

Autentificarea conținutului multimedia

Scopul autentificării

- semnalele multimedia pot fi ușor reproduse, modificate și falsificate
- efect negativ major prin expunerea lor în mass-media, în sălile de judecată, în campanii politice, etc.
- determinarea veridicității conținutului multimedia
- două întrebări:
 - A fost modificat conținutul în vreun fel?
 - Dacă a fost modificat, unde și în ce măsură a fost modificat?

Tehnici de autentificare

- trei abordări:

A. sistemele de autentificare ce utilizează tehnici de watermarking digital (autentificare activă)

- inserarea unui watermark invizibil
- watermark fragil sau semifragil
- lipsa watermark-ului → material neautentic

B. sisteme de autentificare bazate pe semnături digitale

- se extrag caracteristici (features) din imaginea originală
- caracteristici fragile la modificări intenționate
- caracteristici robuste la distorsiuni neintenționate
- funcție hash aplicată caracteristicilor → semnătură digitală
- semnătura transmisă împreună cu imaginea

Tehnici de autentificare

C. sisteme de autentificare bazate pe proprietățile intrinseci ale materialului multimedia (autentificare pasivă)

- nu necesită inserarea de informații adiționale,
- detectează eventuale modificări din proprietățile intrinseci ale conținutului multimedia

Tehnici de autentificare

C. sisteme de autentificare bazate pe proprietățile intrinseci ale materialului multimedia

- tehnici la nivel de pixel;
 - Detectează: clonare, reeșantionare, combinarea mai multor imagini
- tehnici bazate pe formatul materialului multimedia
 - Cuantizare JPEG, dublă compresie JPEG, lipsa artefactelor JPEG

Tehnici de autentificare

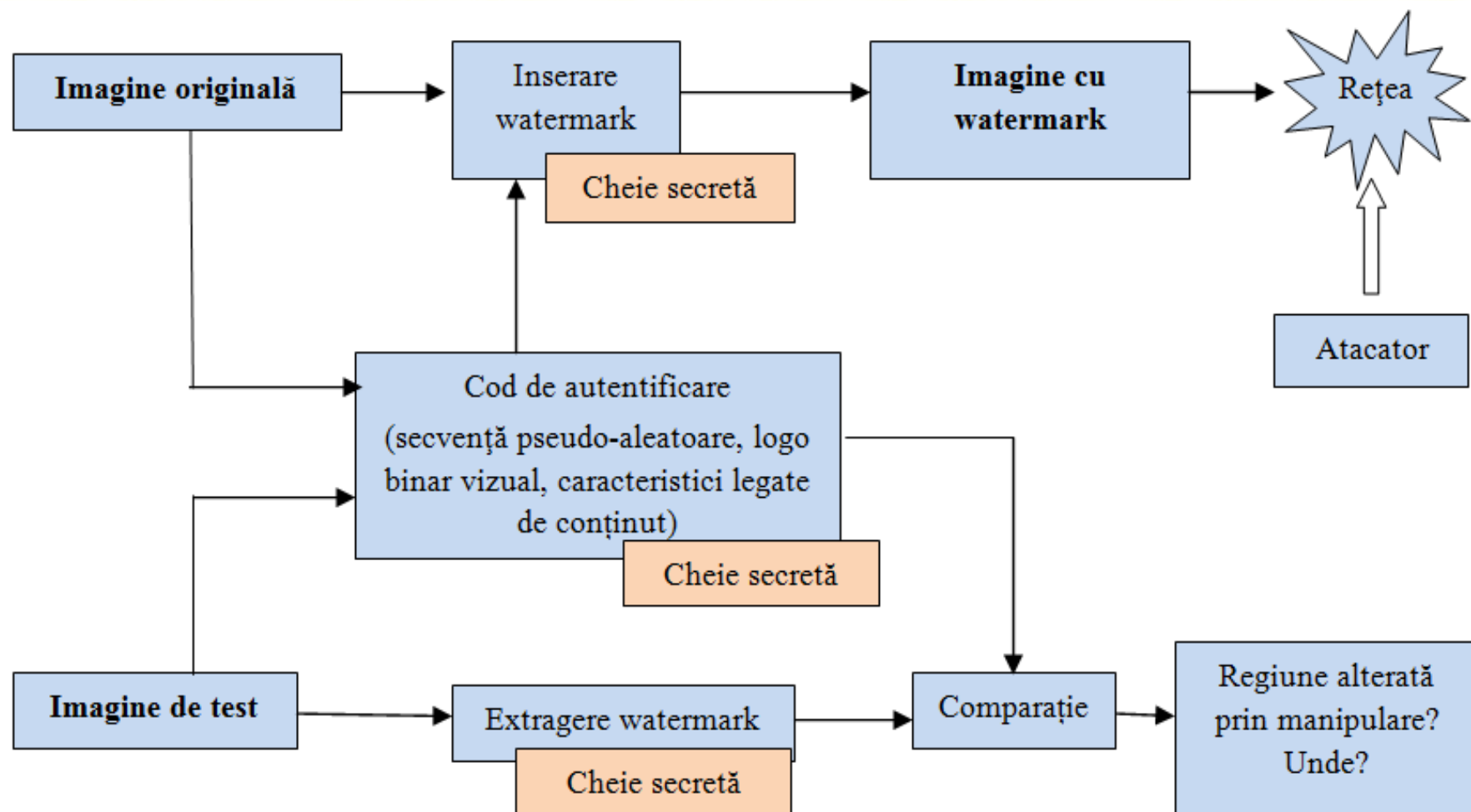
C. sisteme de autentificare bazate pe proprietățile intrinseci ale materialului multimedia

- tehnici bazate pe analiza artefactelor introduse de camerele digitale prin intermediul lentilelor, senzorilor sau procesărilor interne
 - Aberații cromatice, CFA, zgomot
- tehnici ce exploatează anomalii între obiecte fizice, lumină și cameră
 - Condiții de iluminare, direcția luminii
- tehnici geometrice
 - poziția relativă a obiectelor din scenă față de cameră

A. Sistemele de autentificare bazate pe watermarking

- watermark fragil sau semi-fragil
- secvență aleatoare de biți, logo vizual binar sau caracteristici legate de conținut
- inserare imperceptibilă în datele gazdă
- watermark extras la detector
- codul extras comparat cu cel original
- diferențele indică faptul că datele originale au fost manipulate

Sistemele de autentificare bazate pe watermarking



Sistemele de autentificare bazate pe watermarking

Cerințe:

- Transparența perceptuală
- Detecția falsificării
- Localizare
- Compatibilitatea (semi-fragilitatea)
- Portabilitate (date aditionale)
- Securitatea algoritmului

Sistemele de autentificare bazate pe watermarking

- Clasificare în funcție de toleranța la distorsiuni:
 - watermarking pentru autentificare exactă/brută
 - watermark fragil
 - watermarking pentru autentificare selectivă/fină
 - watermark semi-fragil

Watermarking pentru autentificare exactă/brută

- autentificare strictă, fără toleranță pentru manipularea conținutului
- watermark fragil
- un singur bit modificat → date neautentice
- tehnici criptografice pentru a atinge un nivel de securitate înalt
- **Exemplu simplu:** metoda LSB în domeniul spațial
- Dezavantaj?

Watermarking pentru autentificare exactă/brută

Exemplu mai complex (metodă propusă de M. Wu și B. Liu):

- imagine împărțită în blocuri
- în fiecare bloc planul LSB înlocuit cu XOR dintre biții watermark-ului și suma hash a imaginii.
- suma hash obținută din dimensiunile imaginii și restul valorilor pixelilor din bloc mai puțin bitul LSB
- rezultat criptat cu cheia privată a utilizatorului

Watermarking pentru autentificare exactă/brută

Funcție (sumă, cod) hash

- algoritm care atribuie datelor de lungime variabilă un cod de dimensiune fixă
- Proprietățile unei funcții hash criptografice ideale:
 - este **ușor de calculat** valoarea hash pentru orice mesaj
 - este **nefezabil** de generat un mesaj care are o **valoare hash dată**
 - este **nefezabil** de modificat un mesaj fără a **schimba valoarea sa hash**
 - este **nefezabil** de găsit **două mesaje diferite** cu aceeași valoare hash

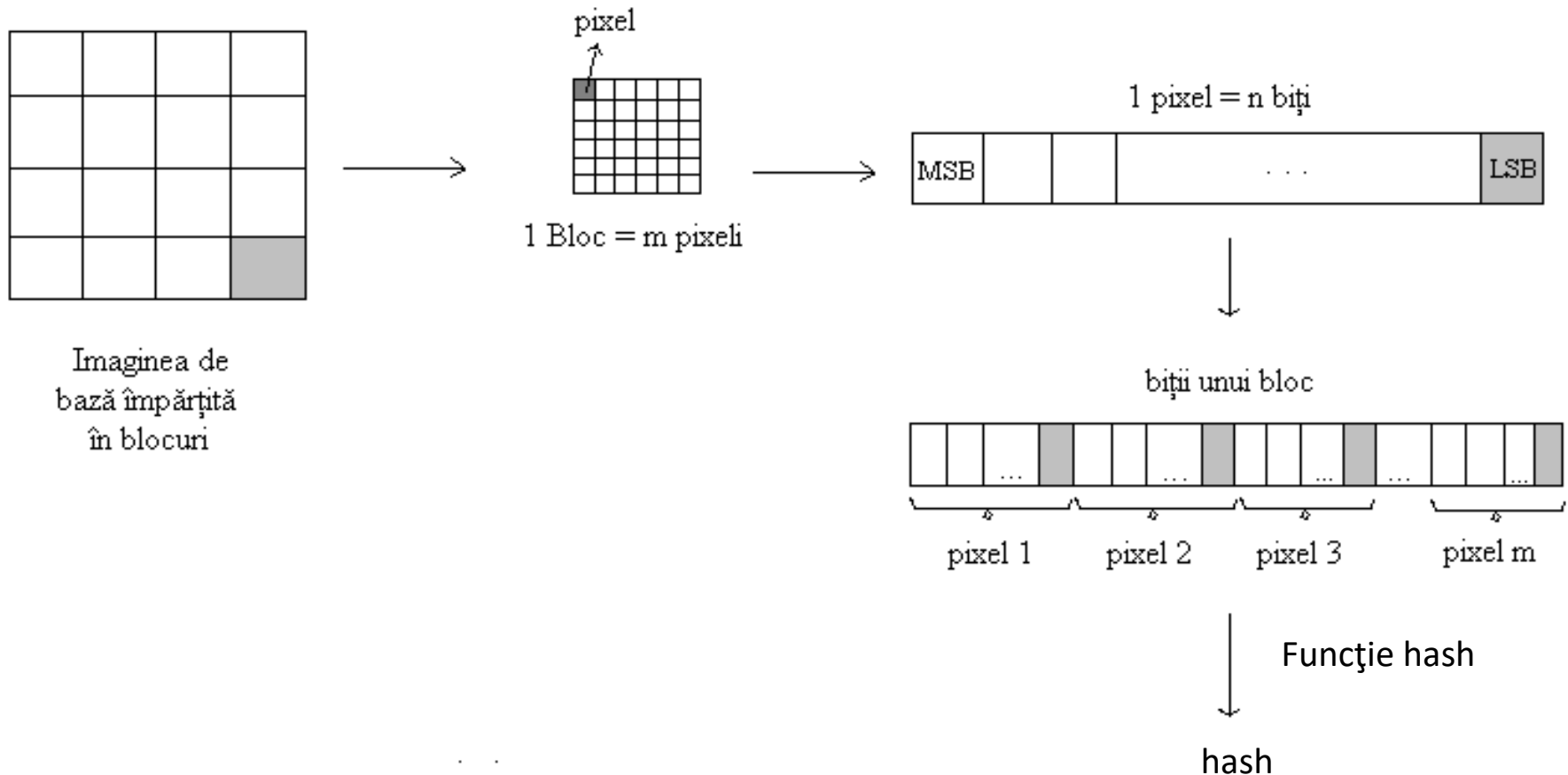
Watermarking pentru autentificare exactă/brută

Funcție (sumă, cod) hash

- Aplicații:
 - detecția de erori la transmisiune și stocare, semnături digitale, coduri de autentificare, amprentare, indexare
- Exemple
 - Bit de paritate
 - CRC (Cyclic Redundancy Check) – 32 biți (Ethernet)
 - MD5 (Message Digest 5) – 128 biți (32 digiți hexa)
 - SHA-1-3 (Secure Hash 1-3) – 160 biți (40 digiți hexa)

Watermarking pentru autentificare exactă/brută

Exemplu (metodă propusă de M. Wu și B. Liu)



Watermarking pentru autentificare exactă/brută

Exemplu (metodă propusă de M. Wu și B. Liu)

Detecție:

- planul LSB decriptat folosind cheia publică
- XOR între rezultatul decriptării și suma hash recalculată din același bloc
- biții watermark-ului recuperat comparați cu cei ai watermark-ului original
- dacă toți sunt identici → bloc autentic
- cel puțin 1 bit diferit → bloc alterat

Watermarking pentru autentificare selectivă/fină

- Tehnici de watermarking semi-fragile
- Robuste la prelucrări neagresive ale conținutului (compresie ușoară, filtrare, îmbunătățirea contrastului, zgomot)
- Fragile la modificarea conținutului (ștergere, înlocuire, inserare de elemente în imagine)
- Tehnici cu spectru împrăștiat sau bazate pe cuantizare în domenii transformate (DCT, Wavelet, etc.)

Watermarking pentru autentificare selectivă/fină

A. Metode bazate pe blocuri

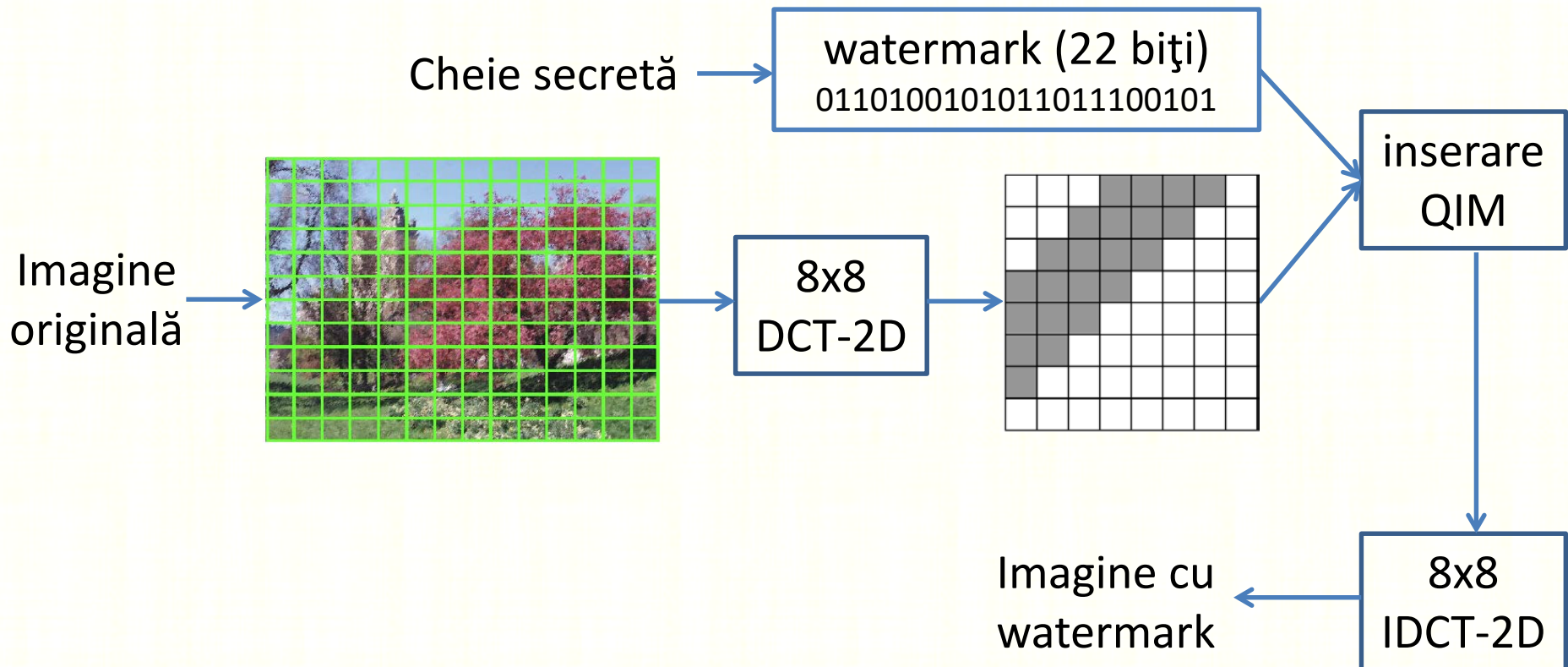
- blocuri din domeniul spațial sau transformat
- Imagine/cadru împărțit în blocuri
- watermark inserat în fiecare bloc
- autentificare a fiecărui bloc

Watermarking pentru autentificare selectivă/fină

A. Metode bazate pe blocuri

– **EXEMPLU:** Autentificare bazată pe blocuri în domeniul DCT

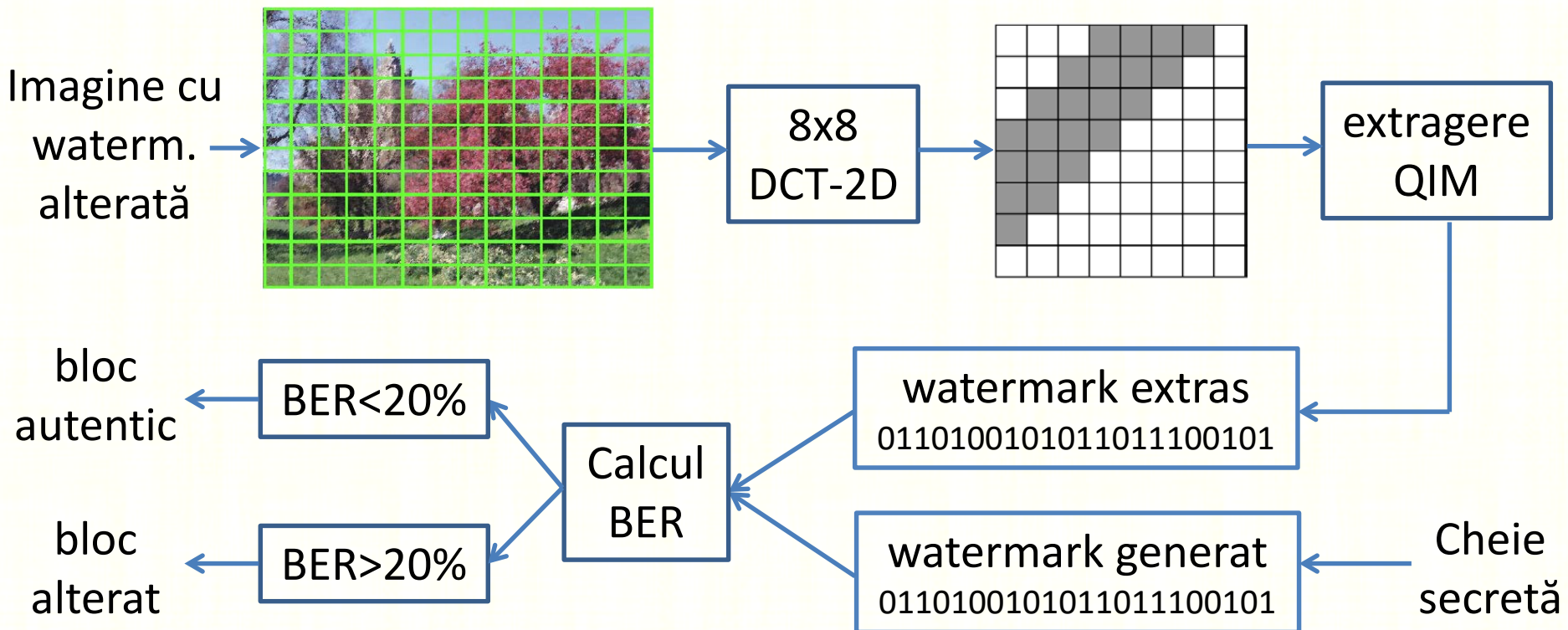
Inserare watermark:



Watermarking pentru autentificare selectivă/fină

A. Metode bazate pe blocuri

- **EXEMPLU:** Autentificare bazată pe blocuri în domeniul DCT
- Extragere watermark și autentificare:



Watermarking pentru autentificare selectivă/fină

A. Metode bazate pe blocuri

Dezavantaje:

- rezoluția minimă de detecție = dimensiunea blocului
- îmbunătățire rezoluție → reducere dim. bloc. → capacitate crescută a watermark-ului → degradarea calității
- dificil de inserat watermark-uri în regiuni uniforme fără a cauza artefacte vizuale importante
- vulnerabilitatea la diferite tipuri de atacuri locale, ca de ex. cuantizarea vectorilor (VQ – Vector Quantization)
- în special dacă watermark-ul și procedeul de inserare sunt independente de bloc
- Soluție: inserarea de watermark-uri diferite în blocuri diferite

Watermarking pentru autentificare selectivă/fină

A. Metode ce nu sunt bazate pe blocuri

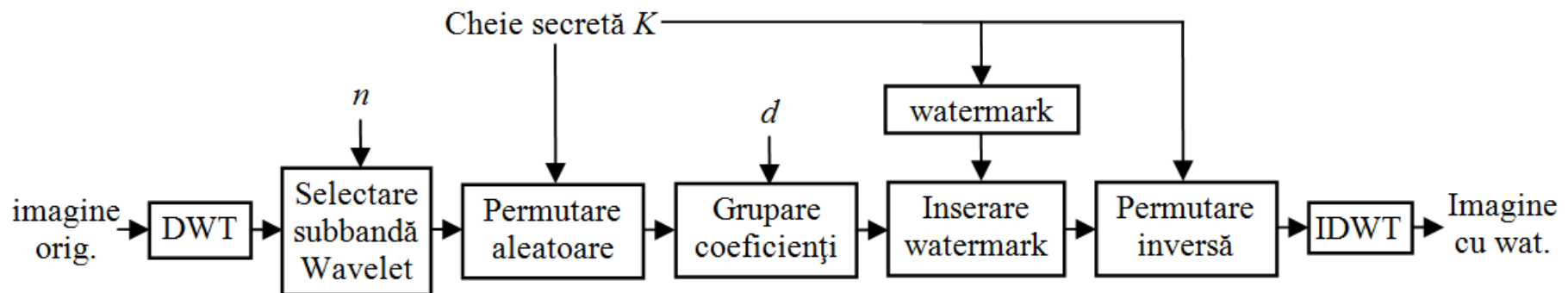
- utilizarea DWT în locul DCT permite o mai bună localizare spațială a modificării
- securitate asigurată prin permutarea coeficienților wavelet în care se inserează watermark-ul
- îmbunătățirea preciziei localizării manipulării
- scăderea capacității watermark-ului
- creșterea calității perceptuale

Watermarking pentru autentificare selectivă/fină

A. Metode ce nu sunt bazate pe blocuri

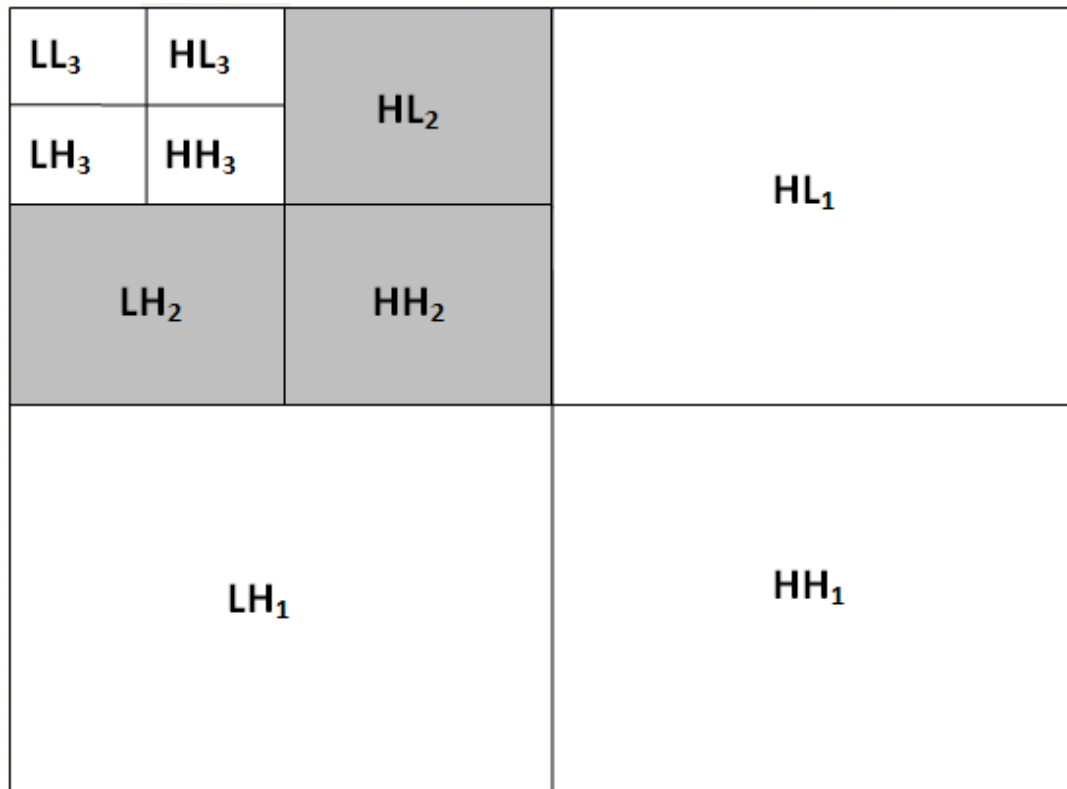
– **EXEMPLU:** Autentificare în domeniul Wavelet

Inserare watermark:



Watermarking pentru autentificare selectivă/fină

- Descompunere Wavelet



Watermarking pentru autentificare selectivă/fină

- Permutare coeficienți Wavelet
 - Concatenare într-un vector
 - Permutare pe baza unei chei secrete
- Grupare coeficienți
 - Grupuri de câte d coeficienți
 - Controlul capacității
- Generare watermark pentru autentificare
 - Secvență pseudo-aleatoare binară pe baza cheii
 - Dim. watermark = nr. grupuri coeficienți
- Se calculează media ponderată a fiecărui grup i de coeficienți permutați

$$m_i = \sum_{j=1}^d (-1)^j |c_i(j)|$$

Watermarking pentru autentificare selectivă/fină

- Inserare watermark
 - Un bit în fiecare grup de coeficienți
 - cuantizarea lui m_i la cel mai apropiat număr par sau impar de pași de cuantizare (QIM)

$$m_i^w = \begin{cases} \lfloor m_i/Q \rfloor \cdot Q & \text{dacă } \text{mod}2(\lfloor m_i/Q \rfloor) = w_i \\ \lfloor m_i/Q \rfloor \cdot Q + Q & \text{dacă } \text{mod}2(\lfloor m_i/Q \rfloor) \neq w_i \end{cases}$$

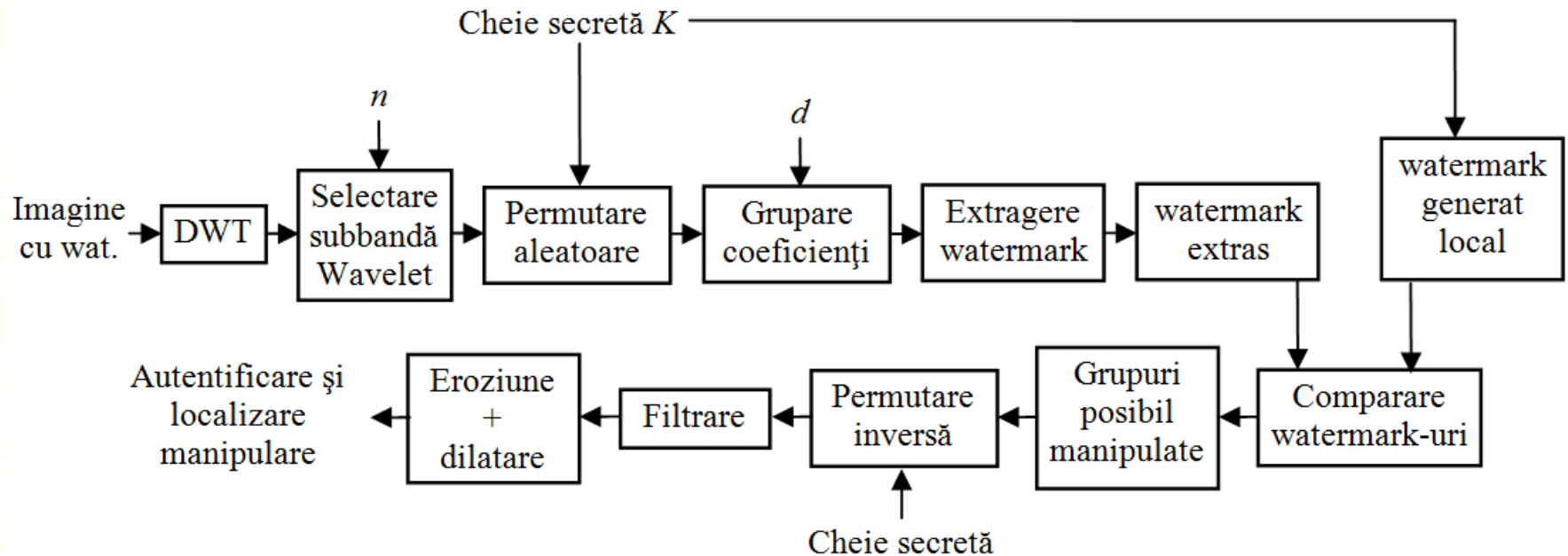
- Schimbarea mediei prin modificarea coeficientului wavelet cu cea mai mare valoare absolută din grup

$$c_{i,\max}^w(j) = c_{i,\max}(j) + (-1)^j \cdot \text{semn}(c_{i,\max}(j)) \cdot (m_i^w - m_i)$$

$$\text{semn}(x) = \begin{cases} -1, & \text{dacă } x \leq 0 \\ 1, & \text{dacă } x > 0 \end{cases}$$

Watermarking pentru autentificare selectivă/fină

Extragere watermark și autentificare



Watermarking pentru autentificare selectivă/fină

- Extragere watermark din medie
 - Un bit de watermark din media fiecărui grup

$$w'_i = \text{mod } 2(\lfloor m'_i / Q \rfloor)$$

- Generare locală a watermark-ului original w pe baza cheii secrete
- Autentificare:
 - Dacă $w' = w \rightarrow$ imagine autentică
 - Dacă $w'_i \neq w_i \rightarrow$ coef. grupului i etichetați ca posibil alterați

Watermarking pentru autentificare selectivă/fină

- Permutare inversă a coeficienților
 - Coef. posibil alterați împrăștiați în subbenzile LH, HL și HH
 - Densitate mare de astfel de coef. în zonele modificate intenționat
 - Coeficienții izolați eliminați prin filtrare
- Matrice de autentificare A de dimensiune $(M / 2^n) \times (N / 2^n)$
 - $A(x,y)=1$, dacă $LH(x,y)$, $HL(x,y)$ sau $HH(x,y)$ sunt posibil alterați
 - Valorile “1” izolate din matricea A eliminate prin filtrare
 - Pozițiile “1” din matricea A sunt mapate în domeniul spațial și indică zonele alterate
 - Fiecare poziție “1” din matricea A corespunde unei regiuni de $2^n \times 2^n$ pixeli din imaginea originală

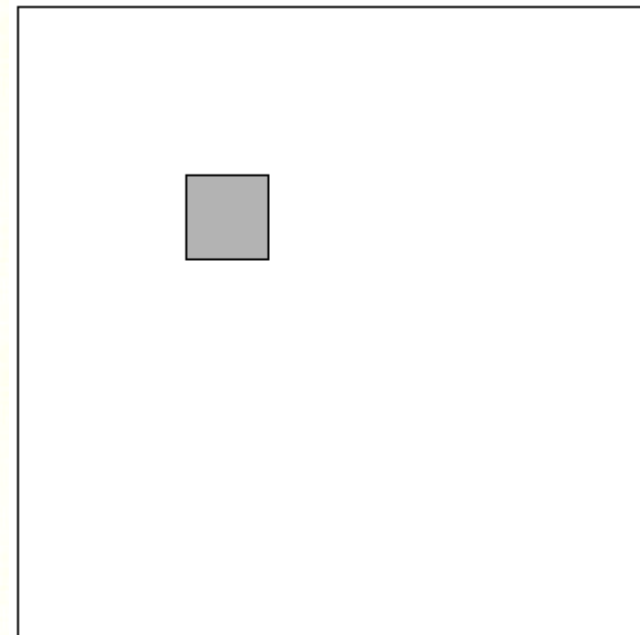
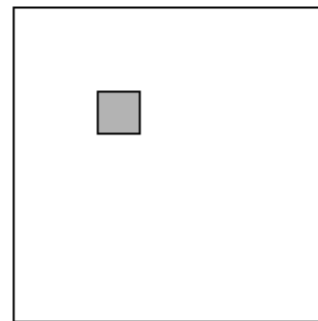
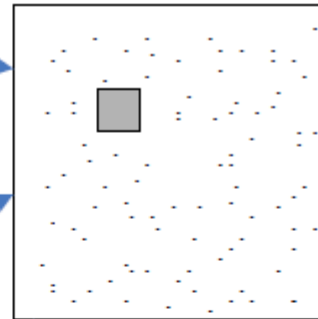
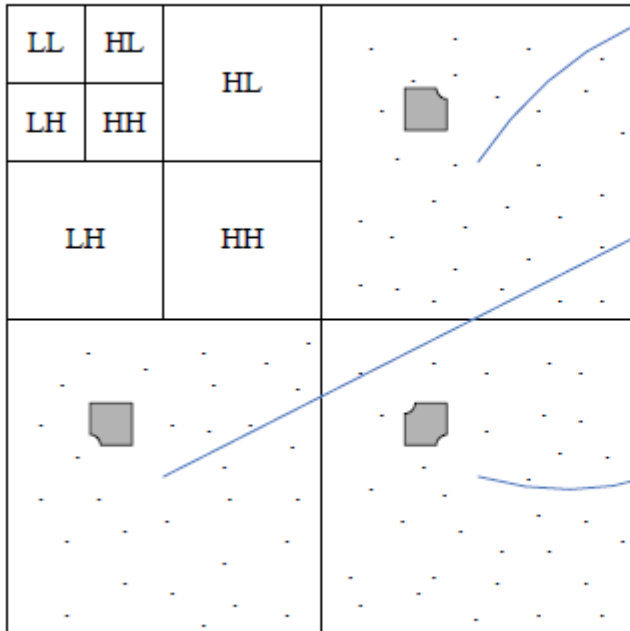
Watermarking pentru autentificare selectivă/fină

Descompunere Wavelet

Matricea A

Filtrare

Rezultatul autentificării



Watermarking pentru autentificare selectivă/fină

- Rezultate



Imaginea originală



Imagine cu watermark
alterată

Watermarking pentru autentificare selectivă/fină

- Rezultate

$n=2$, $d=8$, $Q=8$



înainte de filtrare



după filtrare

Watermarking pentru autentificare selectivă/fină

- Capacitatea watermark-ului pentru o rezoluție de detecție de 4x4 pixeli
 - Blocuri: 1 bit / bloc 4x4
 - Wavelet, $n=2$: 1 bit / grup d coeficienți = $1/d$ biți / coeficient
= $1/d$ biți / bloc 4x4