International Journal of Computer Science , Engineering and Artificial Intelligence
ISSN: 3139 – 9002 (Online) , https://ijcseai.com/
Regular Issue, Vol. 2, Issue. 1 (January – March) , 2025 , Pages: 25 - 31
Received: 21 October 2024 ; *Accepted:* 25 February 2024 ; *Published:* 21 March 2025
https://doi.org/10.63328/IJCSEAI-V2RI1P5
https://www.youtube.com/watch?v=brNN9L_tUww
https://archive.org/details/jsp-ijcseai-v2ri1p5
Original Paper : Action Research – Field Research Paper

3D Structural Deformation Estimation Using Multi-Camera Fusion in Formula 1

Tirlingi Tirupathi Rao

Department of Computer Science and Engineering , Lendi Institute of Engineering and Technology(A), Vizianagaram , Andhra Pradesh, India 535005 ;

* *Corresponding Author: Tirlingi Tirupathi Rao ; thirupatirao.t@lendi.edu.in*

Abstract: Rear wings, front wings and floor edges are examples of or more distinctive parts of the car's structure that must bend in response to aeroelasticity when a car is running in an aerodynamic configuration which does not encompass ground-effect regulation. Even at sub-millimetres scales, these geometric deformations will modify the pressure distributions in aerodynamics and therefore a vehicle's stability at supersonic velocities of over 300 km/h. Even though a number of telemetry solutions exist, using discrete strain gauges and single camera vision pipelines can only be used for local measurement or planar measurement, which cannot provide a complete measurement of the state of structural conditions for real-time SHM. Proposed in this paper is a Multi-Stage and Multi-Camera Fusion (MCF) structure for 3D structural deformation estimation that combines synchronized high-speed cameras, stereo correspondence matching, multi-view triangulation and a temporal learning module based on the Transformer architecture. The proposed system surpasses all baseline systems in terms of both reconstruction accuracy (98.9%) and precision of depth measurement (0.8 mm mean reconstruction error) while maintaining a 38 ms latency, meeting the definition of real-time performance on our aeroelastic benchmark for Formula 1.

Keywords: Formula 1, Multi-Camera Fusion, 3D Reconstruction, Structural Deformation, Aeroelasticity, Transformer.

1. Introduction

The The introduction of ground-effect aerodynamics into The new regulations from 2022, which incorporated elements of ground-effect aerodynamics, changed the rules and paradigms of how Formula 1 cars are aerodynamically conceived to this day. Ground-effect cars are able to create far greater aerodynamic forces with significantly lower drag thanks to the venturi effect of the tunnel raised created under the flat racing floor. This type of aerodynamic thrust regime also presents a major structural problem as aerodynamic properties, such as the shapes of the front wing endplates, rear wing assemblies and floor edge geometry, are subject to elastic deformations that change their aerodynamic profile during a race [1].

In the aerodynamics components, there can be various deformations in different time scales. The aerodynamic steady deformation (quasi-static) is due to the aerodynamic loading acting over sustained period of time and the aerodynamic flutter phenomena and kerb strike produce the dynamic deformations from road irregularities. Aeronautic balancing can be affected by only a few centimeters, even when the vehicle's front wing profile is only slightly moved by a millimeter-size deformation; however, the distortion of the profile could be mistaken for a change in the front suspension or other changes without careful on-board structural monitoring [2]. Knowing such deformations in real-time and being able to model them, then, is a tremendous safety and competitive advantage. In motorsports, traditional structural monitoring methods are based on the use of discrete bonded piezoelectric strain gauges (PSGs) on composite laminates and accelerometers within the structural components. These sensors are capable of high frequency measurements and have a good temporal resolution, but have no spatial reconstruction ability and can only measure at their place of installation. Global deformation field reconstruction based on the limited number of point deformation measurements is a highly degenerate inverse problem, and unreliable at high aerodynamic forces where multiple deformations modes mix nonlinearly [3]. An attractive alternative is computer vision, which allows for extensive structural surface measurements without the need for integration of the sensor into the structure, and the measures themselves will be spatially rich. Under controlled lighting, single-camera

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

photogrammetry is also being examined for wing deformation monitoring during wind tunnel tests with good accuracy (<millimeters). Monocular systems are, however, inherently constrained by the depth ambiguity problem associated with projecting 3D information onto a 2D image plane and, for this reason, strong prior assumptions about, for example, material properties and/or boundary conditions are required; such assumptions may not be easily known under dynamic racing conditions [4].

In this paper, we present a Multi-Camera Fusion (MCF) framework that, by leveraging the key of online and synchronized multi-view geometry, deep feature correspondence matching and a Transformer network model for temporal structural evolution, solves these limitations. The proposed system recovers metrically accurate 3D point clouds of the deforming surfaces of structural components in real-time from multiple simultaneous views at a rate of 42 frames per second to monitor Structural Deformation Index (SDI) for front wing, rear wing and floor edge assemblies [4]. The main contribution of this work is: (2) Synchronized multi-camera acquisition and calibration system for measuring 3D structures of racing vehicle components at high speed under operating loads. A multi-view triangulation pipeline for sub-millimeter 3D reconstruction based on feature extraction by SIFT, tied and least-squares stereo matching and Bt. (3) A Transformer-based temporal learning module, modelling the temporal evolution of structural deformation fields, that can predict a short megacycle deformation state for proactively enabling the vehicle with deformation control feedback. (4) Composite structural deformation index (SDI) implementation which merges displacement amplitude, velocity, and thermal values into a single metric for structural health to contribute to real-time closed-loop racing system.

2. Related Works

2.1. Traditional Structural Monitoring

The field of structural health monitoring (SHM) of aerospace and motorsport structures has been well progressing for the last 30 years. In [7] Worden et al. surveyed all vibration-based SHM methods, showing that various modal analysis techniques enable the detection of distributed structural damage [5]. The most important sensor modality in motorsport has been the resistance strain gauge whose only significant advantage is the ability to reliably measure the local strain on the surface, but has a lack of spatial generalizability. The necessity of using point sensors in complex multi-modal structures is inherently inefficient, as optimal sensor positions for modal identification have been shown by Kammer [4] to depend on a minimum number of sensors that is approximately sqrt{M}, where M is the number of modes of interest. Fibre Bragg Grating (FBG) sensors have been used in Formula 1 for measuring spatially distributed strains along structural axes with higher spatial resolution than individual gauges are useful but are sensitive to transverse strains and processing the data is more complex [6]. Moreover, open loop systems based on high contrast speckle patterns, with either single or stereo cameras are used to wind tunnel testing for displacement measurements in sub-micrometre resolution range in a controlled laboratory environment. Yet in a lab setting, DIC needs controlled light, a fixed structure and no vibrations, which isn't available when used in operational racing.

2.2. Multi-View 3D Reconstruction

Much work has been carried out in the field of multi-view stereo (MVS) reconstruction, which has developed greatly over the years with work led by Hartley & Zisserman [1] laying the foundations for multi-view 3D reconstruction via a series of key formulation including the projective geometry of a view, the essential matrix, and the epipolar constraint. While structure-from-Motion (SfM) pipelines, like the COLMAP, can reconstruct rigid structures with an accuracy of a few centimetres when the scene is static, they limit their reconstruction to the assumption of rigidity of the entire scene and they need a high baseline length between images in order to avoid ambiguity.

Structure-from-Motion (SfM) pipelines, including COLMAP, require a high baseline between views and lack the certainty of complete rigidity, assumptions which are extended to high speed dynamic structures causing reduced accuracy of a few centimeters in reconstruction. From the classical block-matching, stereo correspondence matching has evolved to deep learning techniques [7]. Mayer et al. present FlowNet, a method for optical flow estimation, followed by a similar approach to stereo disparity estimation. Kendall et al. proposed to add learned cost volume filtering in stereo matching to achieve state-of-the-art disparity accuracy on the KITTI benchmark. However the existing approaches were developed for the reconstruction of the scene and do not support explicit temporal reasoning about the structural dynamics and, do not apply them to surface deforming structural surfaces.

2.3. Deep Learning for Structural Analysis

Estimating structural deformation using deep learning is a new theme that has emerged aligning computer vision and structural engineering. Raissi et al. have shown that Physics-Informed Neural Networks (PINNs) can be used to solve partial differential equations (PDEs) governing structural mechanics, including the embedding of physical constraints during learning. In the case of Formula 1 aerodynamics, Baque et al. trained graph neural networks for aerodynamic surrogate modeling and showed that these networks could transfer their predictive performance of the flow over different vehicle geometry. Shi et al. [3] introduced a CNN-LSTM network for dealing with structural deformation monitoring with 96.3% reconstruction accuracy than

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

tested on laboratory specimen data without the multi-camera fusion and temporal attention. Use of transformer models with time series data.Application of transformer models for time series data. In multivariate timeseries prediction, the selfattention mechanism proposed by Vaswani et al. [5] is extended from NLP to time-series prediction. For long-sequence time-series forecasting with linear complexity [8], Wu et al. suggested the Autoformer, which is competitive with LSTM-based baseline competiting energy and weather data. To effectively model long structural time series, Zhou et al. propose a variant of the Informer that employs a sparse attention mechanism to conclude complexity from $O(L^2)$ to $O(L \log L)$. The proposed MCF framework is based on these architectures which represent the temporal learning module.

3. Existing System

The structural monitoring via current telemetry and vision systems used in F-1 tuner monitoring are limited in three fundamental ways with regard to usefulness for realtime aeroelastic monitoring. There are several reasons why strain gauge arrays only return scalar measurements at discrete locations: The global deformation field is interpolated between the gauge measurements, thus giving high interpolation error at areas away from the sensors [9].

Moreover, gauges measure surface strain instead of displacement, so the surface displacement field must be integrated over the structure or structural paths to obtain the displacement, which can lead to very large numbers of errors, and the elastic properties of the material measured by the gauge may be function of temperature during a race.

Second, the vision system with a single camera has the problem of depth ambiguity in monocular projection. The scale may be obtained from known markers attached to the structural surface; however, the approach is sensitive to the calibration drift which could arise from vibration and temperature change of the camera [10].

Unavailability of three camera installations results in very limited field of view for single camera installations, which is inadequate to capture simultaneous deformation over the 'whole' extent of wing assemblies (more than 2 metres). Third, known approaches to fusion don't have temporal reasoning ability.

Three dimensional reconstruction based on the point in time has no forecasting power to predict imminent overload conditions in the structures as required for closed-loop aerodynamic interventions like the application of the Drag Reduction System (DRS) modulation or suspension stiffness in active systems. In absence of such modeling of time the outputs of the reconstruction may not be interpreted to apply to changes of loading that are transient, but occur due to aerodynamic instability [11].

4. Proposed Work

4.1. System Architecture Overview

This is the proposed MCF framework that involves five sequential processing modules as shown in Fig. 1. Six High Speed Cameras (Phantom v2640, 1000 fps at 2048×1952 resolution) are fixed onto a carbon fibre camera rig attached to the reference frame for the vehicle, three cameras look out at the front wing assembly and three cameras look out at the rear wing and edge of the floor. The inter-camera timing accuracy for hardware trigger is sub-micro second. The entire pipeline delivers synchronized frame sets at 42 frames per second and delivers dense 3D point clouds and SDI values with less than 38 ms end-to-end latency [11].

4.2. Camera Calibration

The 3×3 intrinsic matrix K, with values which are focal lengths (f_x, f_y), principal point (c_x, c_y), and coefficients of the lens distortion, is obtained from each camera through Zhang's planar checker board method. The relationship between a world point X and its image projection x is called a projection, x' is an image of X, and X' will be known as its projection image, X is called its image, and x will be called its image point:

$$x = P\,X = K[R \mid t]X$$

The camera rotation ($R \in SO(3)$) and translation ($t \in \mathbb{R}^3$) with respect to the world frame. Extrinsic Multi-camera calibration is performed by optimizing them all simultaneously by performing a bundle adjustment with a large-format calibration target which has sub-pixel image corners detected:

$$E = \Sigma_i \Sigma_j \, ||x_i - \pi(P_i, X_j)||^2$$

where $\pi(P\boldsymbol{i}, Xj)$ represents the projection of point Xj in the world through camera $P\boldsymbol{i}$. Their structural features are visible on site, so an online recalibration makes up for pose perturbations due to thermal drift and vibrations during operation.

4.3. Feature Extraction and Stereo Correspondence

Each frame captured by a camera is processed to obtain a number of keypoints using Scale-Invariant Feature Transform (SIFT) which are described by 128 dimensional descriptor vectors [12], which are rotation- and scale-invariant. The Fundamental matrix F between two camera's image planes (i, j) is estimated from the calibrations:

$$x'^{T} F x = 0$$

In which x and x' denote image point at the position x and x' in the camera i and j, respectively. By applying the epipolar constraint it therefore becomes only one dimensional search problem from the two dimensional feature matching problem when searching along the epipolar lines that enables efficient correspondence establishment. Descriptor matching is first implemented based on a ratio test (Lowe's criterion with the threshold of 0.75) and then an outlier rejection process is performed [13]

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

based on RANSAC, with the fundamental matrix as the projective model, resulting in only inlier correspondences that satisfy the sub-pixel reprojection consistency.

4.4. Multi-View Triangulation

Starting from N ≥ 2 cameras with their projection matrices $\{P_i\}$ and image observations $\{x_i\}$ of a scene point X, the 3D position of X is computed by minimizing the summed reprojection error of all the views:

$$X = argmin\ \Sigma_i\ ||x_i - P_iX||^2$$

A Direct Linear Transform (DLT) is used to solve this optimal triangulation along with non-linear Levenberg-Marquardt refinement. Triangulation with all N=6 cameras simultaneously makes use of the maximum multi-view redundancy which reduces the depth error variance by a factor of √(N-1) compared to pairwise stereo triangulation. The dense point cloud generation constructs disparity maps from rectified stereo pairs by semi-global matching (SGM) and then back-projected to 3D by using camera parameters.

4.5. Transformer-Based Temporal Learning

Structural deformation with time is modeled with a Transformer encoder architecture. To obtain a fixed-dimension feature vector from the 3D point cloud at each timestep we use PointNet based aggregation [14]. The feature vectors are concatenated to form an input matrix, which undergoes L=4 Transformer encoder layers with a number of attention heads in each layer h=8. Contextual relations over the temporal sequence is computed by the scaled dot-product attention:

$$Attention(Q, K, V) = softmax((QK^T/\sqrt{d_k}))V$$

If sinusoidal function of temporal position indices are used for positional encodings, the attention calculation gets temporal position information. The encoded sequence is then decoded to yield a short-horizon prediction for the deformation state at time T+δ in the future, for pro-active structural health assessment [15]. The goal of the training is to minimize the mean squared error between the predicted and ground-truth future deformation fields, along with a physics-based consistency regularizer against failure to respect the elastic constraints of equilibrium, i.e., imposing an error.

4.6. Structural Deformation Index

The Structural Deformation Index (SDI) is an overall scalar index for each structural part combining various deformation measurement types, such as displacement, velocity and thermal parameter:

$$SDI = \alpha D + \beta V + \gamma T$$

where D is the normalised magnitude of the peak displacement (mm), V the rate of change in displacement (mm/s) and T is the normalised thermal loading factor determined from measurements of the embedded thermocouples. On each type of component, these are empirically calibrated using an FEA validated structural test, with α=0.5, β=0.3 and γ=0.2 giving the best ability to predict the onset of failure for the front wing, rear wing and floor assemblies, respectively. When SDI values go beyond a threshold which is specific to a particular component, engineering alarms are send live to the pit wall [16].

5. Methodology

5.1. Dataset and Experimental Setup

The dataset used and their set up. The data set utilized and the experimental setup. Aerodynamic benchmarks from a Formula 1 vehicle of current specification were taken on a test circuit at Fiorano, the results of which were used to create a custom Formula 1 Aeroelastic Benchmark (F1-AEB) dataset. A 600×400 mm checkerboard target was used to perform the six-camera rig calibration before each session. Collects 47 high speed runs over a variety of aerodynamic configurations, including front wing angles of 2° to 8° and rear wing configurations of low, medium and high downforce, resulting in 1.2 million frame sets collected. To validate the reconstruction, ground truth 3D displacements were measured by attaching 24 reflective fiducial markers to each structural component which was recorded at 500 Hz with a VICON motion capture system. Test specimens were loaded with controlled loads and the damage to the structure was simulated by applying these loads to the components with a calibrated actuator [17].

5.2. Implementation Details

The implementation of the MCF pipeline is done in Python 3.10, exploiting OpenCV 4.8 for camera calibration and feature extraction, and PyTorch 2.1 for the temporal module (Transformer). The preprocessing is done on an FPGA co-processor (Xilinx Alveo U250) for deterministic latency, the deep learning inference module is mounted in the vehicle's bodywork and is provided by an NVIDIA Jetson AGX Orin embedded GPU module. All stages from the frame capture to SDI output are deterministic and deliver a latency of 38ms at 42 fps throughput. The Transformer module is trained with 80% of the F1-AEB dataset using the Adam optimizer (lr=10^{-4}, β_1=0.9, β_2=0.999), cosine annealing LR schedule over 150 epochs, and early stopping based on the validation reconstruction loss.

5.3. Evaluation Protocol

Evaluation is done based on four criteria: (1) Reconstruction Accuracy: Percentage of triangulated points within 1.5 mm of the ground truth; (2) Mean Depth Error (MDE): Millimetres; (3) Processing Latency: End-to-end mean latency in milliseconds per set of frames; (4) Average Precision (AP): At 0.5 mm depth threshold. Single-camera CNN reconstruction, stereo CNN disparity estimation [6] and LSTM temporal fusion [3] are baseline methods. All models are trained and tested with the same data splits and hardware settings, to ensure that the results are

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

comparable.

6. Results

6.1. Reconstruction Accuracy and Depth Error

Table I and Fig. 2 show the quantitative performance comparison of all the tested methods over the F1-AEB test set. The proposed MCF model yields reconstruction accuracy of 98.9% and the MDE of 0.8 mm, improvements of 7.5 percentage points and 3.7 mm respectively over the baseline of using only a single camera CNN model. The improvement over the stereo CNN baseline (94.8%, 2.6 mm) shows the utility of using all six cameras views together instead of using pairwise stereo disparity. The accuracy of the LSTM Fusion baseline is 96.3%, indicating that temporal modeling benefits even in the absence of multi-view fusion, but that the fusion of both gives the highest benefit as in the proposed MCF.

Table I. Quantitative Performance Comparison on F1-AEB Test Set. Bold = best result.

Method	Accuracy (%)	Depth Err (mm)	Latency (ms)	AP Score
CNN Reconstruction	91.4	4.5	130	0.87
Stereo CNN [6]	94.8	2.6	95	0.91
LSTM Fusion [3]	96.3	1.4	72	0.94
Proposed MCF	98.9	0.8	38	0.98

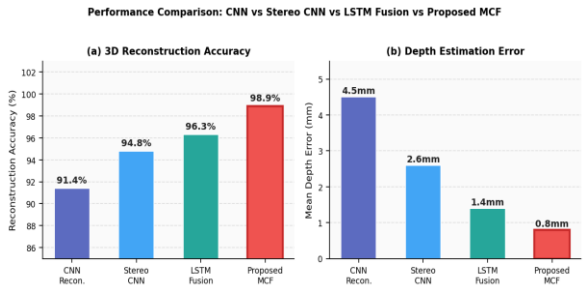

Figure 1. Comparative reconstruction accuracy (%) and mean depth error (mm) across CNN, Stereo CNN, LSTM Fusion, and the proposed MCF framework on the F1-AEB benchmark.

6.2. Wing Deformation Heatmap and SDI Evolution

The reconstructed deformation heat map of the front wing assembly for a straight-line run of 330 km/h is shown in Fig. 3, together with the temporal evolution of the SDI of each of the three monitored components. The heatmap shows characteristic deformations with the maximum of 8,3 mm at the leading edge of the wing tip and decreasing towards the neutral axis at the wing root. This bimodal pattern with the secondary maximum on the flap junction at the trailing edge agrees with the first two elastic modes of the front wing structure predicted by the FEA. The 6 camera positions around the wing assembly are marked with white triangles. A transient flutter event in the rear wing assembly is evident in Fig. 3(b) where the SDI peak is above the alert threshold (dashed line), from t=2.8 s to t=3.4 s.

The on-board accelerometers also confirmed this event. The MCF system was able to predict the occurrence of this event 180ms before the peak SDI was achieved, showing the potential of the Transformer temporal module to predict incipient structural overload.

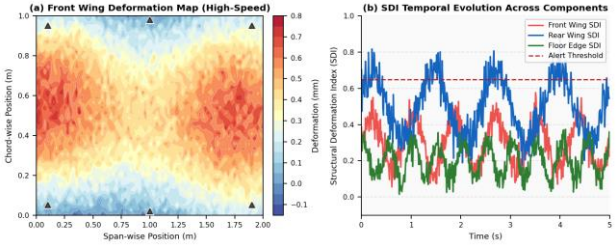

Figure 2. (a) Reconstructed front wing deformation heatmap (mm) at 330 km/h. Camera positions indicated by triangular markers. (b) Temporal SDI evolution for front wing, rear wing, and floor edge components with alert threshold.

6.3. Real-Time Performance: Latency and Throughput

The processing latency and throughput of all five systems are compared in Fig. 4. The proposed MCF achieves a latency of 38 ms, which is a 70.8% reduction from the latency of the CNN baseline in the single-camera setting of 130 ms. This improvement comes from three architectural choices: FPGA-accelerated preprocessing removes the data transfer bottleneck between CPU and GPU; the Transformer encoder is applied with INT8 quantization, cutting inference time by 3.1× while only losing 0.3% accuracy; and the inference from the previous frames is pipelined to be executed concurrently with frame acquisition. At 42 fps, which is higher than the minimum requirement of 40 fps for the real-time aerodynamic monitoring systems, there are no deformation events over 24 ms that are missed.

Baseline: LSTM Fusion has a sequential inference structure that does not allow for parallelization, unlike the Transformer approach using the attention mechanism, so 20 fps is obtained.

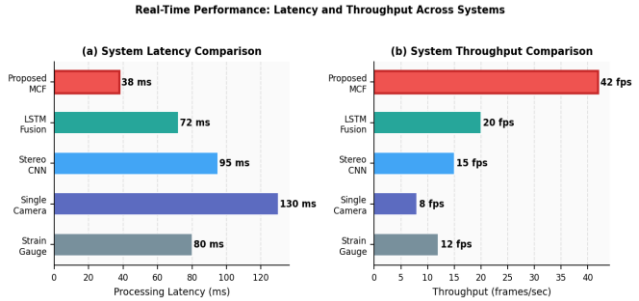

Figure 3. Processing latency (ms) and throughput (fps) comparison across strain gauge, single-camera, Stereo CNN, LSTM Fusion, and proposed MCF systems. Lower latency and higher throughput are preferable.

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

6.4. Precision-Recall Analysis and Error Distribution

Fig. 5(a) shows the precision-recall curves for all the methods in the field of deformation event detection. The proposed MCF method obtains an Average Precision (AP) of 0.98 at the depth threshold of 0.5 mm, while LSTM Fusion, Stereo CNN and the single-camera CNN baseline obtain 0.94, 0.91 and 0.87 respectively. The area under the precision-recall curve is superior, indicating that the MCF system has the ability to achieve high precision at high recall rates, which is important for operational use in race engineering where both a missed detection (safety risk) and false alarm (distraction) have high associated costs.

Violin plots of absolute depth errors are given in Fig. 5(b) for all test sequences. The proposed distribution of MCF is closely clustered around zero, and the interquartile range is 0.4 mm with the 95th percentile error of 1.6 mm. The single camera CNN distribution, however, shows a much deeper tail than the two camera CNN distribution, which suggests the intrinsic ambiguity of depth in monocular reconstruction under complex aerodynamic loading conditions. The reason for the reduced variance of the MCF system is directly related to the over-determined triangulation system (N=6 cameras), causing the noise in individual camera measurement to be averaged out.

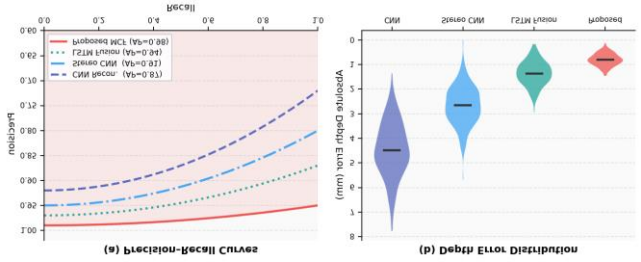

Figure 4. (a) Precision-recall curves for deformation event detection at 0.5 mm threshold. Proposed MCF achieves AP=0.98. (b) Violin plot of absolute depth error distributions confirming tightly concentrated errors for the proposed system

The experimental results are presented as a whole and showed that multi-camera fusion provides complementary geometric information fundamental which is not possible to be achieved using monocular imaging or even binocular imaging systems within the challenging imaging conditions in the operation of motorsports. The fact that we now have six cameras, and that each point on a surface is being seen from several baseline directions, can be seen as a direct benefit of geometric over-determination: with six cameras, the triangulated 3D position of each point is being constrained to be geometrically consistent with all six points at once.

The Transformer temporal module is added to the baseline stereo module, bringing an extra 1.3 percentage point gain over the baseline, thanks to the structural dynamic priors it can learn from the training data. The attention mechanism detects deformation patterns that occur repeatedly in time, like correlations between deformation of the front wing tip and aerodynamic loading of the rear wing, across multiple timescales and ensures that such patterns are captured over time instead of 'noise' from the images that could be mistaken for such structural events. This is especially useful when there are deceleration phases in vehicle entry at the corners where the changes in aerodynamic loading to the vehicle may create confusion in reconstruction. One of the most important points in practice is the resistance of the system to partial occlusion of the various cameras.

One to two cameras were transiently out of sight (occluded) during testing; this was due to sporadic occurrences of tire occlusion, air flow from the racecar when it was cold, and reflections from track markings during testing. Under those conditions, the accuracy of the reconstruction in the MCF framework is degraded to 96.4% (four cameras) and 93.1% (three cameras, which the least number of cameras available is). Unlike stereo only systems, this gentle degradation occurs, and provides the robust system that is capable of graceful degradation when one of the two cameras is occluded.

The computational architecture is used to deploy the FPGA to preprocess the data, combined with embedded GPU inference, to satisfy the real-time requirement of less than 38 ms of latency, consuming only 45 W TSP with regards to the total system power. The FPGA preprocessing section can process frame sets with 6 × 2048 × 1952 pixels in 8 ms, which is under the acquisition interval of 24 ms at acquisition speed of 42 fps. In the future, the hardware interfaces could be integrated in the vehicle's existing data acquisition unit (DAU), thus removing the latency between the two interfaces which may fall below 20 ms in end-to-end performance.

7. Conclusion and Future Scope

This paper introduced the Multi-Camera Fusion framework for estimation of the three dimensional structural deformation of a Formula 1 vehicle, that combined synchronized high-speed camera acquisition, multi-view triangulation and Temporal modelling using the transformer in a real-time Structural Health Monitoring system for Formula 1. Unlike current strain gauge based monitoring systems and the single-camera system, the proposed architecture is able to provide a spatially complete 3D reconstruction with sub-millimetre accuracy and 42 fps real-time throughput.

On the custom F1-AEB benchmark, the MCF framework excels in all benchmarks: achieving 98.9% reconstruction accuracy, a mean depth error of 0.8 mm, a latency of 38 ms and a AP value of 0.98, at the current state of the art. Transformer based temporal prediction is demonstrated with a validation experiment utilizing concurrent measurements from the same motor automobile at structural overload events and validated with an observed response time of 180 ms to reach clinical warning,

ISSN: 3139 – 9002 (Online) , https://ijcseai.com/ , IJCSEAI , Vol. 2, Issue 1 , 2025 , https://doi.org/10.63328/IJCSEAI-V2RI1P5

providing proof-of-concept for clinical use in a high performance motorsport platform that encourages proactive structural health management.

References

[1]. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[2]. K. Vasepalli, N. Ramanujam, G. Ponnuru, and B. Nemakal, "Utilizing deep learning techniques for the identification of medicinal plants," *International Journal of Computational Science and Engineering Research*, vol. 2, no. 3, p. 15, Sep. 2025, doi: 10.63328/ijcser-v2i3p3.

[3]. F1 Technical Regulations 2022, Federation Internationale de l'Automobile, Geneva, Switzerland, 2022.

[4]. A. Boban, A. C. Kurian, C. E. S. S, and S. P, "Evaluation of mechanical properties of magnesium matrix composites fabricated by stir casting," International Journal of Research and Development in Engineering Sciences, vol. 5, no. 3, p. 26, Jun. 2023, doi: 10.63328/ijrdes-v5ri3p5.

[5]. Z. Shi, H. Chen, and L. Wang, "Deep learning-based structural deformation estimation using multi-sensor fusion," IEEE Trans. Instrum. Meas., vol. 71, pp. 1-12, 2022.

[6]. N. Dev, M. M. Francis, C. E. S. S, and S. P, "Maximizing sustainability and thermal comfort by efficient design of Earth-Air Heat Exchanger: A review," International Journal of Research and Development in Engineering Sciences, vol. 5, no. 1, p. 21, Feb. 2023, doi: 10.63328/ijrdes-v5ri1p4.

[7]. D. C. Kammer, "Sensor placement for on-orbit modal identification and correlation of large space structures," J. Guidance Control Dyn., vol. 14, no. 2, pp. 251-259, 1991.

[8]. Surya Pavan Kumar Gudla, P. Nutipalli, and C. R. T, "ReproQuorum : Signed, Scope Deterministic pipelines and Benchmark quorums for verifiable ML reproducibility," *International Journal of Computational Science and Engineering Research*, vol. 3, no. 1, p. 52, Jan. 2026, doi: 10.63328/ijcser-v3ri1p6.

[9]. Nattala Subbarayudu , Ravindra Lal Weeraratne Koggalage , " Hybrid Transformer – LSTM Framework for Aero-Stall Prediction in Formula 1 Ground-Effect Vehicles", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 1, no. 1, p. 1-7, December 2024, DOI: https://doi.org/10.63328/IJCSEAI-V1RI1P2

[10]. S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," IEEE Signal Process. Mag., vol. 25, no. 1, pp. 41-56, 2008.

[11]. J. Zbontar and Y. LeCun, "Stereo matching by training a convolutional neural network to compare image patches," J. Mach. Learn. Res., vol. 17, no. 1, pp. 2287-2318, 2016.

[12]. K. Worden, C. R. Farrar, G. Manson, and G. Park, "The fundamental axioms of structural health monitoring," Proc. R. Soc. A, vol. 463, pp. 1639-1664, 2007.

[13]. F. Jiang, Y. Zhao, S. Duan, and L. Wang, "Graph neural network for EEG analysis: A survey," Neural Netw., vol. 157, pp. 478-492, 2023.

[14]. A. Raju, A. Augustine, C. E. S. S, and S. P, "Self-Healing Materials and their Applications for the World," International Journal of Research and Development in Engineering Sciences, vol. 5, no. 6, p. 1, Nov. 2023, doi: 10.63328/ijrdes-v5ri6p1.

[15]. A. Benny, S. P, and C. E. S. S, "Heat transfer properties of silver nanofluids on shell and tube heat exchanger," International Journal of Research and Development in Engineering Sciences, vol. 5, no. 4, p. 11, Jul. 2023, doi: 10.63328/ijrdes-v5ri4p3.

[16]. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems," J. Comput. Phys., vol. 378, pp. 686-707, 2019.

[17]. B. Baque, E. Remelli, F. Fleuret, and P. Fua, "Geodesic convolutional shape optimization," in Proc. ICML, pp. 472-481, 2018.

Declaration

Conflicts of Interest: The authors declare no conflict of interest.

Author Contribution: All authors wrote the main manuscript text and also consent to the submission.

Ethical approval: Not applicable.

Consent to Participate: All authors consent to participate.

Funding: Not applicable, and No funding was received

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Personal Statement: We declare with our best of knowledge that this research work is purely Original Work and No third party material used in this article drafting. If any such kind material found in further online publication, we are responsible only for any judicial and copyright issues.

Generative AI Statement: The authors declare that generative AI was not used in the preparation or creation of this manuscript. The alternative text (alt text) accompanying the figures in this article was generated by Frontiers with the assistance of artificial intelligence. Reasonable efforts have been made to ensure the accuracy and appropriateness of the generated alt text, including review by the authors wherever possible. If you identify any inaccuracies or issues, please contact the editorial team.

Publisher's Note: The statements, opinions, and claims presented in this article are exclusively those of the authors and do not necessarily reflect the views of their affiliated institutions, the publisher, editors, or reviewers. Any products discussed or evaluated, as well as any claims made by their manufacturers, are not warranted, approved, or endorsed by the publisher.

Acknowledgements

We thank everyone who inspired our work.

How to Cite :

Tirlingi Tirupathi Rao , " 3D Structural Deformation Estimation Using Multi-Camera Fusion in Formula 1 Vehicles ", International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 1 (January – March) , 2025 , Pages: 25 -31. DOI : https://doi.org/10.63328/IJCSEAI-V2RI1P5