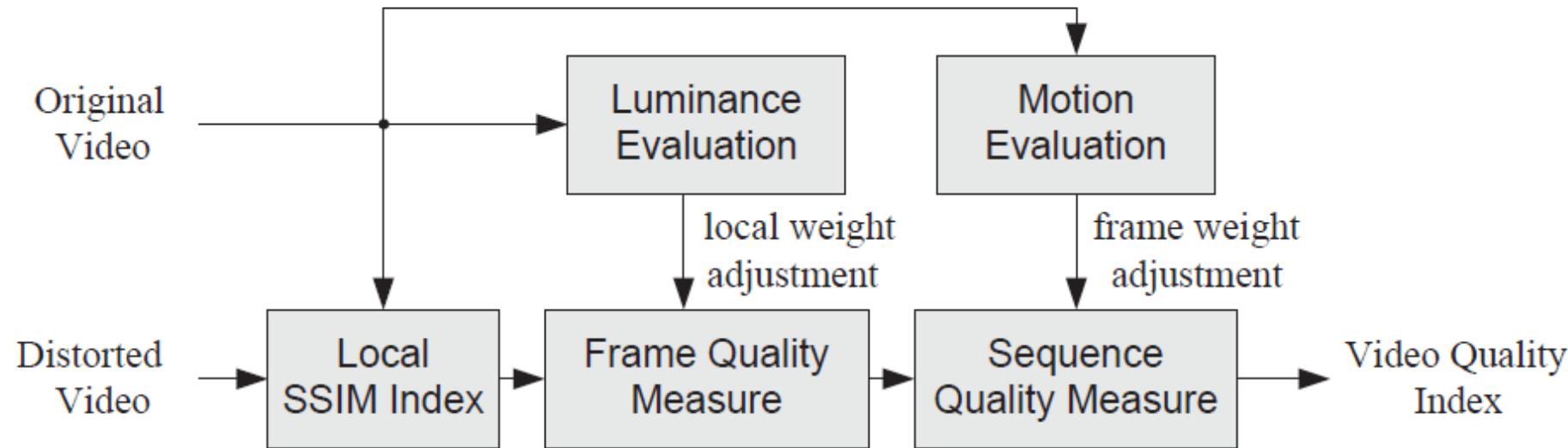


Évaluation objective de la qualité des vidéos

Métrique a référence complète

Schéma VSSIM



Niveau local

1. les zones d'échantillonnage locales sont extraites du Frame dans les emplacements spatiaux correspondants dans les séquences vidéo originales et déformées, respectivement. Les zones d'échantillonnage sont des fenêtres 8*8 choisies au hasard. Le nombre de fenêtres d'échantillonnage par image vidéo R_s permet de représenter la densité d'échantillonnage.
2. L'approche d'indexation SSIM est alors appliquée sur la luminance Y , et les deux composantes de la chrominance Cb et Cr , d'une façon indépendante, et ensuite combinées avec une somme pondérée pour avoir mesure la qualité local

$$\begin{aligned} \text{SSIM}_{ij} = & W_Y \text{SSIM}_{ij}^Y + W_{Cb} \text{SSIM}_{ij}^{Cb} \\ & + W_{Cr} \text{SSIM}_{ij}^{Cr}, \end{aligned}$$

Les poids de la relation sont fixés empiriquement comme suit : $W_Y = 0,8$, $W_{Cb} = 0,1$, $W_{Cr} = 0,1$

Niveau Frame

- ❖ C'est le second niveau de mesure de qualité, dans lequel les différent mesure locales, issues de la première étape sont combiner ensemble, pour avoir un indexe de qualité par frame, selon la formule suivante:

$$Q_i = \frac{\sum_{j=1}^{R_i} w_{ij} \text{SSIM}_{ij}}{\sum_{j=1}^{R_i} w_{ij}},$$

Wij est le poids donné pour les SSIM calculer localement, il est ajusté en utilisant l'observation que les régions sombres attirent moins l'attention que les région lumineuses, donc il est calculé en utilisant la moyen de luminance dans les blocks, avec la relation suivante :

$$w_{ij} = \begin{cases} 0 & \mu_x \leq 40, \\ (\mu_x - 40)/10 & 40 < \mu_x \leq 50, \\ 1 & \mu_x > 50. \end{cases}$$

Niveau vidéo

- ❖ C'est le troisième et le dernier niveau, où les indexe niveau frame sont combinés, pour avoir un indexe global de toute la vidéo. La formule qui permet cette combinaison est :

$$\mathbf{Q} = \frac{\sum_{i=1}^F \mathbf{W}_i \mathbf{Q}_i}{\sum_{i=1}^F \mathbf{W}_i},$$

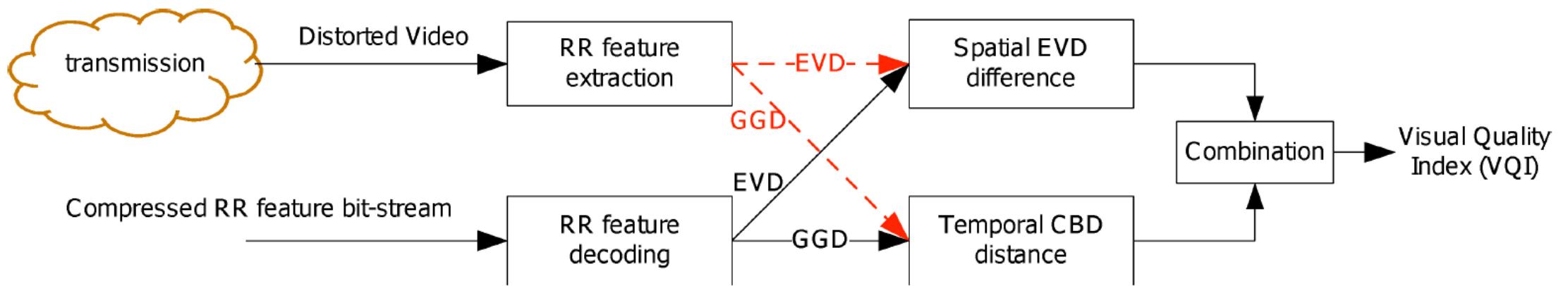
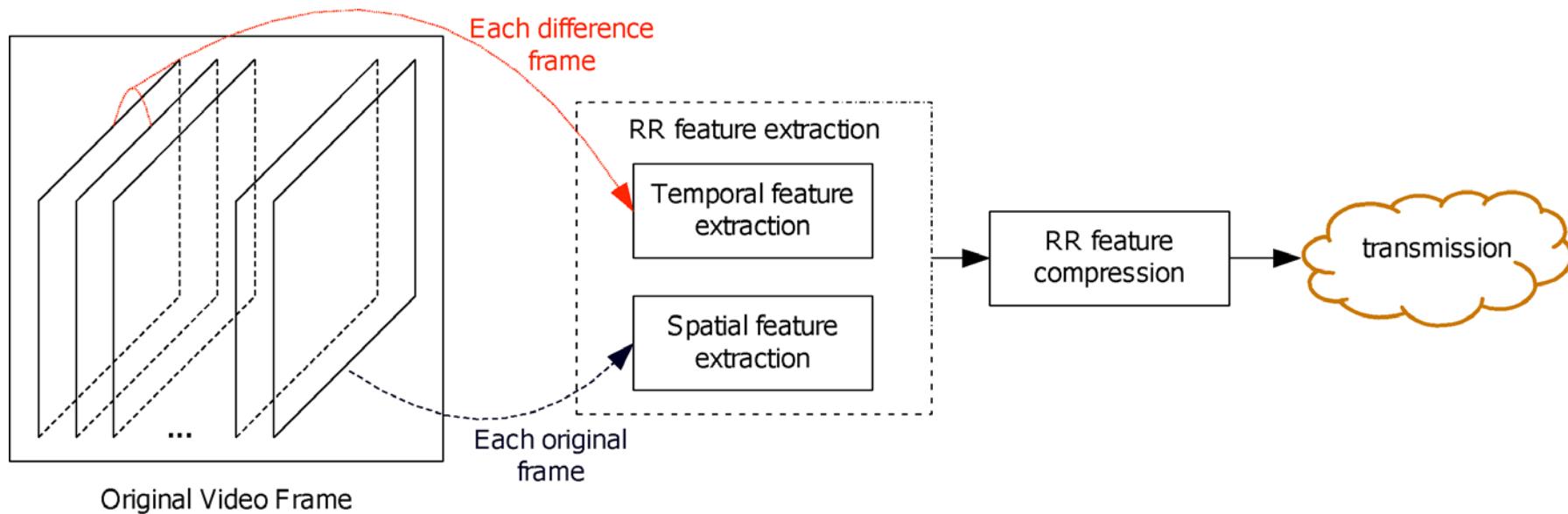
Pour calculer le poids \mathbf{W}_i , il faut (pour chaque fenêtre d'échantillonnage) trouver le vecteur de mouvement \mathbf{m}_{ij} , en utilisant la méthode du **block matching** avec recherche dans le frame suivant. Ensuite le degrés de mouvement dans le frame et calculer par la formule :

$$\mathbf{M}_i = \frac{(\sum_{j=1}^{R_s} m_{ij}) / R_s}{K_M}, \quad Km = 16$$

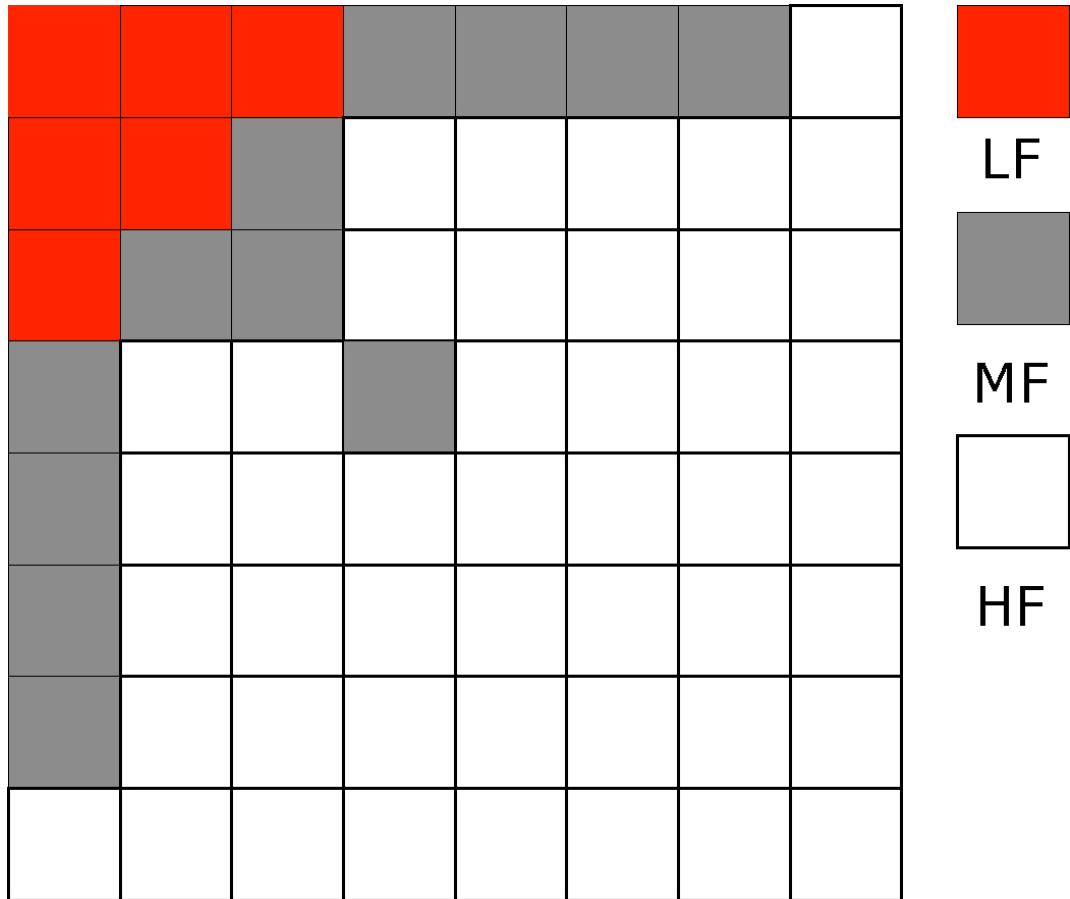
$$\mathbf{W}_i = \begin{cases} \sum_{j=1}^{R_s} w_{ij} & \mathbf{M}_i \leq 0.8, \\ ((1.2 - \mathbf{M}_i) / 0.4) \sum_{j=1}^{R_s} w_{ij} & 0.8 < \mathbf{M}_i \leq 1.2, \\ 0 & \mathbf{M}_i > 1.2. \end{cases}$$

Le poids de chaque frame est calculé alors par :

Métrique a référence réduite

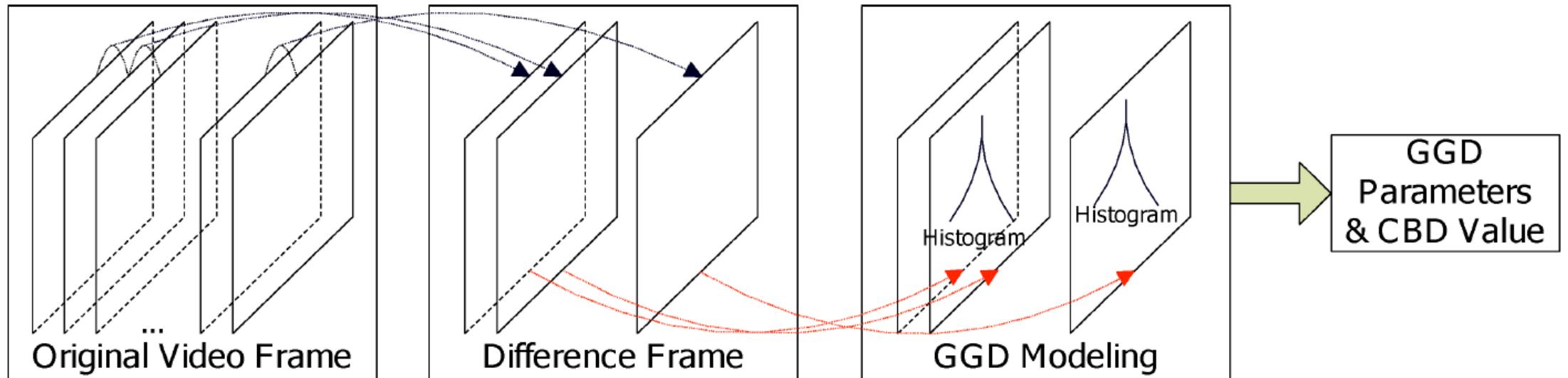


Extraction des caractéristiques spatiales



$$EVD = \frac{(M + H)}{L}.$$

Extraction des caractéristiques temporales



$$d_{\text{CBD}}(p, p_{\alpha, \beta}) = \sum_{i=1}^{h_L} |p(i) - p_{\alpha, \beta}(i)|$$

Comparaison des caractéristiques spatiales

$$EL = |EVD_{\text{ori}} - EVD_{\text{pro}}|$$

$$EL_V = \frac{EL}{EVD_{\text{ori}}} = \frac{|EVD_{\text{ori}} - EVD_{\text{pro}}|}{EVD_{\text{ori}}}$$

Comparaison des caractéristiques temporales

$$d_{\text{CBD}}(p, p_d) = \sum_{i=1}^{h_L} |p(i) - p_d(i)|$$

$$d_{\text{CBD}}(p, p_d) \triangleq |d_{\text{CBD}}(p_{\alpha, \beta}, p_d) - d_{\text{CBD}}(p, p_{\alpha, \beta})|$$

Combinaison des indexes pour avoir l'indexe vidéo

$$Q_s = EL_V \times \log_{10} \left(1 + \frac{d_{\text{CBD}}(p, p_d)}{c} \right).$$

$$\text{VQI} = \sum_{i=1}^N Q_s(i)/N$$