

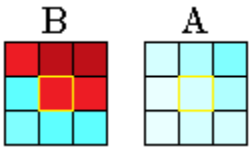
אם נרצה להסביר את זה בצורה פשוטה, קשה יותר לעין שלנו למצוא פרח בצבע מסויים בתוך שדה עם מלא פרחים בצבעים שונים, מאשר למצוא ענן אפור בשמיים כחולים לגמרי. שימו לב לדוגמא:



איזה קו יותר קל לזהות? זה שבמרכז או זה שבפינה השמאלית התחתונה?

צורת הניקוד כאן טיפה יותר מסובכת , אנחנו מחלקים כל פיקסל על ידי שיכלול של כל שמונת הפיקסלים סביבו, מבצעים SNR מקומי לכל פיקסל, ולפי ה-MSE (המקומי) נותנים ציון לפיקסל. כאן זה הפוך- ככל שה-MSE המקומי גדל כך הרלוונטיות להחביא שם מידע גדלה גם היא

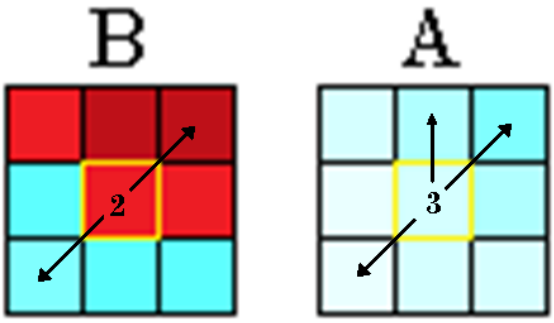
הסתכלו על המצבים הבאים:



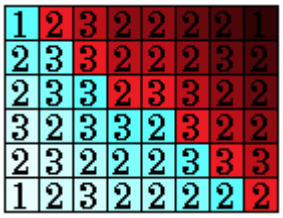
אם ננסה ליצור PSNR בפיקסל המסומן ב-A נקבל MSE גבוה בהרבה מיצירת PSNR בפיקסל המסומן ב-B. למה? כי ה-SNR המקומי ביחס לפיקסל המסומן ב -A הרבה יותר נמוך מה -SNR המקומי ביחס לפיקסל המסומן ב-B.

מהלך המיפוי מתנהל כך: בוחרים פיקסל, בודקים בכמה פיקסלים שונים ממנו הוא "נוגע" ולפיכך נותנים לו ציון (אם היינו רוצים להיות ממש יעילים , היינו יוצרים פונקציה שבודקת מה רמת הניגוד בכל פיקסל ביחס לפיקסלים שלידו).

לדוגמא:

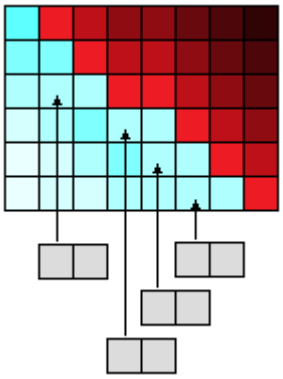


הפיקסל המסומן ב-A נוגע בשלושה גוונים השונים ממנו , הפיקסל המסומן ב-B נוגע רק בשניים . לכן, המיפוי של התמונה שלנו יראה כך:



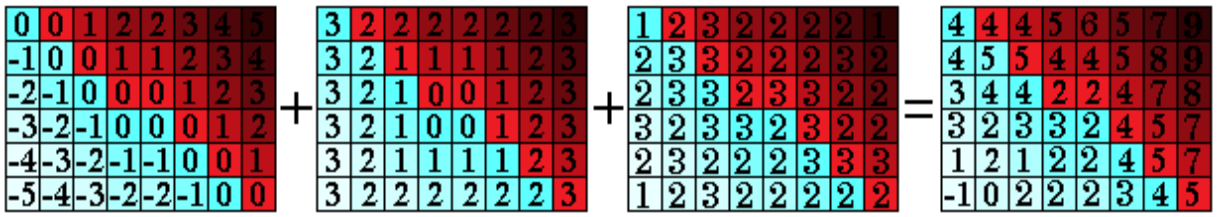
הבה נמחיש זאת, במידה ונרצה להכניס את התמונה הקודמת את המידע הבא:

לאור הבנת העובדה השלישית , נעדיף למקם את המידע איפה שה-SNR הכי גבוהה, כך נקבל מינימום של MSE. לכן, נעדיף לבצע את ההטמעה באופן הבא:



(שימו לב שגם בפיקסלים כאלה מדובר בהטמעה כמעט מושלמת)

כאן ננסה לראות איך המיפוי יראה כשנתחשב בשלושת העובדות שלמדנו



לאחר שקיבלנו את מפת הרלוונטיות המשכללת בתוכה את שלושת העובדות, נוכל בקלות לדעת היכן הכי כדאי לנו למקם את המידע שלנו.

סיכום

יש עוד המון "עובדות" או לוגיקות מהן אפשר להגיע לעוד דרכים ליעול הטמעה וכך להגיע לרמות סטגנו מאוד גבוהות. הרעיון הכללי הוא להבין ממי אנחנו רוצים להסתיר את המידע (במקרה שלנו- העין האנושית, אך יש המון "גורמים עויינים" כגון אוזן, תוכנת מחשב וכו'). מימוש אלגוריתם שכזה איננו קשה במיוחד וניתן לביצוע גם אם נתבסס רק על שלוש עובדות אלו. כמו כן, נוכל ליעל אותו יותר אם נתייחס לנקודות הבאות:

- התייחסות לגוונים נוגדים המבלבלים את העין.
- התייחסות לגוונים משלימים.
- התייחסות לרעשים נוספים ולאו דווקא לממוצע ה-SNR.
- שינוי הניקוד על פי איזורים קטנים יותר בתמונה.
- שינוי הניקוד ע ל פי שטחים בעלי SNR ממוצע לשמונת גזרות SNR קרובות (ולא רק ברמת הפיקסל).