**Introduction**

Aerial Video Surveillance (AVS) is employed over wide area, for hours at a time resulting in immense data collections.

**Computational objective**

Disambiguate independent and structural motion in Wide Area Motion Imagery (WAMI) efficiently and robustly.

**Applications**

- Increase video compression rate via object-oriented video compression
- Robustly identify moving objects (vehicles, persons, sudden change in scenery) for end users in what can often be sparse terrain

**Signal decomposition**

Let \( f(x) \) be the intensity normalized 1D vector for a fixed time period \( n \) at location \( x = (x, y) \).

Owen's equation for signal decomposition is as follows:

\[
\begin{align*}
\hat{f}(x) &= f_s(x) \times f_e(x) \\
\hat{f}(x) &= f_s(x) \cdot f_e(x) - \mathcal{H}(f_s(x)) \cdot \mathcal{H}(f_e(x))
\end{align*}
\]

where \( \mathcal{H}(f) \) is the Hilbert transform of signal \( f \), and the smooth and local energy components are derived below, respectively.

\[
f_s(x) = \frac{\hat{f}(x) \cdot f_e(x) + \mathcal{H}(\hat{f}(x)) \cdot \mathcal{H}(f_e(x))}{f_e^2(x) + \mathcal{H}^2(f_e(x))}
\]

\[
f_e(x) = \sqrt{\hat{f}^2(x) + \mathcal{H}^2(\hat{f}(x))}
\]

The reconstruction of \( f(x) \) with the Direct Current (DC) component nullified is \( \hat{f}(x) \).

**Time-frequency analysis**

The power spectrum of \( f_s(x) \) is calculated for every pixel location. The metric \( g(x) \) summates over the normalized powers which peak along the period axis. Considering only the maximal values of periodic power, there is no need for any windowing function, which in practice may dampen the values we seek and increase the minimum length of the time window \( n \) required.

Let \( P = \{p_1(x), p_2(x), \ldots, p_k(x)\} \) be the sorted set of peaks of normalized power in descending order where \( \tau < |p/2| \). Then the motion classification metric is defined as follows:

\[
g(x) = \sum_{i=1}^{k} p_i(x), \quad k \leq \tau
\]

The typical ranges of the metric's parameters are \( 3 \leq k \leq 10, \alpha \in [0.05, 0.30], \) and \( \beta \in [0.50, 0.90] \). The time window \( n \) should observe at least a single cycle of precession.

**Experiments and results**

Results are shown for a qualitative view structural motion detection using artificial data (first row) and for independent motion detection using proprietary, real WAMI showing ROC and PR curves (color graphs). The run-time of the method per pixel is \( O(n \log n) \).

In the first row, regions undergoing precession are tinted blue with \( \beta = 0.5 \). Like many pixel-based detectors, this does not guarantee detection for region interiors due to surface homogeneity. In real data, surfaces glean in the sensor appearing inhomogeneous over time.

For WAMI, precision increases with \( n \) and peaks for specific combinations of \( k \) and \( \alpha \), across all blocks. Precision confidence for all curves increases with the amount of independent motion. Gains for all sequences diminish once two cycles of precession have been observed (\( n = 600 \)).