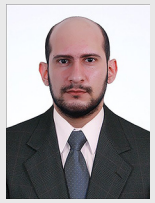
Atmospheric Turbulence Mitigation Using a Multi-Frame Transformer Architecture



This work investigates the use of a Transformer-based deep learning model, the Turbulence Mitigation Transformer (TMT), for restoring image sequences degraded by atmospheric turbulence. The architecture incorporates temporal-channel joint attention, multi-scale supervision, and an encoder–decoder framework to jointly model spatial and temporal features. A temporally averaged frame, computed as the pixel-wise mean of the input sequence, is added to the model input to provide a stable structural reference during reconstruction. Training is performed using a composite loss function that combines Structural Similarity Index (SSIM), Charbonnier, and Mean Squared Error (MSE) terms, ensuring a balance between perceptual quality and pixel-level accuracy. Image sequences are grouped according to turbulence strength using the refractive index structure parameter Cn2, enabling evaluation across a range of distortion conditions. The inclusion of the average frame improves restoration quality at all turbulence levels. The most significant improvements in SSIM occur under moderate turbulence, where the model most effectively recovers structural details.

Juan Carrillo *and Jerome Gilles*

This research is supported by the Air Force Office of Scientific Research grant (FA9550-21-1-0275), the Department of Mathematics & Statistics at San Diego State University, and the Computational Science Research Center (CSRC) at San Diego State University

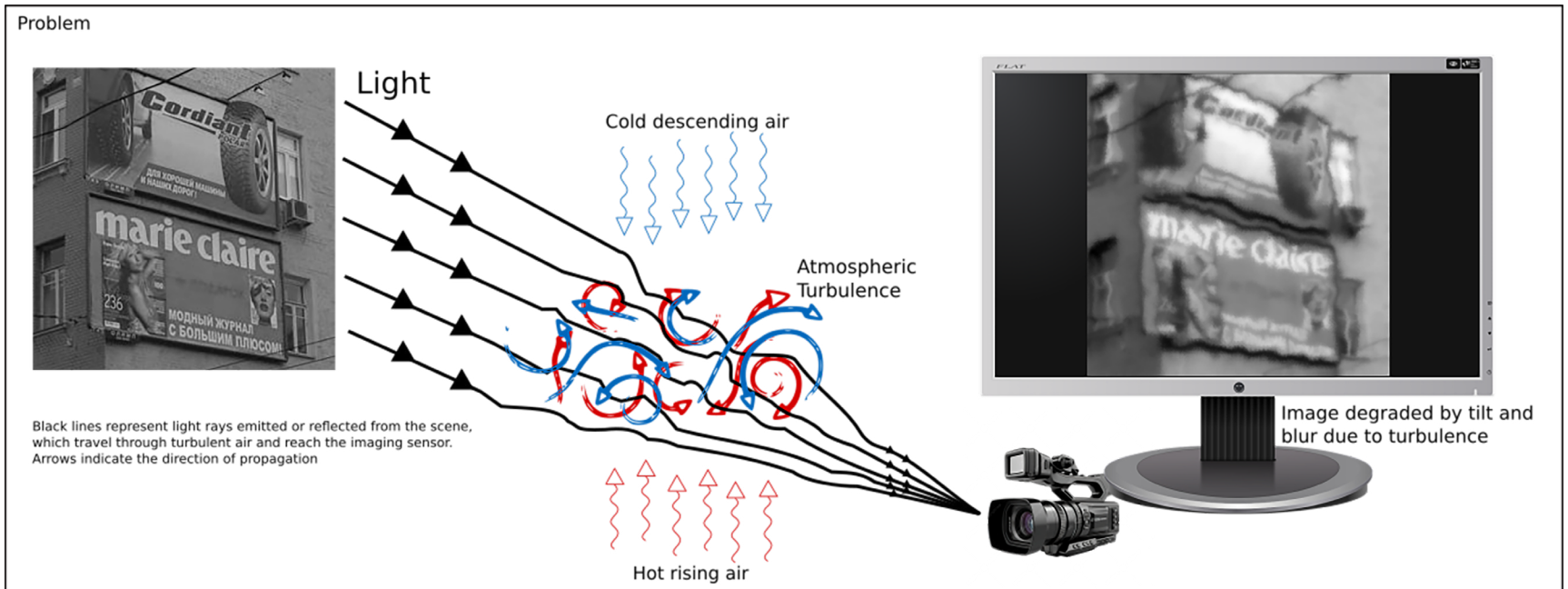


Figure 1. Illustration of atmospheric turbulence affecting image formation. Light rays emitted or reflected from the scene (left) travel through turbulent air layers caused by hot rising and cold descending air flows, which introduce refractive index fluctuations. These fluctuations result in image distortion characterized by geometric warping (tilt) and loss of sharpness (blur), as seen in the degraded image on the display. Arrows on the light paths indicate direction of propagation toward the imaging sensor.

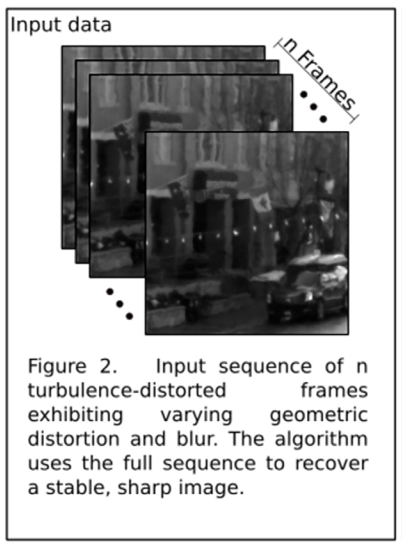


Figure 2. Input sequence of n turbulence-distorted frames exhibiting varying geometric distortion and blur. The algorithm uses the full sequence to recover a stable, sharp image.

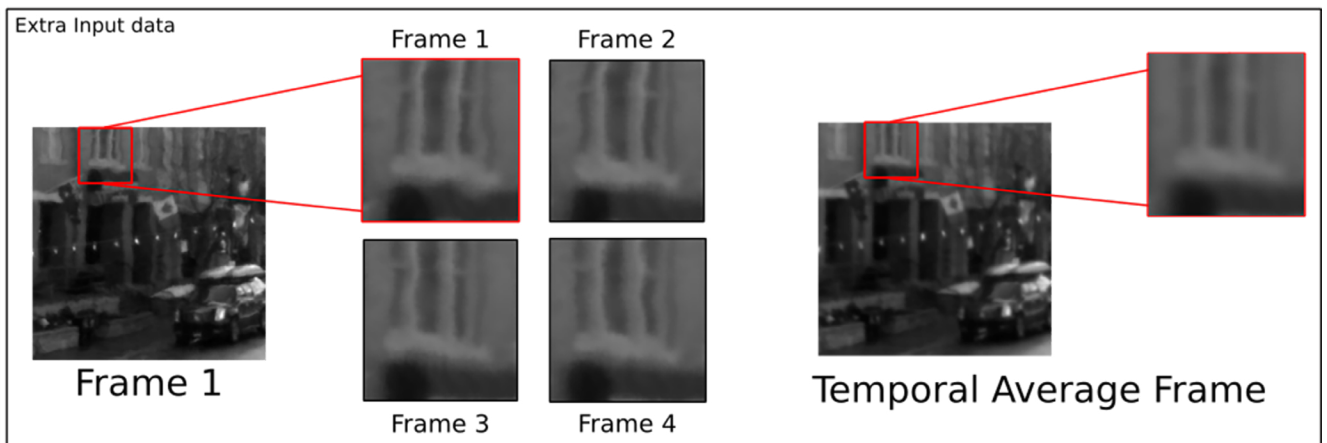


Figure 3. Temporal degradation in turbulence-distorted video frames. The left panel shows four individual frames affected by atmospheric turbulence, where local geometric distortions vary over time. On the right, the temporal average frame reduces temporal fluctuations but results in a blurred image due to the accumulation of misaligned distortions.

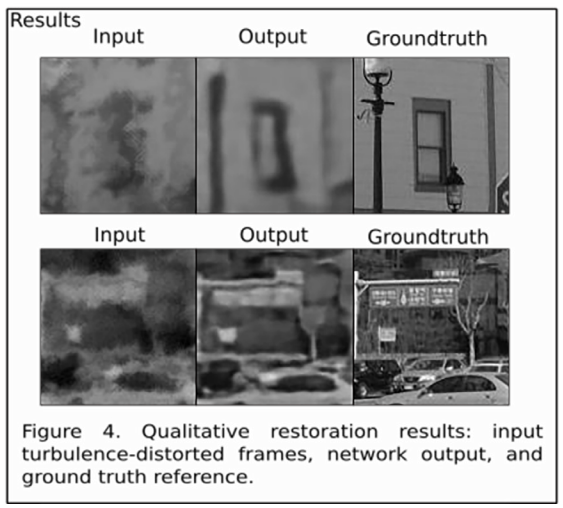


Figure 4. Qualitative restoration results: input turbulence-distorted frames, network output, and ground truth reference.