



# **Estegoanàlisi d'imatges en el domini espacial**

**Daniel Lerch**

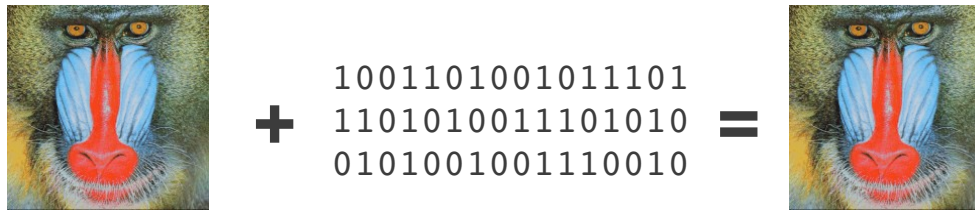
**22 de Novembre del 2011**



# **Introducció a la esteganografia i l'estegoanàlisi d'imatges**

# Esteganografia i Estegoanàlisi

- **Esteganografia:** Del grec *στέγανός* (**steganos**, encobert). Ciència que estudia la comunicació de missatges de manera que la seva existència **no sigui detectada**.



- Cal diferenciar-la de la criptografia, del grec *κρυπτός* (**krypto**, ocult, secret). Mentre que la **criptografia** tracta d'ocultar el contingut d'un missatge, la **esteganografia** tracta d'ocultar-ne la seva existència.

$$\text{Encrypt} \begin{pmatrix} 1001101001011101 \\ 1101010011101010 \\ 0101001001110010 \end{pmatrix} = \begin{matrix} 0101111101010001 \\ 1011000100100011 \\ 0111011101110100 \end{matrix}$$

- **Estegoanàlisi:** Ciència que estudia la **detecció de missatges ocults** fent servir esteganografia.

# Formes d'esteganografia en Imatges

Hi ha dues maneres habituals d'inserir informació a una imatge:

- Inserció amb **substitució del LSB**.
- Inserció amb **increment/decrement del valor del píxel ( $\pm 1$ )**.

I:	101	100	99	100	102
M:	0	1	1	0	1

Substitució LSB



<b>100</b>	<b>101</b>	99	100	<b>103</b>
------------	------------	----	-----	------------

I:	101	100	99	100	102
M:	0	1	1	0	1

$\pm 1$

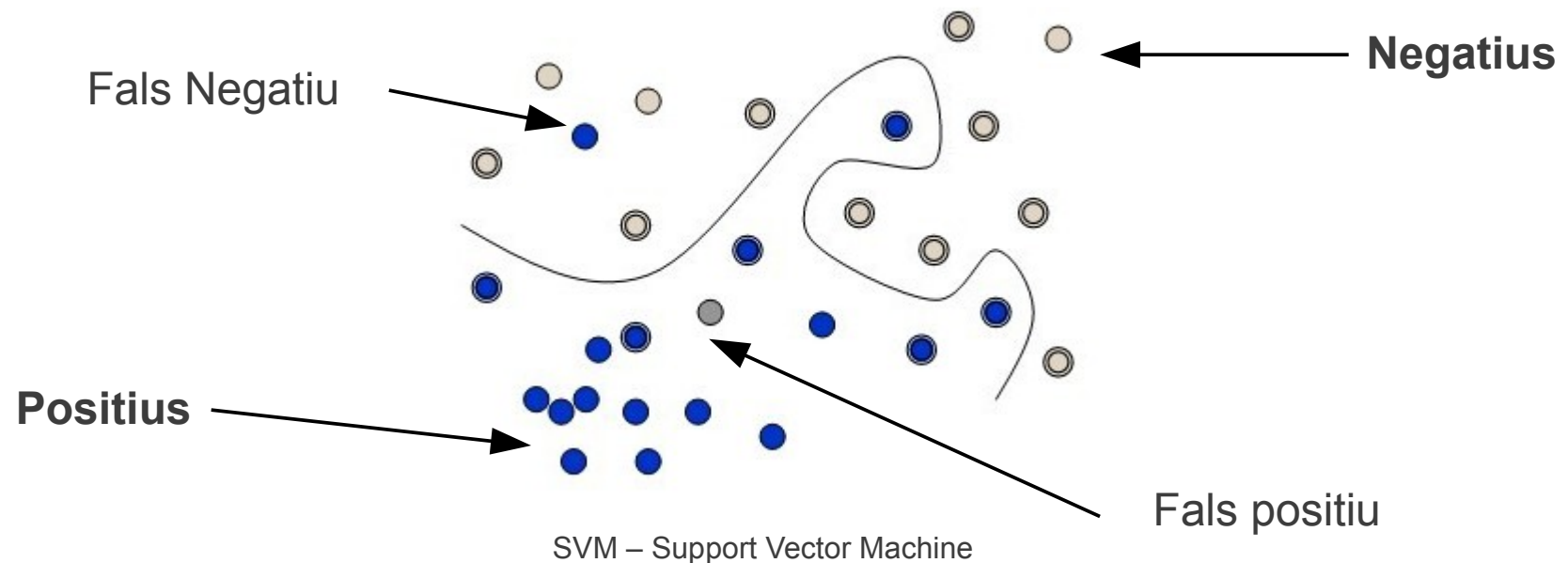


<b>100</b>	<b>101</b>			<b>101</b>
<b>102</b>	<b>99</b>	99	100	<b>103</b>

# Procediment d'estegoanàlisi en l'estat de l'art

## Procediment d'estegoanàlisi en l'estat de l'art:

- 1) Extracció de característiques d'una BD d'imatges.
- 2) Entrenament d'un classificador.
- 3) Avaluació del mètode amb un conjunt de test.

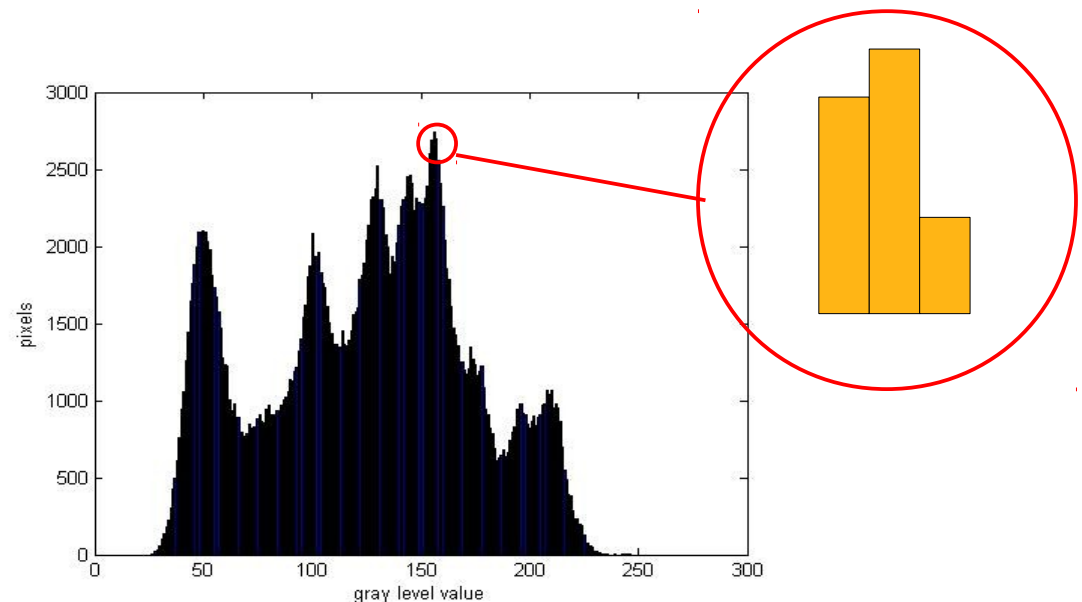




# **Anàlisi de l'histograma d'una imatge**

# Anàlisi d'imatges mitjançant l'histograma

- **Histograma:** Representació gràfica d'una variable en forma de barres. L'alçada de cada barra és proporcional a la freqüència dels valors representats.



- Cada barra representa la intensitat del píxel corresponent.
- L'anàlisi de l'histograma és una eina fonamental per a la detecció d'alteracions a la imatge.

# Substitució del LSB (LSB Replacement)

101	100	99	100	102
-----	-----	----	-----	-----

0	1	1	0	1
---	---	---	---	---

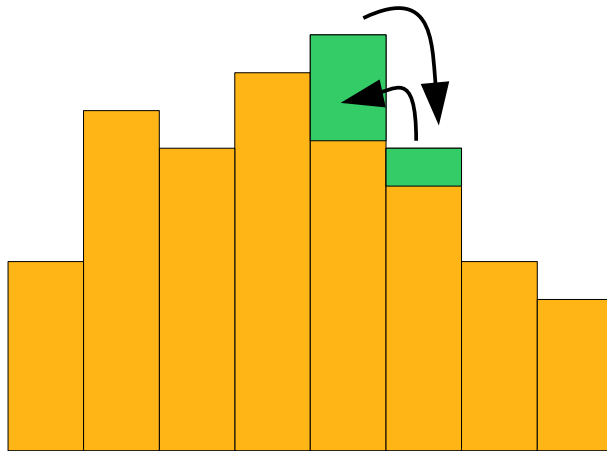
Substitució LSB



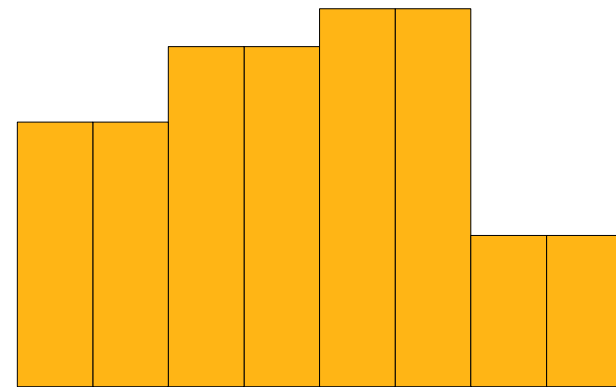
100	101	99	100	103
-----	-----	----	-----	-----

- Els nombres **parells** només **incrementen** el seu valor.
- Els nombres **senars** només **decrementen** el seu valor.
- No es fa servir en mètodes de l'estat de l'art degut a l'**atac RS** [Fridrich 2001], que detecta insercions del **3%**.

Les insercions fan que les parelles de barres consecutives tendeixin a tenir la mateixa alçada



Histograma original



Histograma amb missatge ocult



# Increment/Decrement del valor del píxel ( $\pm 1$ / LSB Matching)

101	100	99	100	102
-----	-----	----	-----	-----

0	1	1	0	1
---	---	---	---	---

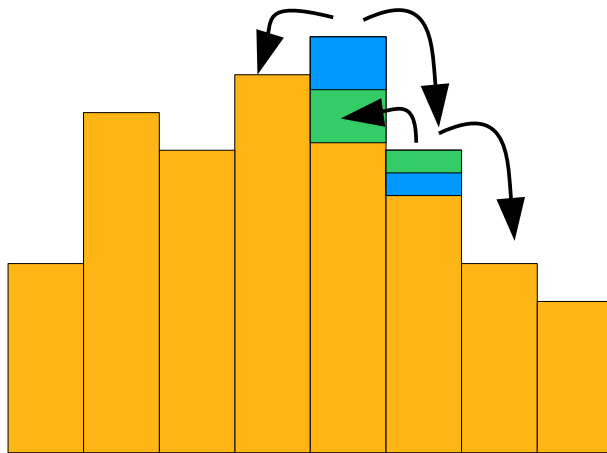
$\pm 1$   
↓

100	101	99	100	101
102	99			103

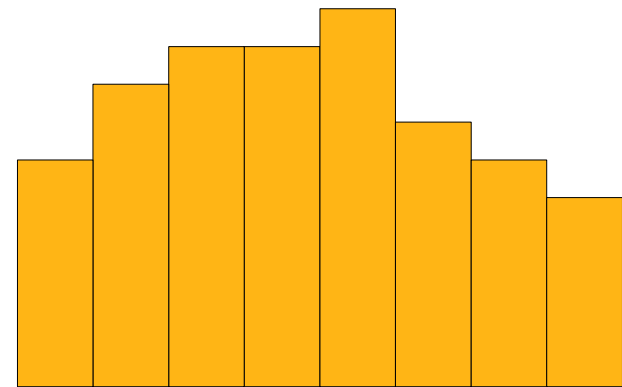
- Quan s'ha de modificar un píxel es tria aleatòriament si fer-ho amb +1 o amb -1.

- Utilitzat en els mètodes de l'estat de l'art.

**Les insercions fan que l'histograma es torni més suau, però aquesta suavització la provoca també el soroll, cosa que la fa difícil de detectar.**



Histograma original



Histograma amb missatge ocult

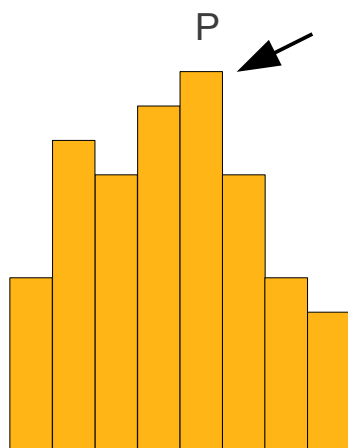


# **Sistemes d'esteganografia basats en el desplaçament de l'histograma**

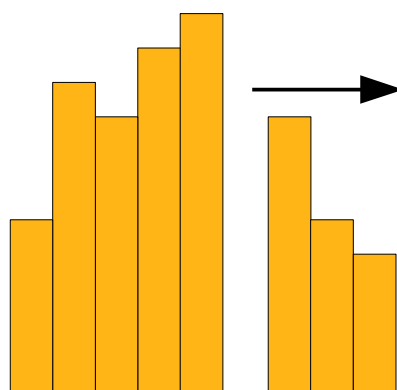
# Desplaçament de l'Histograma: [Ni 2003]

## ▪ Procediment d'inserció:

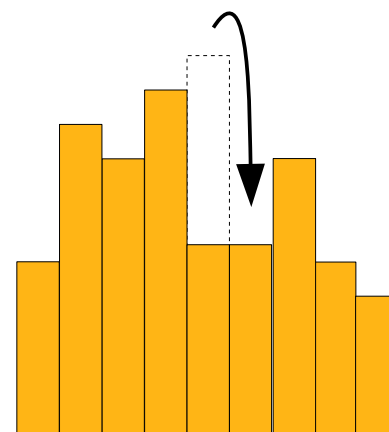
- 1) Localització d'una barra de l'histograma prou alta (píxel P).
- 2) Desplaçament de l'histograma (sumant 1 a tots els píxels amb valor més gran que P).
- 3) Inserir informació als píxels amb valor P. Es deixarà P per a emmagatzemar un 0 i P+1 per emmagatzemar un 1.



Histograma original



Histograma desplaçat

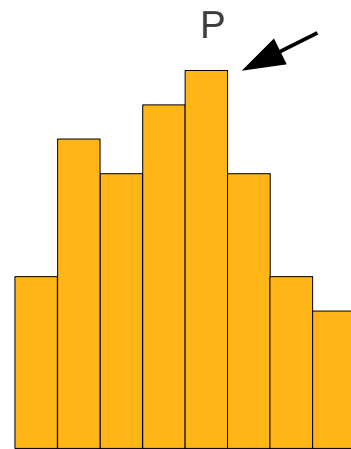


Histograma amb missatge ocult

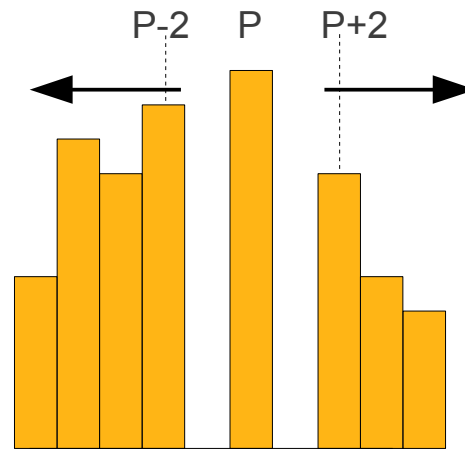
# Desplaçament de l'Histograma: [Mohsenzadeh 2009]

## ▪ Procediment d'inserció:

- 1) Localització d'una barra de l'histograma prou alta (píxel P).
- 2) Desplaçament de l'histograma cap a la dreta (sumant 1 a tots els píxels amb valor més gran que P) i cap a l'esquerra (restant 1 a tots els píxels amb valor més gran que P).
- 3) Recórrer la imatge cercant píxels amb valor P+2 i P-2. Per inserir un zero substituir un dels seus veïns per P-1, per a inserir un 1, per P+1. El veí a triar dependrà d'una clau compartida entre emissor i receptor.



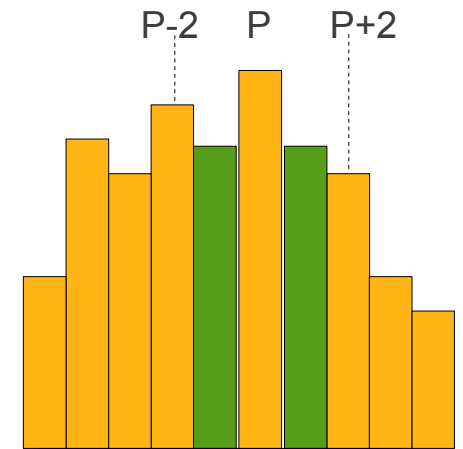
Histograma original



Histograma desplaçat

a	b	c
d	<b>P±2</b>	e
f	g	h

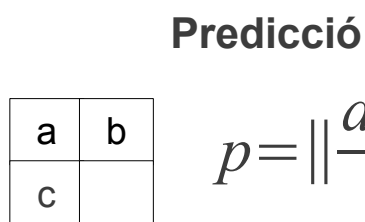
Veïns de P±2



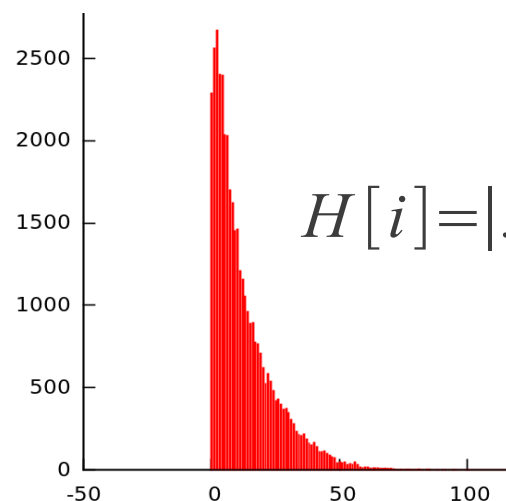
Histograma amb missatge ocult

# Desplaçament de l'Histograma: HSPE I

- **Mètode HSPE:** Desplaçament de l'histograma de predicció d'errors (*Histogram Shifting of Prediction Errors*).
- Es poden fer servir una gran quantitat de fórmules per a predir el valor d'un píxel veí.



$$p = \left\| \frac{a + b + c}{3} \right\|$$



99	101	101	100	101
98	100	100	103	100
96	98	101	102	105
97	99	102	104	103
95	96	99	101	103

Matriu de píxels

**Predicció**

	99	101	100	102
	98	99	101	102
	97	99	102	104
	97	99	102	103

Matriu de prediccions

**Diferència**

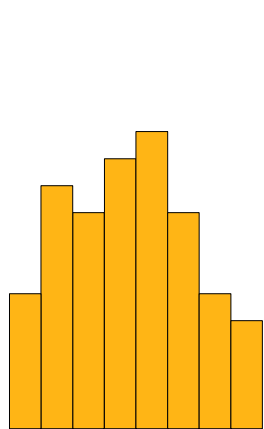
	1	1	3	2
	0	2	1	3
	2	3	2	1
	1	0	1	0

Matriu d'errors de predicció

# Desplaçament de l'Histograma: HSPE II

## ▪ Procediment d'inserció:

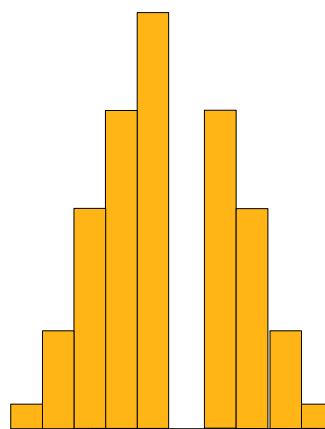
- 1)Necessitem una fórmula que ens permeti realitzar una predicció del valor d'un píxel partint dels seus píxels veïns.
- 2)Generem una matriu formada per la diferència entre el valor del píxel i la seva predicció.
- 3)Generem l'histograma HSPE fent servir aquesta matriu.
- 4)A continuació se segueix el mateix procediment que al mètode anterior [Ni 2003].



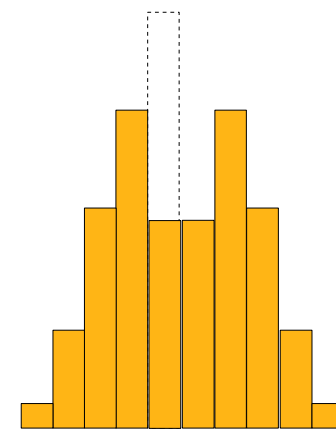
Histograma original



Histograma HSPE



Histograma desplaçat



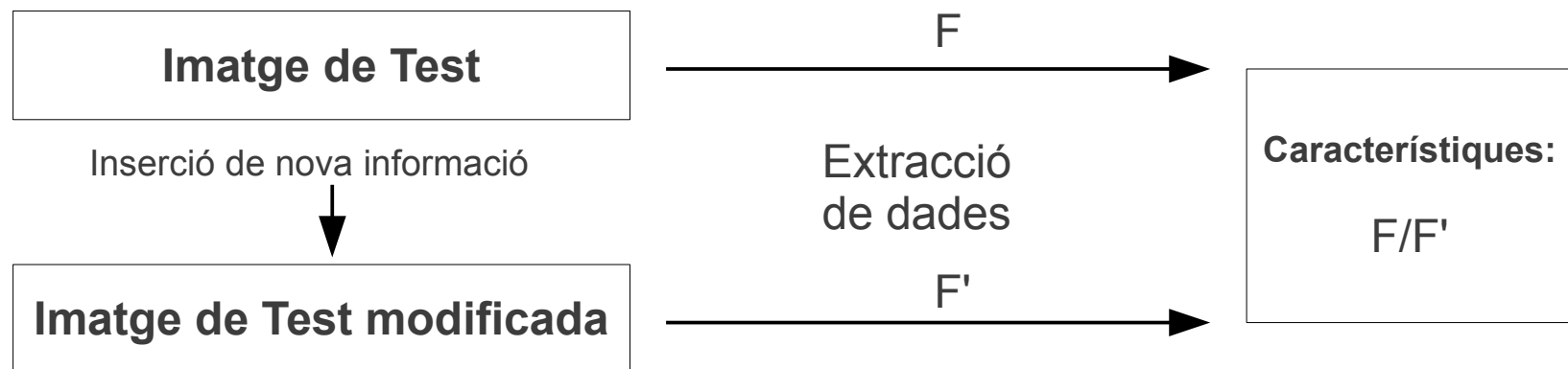
Histograma amb missatge ocult



# **Mètodes d'estegoanàlisi**

# Mètodes Proposats

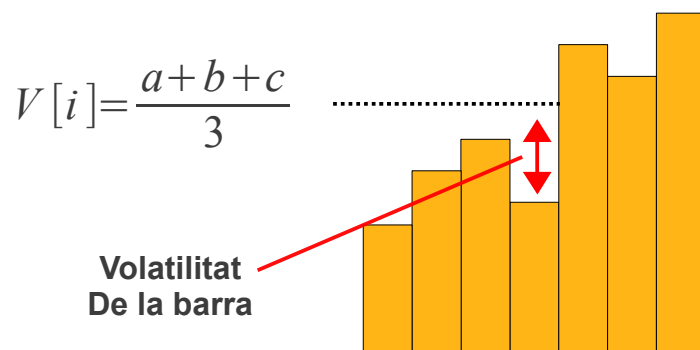
- **Mètode I:** La volatilitat com a característica
- **Mètode II:** Patrons de píxels veïns.
- Basats en sistemes de classificació:
  - Extracció de característiques
  - Entrenament
  - Test
- Procediment per a generar característiques:
  - Els mètodes actuals obtenen les característiques de la imatge analitzada.
  - Els mètodes presentats obtenen les característiques com a relacions entre les dades abans i després d'inserir informació.





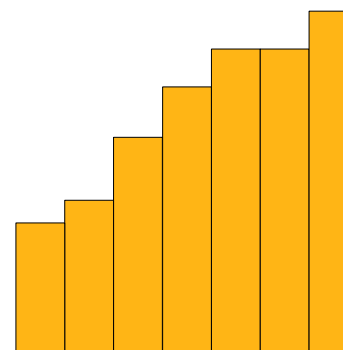
# Mètode I: Volatilitat de l'histograma

- La volatilitat mesura quant diferents son les barres dels histogrames respecte dels seus veïns.



$$V[i] = \frac{a+b+c}{3}$$

Histograma original



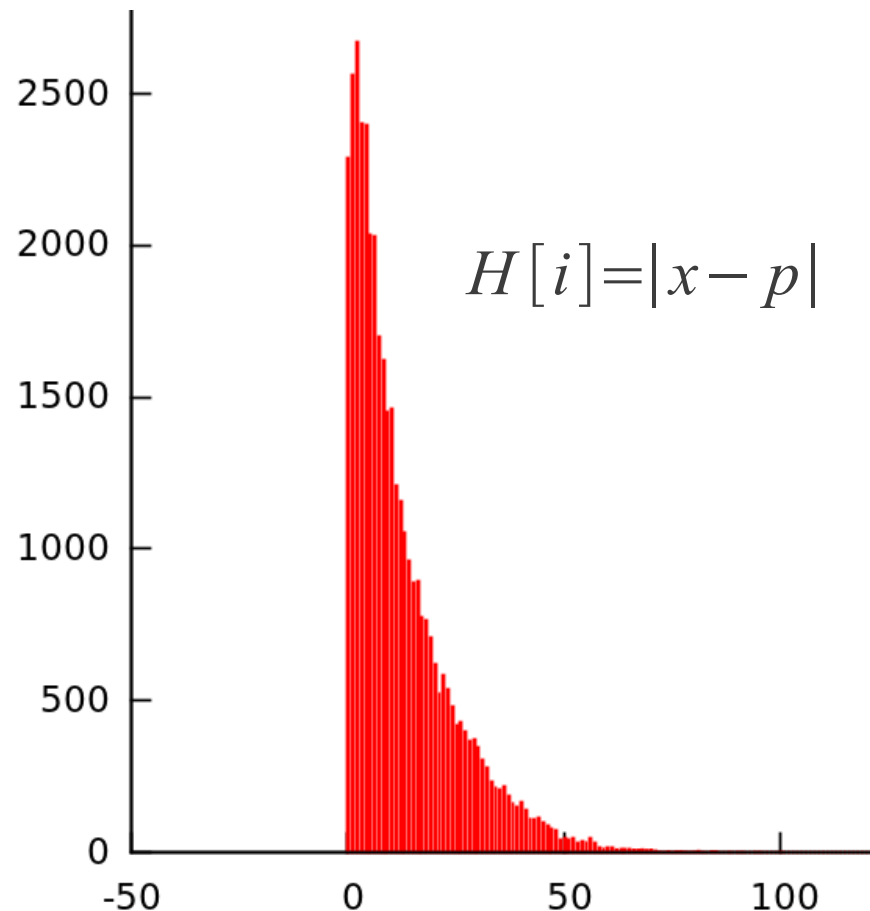
Histograma modificat.  
La volatilitat disminueix.

- Volatilitat:** Suma de diferències entre una barra i la mitja d'ella amb els seus veïns.

$$V = \sum_{i=1}^{255} \left| \frac{H[i-1] - 2H[i] + H[i+1]}{3} \right|$$

## Mètode I: Volatilitat de l'histograma de predicció d'errors

- Algunes mètodes com [Mohsenzadeh 2009] estan dissenyats per a no modificar l'Histograma. Altres com [Ni 2003] no el modifiquen suficient perquè el càlcul de volatilitat sigui fiable.



$$H[i] = |x - p|$$

a	b
c	

$$p = \left\| \frac{a + b + c}{3} \right\|$$

Una alternativa es fer servir l'**histograma de predicció**.

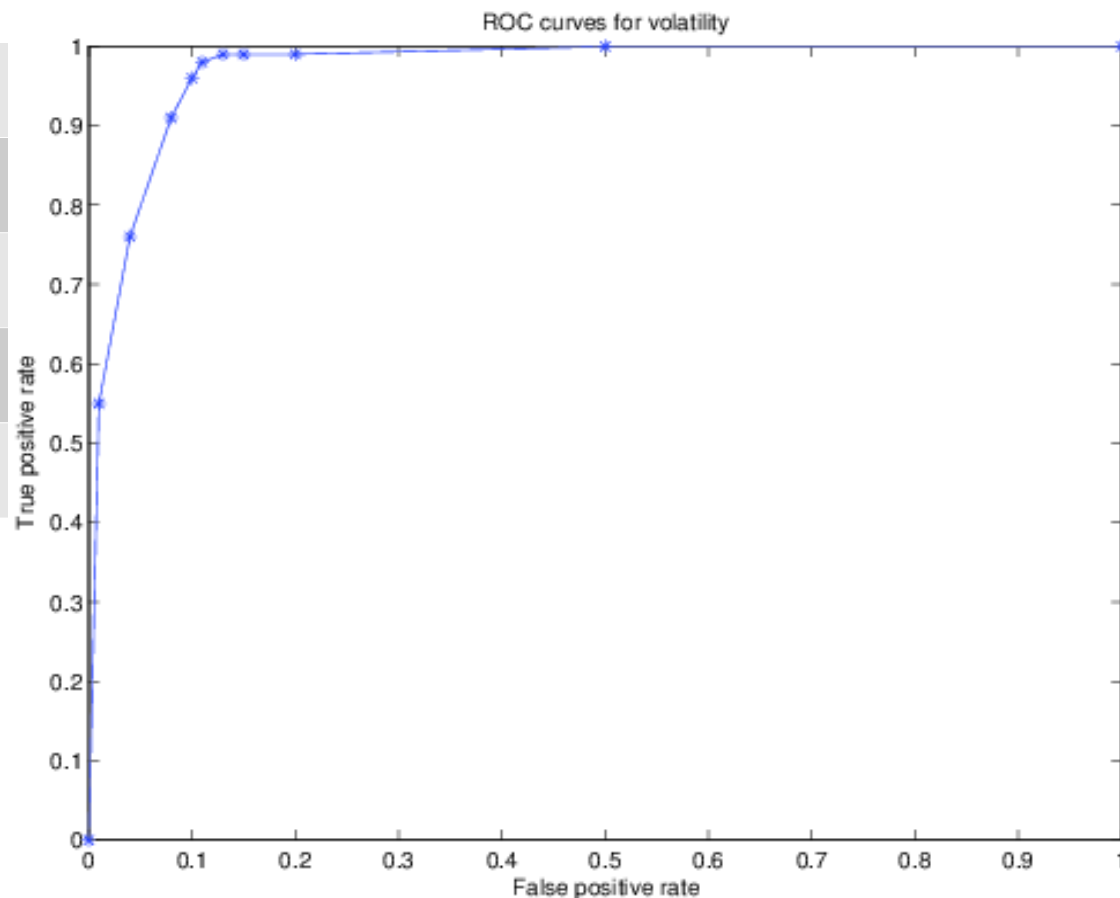
L'histograma de predicció és poc volàtil. La seva modificació en inserir dades, **augmenta la seva volatilitat**.

# Mètode I: Experiments i Resultats

Per a les proves es fan servir imatges de la NRCS (Natural Resources Conservation Service), àmpliament utilitzades en esteganografia i estegoanàlisi.

## Resultats:

<b>TOTAL</b>	<b>86.05%</b>
Positiu	46.15%
Negatiu	39.90%
Falsos Positiu	10.10%
Falsos Negatiu	03.85%



## Mètode II: Patrons de píxels veïns

- Generar tots els possibles patrons amb tots els píxels veïns no és viable, doncs es generarien massa patrons.
- Si fem servir diferències de píxels podem reduir la potència en una unitat.
- Els píxels veïns solen ser similars, per la qual cosa es pot establir un límit  $L$  per a les diferències sense perdre patrons importants.
- La simetria de les imatges fa que certes dades es puguin considerar redundants.

$$L=[-2,2]$$

100	101	99
103	<b>100</b>	101
105	101	102

$256^9$  patrons

0	-1	1
-3	<b>0</b>	-1
-5	-1	-2

$256^8$  patrons

0	-1	1
<b>-2</b>	<b>0</b>	-1
<b>-2</b>	-1	-2

$5^8$  patrons

	-1	1
	<b>0</b>	-1
		-2

$5^4$  patrons

# Mètode II: Referència a màxims i mínims

## Referència al mínim:

	101	<b>99</b>
	100	101
		102

	2	<b>0</b>
	1	1
		3

	3		1
1		<b>0</b>	2

Patró al mínim

## Referència al màxim:

	101	99
	100	101
		<b>102</b>

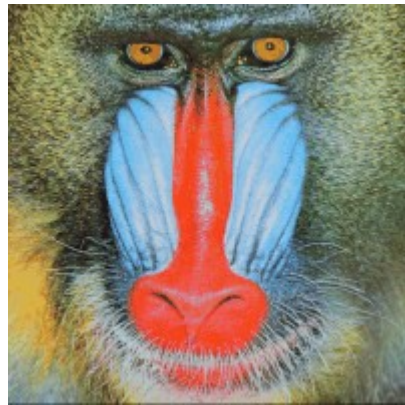
	2	3
	2	1
		<b>0</b>

	2		3
2		<b>0</b>	1

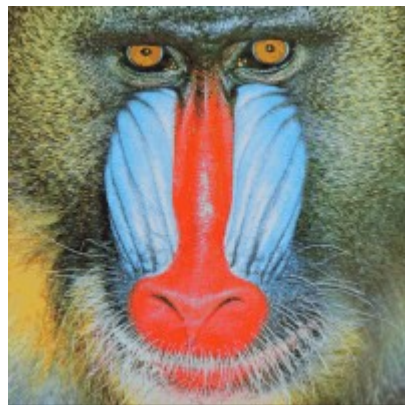
Patró al màxim

Permet reduir el nombre de patrons optimitzant L i aprofita la informació dels màxims, no utilitzada pel model anterior.

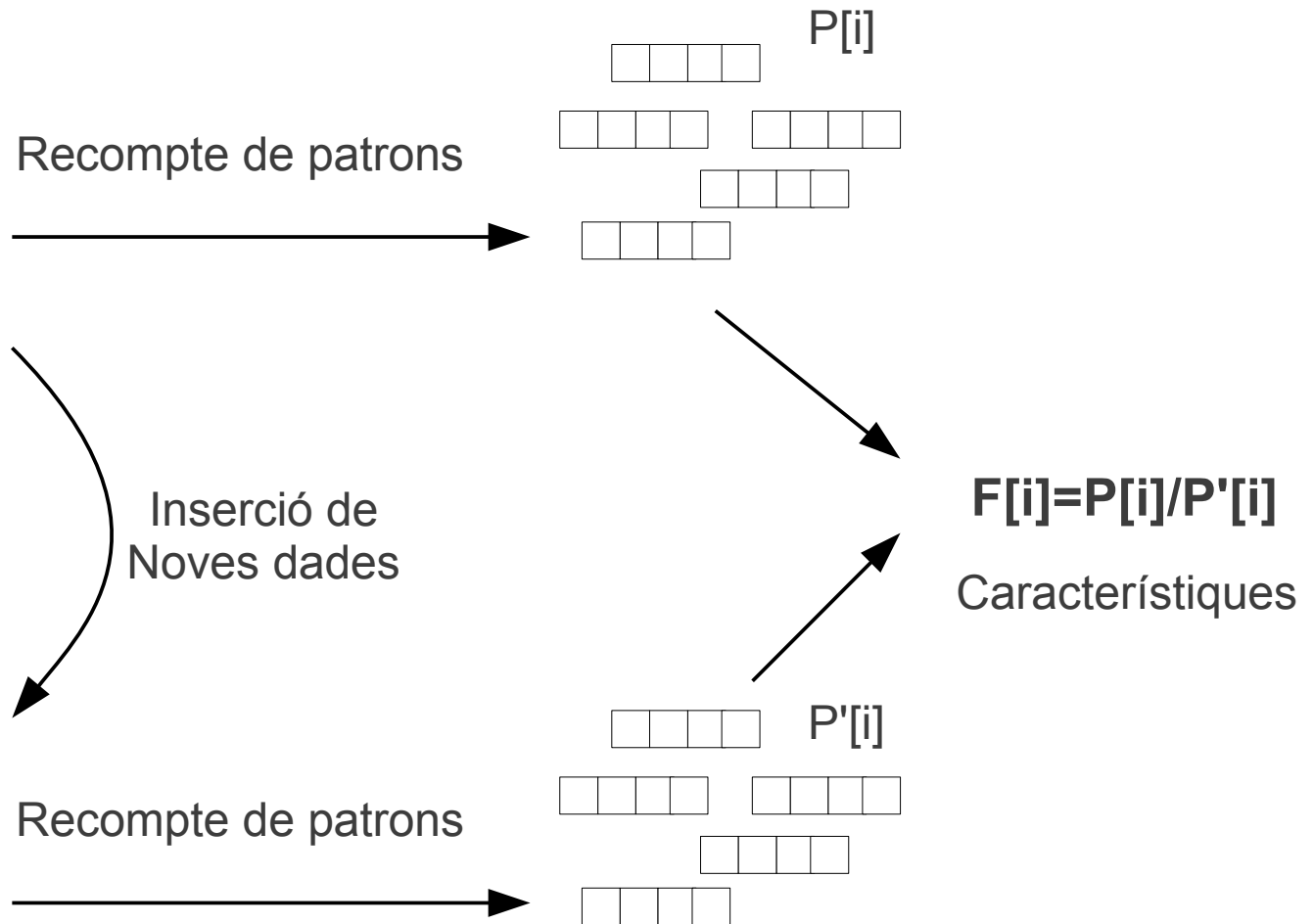
# Mètode II: Generació de característiques



Imatge de test



Imatge amb noves dades inserides



# Mètode II: Experiments i Resultats

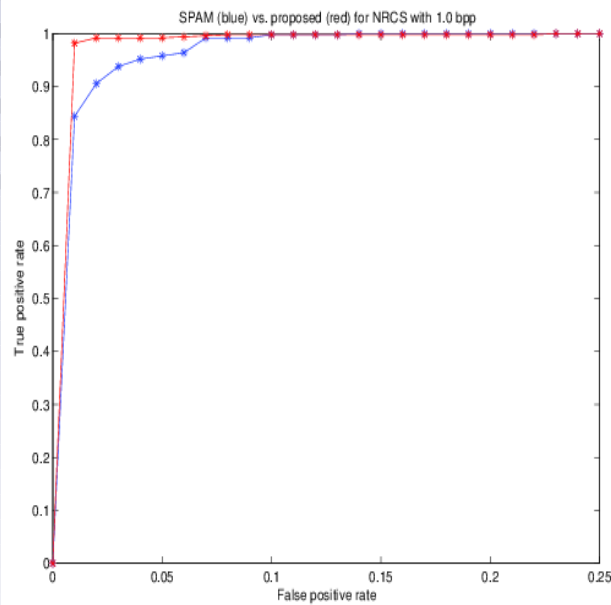
- Es fan servir dues bases de dades d'imatges, la del NRCS (Natural Resources Conservation Service) y la del BOSS (Break Our Steganographic System).
- Es compara amb SPAM el mètode de l'estat de l'art amb millors resultats.

	MÈTODE PROPOSAT 256 característiques					SPAM 687 característiques				
	%	P	FP	N	FN	%	P	FP	N	FN
NRCS 100%	<b>99.00%</b>	496	6	494	4	<b>95.80%</b>	495	37	463	4
NCRS 50%	<b>90.90%</b>	460	51	449	40	<b>83.10%</b>	444	113	387	55
NRCS 25%	<b>79.10%</b>	393	102	398	107	<b>67.90%</b>	387	208	292	112
BOSS 100%	<b>100.0%</b>	500	0	500	0	<b>99.60%</b>	499	3	497	1
BOSS 50%	<b>99.90%</b>	500	1	499	0	<b>98.80%</b>	498	10	490	2
BOSS 25%	<b>98.90%</b>	495	6	494	5	<b>96.90%</b>	500	31	469	0

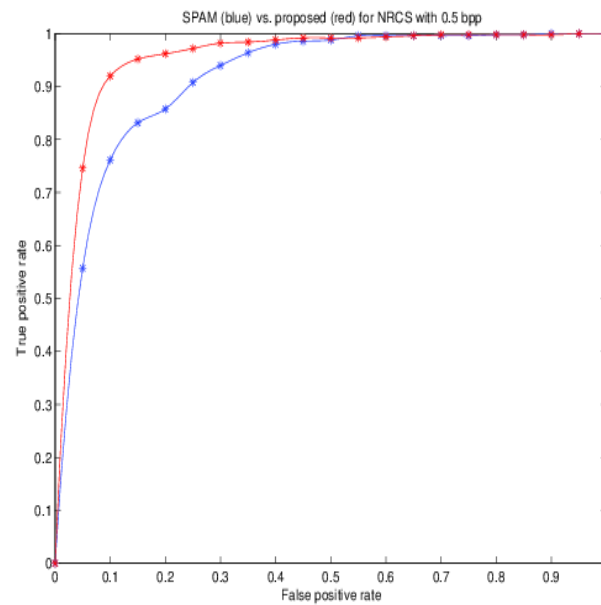
# Mètode II: Experiments i Resultats II

Comparació amb corbes ROC per a la base de dades NCRS

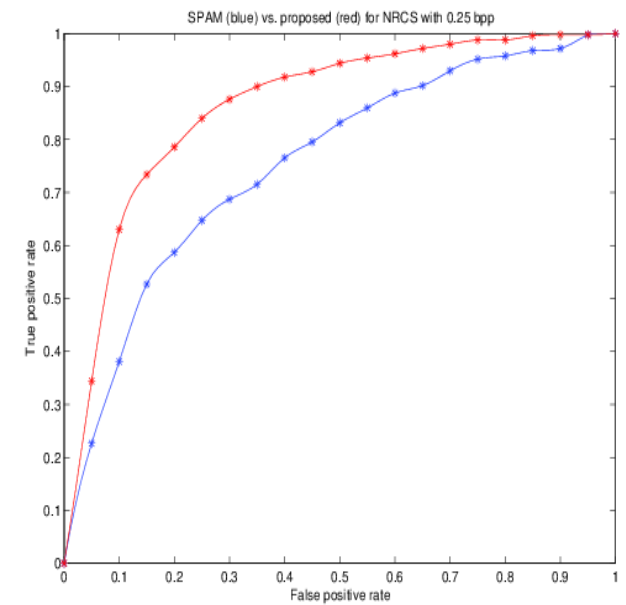
Bitrate 100%



Bitrate 50%



Bitrate 25%



■ Mètode proposat

■ SPAM





# **Estegoanàlisi d'imatges en el domini espacial**

**Gràcies per la seva atenció.**

**Preguntes?**