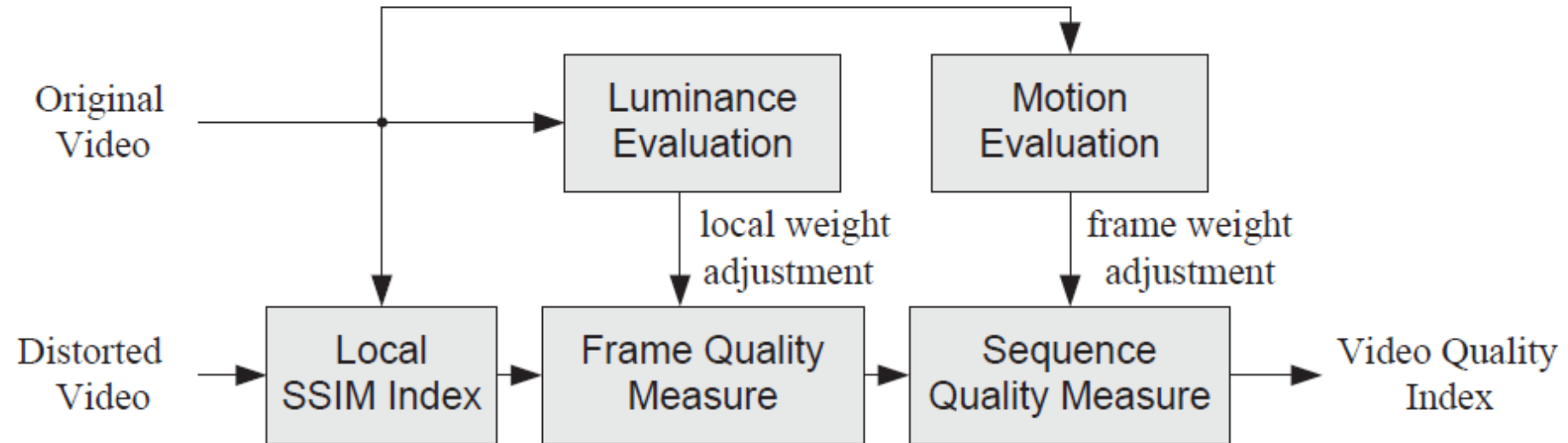


# Évaluation objective de la qualité des vidéos

Métrieue 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 Rs 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

$$SSIM_{\vec{y}} = W_Y SSIM_{\vec{y}}^Y + W_{Cb} SSIM_{\vec{y}}^{Cb} + W_{Cr} SSIM_{\vec{y}}^{Cr},$$

Les poids de la relation sont fixés empiriquement comme suit : **Wy = 0,8, Wcb = 0,1, Wcr = 0,1**

# Niveau Frame

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

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

**$w_{ij}$**  est le poids donné pour les SSIM calculés localement, il est ajusté en utilisant l'observation que les régions sombres attirent moins l'attention que les régions lumineuses, donc il est calculé en utilisant la moyenne 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 est le dernier niveau, où les indexes niveau frame sont combinés, pour avoir un indexe global de toute la vidéo. La formule qui permet cette combinaison est :

$$Q = \frac{\sum_{i=1}^F W_i Q_i}{\sum_{i=1}^F W_i},$$

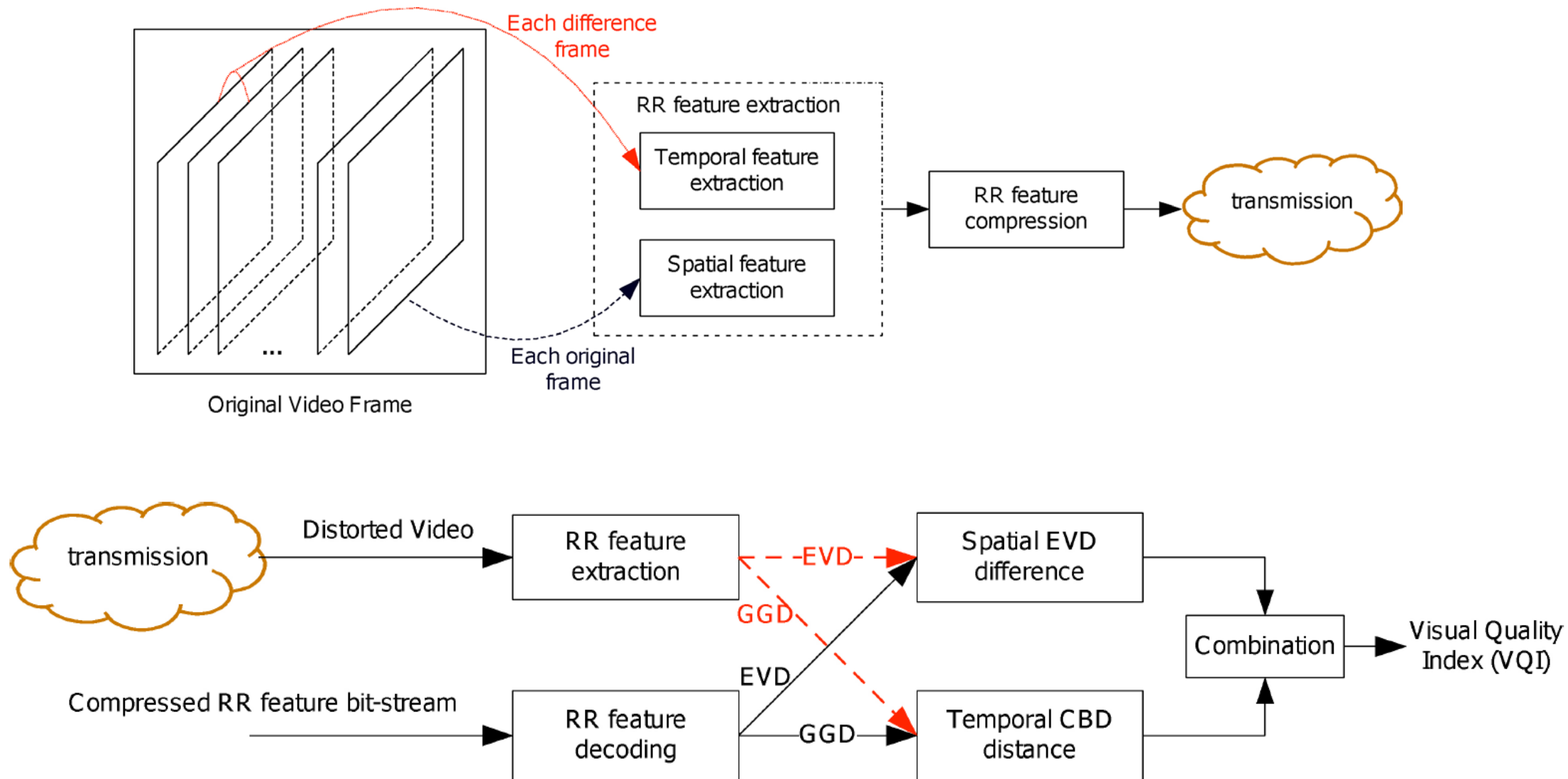
Pour calculer le poids **W<sub>i</sub>**, il faut (pour chaque fenêtre d'échantillonnage) trouver le vecteur de mouvement **m<sub>ij</sub>**, en utilisant la méthode du **block matching** avec recherche dans le frame suivant. Ensuite le degrés de mouvement dans le frame et calculer par la formule :

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

Le poids de chaque frame est calculé alors par :

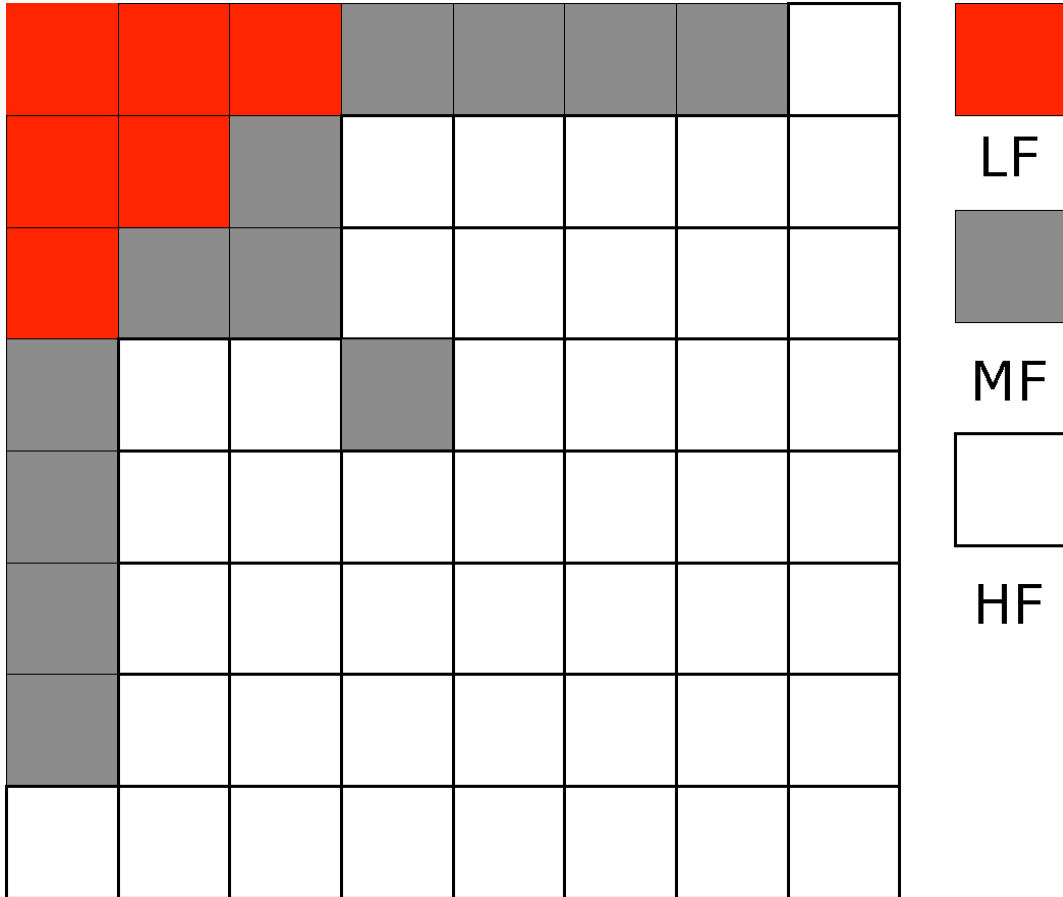
$$W_i = \begin{cases} \sum_{j=1}^{R_s} w_{ij} & M_i \leq 0.8, \\ ((1.2 - M_i)/0.4) \sum_{j=1}^{R_s} w_{ij} & 0.8 < M_i \leq 1.2, \\ 0 & M_i > 1.2. \end{cases}$$

# Métrique a référence réduite



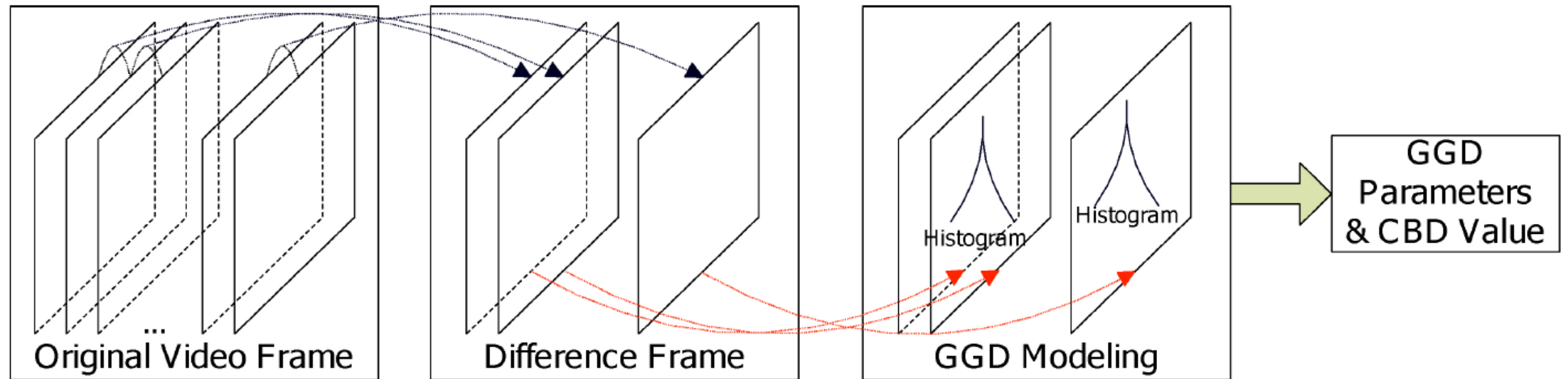


# Extraction des caractéristiques spatiales



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

# Extraction des caractéristiques temporeles



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

# 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 temporelles

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

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

Combinaison des indexes pour avoir l'indexe vidéo

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

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