Temporal Resolution Enhancement in Compressed Video

Mark A. Robertson and Robert L. Stevenson

Laboratory for Image and Signal Analysis

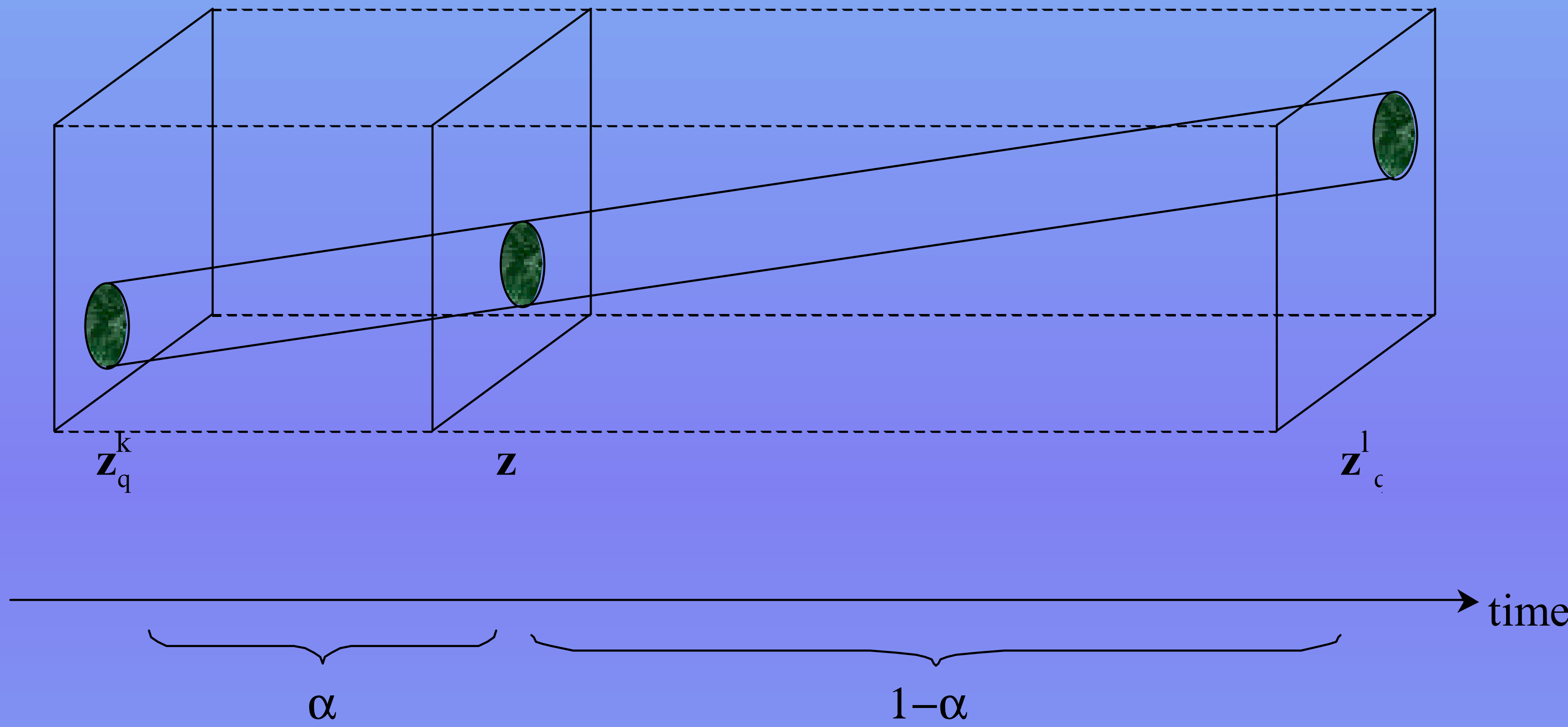
University of Notre Dame, IN 46556

Motivation

1. Sometimes, we may want to increase video frame rate to avoid jerky, uneven motion:
   (a) Frames can be "dropped" during encoding
   (b) Might want a "slow-motion" effect
2. Frame rate is increased by inserting frames between received frames.
3. Want to prevent propagation of compression artifacts from received frames to inserted frames.

Notation:

$\mathbf{z}_q^k$	Decoded frame at time k (with quantization noise)
$\mathbf{z}_q^l$	Decoded frame at time l (with quantization noise)
$\mathbf{z}$	Frame to insert between $\mathbf{z}_q^k$ and $\mathbf{z}_q^l$



Distance between $\mathbf{z}_q^k$ and $\mathbf{z}_q^l$ is normalized to one

Assume a MAP criterion for estimation:

$$\begin{aligned}\hat{\mathbf{z}} &= \arg\max_{\mathbf{z}} p\left(\mathbf{z}|\mathbf{z}_q^k, \mathbf{z}_q^l\right) \\ &= \arg\max_{\mathbf{z}} p\left(\mathbf{z}\right) p\left(\mathbf{z}_q^k, \mathbf{z}_q^l|\mathbf{z}\right)\end{aligned}$$

$p\left(\mathbf{z}_q^k, \mathbf{z}_q^l|\mathbf{z}\right)$ relates observations to image we are estimating

$p\left(\mathbf{z}\right)$ is a prior image model that encourages smoothness

Need to relate observations $\mathbf{z}_q^k$ and $\mathbf{z}_q^l$ with frame to insert $\mathbf{z}$.

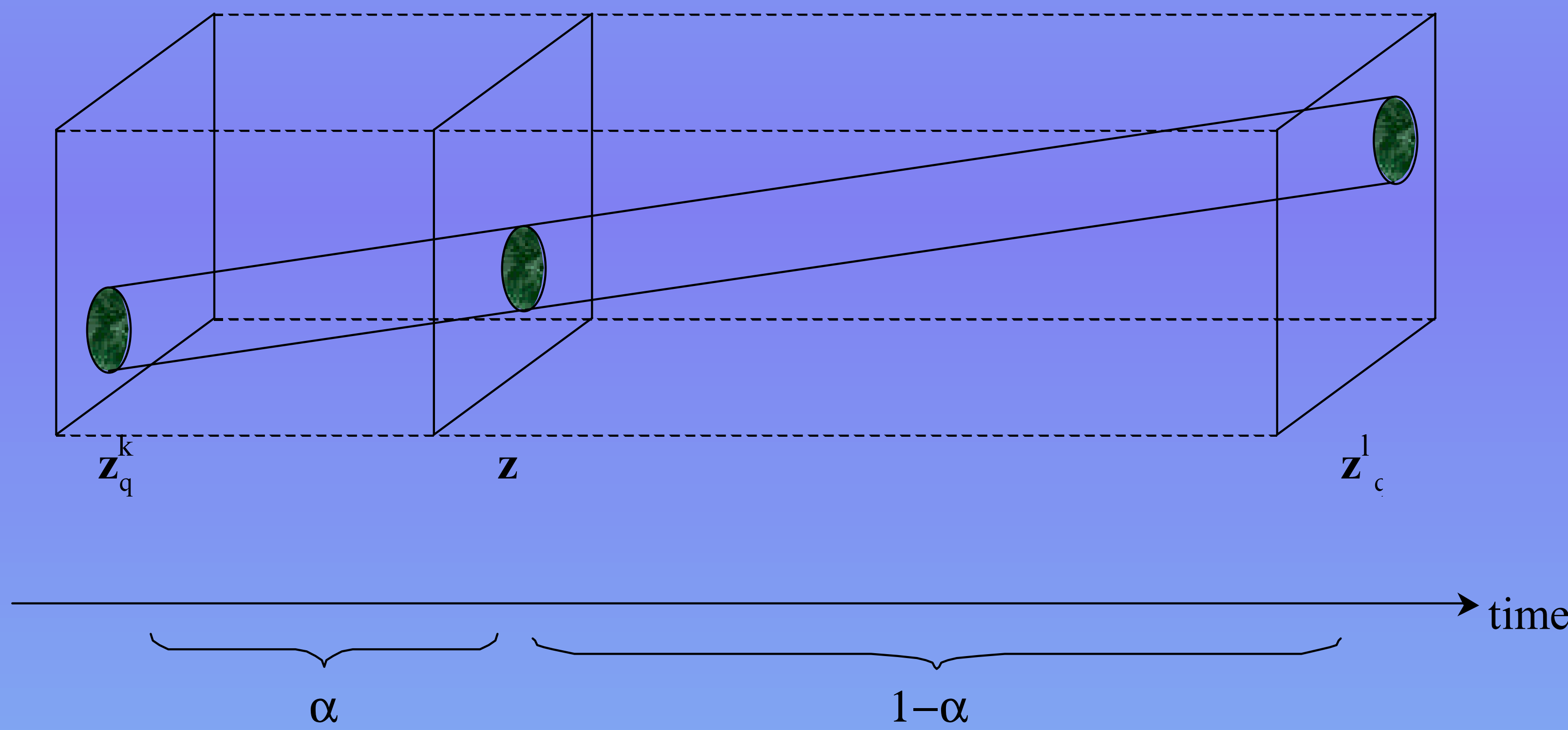
$\Longrightarrow$ start with original images (include compression noise later):

$$\mathbf{z}^k = \mathbf{A}^k \mathbf{z} + \mathbf{n}_m^k$$

$\mathbf{A}^k$ is the *motion-compensation* matrix.

$\mathbf{n}_m^k$ is the error in predicting $\mathbf{z}^k$ from $\mathbf{z}$

Motion compensation matrix from $\mathbf{z}^k$ to $\mathbf{z}$ is determined based on motion from $\mathbf{z}_q^k$ to $\mathbf{z}_q^l$ assuming constant velocity:



Now, include quantization noise in the model:

$$\begin{aligned}\mathbf{z}_q^k &= \mathbf{z}^k + \mathbf{e}_{\mathbf{z}}^k \\ &= \mathbf{A}^k\mathbf{z} + \mathbf{n}_m^k + \mathbf{e}_{\mathbf{z}}^k \\ &= \mathbf{A}^k\mathbf{z} + \mathbf{n}^k\end{aligned}$$

$\mathbf{n}^k$ is the noise (error) between $\mathbf{z}_q^k$ and its motion compensation from $\mathbf{z}$.

$$\begin{aligned}\mathbf{e}_{\mathbf{z}}^k &\Longrightarrow \text{assume Gaussian with } \mathbf{K}_{\mathbf{e}_{\mathbf{z}}}^k \\ \mathbf{n}_m^k &\Longrightarrow \text{assume Gaussian with } \mathbf{K}_m^k\end{aligned}$$

$\mathbf{e}_{\mathbf{z}}^{k}$ Quantization noise is characterized in the spatial domain by explicitly considering the quantization intervals used on DCT-domain signal. In general, *not* IID.

$\mathbf{n}_{m}^{k}$ Motion compensation noise allows high variance to be assigned to areas of low confidence—e.g., appearing or disappearing objects, lighting changes, motion estimation errors.

(more details in paper)

Assuming conditional independence

$$p\left(\mathbf{z}_q^k, \mathbf{z}_q^l|\mathbf{z}\right) = p\left(\mathbf{z}_q^k|\mathbf{z}\right) p\left(\mathbf{z}_q^l|\mathbf{z}\right)$$

where

$$p\left(\mathbf{z}_q^k|\mathbf{z}\right) = \frac{(2\pi)^{-\frac{N}{2}}}{|\mathbf{K}^k|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{A}^k\mathbf{z} - \mathbf{z}_q^k)^t {\mathbf{K}^k}^{-1}(\mathbf{A}^k\mathbf{z} - \mathbf{z}_q^k)\right\}$$

and

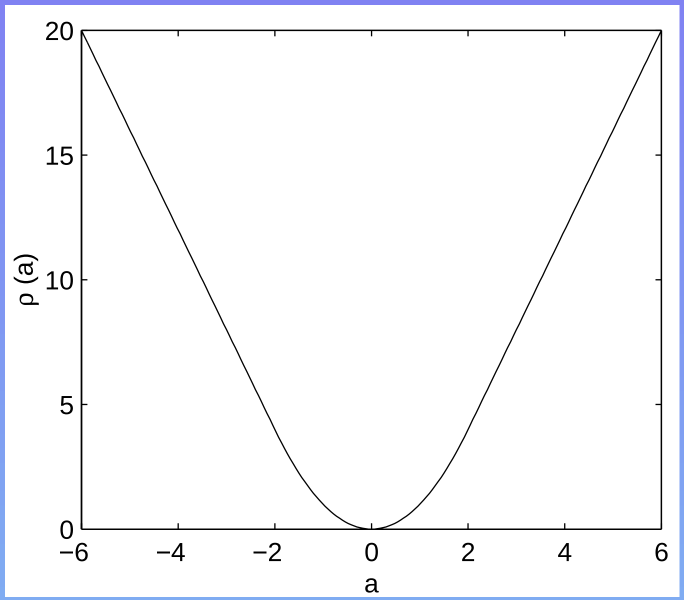
$$\mathbf{K}^k = \mathbf{K}_{\mathbf{e_z}}^k + \mathbf{K}_m^k$$

Prior Term

Huber Markov Random Field (HMRF):

$$p(\mathbf{z}) = \frac{1}{G} \exp\left\{ -\lambda_I \sum_{c \in \mathcal{C}} \rho_T\left(\mathbf{d}_c^t \mathbf{z}\right) \right\}$$

where Huber function $\rho_T(\cdot)$ looks like



Prior term has two main effects:

- Smooths artifacts such as blocking and ringing
- Fills in "gaps" in $\mathbf{z}$, to which no motion vectors from $\mathbf{z}_q^k$ or $\mathbf{z}_q^l$ point

Resulting optimization:

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \left[\lambda_F r(\mathbf{z}) + s(\mathbf{z})\right]$$

where $r(\mathbf{z})$ is from the prior smoothness term, and $s(\mathbf{z})$ is from the likelihood term.

Solve using gradient descent:

Gradient is $\mathbf{g}(\mathbf{z}) = \lambda_F \nabla r(\mathbf{z}) + \nabla s(\mathbf{z})$

Estimate at k^{th} iteration is updated as

$$\mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} - \tau^{(k)} \mathbf{g}\left(\mathbf{z}^{(k)}\right)$$

received frame 1: ■ □ □ □

received frame 2: □ □ □ ■

received frame 1: ■ □ □ □

inserted frame 1: □ ■ □ □

inserted frame 2: □ □ ■ □

received frame 2: □ □ □ ■

received frame 1: ■ □ □ □ □

received frame 2: □ □ □ □ ■

received frame 1: ■ □ □ □ □

inserted frame 1: □ ■ □ □ □

inserted frame 2: □ □ ■ □ □

inserted frame 3: □ □ □ ■ □

received frame 2: □ □ □ □ ■

Conclusion

Frame rate adjustment in compressed video that:

- Uses Bayesian estimation
- Explicitly considers the effect of compression noise in received video frames
- Uses prior image model to smooth compression artifacts, as well as “fill in” motion gaps in inserted frames