

Utilisation des distances tangentes pour la compensation de mouvement

– Application au codec Theora –

Jonathan Fabrizio



Séminaire LRDE
18/05/11





Utilisation des distances tangentes pour la compensation de mouvement

– Application au codec Theora –

- Rappel sur la distance tangente
- Estimation de mouvement
- Compensation de mouvement
 - Principe
 - Résultats
- Theora
 - Rappel sur Theora
 - Principe
 - Résultats



Rappel sur la distance tangente



La distance tangente

- Introduite par Simard et al. dans les années 90
- Historiquement utilisée pour la reconnaissance de caractères

Simard, P., Victorri, B., Le Cun, Y., and Denker, J. (1992). « Tangent Prop - A formalism for specifying selected invariances in an adaptive network ». In *Advances in Neural Information Processing Systems 4*, pages 895-903.

Schwenk, H. and Milgram, M. « Constraint tangent distance for on-line character recognition ». *Proceedings of the 13th International Conference on Pattern Recognition – 1996*.

Patrice Y. Simard , Yann A. Le Cun , John S. Denker and Bernard Victorri, « Transformation Invariance in Pattern Recognition - Tangent Distance and Tangent Propagation ». *Neural Networks : Tricks of the trade*, 1998.



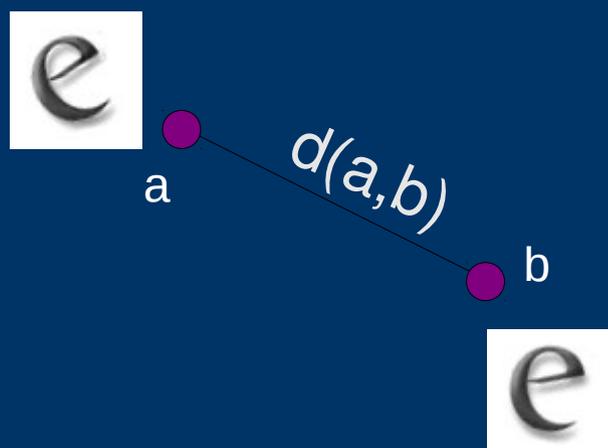
La distance tangente

- Comparaison de deux motifs



La distance tangente

- Comparaison de deux motifs
 - Par la distance euclidienne :
 - $d(a,b)$

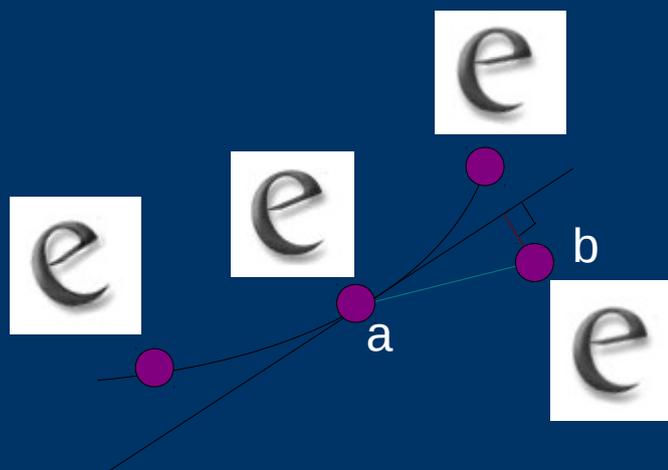


La distance tangente

- Comparaison de deux motifs
 - On aimerait tenir compte de la transformation
 - $d(R(a), b)$
 - Par la distance tangente :
 - $d(a + \lambda_1 v_1 + \dots, b)$
 - Il faut trouver les λ qui minimisent la distance



La distance tangente

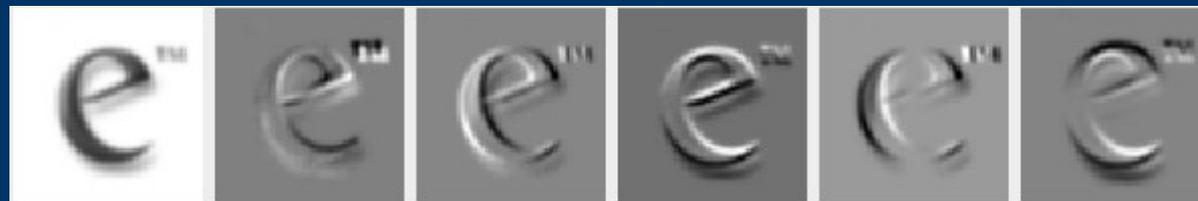


— Distance tangente
— Distance euclidienne



La distance tangente

- Différentes transformations peuvent être modélisées
 - Rotation, translation...
 - Changement d'illumination
 - ...



Rotation

Translations

Etirements

La distance tangente

- Démo

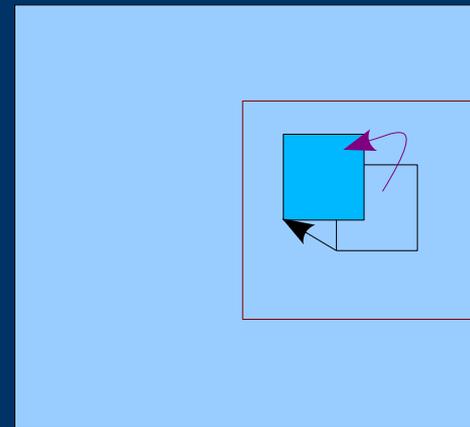
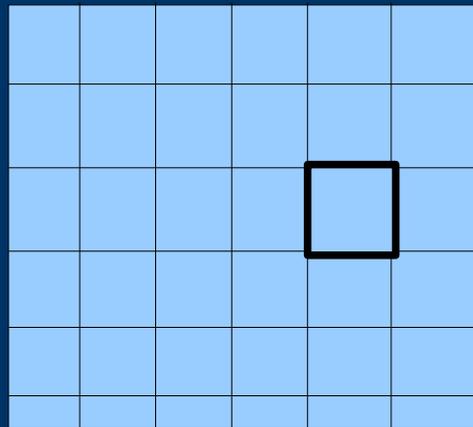


Estimation de mouvement



Estimation de mouvement

- Amélioration du « block matching »
 - Découpage de l'image i en blocs
 - Recherche le déplacement de chacun des blocs dans l'image $i+1$

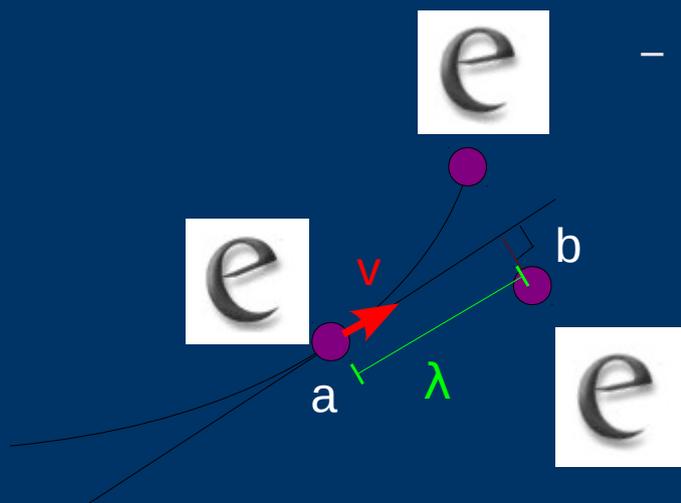


Estimation de mouvement

- Amélioration du « *block-matching* »
 - Utilisation des distances tangentes pour sélectionner le bloc le plus ressemblant

- $d(a + \lambda_1 v_1 + \lambda_2 v_2 + \dots, b)$

- La comparaison est plus robuste
 - Les λ peuvent fournir des informations sur la transformation



Estimation de mouvement

- Résultats

Jonathan Fabrizio, Séverine Dubuisson, "Motion Estimation using Tangent Distance".
International Conference on Image Processing 2007.



Compensation de mouvement

- Application : compression vidéo



Compensation de mouvement

- Application : compression vidéo
 - Pour gagner de la place on cherche les redondances
 - Spatiales
 - Temporelles



Compensation de mouvement

- Exploitation de la redondance temporelle (Encodage)
 - Découpage de l'image i en blocs
 - Recherche le déplacement de chacun des blocs dans l'image *de référence*

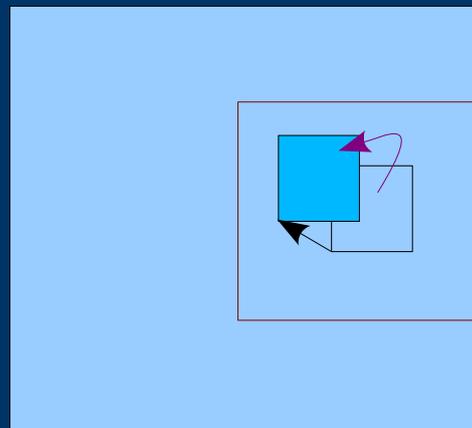


Image de référence

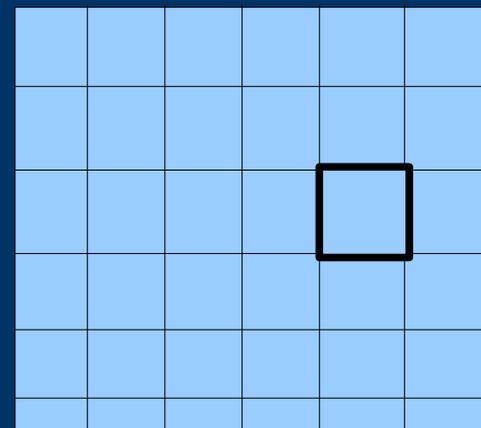


Image i

Compensation de mouvement

- Exploitation de la redondance temporelle (Décodage)
 - Génération de la prédiction de l'image i
 - On applique le vecteur translation à chaque bloc

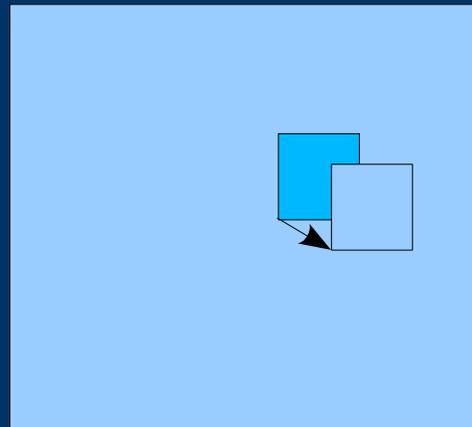
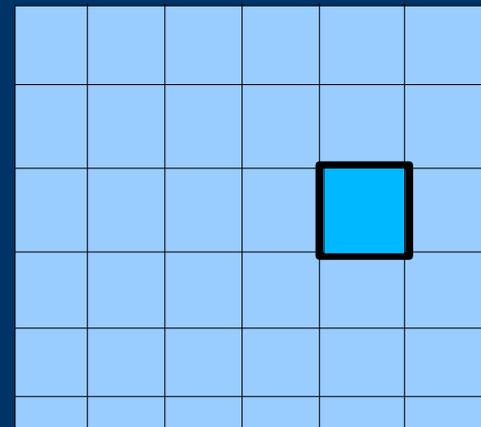


Image de référence



Prédiction de l'image i

Compensation de mouvement

- Le « *block-matching* »
 - Dans les codecs, c'est souvent l'algorithme utilisé pour générer les prédictions
- Amélioration du « *block-matching* »
 - Notre algorithme permet de prédire :
 - l'évolution spatiale du bloc
 - l'évolution temporelle



Compensation de mouvement

- Amélioration du « *block-matching* »
 - Notre algorithme permet de prédire :
 - l'évolution spatiale du bloc
 - l'évolution temporelle
 - Information ajoutée :
 - Les vecteurs déplacements
 - Les coefficients



Compensation de mouvement

- Résultats



(A)



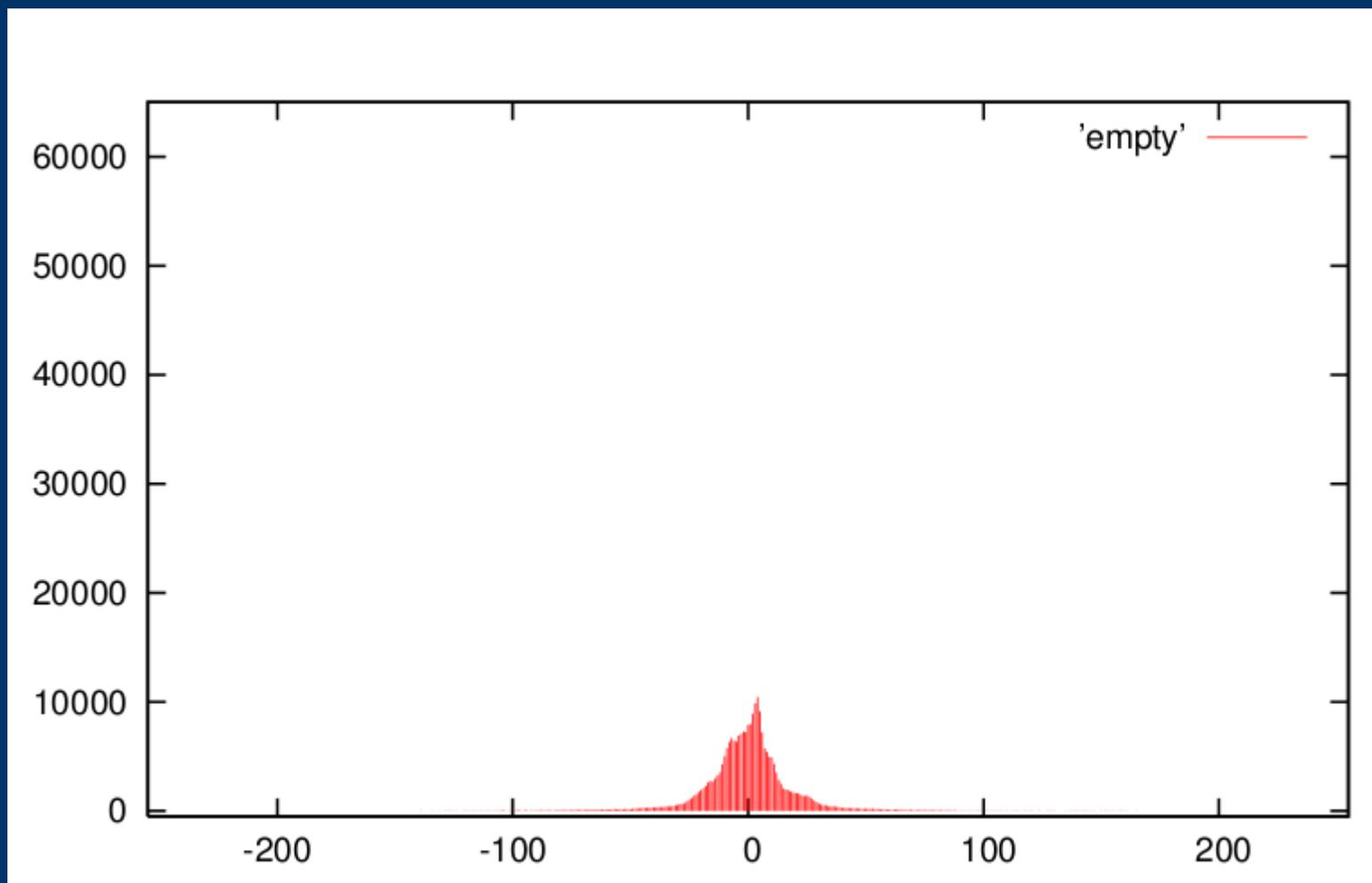
(B)

Compensation de mouvement

- Paramètres :
 - Taille des blocs : 8x8
 - Transformations modélisées :
 - Etirement horizontal
 - Etirement vertical
 - Changement d'illumination



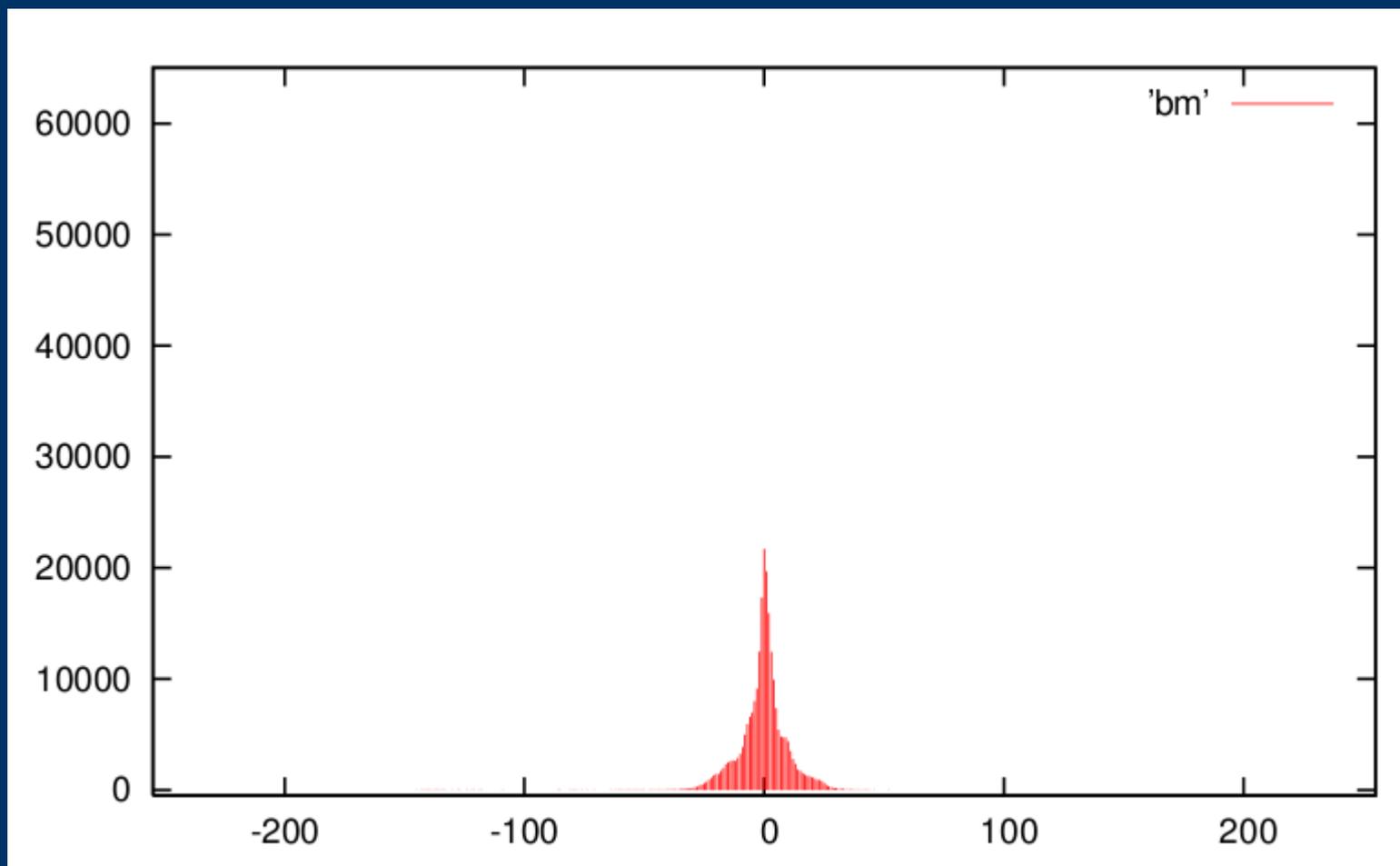
Compensation de mouvement



Prédiction sans compensation de mouvement



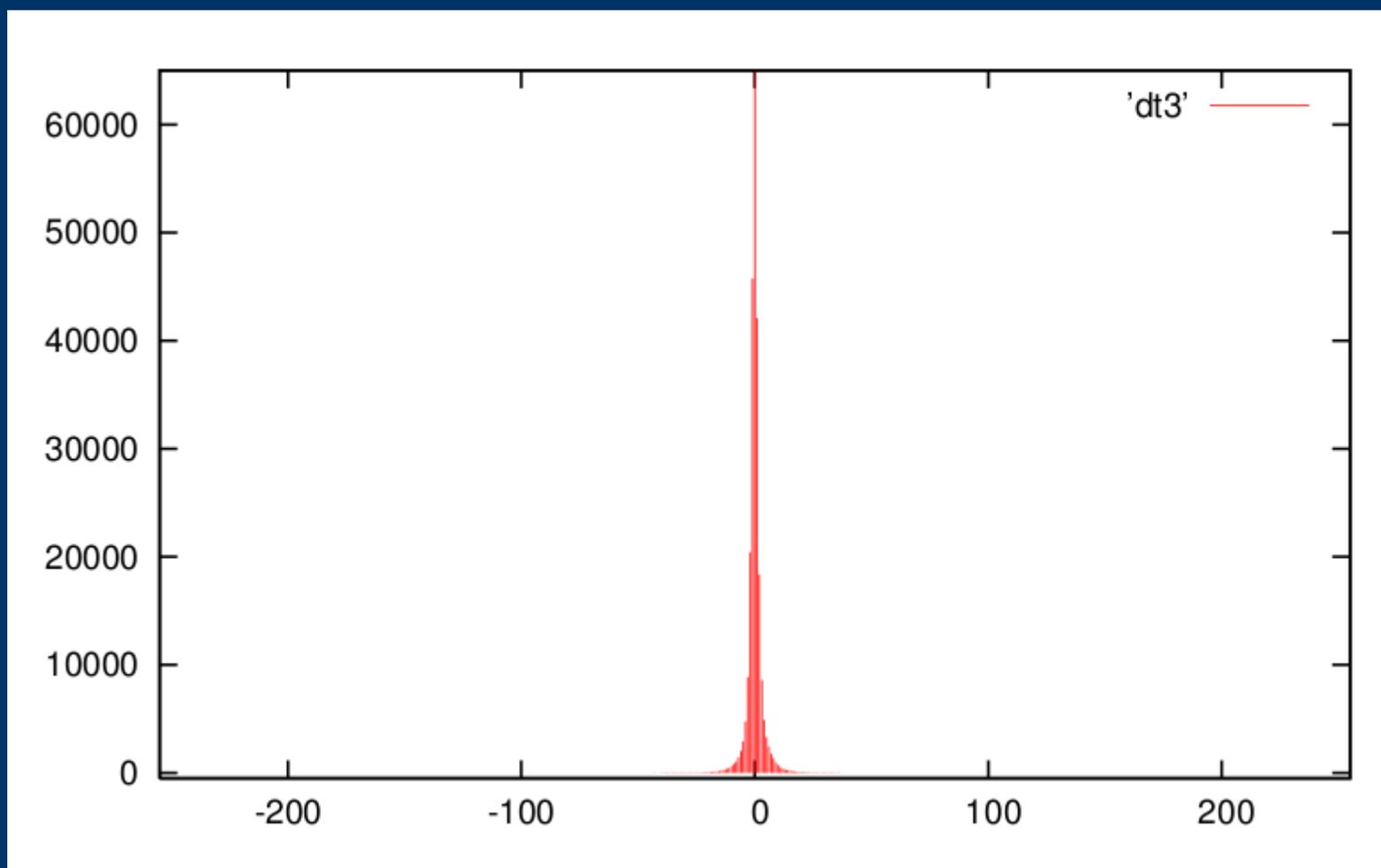
Compensation de mouvement



Prédiction avec compensation de mouvement par *block-matching*



Compensation de mouvement



Prédiction avec compensation de mouvement par notre algorithme



Compensation de mouvement

- Peut-on espérer gagner de la place ?



Compensation de mouvement

- Peut-on espérer gagner de la place ?
 - Oui !
 - Taille utilisée par notre encodage
 - 113839 octets
 - Taille utilisée pour la prédiction par BM
 - 160160 octets

J. Fabrizio, S. Dubuisson, "Utilisation des distances tangentes pour la compensation de mouvement", CORESA 2007



Compensation de mouvement

- Peut-on espérer gagner de la place ?
 - Bémol : l'encodage réalisé par les codecs est plus riche. Notre méthode aura-t-elle autant d'impact ?



Theora – Rappels



Theora – Rappels

- Origines

- Suite de VP3
- Xiph.org
Foundation

- Pourquoi Theora

- Libre
- Simple
- A été recommandé dans la norme HTML 5
- Beaucoup de navigateurs le supportent
- Utilisé par des sites importants (Wikipedia, Dailymotion)



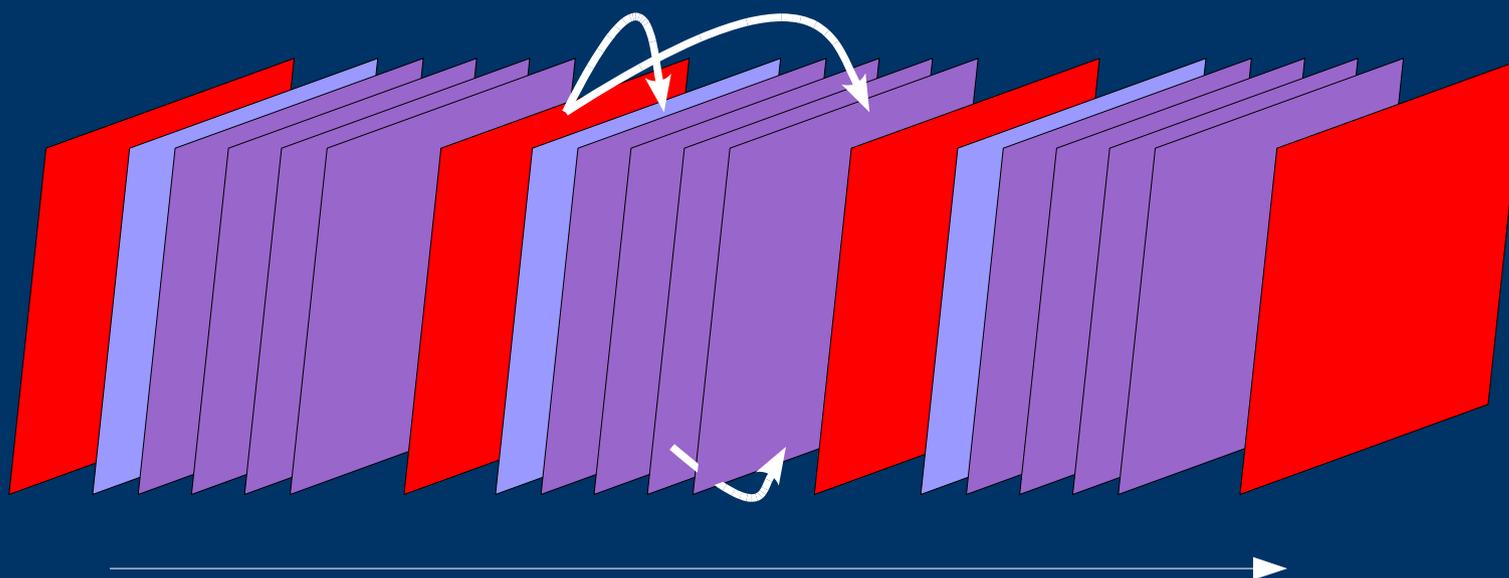
Theora – Rappels

- Espace de couleur YUV
- Sous échantillonnage possible de la chrominance



Theora – Rappels

- Prédiction avant :

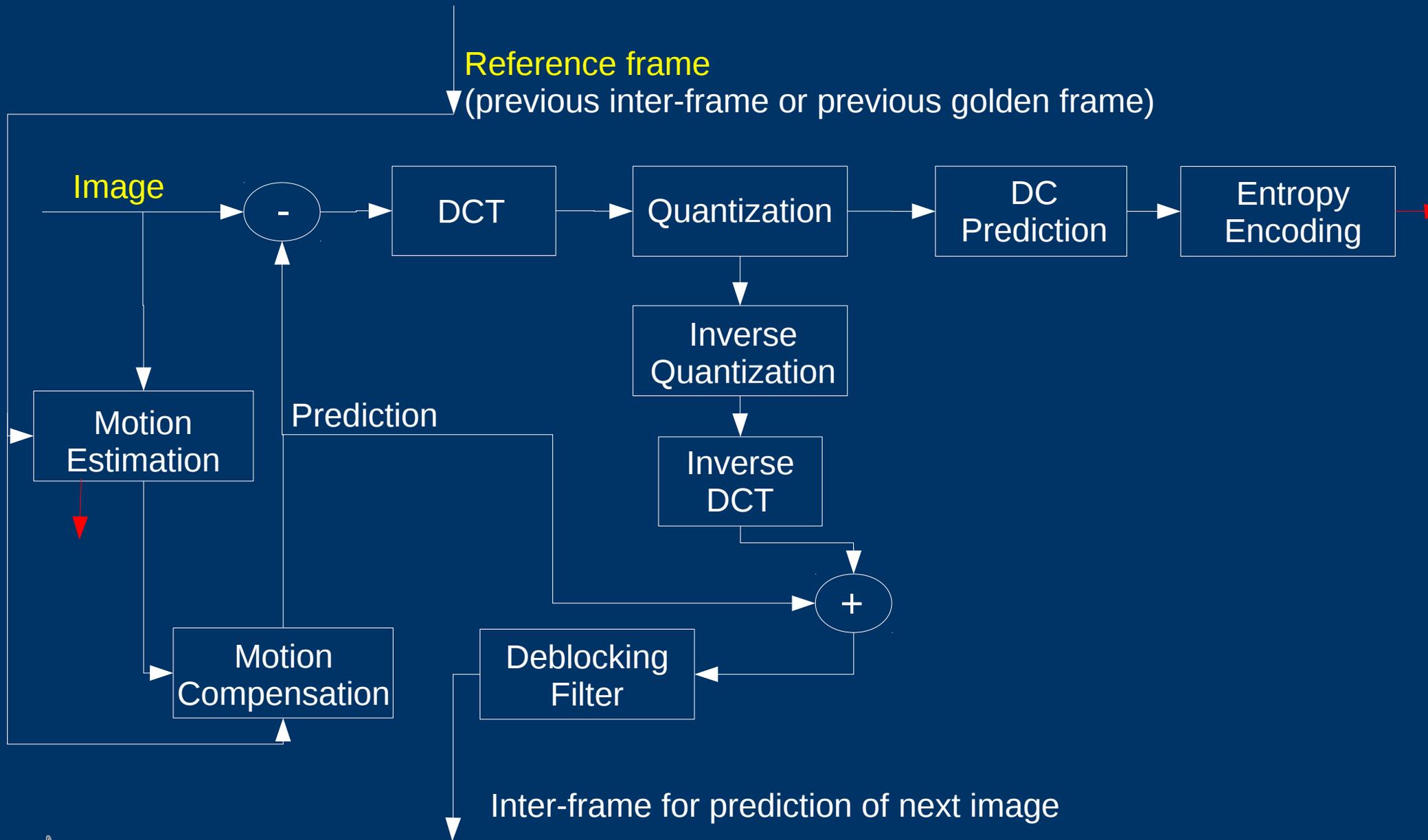


Intra-frame (golden frame)

Inter-frame

Inter-frame

Theora – Rappels



Theora – Rappels

- Compensation de mouvement
 - Recherche dans un voisinage -15, +15
 - Précision au $\frac{1}{2}$ pixel
- Les vecteurs sont enregistrés
 - Soit sur 2x5 bits + 1 bit pour le signe
 - Soit par un code de Huffman



Theora – Intégration

- Intégrer notre méthode
 - Il faut trouver un encodage !
- Problème :
 - Il faut que l'encodage soit moins grand que le gain apporté sur l'encodage de l'erreur de prédiction



Theora – Intégration

- 1ere étape : L'utilisation directe
 - n'a donné (presque) aucun résultat intéressant !



Theora – Intégration

- 2nd étape :
 - Enregistrer seulement les blocs où le gain est significatif
 - Que veut dire significatif ?
 - Il faut que la différence entre l'erreur commise par BM et l'erreur commise par DT soit suffisamment grande.



Theora – Intégration

- 2nd étape :
 - Enregistrer seulement les blocs où le gain est significatif
 - Que veut dire significatif ?
 - Il faut que la différence entre l'erreur commise par BM et l'erreur commise par DT soit suffisamment grande.
- ECHEC !

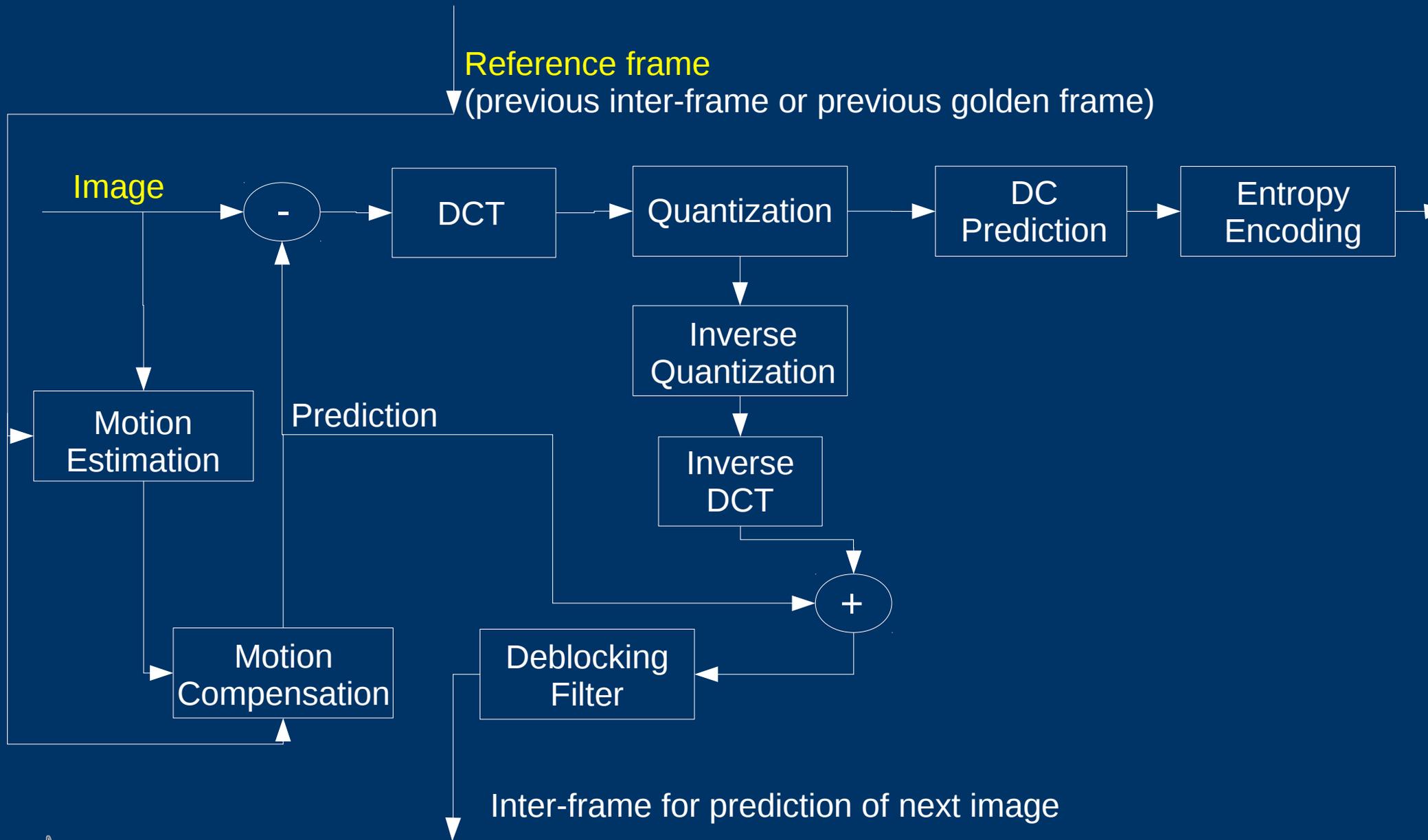


Theora – Intégration

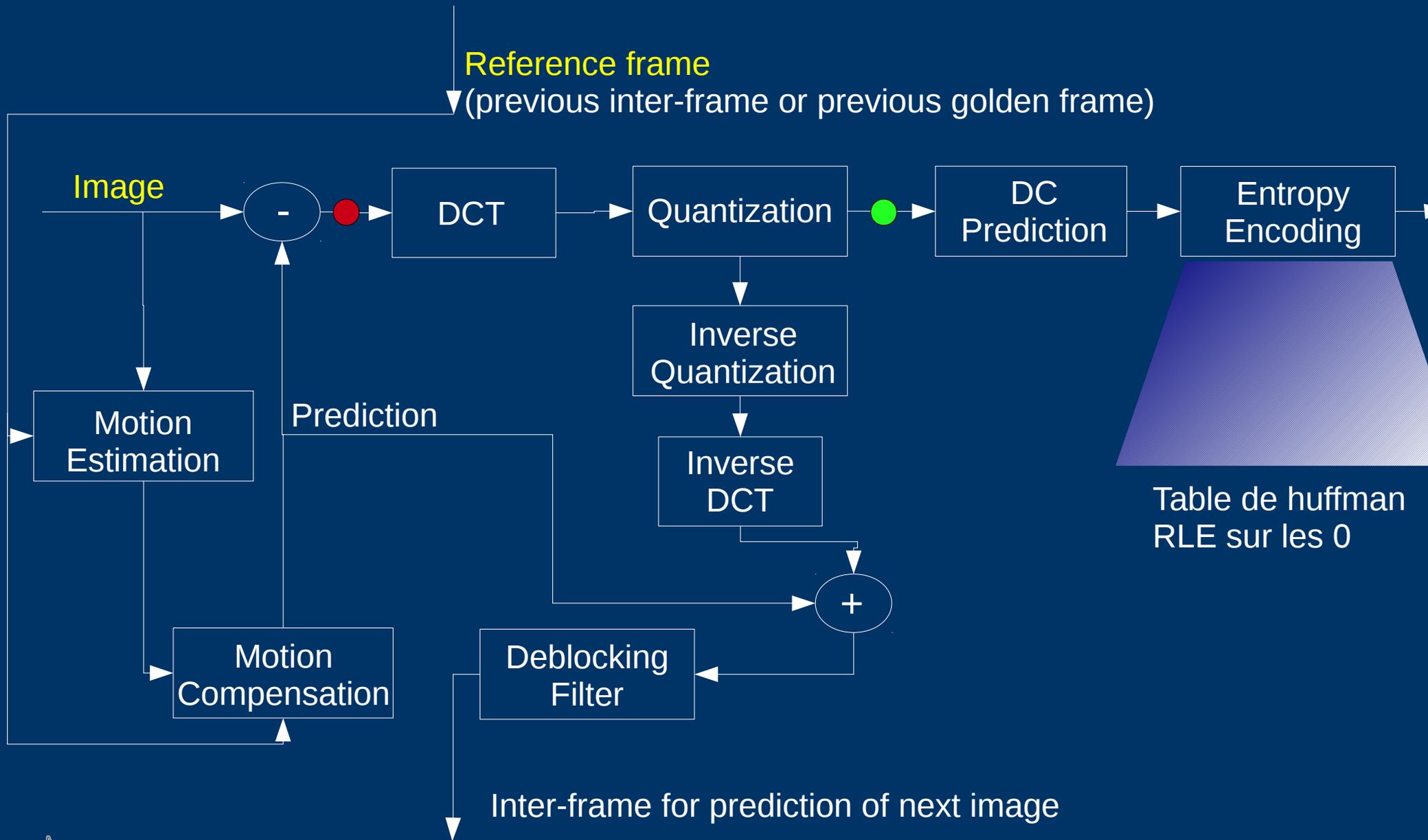
- 3ième étape :
 - Enregistrer seulement les blocs où le gain est significatif
 - Que veut dire significatif ?
 - Effectuer la comparaison dans le domaine fréquentiel !



Theora – Rappels



Theora – Rappels



Theora – Intégration

- 3ième étape :
 - Enregistrer seulement les blocs ou le gain est significatif
 - Elaboration d'une fonction de coût !



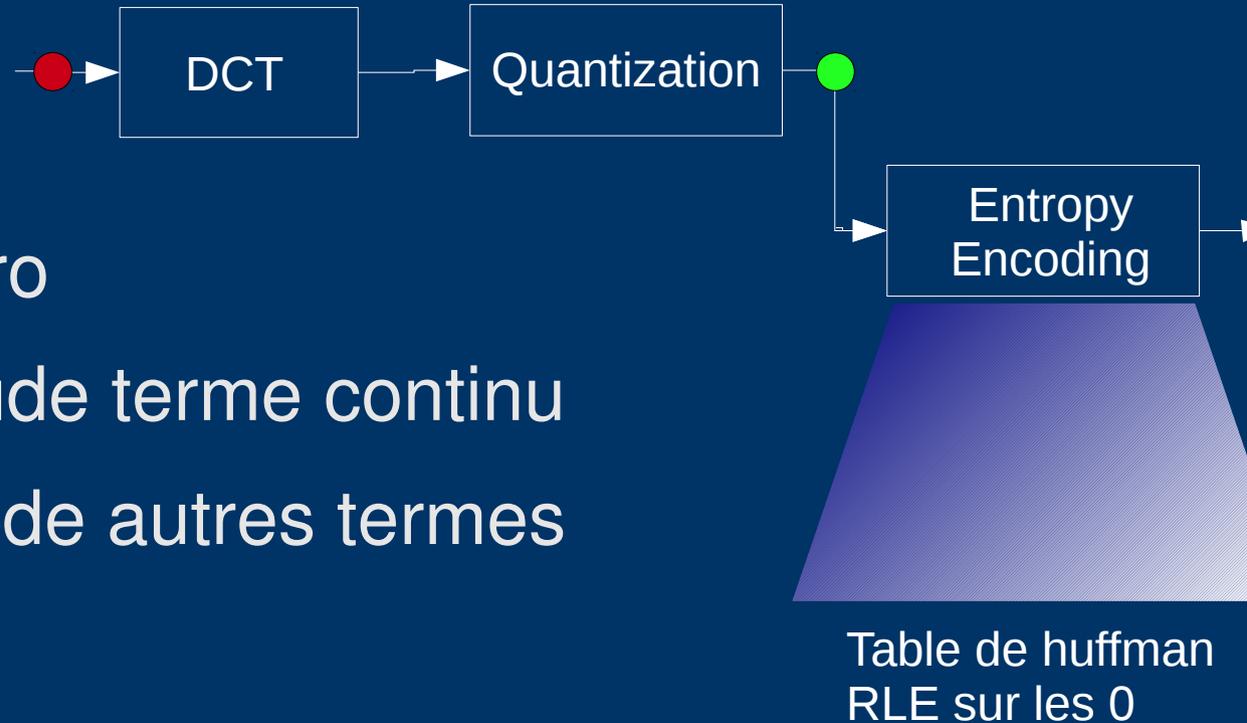
Theora – Intégration

- 3ième étape :

- Z : nb de zero
- DC : Amplitude terme continu
- AC : Amplitude autres termes

$$Kz * (64 - Z) +$$
$$DC / Kdc +$$
$$(AC / Kac1)^2 * Kac2$$

$$Kz = 2$$
$$Kdc = 10$$
$$Kac1 = 1500$$
$$Kac2 = 100$$



Theora – Intégration

- 3ième étape :
 - IMA : erreur cumulée dans le domaine spatial

$$Kz*(64-Z) +$$

$$DC/Kdc +$$

$$(AC/Kac1)^2*Kac2 +$$

$$(IMA/Kima1)^2*Kima2$$

$$Kz = 2$$

$$Kdc = 10$$

$$Kac1 = 1500$$

$$Kac2 = 100$$

$$Kima1 = 1532$$

$$Kima2 = 100$$



Theora – Intégration

- 3ième étape :
 - DT : Nombre de coefficients à ajouter

$$Kz * (64 - Z) +$$

$$DC / Kdc +$$

$$(AC / Kac1)^2 * Kac2 +$$

$$(IMA / Kima1)^2 * Kima2 +$$

$$DT * Kz$$

$$Kz = 2$$

$$Kdc = 10$$

$$Kac1 = 1500$$

$$Kac2 = 100$$

$$Kima1 = 1532$$

$$Kima2 = 100$$



Theora – Intégration

- 3ième étape :
 - Un code RLE qui va déterminer si on garde les coefficients
 - Un code de Huffman pour enregistrer la partie entière
 - Un code de Huffman pour enregistrer la partie décimale (arrondi à 10^{-1})
 - Un bit pour le signe



Theora – Intégration

- Résultats



Théora – Resultats

Sequence (Quality x)	Method	Prediction error		Motion vectors	Image size	Gain/ image	MSE
Coastguard	BM DT						
Container	BM DT						
Football	BM DT						
Foreman	BM DT						
Garden	BM DT						
Hall monitor	BM DT						
Mobile	BM DT						
Tennis	BM DT						



Théora – Resultats

Sequence (Quality 0)	Method	Prediction error		Motion vectors	Image size	Gain/ image	MSE
Coastguard	BM	373	-	2035	2463	-	62.6
	DT	68 (304)	739	1809	2672 (-208)	-6.6%	41.0
Container	BM	289	-	1956	2301	-	72.1
	DT	66 (223)	436	1495	2053 (248)	13.2%	68.2
Football	BM	888	-	1752	2687	-	125.3
	DT	272 (615)	1411	1610	3342 (-654)	-23.3%	96.3
Foreman	BM	317	-	2195	2568	-	39.8
	DT	51 (265)	450	1843	2401 (166)	8.0%	29.0
Garden	BM	1252	-	1603	2902	-	297.7
	DT	577 (675)	1748	1473	3845 (-943)	-29.7%	246.9
Hall monitor	BM	398	-	2083	2536	-	57.3
	DT	99 (299)	474	1594	2223 (313)	14.2%	31.0
Mobile	BM	1523	-	1438	3008	-	373.1
	DT	661 (862)	2018	1336	4062 (-1054)	-30.5%	301.8
Tennis	BM	425	-	1830	2303	-	104.4
	DT	146 (279)	573	1533	2300 (2)	3.2%	84.1



Théora – Resultats

Sequence (Quality 25)	Method	Prediction error		Motion vectors	Image size	Gain/ image	MSE
Coastguard	BM	3760	-	1773	5588	-	17.9
	DT	2843 (917)	607	1699	5205 (383)	6.9%	19.6
Container	BM	1557	-	1701	3313	-	16.6
	DT	1177 (380)	182	1326	2740 (572)	17.2%	17.6
Football	BM	6858	-	1445	8351	-	41.6
	DT	5485 (1373)	1370	1449	8352 (-1)	0.0%	43.2
Foreman	BM	2384	-	2004	4444	-	13.8
	DT	1716 (668)	382	1741	3894 (549)	13.1%	15.2
Garden	BM	9225	-	1426	10699	-	72.0
	DT	8208 (1017)	1205	1341	10802 (-103)	-0.8%	72.9
Hall monitor	BM	2028	-	1953	4037	-	11.7
	DT	1468 (559)	270	1573	3368 (669)	16.9%	12.7
Mobile	BM	11705	-	1159	12911	-	98.4
	DT	10335 (1369)	1399	1125	12907 (4)	0.2%	99.1
Tennis	BM	3022	-	1541	4611	-	33.4
	DT	2206 (816)	444	1381	4078 (532)	12.8%	35.1



Théora – Resultats

Sequence (Quality 45)	Method	Prediction error		Motion vectors	Image size	Gain/ image	MSE
Coastguard	BM	9876	-	1656	11587	-	6.8
	DT	8754 (1121)	808	1658	11276 (311)	2.7%	7.3
Container	BM	4398	-	1464	5917	-	14.4
	DT	3656 (742)	226	1301	5238 (679)	11.3%	6.5
Football	BM	17425	-	1348	18821	-	16.5
	DT	15694 (1731)	1696	1361	18798 (22)	0.1%	17.0
Foreman	BM	7006	-	1869	8930	-	5.9
	DT	5883 (1122)	620	1721	8279 (651)	7.9%	6.5
Garden	BM	22061	-	1388	23496	-	23.2
	DT	20551 (1509)	1496	1323	23417 (78)	0.3%	23.6
Hall monitor	BM	6733	-	1894	8682	-	5.0
	DT	5237 (1495)	443	1709	7445 (1237°)	14.3%	5.1
Mobile	BM	27774	-	1099	28920	-	30.4
	DT	25546 (2227)	1841	1091	28525 (394)	1.4%	30.6
Tennis	BM	8666	-	1338	10052	-	13.0
	DT	7731 (935)	613	1282	9673 (378)	4.2%	13.7



Théora – Resultats

Sequence (Quality 63)	Method	Prediction error		Motion vectors	Image size	Gain/ image	MSE
Coastguard	BM	2137	-	1623	23006	-	2.0
	DT	19934 (1393)	1058	1637	22684 (321)	1.4%	2.0
Container	BM	10875	-	1316	12246	-	1.0
	DT	10077 (797)	385	1238	11756 (489)	4.0%	1.0
Football	BM	40195	-	1338	41580	-	3.1
	DT	37676 (2518)	2097	1354	41175 (404)	1.0%	3.2
Foreman	BM	18381	-	1774	20211	-	1.5
	DT	16500 (1881)	1057	1693	19305 (905)	5.0%	1.9
Garden	BM	48064	-	1375	49487	-	3.8
	DT	45892 (2172)	1800	1335	49074 (412)	0.8%	3.9
Hall monitor	BM	20473	-	1845	22374	-	2.0
	DT	17622 (2851)	958	1805	20440 (1933)	8.6%	2.0
Mobile	BM	60977	-	1073	62096	-	4.2
	DT	57650 (3327)	2275	1079	61051 (1045)	1.7%	4.5
Tennis	BM	22405	-	1286	23738	-	2.5
	DT	21061 (1343)	891	1274	23274 (464)	2.0%	2.7



Théora – Résultats

- Résultats encourageants !
 - On peut faire en sorte de ne jamais perdre de la place par rapport à l'approche par BM car au pire on met les coefficients à 0...
- Impact sur le nombre de « golden frame » ?



Théora – Résultats





Conclusion



Conclusion

- Bilan :
 - Nous avons mis en place une méthode efficace de compensation de mouvement
 - Nous avons prouvé son efficacité à la fois de manière théorique mais aussi de manière pratique en l'intégrant dans un codec (Theora)



Conclusion

- Améliorations possibles :
 - Sélectionner les transformations modélisées
 - « Creuser » la comparaison dans le domaine fréquentiel
 - revoir la fonction de score ?
 - minimiser dans le fréquentiel ?
 - Revoir l'encodage des valeurs
 - Etudier la vitesse d'exécution
 - ...

