Fig. 2: Basic decomposition scheme of the CA-WL for $i_{\max}{=}3$ decomposition levels resulting in one BL and three ELs. Depending on the underlying motion in the original sequence, the local depth of the decomposition differs.

box in Fig. 1. In the first step, splitting is performed by decomposing the input signal into even- and odd-indexed frames l_{2t} and l_{2t-1}. In the second step, the even frames are predicted from the odd frames by a prediction operator $\mathcal{P}$. Subtracting the predicted values $\mathcal{P}(l_{2t-1})$ from the even frames, results in the HP coefficients h_{2t}. In the third step, the HP coefficients are filtered by an update operator $\mathcal{U}$ and are added back to the odd frames, resulting in the LP coefficients. To get coarser temporal resolutions, the lifting scheme can be iterated on the LP subband by $i_{\max} = \log_2(T)$ decomposition levels, where T equals the total number of original frames.

Since the lifting structure offers a flexible framework, it can be modified in multiple ways. By introducing rounding operators as introduced in [8], integer to integer transforms can be achieved, which yield perfect reconstruction. This makes the lifting structure of the WT highly attractive for many professional applications by offering a scalable representation and lossless reconstruction at the same time. However, due to temporal displacements in the sequence, blurriness and ghosting artifacts will appear in the LP subband. These can be alleviated by incorporating MC methods directly into the lifting structure without losing the property of perfect reconstruction. This is called motion compensated temporal filtering (MCTF) [9] and can be achieved by realizing the prediction operator $\mathcal{P}$ by the warping operator $\mathcal{W}_{2t-1\to 2t}$. Instead of the original odd frames, a compensated version is subtracted from the even frames. In case of the Haar wavelet filters, the prediction step is given by

$$h_{2t} = l_{2t} - \lfloor \mathcal{W}_{2t-1\to 2t}(l_{2t-1}) \rfloor. \tag{1}$$

However, to achieve an equivalent wavelet transform, the compensation has to be inverted in the update step [10]. By reversing the index of $\mathcal{W}$, the LP coefficients of the Haar transform can be calculated by

$$l_{2t-1} = l_{2t-1} + \left\lfloor \frac{1}{2}\mathcal{W}_{2t\to 2t-1}(h_{2t}) \right\rfloor. \tag{2}$$

$\mathcal{W}$ can be realized by different approaches of MC. In this work, we will employ a block-based approach.

3. CONTENT ADAPTIVE WAVELET LIFTING

Considering video sequences from surveillance systems or medical data sets, which comprise a temporal acquisition of images with contrast medium, there will be parts, where almost no motion occurs over time. In this case, an adaptive temporal scaling is advantageous, which performs iteratively a further decomposition, if subsequent frames are similar enough. If there are no changes over several frames, they shall be represented by only one LP frame. For significant changes among subsequent frames, for example when the contrast medium gets visible, these changes shall be represented in the LP subband with finer temporal resolution.

Fig. 2 shows the basic approach of our proposed content adaptive wavelet transform. Index i indicates the number of the current decomposition level. For $i{=}0$, no decomposition has been done so far, which corresponds to the original video sequence. In the first row, a schematic video sequence is given, which consists of three sections, each with a different amount of moving content. The corresponding amount of motion is described by the legend on the right side of Fig. 2. While in this example the first decomposition is performed for the entire sequence, the second decomposition is performed only on the frames with no or low motion. The third decomposition is exclusively done on frames with no motion. The resulting BL and ELs are marked at the right side. Since the maximum decomposition level $i_{\max}$ is equal to 3, three ELs are generated. By combining the ELs with the BL at the decoder side, the original sequence can be reconstructed step by step without any loss.

3.1. Calculation of the Stopping Criterion

Haar WTs can be represented with tree structures [4]. For 3-D SBC, the tree structure is given by decomposing two subsequent frames into LP and HP frames. To realize the CA-WL, the costs of the single nodes in every tree have to be considered. If the combined costs of the child nodes exceed the costs of the parent node, this means for an arbitrary signal s, if

$$\mathcal{C}(s_{i,[2t-1,2t]}) \leq (\mathcal{C}(s_{i+1,2t-1}) \cup \mathcal{C}(s_{i+1,2t})) \tag{3}$$

holds, then the child nodes shall be pruned from the tree. Here, $\mathcal{C}(\cdot)$ describes a cost functional, which represents the coding costs, such as entropy [11] or rate-distortion [12]. In this work, every decomposition level is performed for the entire input sequence in advance, before a retrospective evaluation of the resulting costs is done. Further, we decided to use a rate-distortion-based approach for calculating the coding costs. Therefore, we formulate the Lagrangian cost