World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

A novel design concept of cost-effective permanent rail-track monitoring system

Mustafa M. Amami *

Department of Civil Engineering, Benghazi University, Benghazi, Libya.

World Journal of Advanced Research and Reviews, 2022, 13(03), 451–473

Publication history: Received on 18 February 2022; revised on 20 March 2022; accepted on 22 March 2022

Article DOI: https://doi.org/10.30574/wjarr.2022.13.3.0255

Abstract

This paper presents a novel design concept of low-cost permanent monitoring system for rail-track problems. The introduced design is theoretically overcome the currant in-use designs in terms of providing permanent accurate observations for detecting rail-track irregularities for all railway network during service time using low-cost integrated sensors. The paper investigates the rail-track common irregularities by identifying the issues, presenting challenges, and highlighting the limitations of currently in-use rail-track monitoring techniques. Then, the potential of using cost-effective sensors and intelligent techniques for detecting the rail-track geometric parameters and rail irregularities is theoretically studied. This includes using low-cost GPS, MEMS-INS & high frequency single-point and multi-spot laser distance sensors for detecting rail surface cracks, fractures, cross-level, wrap, gauge and vertical & horizontal rail alignments. Different levels of GPS/MEMS-INS integration are covered, illustrating the advantages of each solution and the optimal utilization in the system. Vision-based monitoring using low-cost HD non-metric digital cameras is also investigated, showing the optimal specifications required for fulfill the system requirements. Data logging and GPS-time-based synchronization method is presented as a part of the design concept of the suggested system. The processing techniques and unit outputs are illustrated in details, showing how each unit works for detecting the rail-track irregularities, providing the geodetic positioning and the vision-based assessing. The main part of the system, which is Automatic Adjusted Wheeled Carrier Frame (AAWCF) is discussed in more details, showing simple and clear different views for all details and integrated sensors. It is recommended to carry out this idea, by building up and evaluating the performance of AAWCF in real railway environments, evaluating the design concept of each unit individually to know the weaknesses of each technique, and then evaluating the whole system in all expected railway cases to find out the advantages and limitations of this design concept.

Keywords: Rail-track monitoring system; Railway's irregularities; Cost-effective; Low-cost GPS/INS; Laser and vision sensors

1 Introduction

Railways are regarded as one of the most important transportation methods and classified as the most cost-effective frequently-used transport mechanism for both passengers and goods carriage over long, as well as, short distances. In terms of reliability and safety, railways are one of the most dependable transport mechanisms, where it is the least affected by usual weather fluctuations that might concern the other transport mediums. Railway services include passenger railways, under & over ground urban metro and goods carriages. Rail transport has many advantages over the other transportation methods, where it is environmental-friendly with low emission of pollutants, low-energy consumption with low-rolling resistance between the wheel and rail, low-cost, well organization, stable routes and schedules, high speed over long distances, suitability for huge and heavy goods, significant capacity, and certain, constant and systematic services. Additionally, rail transport is an enabler of economic progress and a financial support of many countries and governments across the world, especially in crowded areas, in which railways is used extensively

* Corresponding author: Mustafa M. Amami
Department of Civil Engineering, Benghazi University, Benghazi, Libya.

every single day, leading to considerable financial returns. Therefore, railway industries across the world work continuously to make the rail systems more attractive, reasonably-priced and available transport method for freight and passengers. In crowded countries, railways carry the majority of the distant passenger traffic, as well as, the Import and export goods. Also, many heavy Industries, such as iron and steel industry, exclusively depends upon the railways for transferring basic and raw manufacturing materials and distributing products on the markets. Railway sectors help to fight poverty and provide social peace by offering occupation for a significant number of peoples across the world. On the other hand, there are also disadvantages associated with railways, including the heavy initial construction cost and time, particularly when a whole rail network is to be constructed. In addition to the railway's overhead expenses, the cost of regular maintenance is considerable compared to other transport mechanisms. This is, of course, beside the disability of rail transport to provide door-to-door service which leads usually to wasting time and increasing the cost due to the in-between services [1].

Rail transport system is more complicated than others, where it includes different systems that should work together perfectly to provide the required quality level of timekeeping, regularity, dependability, less crowding, safety and comfort. One of the main vital systems in rail transport is the monitoring and maintenance system, where whose neglect leads to horrific accidents and sometimes to the destruction of the whole rail system. To avoid increasing the railways irregularities with time, the rate of regular inspection and monitoring must be increased, which may not be easily achievable with the great expansions of train networks and with the limitations of the current in-use monitoring systems. The main irregularities and problems that must be regularly inspected and monitored are those related to the rails, including rails problems (fracturing and cracking) and the track geometric parameters (gauge, cross-leveling, wrap and the smoothness of vertical and horizontal alignments). Also, to reduce the noise and vibration level suffered throughout rail-used travel, track irregularities must be carefully monitored and kept within the allowable specification ranges. The quality of rail-track geometry has a critical effect on the dynamic performance of the vehicle, reducing the vertically and laterally vibrations and the smoothness of the high-speed rail-track is a main factor for providing the required level of safety and passenger comfort. The main reasons behind degrading the track condition and causing track irregularities are the everyday routing of heavy trains (wheel-rail interaction force) and the settlement of the track foundation, which lead to drifting or moving in the rails away from its designed geometric position. Rail track irregularities are the common causes behind safety problems, accidents and additional track deterioration. For-example, gauge irregularity leads to reducing the rail and train-wheels service life causing running accidents when falling the train-wheels off rail, especially in high speeds. Consequently, it is essential to employ good strategies for the maintenance of railway networks to avoid a disruption to services and ensure the safety of the system. Also, regular inspection and precise measurements must be carried out to detect the problems in initial stages and tack the early suitable reactions [2]. The main parameters that can define the rail-track irregularities are rail fractures and cracks, gauge, cross level, twist and the smoothness level of the vertical and horizontal track. Rail fracture tends to state as a small deep crack, then becomes a visual surface crack and the rate of moving between these three steps are unexpected. Limitations in the manufacturing material is the main reason behind the rail fractures and the other important reason is the unintentional crack formation during the constructional and maintenance operations. Accordingly, rails must have a high resistance to crack initiation and crack propagation and rail-cracks must be under nearly-permanent inspection. Track gauge is the spacing between the inner faces of the rail, which has to stay within specific ranges to ensure that the train-wheels do not slip off the track or fall inside it, causing accidents. Cross-level measures the difference in height between the top surface of the two rails at the same location across. Zero cross-level indicates that both rails at the same level and reverse cross-level points to the case when the external rail of curved track has lower height than the internal rail. The track cross-level takes into consideration when specifying the train design speed limits. In straight track, it is preferred to use zero cross-level and in curved track, super-elevation between the two rail sides is used for higher speeds. Warp is the difference between two cross-levels in a certain length of track. Warp points to the changing rate (value & sign) of cross-levels along a specific distance. Therefore, even though the cross-levels are in the allowed ranges, wrap parameter can be out of the permissible limits. The danger of this parameter comes from facing sequential wrap changes from a positive to a negative by changing the sign of cross-level in short distances. High wrap rate results in abnormal vehicle movements left and right and in high speeds, train wheels might derail, causing accidents. As for the tack alignment parameter, it can be divided into horizontal and vertical alignments. Alignment parameter points to the applicability of the in-use rail track to the design track, in both vertical and horizontal dimensions. However, it is unexpected for the whole or even considerable parts of vertical or horizontal track alignments to move or change. Therefore, studying the vibration smoothness level in vertical and cross rail directions gives clear indications about the rail alignment parameter [3]. Figure (1) illustrates the main rail-track irregularities.

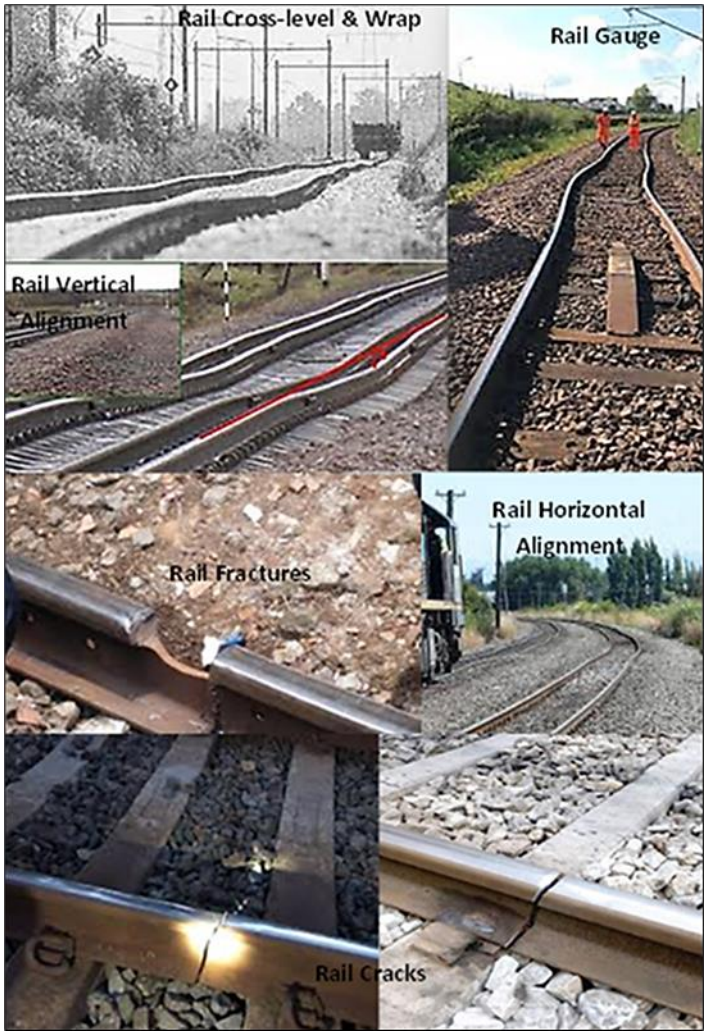

Figure 1 Main rail-track irregularities

Visual walking inspection is one of the most common track monitoring methods, which can be used for rail fractures, surface noteworthy cracks, significant changes in gauge, cross-level, and wrap. However, this method is costly, time-consuming, subjective, and limited in terms of small rail-track irregularities detection. Precise rail-track monitoring can also be applied for measuring rail-geometry parameters using typical surveying equipment, such as robotic total stations, digital levels and smart stations, which tend to be fixed on railways manual trolley. Precise surveying equipment are frequently designed to be used as stationary and away from vibrations, thus such monitoring system is not suitable to be used as a rail-contact, especially in high speed. The surveying equipment-based rail monitoring is expensive, time-consuming, limited coverage per time, and traffic interrupted. Precise and high frequency track geometry advanced surveying sensors, such as Global Navigation Satellite Systems (GNSS) and Inertial Measurements Unite (IMU) can also be used for measuring the rail-track geometric parameters without interrupting the railway traffic. The integration of differential carrier phase GNSS/Inertial Navigation System (INS) can provide high-rate absolute precise positioning reaching a few millimeters level. This advanced geodetic surveying equipment tend to be fixed with other non-rail-contact techniques, such as computer vision, camera, laser scanners, and ultrasonic devices on a track recording vehicle. This vehicle travels from time to time over the rail network to detect the rail irregularities and measures the track geometric parameters, including longitudinal level of both rails, horizontal alignment of both rails, track gauge, cross level, twist and gradient. However, these vehicles are high expensive and not always available for covering the whole railways network. Laser scanning is high-cost technique may involve a high maintenance cost. Vision-based technique is cost-effective, but complicated data processing and filtering are required, providing undependable results. Moreover, the load and speed of these vehicles may not be able to simulate the actual rail condition under the real train load and speed. This difference in travelling speed between these vehicles and trains on the network leads sometimes to disturbing systematic services. Train-by monitoring is a monitoring technique in which sensors are fixed on the train and the recorded data are used for monitoring the condition of the rail-track. This monitoring technique can provide more reliable results, where it is applied under the actual train load and speed. Also,

it works with the service, thus no need for stopping or disrupting the service. With monitoring system in each train, permanent rail irregularities monitoring can be applied, which helps to accelerate and facilitate maintenance operation. However, it is extremely expensive to provide each train with advanced precise geodetic surveying and optic sensors. On the other hand, train-by monitoring using imprecise sensors may not be able to achieve the required purpose, resulting big amount of unhelpful and imprecise observations and measurements. Figure (2) illustrates a number of common rail-track monitoring systems.

Figure 2 A number of common rail-track monitoring systems

Aim & Objectives

The aim of this paper is to introduce a novel design concept of cost-effective permanent monitoring system for rail-track irregularities. The suggested design is expected to overcome the previous designs in terms of providing every-day precise measurements and irregularities detection for all in-use rails via low-cost integrated sensors. To meet the above aim, the following objectives have been set:

- Investigate the rail-track main irregularities by identifying the issues, presenting challenges, and highlighting the limitations of currently in-use rail-track monitoring techniques.
- Studying theoretically the potential of using low-cost sensors and intelligent techniques for detecting the rail-track geometric parameters and rail irregularities, including the followings:
  - Low-cost high frequency multi-spot Laser distance sensors for rail surface cracks and fractures detection.
  - Low-cost high frequency single-point laser distance sensors for rail cross-level measurements.

- Low-cost high frequency single-point laser distance sensors for rail gauge measurements.
- The integration of Stand-Alone GPS (SA-GPS), Carrier Phase delta Positioning (CP-DP) and Micro-Electro-Mechanical Systems (MEMS)-INS to provide Enhanced Absolute Navigation Solution (EA-NS) for wrap measurements.
- Low-cost HD non-metric digital cameras with high frames/second video recording for illustrating photos of irregularities detected by the mentioned techniques.
- Raw MEMS-INS of rail-track horizontal and vertical alignments irregularities.
- Logging all the recorded data and synchronizing in GPS time for easy matching and fast processing.

- Designing the main carrier frame of the system and assembling the components and units together.

2 Design Concept, Workflow & System Units Description

The introduced rail monitoring system in this paper consists of six units. Each unite works for specific target with individual design concept. In this part of the paper, each unite will be described, illustrating aim, design concept and workflow. Additionally, the used sensors will be described individually, before showing the system units assembly and the general workflow.

2.1 GPS/INS Navigation Unit (NU)

2.1.1 NU: Description & Aims

This unit consists of low-cost single frequency GPS receiver and low-cost MEMS based IMU, which are integrated together for providing EA-NS and precise Relative Navigation Solution (RNS). SA-GPS using low-cost receivers can provide a few meters accuracy in open-sky and tens of meters in narrow-sky and high-multipath environments. Therefore, an integration between SA-GPS and CP-DP will be carried out for providing Enhanced Absolute Stand-Alone Positioning (EA-SAP), which will be used in the introduced system for determining the absolute position of each rail irregularity detected by the other system units. EA-SAP will be used for specifying the maintenance area, thus, constant plane accuracy of one to two meters will be acceptable [4, 5]. INS is a navigation aid system that uses a computer and IMU. IMU includes motion sensors and rotation sensors to continuously calculate the position, orientation, and velocity relative to a known starting point. The accuracy of INS depends mainly on the initial state accuracy, inertial sensor quality, such as accelerometers and gyros, stability and reliability of inertial sensors, and the correction models used. INS is a self-contained navigation passive, worldwide, and easy to operate and independent system. In addition, INS can be used in all weather and attitude. However, INS should be provided with initial position and rotations for achieving absolute orientation and when it has been initialized, no more help is needed for navigation. INS has become a necessary request in a great deal of application, such as the aircraft navigation, submarines and ships, tactical and strategic missiles and space craft. INS suffers from different type of errors, some of them can be bounded, such as those of acceleration, velocity and initial tilt, and others hard to be bounded including azimuth misalign, leveling gyro drift and azimuth gyro drift. Small errors in the acceleration and angular velocity measurement are cumulated with time to be great errors in position where each position is calculated from the previous calculated position. Therefore, the position must be regularly updated from another navigation system and the updating level depends on the quality of the sensors used and the accuracy required from the system. Current developments in the MEMS construction of devices lead to manufacturing undersized and light IMUs. MEMS-INS has become commonly used due to its significant low-cost, tiny size and the spinning-wheel less [6, 7].

The integration of GPS/INS can help to overcome the limitations of the two systems providing integrated system better than either on a stand-alone basis. Loosely coupled integration is the typical integration of GPS/INS, in which GPS is run autonomously and, at the same time, GPS/INS integrated solution is enabled. When GPS is hidden, INS stand-alone solution smoothed by Kalman Filter (KF) is used to fill in the gap. However, the integration of low-cost GPS with low-cost INS is not useful, where the precision of gyro measurements is degraded by the fluctuated quality of GPS, no improvements to the quality of individual GPS positioning can be seen, and the obtained navigation solution degrades very quickly when GPS signals are hidden [6]. Loosely coupled integration of CP-DP with MEMS-INS will be used in this system instead of conventional integration of low-cost GPS/INS sensors, providing RNS. This integration navigation solution takes the advantages of the two systems, providing precise, high frequency, and reliable low-cost relative navigation solution even in week GPS areas and high multipath environments. This can be attributed to utilizing carrier phase observables in CP-DP, which are precise and extremely less affected by multipath than code observables. RNS is

significantly capable to deal with cycle slips, but it is still limited in GPS-off areas, where the performance of MEMS-INS deteriorates without GPS supporting. In this paper, RNS will be used in the railway horizontal and vertical alignment monitoring by studying the vertical and cross vibrations behavior. Additionally, it will be used in determining the precise distances between each two measured cross-levels, which will be used for determining the rail-track geometric wrap parameter. The third integration will be used in this system is uncoupled integration between the EA-SAP and RNS, providing EA-NS. In this integration, RNS is reset at regular intervals of time using EA-SAP. EA-NS will just be used in the system when GPS signals are hidden, where the system will depend on RNS based on the last absolute GPS position to determine the position of the detected irregularities [5]. The following figure shown the main workflow of the NU [5].

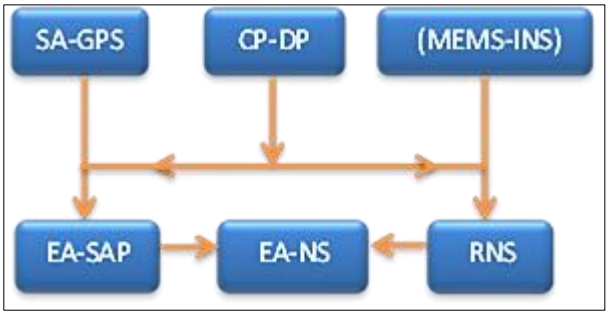

Figure 3 NU main workflow of the

The aims of NU can be summarized in the following table, illustrating the function of each positioning or navigation solution.

Table 1 Aims of NU

Positioning or navigation solution	Function	Rail-track Irregularities targeted
EA-SAP	Determining the absolute geodetic coordinates of any detected rail-track irregularities in expected plane accuracy of 1 to 2 m	Fracture, Crack, Cross-level, Wrap, Gauge, Vertical & Horizontal Alignment
RNS	Determining the precise relative distances between every two recorded cross-level measurements in a few millimetres' accuracy Note: as 2 IMU sensors are used one over each rail, average of the two RNS solutions can be obtained.	Wrap measurements
MEMS-INS	Determining high frequency vibration measurements in vertical and cross directions Note: as 2 IMU sensors are used, 2 MEMS-INS solutions can be obtained, one for each rail.	Vertical and horizontal rail-track alignment irregularities
EA-NS	Determining the absolute geodetic coordinates of any detected rail-track irregularities when GPS is off in tunnels, dins forests, and underground rail-tracks Note: the averaged RNS solutions is used in this integration	Fracture, Crack, Cross-level, Wrap, Gauge, Vertical & Horizontal Alignment

2.1.2 The Integration of SA-GPS, CP-DP & MEMS-INS [8, 9]

- CP-DP derivative

The GPS carrier phase observable in meters can be written as:

$$CP_{(s,r)(k)} = p_{(s,r)(k)} + c(dT_{(s)(k)} - dt_{(r)(k)}) + dion_{(s,r)(k)} + dtrop_{(s,r)(k)} + dor_{(s)(k)} + E_{(s,r)(k)} + L*N$$

Where

$CP_{(s,r)(k)}$	carrier phase observation (m);
$p_{(s,r)(k)}$	true range between receiver *(r)* and satellite *(s)* at epoch *(k);*
c	speed of light;
$dT_{(s)(k)}$	clock error of satellite *(s)* at epoch *(k);*
$dt_{(r)(k)}$	clock error of receiver *(r)* at epoch *(k);*
$dion_{(s,r)(k)}$	ionospheric delay error between receiver *(r)* and satellite *(s)* at epoch *(k)* (m);
$dtrop_{(s,r)(k)}$	tropospheric delay error between receiver *(r)* and satellite *(s)* at epoch *(k)* (m);
$dor_{(s)(k)}$	orbit error of satellite *(s)* at epoch *(k)* (m);
$E_{(s,r)(k)}$	measurement noise including multipath between receiver *(r)* and satellite *(s)* at epoch *(k);*
L	carrier wavelength (m);
N	unknown integer ambiguity (cycle).

The true range between receiver *(r)* and satellite *(s)* at epoch *(k)* can be written as:

$$p_{(s,r)(k)} = ((X_{(s)} - X_{(r)})^2 + ((Y_{(s)} - Y_{(r)}) + ((Z_{(s)} - Z_{(r)}))^{0.5}$$

Where

X, Y and Z	satellite and receiver Cartesian coordinates.

In CP-DP, the first single differencing is formed between one receiver (r), one satellite (s) and two adjacent epochs *((k)* & *($k+1$))*. The single differencing equation can be written as:

$$SD_{(s)(r)(k+1,k)} = CP_{(s,r)(k+1)} - CP_{(s,r)(k)} =$$

$$[\, p_{(s,r)(k+1)} + c(dT_{(s)(k+1)} - dt_{(r)(k+1)}) + dion_{(s,r)(k+1)} + dtrop_{(s,r)(k+1)} + dor_{(s)(k+1)} + E_{(s)(r)(k+1)} + L\,N \,] -$$

$$[\, p_{(s,r)(k)} + c(dT_{(s)(k)} - dt_{(r)(k)}) + dion_{(s,r)(k)} + dtrop_{(s,r)(k)} + dor_{(s)(k)} + E_{(s)(r)(k)} + L\,N \,]$$

Where

$SD_{(s)(r)(k+1,k)}$	single differencing between receiver (r), one satellite (s) and two adjacent epochs ((k) & ($k+1$)).

From the single differences, the double difference ambiguity is removed as long as the integer ambiguity remains constant and the receiver keeps lock the satellite signal. Satellite clock error is reduced based on the stability of the satellite clock over transmission times. Satellite orbit errors are reduced significantly based on the high correlation between the satellite orbit errors over time. Ionosphere and troposphere errors are reduced to the change in delay across the interval. Multipath remains and can be reduced based on the multipath correlation over time. However, receiver clock error is doubled and receiver measurements noise increases. The final formula of single differences equation can be written as:

$$SD_{(s)(r)(k+1,k)} = p_{(s)(r)(k+1,k)} - c\,dt_{(r)(k+1,k)} + E_{(s)(r)(k+1,k)}$$

The second differencing in CP-DP is carried out between two single differences, similar to that in equation (3), cross two satellites *(s)* and *(j)*. This can be written as:

$$DD_{(s,j)(r)(k+1,k)} = SD_{(s)(r)(k+1,k)} - SD_{(j)(r)(k+1,k)} =$$

$$[\, p_{(s)(r)(k+1,k)} - c\,dt_{(r)(k+1,k)} + E_{(s)(r)(k+1,k)} \,] - [\, p_{(j)(r)(k+1,k)} - c\,dt_{(r)(k+1,k)} + E_{(j)(r)(k+1,k)} \,]$$

Where

$DD_{(s,j)(r)(k+1,k)}$	double differences between one receiver *(r)*, two epochs *(k+1)* & *(k)*, and two satellites *(s)* & *(j)*.

Receiver clock error is cancelled out in the double differences. This is extremely important for getting accurate results where the oscillators in low-cost receivers vary in frequency with temperature and pressure making the receiver clock unreliable. The final formula of double differences equation can be written as:

$$DD_{(s,j)(r)(k+1,k)} = p_{(s,j)(r)(k+1,k)} + E_{(s,j)(r)(k+1,k)}$$

The only unknowns in this equation are the receiver Cartesian coordinates in the two epochs *(k)* & *(k+1).* The changes in the receiver positions between the two epochs can be determined by fixing the coordinated of the receiver at epoch *(k)* (as zeros for example) and solving for the receiver coordinated at epoch *(k+1).* To determine the relative position of the receiver at epoch *(k + 1)* from *(k),* the double differences equation should be written as:

$$b = A\,X + v$$

Where

b the measurement victor with a size of (number of epochs -1, 1);
X the parameter victor with a size of (number of epochs * 3, 1) which include the change in Cartesian position across the interval;
A matrix with a size of (number of epoch -1, number of epochs * 3) which relates the parameters to the states;
v a vector of random noise with a size of (number of epochs -1, 1).

This equation can then be solved by least squares as:

$$X = (A^T w\, A)^{-1} A^T w\, b$$

Where

w the weight matrix with a size of (number of epochs -1, number of epochs -1) which is based on the average satellite residuals obtained from the stand-alone code positioning calculations.

- The Integration of SA-GPS/CP-DP [5]

Kalman filter (KF) filters measurements based on the expected changes of these measurements over time and the statistical properties of the system measurement errors. The filter determines the minimum error estimate of the states based on the linear relation between the measurements and these states. The states are composed of number of values that adequate to define the system motion. KF consists of measurement model and dynamic model which will be illustrated here to define the basic elements as related to the integration of CP-DP with code positioning. As for the equations of propagation and update steps, they are well documented in different sources and there is no point for repeating here. The measurement model defines the mathematical linear relationship between the measurements and the filter states. The discrete measurement model at the epoch (k) can be defied as:

$$Z_{(k)} = H\,U_{(k)} + v_{(k)}$$

Where

$Z_{(k)}$ the vector of measurements at the epoch *(k);*
$U_{(k)}$ the system state vector at the epoch *(k);*
H the design matrix measurement which defines the linear relationship between the states and the measurements;
$v_{(k)}$ the measurement residual vector.

The dynamic model describes the change in the state vector parameters over time. The discrete dynamic Model between epochs *(k+1)* & *(k)* can be given as:

$$U_{(k+1)} = M\,U_{(k)} + W_{(k)}$$

where

M the state transition matrix that defines the relation between state vector parameters over time;

$W_{(k)}$ the system noise is approximated based on the is the sampling interval, the spectral density matrix and standard deviations of the driving noise of the system.

In the case of integrating code positioning with CP-DP, the vector of measurements includes six observations; three observations from the absolute code positioning of epoch (k) ($X_{(K)}$, $Y_{(K)}$, $Z_{(K)}$) and three values describing the 3D changes in positioning between epochs *(k) & (k+1)* ($dX_{(K, K+1)}$, $dY_{(K, K+1)}$, $dZ_{(K, K+1)}$). As there are no unknowns, the $Z_{(k)}$ and $U_{(k)}$ are the same vector, H is a (6*6) unit matrix and $v_{(k)}$ equal zero.

$$Z_{(k)} = U_{(k)} = | X_{(K)}\ Y_{(K)}\ Z_{(K)}\ dX_{(K, K+1)}\ dY_{(K, K+1)}\ dZ_{(K, K+1)} |^{T}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

As for the dynamic model, the system state vector at the epoch *(k+1)* includes three observations from the absolute code positioning for this epoch ($X_{(K+1)}$, $Y_{(K+1)}$, $Z_{(K+1)}$) and three values describing the 3D changes in positioning between this epoch and the previous one ($dX_{(K, K+1)}$, $dY_{(K, K+1)}$, $dZ_{(K, K+1)}$). The system state vector can be defined as:

$$U_{(k+1)} = | X_{(K+1)}\ Y_{(K+1)}\ Z_{(K+1)}\ dX_{(K, K+1)}\ dY_{(K, K+1)}\ dZ_{(K, K+1)} |^{T}$$

The smoothed code positioning in epoch *k+1* can be defined as:

$$X_{(K+1)} = X_{(K)} + dX_{(K, K+1)}$$
$$Y_{(K+1)} = Y_{(K)} + dY_{(K, K+1)}$$
$$Z_{(K+1)} = Z_{(K)} + dZ_{(K, K+1)}$$

To relate the state vectors in epoch *k+1* with that in epoch k, the state transition matrix *(M)* should be written as:

$$M = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The standard deviations for the carrier phase measurements used in the system noise vector *W(k)* in the dynamic model is computed using the signal to noise ratio obtained directly from the receiver RXMRAW message. The standard deviation of carrier phase measurements tends to round about millimeters in multipath-free environment and reaches the quarter of the GPS wavelength (5 cm) as a maximum value. This is not the case with stand-alone code measurements where the precision level fluctuates between a few and tens of meters. In this paper, the differences between the computed smoothed position at epoch *(k+1)* and the code position at the same epoch will be used as an indication on the quality of the code measurements at the next epoch. As the GPS carrier phase measurements are a part of the integration solution, dealing with cycle slips is essential where the connection between satellite and receiver tends to be lost, especially in urban areas. In this case, a random integer number of cycles is added to the carrier phase measurements in the GPS data file. Cycle slips can be detected by comparing the differences between the code and carrier phase measurements of two adjacent epochs which is the method used in this paper. Code positioning provides continuous solution as long as four satellites can be detected and it is not necessary for the same satellites to be detected for adjacent epochs, where code positioning in each moment is independent solution. When adequate number of satellites is available, the carrier measurement including cycle slip is removed for more precise solution. As the receiver clock error is removed with CP-DP, the minimum number of satellites required to get the relative positioning is 3. Based on that, if the number of free-cycle slips phase measurements is less than 3, pseudo-range measurements have to be used for fixing the gab in phase measurements to get the relative solution. With just 3 satellites or less, evaluating the quality of code measurements may not be easy as no solution can be obtained. In this case, the integration between CP-DP/MEMS-INS can be useful for fixing cycle slips, and INS measurements should be provided with bigger weights for more precise solution. Satellite residuals should be continually investigated to remove the outliers using Data Snooping Method.

- The Integration of CP-DP/MEMS-INS [5]

The KF vector of measurements in the case of integrating CP-DP/MEMS-INS includes six observations; three observations from the relative or absolute positioning of epoch *(k+1):*

$$(X_{CP\text{-}DP(K+1)}, Y_{CP\text{-}DP(K+1)}, Z_{CP\text{-}DP\ (K+1)})$$

The three values describing the 3-dimensional changes in positioning between epochs (k+1, k+2) determined by MEMS-INS:

$$(\Delta X_{INS\ (K+1,\ K+2)}, \Delta Y_{INS\ (K+1,\ K+2)}, \Delta Z_{INS\ (K+1,\ K+2)})$$

As there are no unknowns, the *Z(k)* and *U(k)* are the same vector, *H* is a (6*6) unit matrix, and v(k) equals zero as illustrated in equations (15) & (16).

$$Z_{(k+1)} = U_{(k+1)} = |X_{CP\text{-}DP\ (K+1)}\ \ Y_{CP\text{-}DP\ (K+1)}\ \ Z_{CP\text{-}DP\ (K+1)}\ \Delta X_{INS\ (K+1,\ K+2)}\ \Delta Y_{INS\ (K+1,\ K+2)}\ \Delta Z_{INS\ (K+1,\ K+2)}|^{T}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Where

Z(k) vector of measurements at epoch (k);
U(k) system state vector at epoch (k);
H design matrix measurement which defines the linear relationship between the states and the measurements;
v(k) measurement residual vector;

As for the dynamic model, the system state vector at the epoch (k+2) includes three observations from the absolute code positioning for this epoch:

$$(X_{CP\text{-}DP\ (K+2)}, Y_{CP\text{-}DP\ (K+2)}, Z_{CP\text{-}DP\ (K+2)})$$

And the three values describing the 3D changes in positioning between this epoch and the previous epoch and the system state vector can be defined as can be described as:

$$(\Delta X_{INS\ (K+1,\ K+2)}, \Delta Y_{INS\ (K+1,\ K+2)}, \Delta Z_{INS\ (K+1,\ K+2)})$$

$$U_{(k+1)} = |X_{CP\text{-}DP\ (K+2)},\ \ Y_{CP\text{-}DP\ (K+2)},\ \ Z_{CP\text{-}DP\ (K+2)},\ \ \Delta X_{INS\ (K+1,\ K+2)}, \Delta Y_{INS\ (K+1,\ K+2)}, \Delta Z_{INS\ (K+1,\ K+2)}|^{T}$$

The smoothed integrated positioning in epoch *(k+2)* can be defined as:

$$(X, Y, Z)_{(K+2)} = (X, Y, Z)_{CP\text{-}DP\ (K+2)} + (\Delta X, \Delta Y, \Delta Z)_{INS\ (K+1,\ K+2)}$$

To relate the state vectors in epoch *(k+1)* with that in epoch *(k)*, the state transition matrix *(M)* takes the size of (6 *6) and can be written as:

$$M = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Where

M state transition matrix that defines the relation between state vector parameters over time;

W(k) system noise is approximated based on the sampling interval, the spectral density matrix and standard deviations of the driving noise of the system.

The standard deviations of CP-DP measurements used in the system noise vector W(k)in the dynamic model can be computed using signal to noise ratio obtained directly from GPS receiver "RXMRAW" message. The standard deviation of CP-DP measurements rounds about millimeters in multipath-free environment and in the case of multipath environments, the quality of GPS carrier phase measurements reaches the quarter of GPS wavelength (5 cm) as a maximum value. As CP-DP technique depends on the differencing between epochs, the quality of measurements might be increased or decreased based on the correlation between the directions of the reflected signals. This means that the quality of the navigation solution might be within millimeters level, and can reaches nearly one decimeter in the worst cases. This is not the case with MEMS-INS measurements, where the precision level is nearly a few decimeters. LC integration of CP-DP/MEMS-INS has been implemented by the author in Matlab. Note: as 2 IMU sensors are used in the system, 2 RNS solutions can be obtained, and the averaged solution is used for more reliable results.

- Uncoupled Integration of EA-SAP/ RNS [5]

UC integration is the simplest level of integration as the averaged RNS is reset at regular intervals of time using EA-SAP, providing EA-NS, which can be obtained using the following formulas:

$$(X, Y, Z)_{EA\text{-}NS(k+2)} = (X, Y, Z)_{EA\text{-}SAP(k+1)} + (\Delta X, \Delta Y, \Delta Z)_{RNS(k+1,\, k+2)}$$

Where

$X, Y, Z_{EA\text{-}NS(k+2)}$ X, Y, Z coordinates of EA-NS at epoch $(k+2)$;
$X, Y, Z_{EA\text{-}SAP}$ X, Y, Z coordinates of EA-SAP at epoch $(k+1)$;
$\Delta X, \Delta Y, \Delta Z_{RNS\,(k+1,\, k+2)}$ RNS between epoch $(k+1)$ and $(k+2)$.

As mentioned above, EA-NS will be used in GPS-off areas, such as tunnels, dins forests, and underground rail-tracks. When GPS signals are hidden, the quality of this positioning solution decreases rapidly as the navigation solution will depend completely on the MEMS-INS solution. However, such GPS-off areas are inconsiderable compared to the whole rail-network, and even though the uncoupled navigation solution cannot provide enough accurate positioning for the detected rail irregularities, the view inspection operation will be confined in limited areas across the whole rail network.

2.1.3 UN: Sensors Description

- u-blox EVK-6H GPS Receiver

u-blox GPS receiver is a L1 frequency (C/A Code) GPS receiver with estimated accuracy (Selective Availability off) of 2.5 and 5 m in plan and height, respectively. This small and light receiver, with size of 25.4 mm x 25.4 mm x 3 mm and weight of just 3 grams, has several advantages, such as excellent navigation accuracy even at low signal level, powerful multipath detection and removal, fast time to first fix, high acquisition and tracking sensitivity, ultra-low power consumption and industrial operating temperature range between -40 and 85°C. This is in addition to its ability to be supported by DGPS and SBAS networks, such as WAAS and EGNOS providing comparable accuracy level. This receiver performs the whole GPS signal processing in one receiver with Patch antenna providing uncomplicated and effortless integration with short time, low-cost and minimum design risks. However, the receiver can be provided with other types of active antennas, such as Helix antenna or choke ring antenna which should be powered by an external power supply. Antenna splitter can be used to connect the antenna with external battery from one side and the antenna with the receiver from another side [9].

The design concept of the introduced system in this paper uses one U-blox EVK-6H GPS receiver with patch antenna, which is recommended to be fixed on an open-sky place on the train roof. Also, it is recommended for the patch antenna to be fixed on the center of the train carriage, where the roof metal surface works as a chock rings frame, reducing the effect of the ground reflected multipath GPS signals on the positioning quality [10]. This receiver can provide 1Hz frequency data, which is linearly interpolated as 1 KHz in the Data Logging & Synchronizing Unit to match the trains speed. This is important as 1Hz GPS positioning can just determine geodetic coordinates for nearly 45.5 m with train travels in 160 Km/h speed and this is reduced to less than 5 cm with 1 KHz data frequency. Linear interpolation is sufficient where horizontal changes in the rail-track alignment tend to be limited and the curves are designed with big radii.

- IMU Micro-strain 3DM-GX3®-25

The 3DM-GX3®-25 is a tiny attitude heading reference system, using MEMS technology. It consists of temperature sensors, tri-axial accelerometers, gyros and magnetometers, and an on-board processor running a sensor fusion algorithm to provide both orientation and inertial measurements. This sensor is able to provide inertial measurements, such as acceleration, angular rate with magnetic field, and delta-Theta and delta-Velocity vectors. This is in addition to the orientation estimates including Euler angles, rotation matrix and quaternion. All these outputs are completely temperature compensated and precisely aligned to an orthogonal coordinate system. According to the manufacturer, gyro drift in this sensor can be eliminated using magnetic North and Earth's gravity and compensating for gyro bias. Also, this sensor is able to provide: precise attitude estimations, high-speed sample rate and high performance under vibration and dynamic conditions. Moreover, the sensor can be used in wide range of applications, such as inertial aiding of GPS, UAV, platform stabilization, antenna and camera pointing. The main advantages of this sensor can be summarized as following: Attitude heading range 360° about all 3 axes and gyroscope range ±300°/sec. Static accuracy ±0.5° pitch, roll, heading typical for static test conditions. Dynamic accuracy ±2.0° pitch, roll, heading for dynamic (cyclic) test conditions and for arbitrary angles. Drift eliminated by complimentary filter architecture. Repeatability 0.2°, resolution <0.1°and data output rate up to 1 KHz [9]. Two Micro-strain 3DM-GX3®-25 IMU are used in the suggested rail monitoring system, which are fixed on the system carrier frame (see section (3.6)) in certain places that examine the highest vibrations on the train and away from the vibration dampers to reflect the real movements of the vehicle on the railways. These sensors with 1 KHz data rate can record reading every 5 cm when train travels on up to 160km/h, which is more than sufficient for studying the smoothness of the horizontal and vertical rail alignments. As just the sudden changes in the vibration behavior are required for evaluating the alignment irregularities, the well-known smooth gyro drifting with time, especially in heading can be neglected.

2.2 Vision Recording Unit (VRU)

2.2.1 VRU: Description & Aim

VRU consists of two off-the shelf non-metric cameras, one for each rail, with capability of providing full-HD 120 frames/second video. The cameras are fixed perpendicularly facing the rail top surface at a certain height, which depends on the camera focal length, image resolution, pixel size, frame rate, and the maximum train speed. Figure (9) shows how the cameras are suggested to be fixed. For example, with 120 frames/second video and train speed of 160 km/h, one frame can be taken every nearly 0.5 m. knowing the camera focal length, pixel size and image resolution, the camera height from the rail surface can be determined using the basics of photogrammetry. The recommended camera fixing height is directly proportional to focal length and train speed, and inversely proportional to frame rate, pixel size and image resolution. To guarantee covering the whole rail without gapping, cameras are recommended to be fixed on double the calculated height after taking into account all factors mentioned above. The cameras are connected to the Data Logging & Synchronizing Unit (DLSU) using wires for recording video frames in GPS time and providing permanent power source. The recorded videos are processed in the system Data Processing Unit and automatically extracted to single geo-referenced frames. These photos are used as additional source for visually assessing the detected rail irregularities and taking the final decision. Figure (4) shows an example of the recommended camera and how to be fixed in rail running and cross directions, where H is the camera-to-rail height, and D1 & D2 are the camera ground coverage in cross and running directions, respectively.

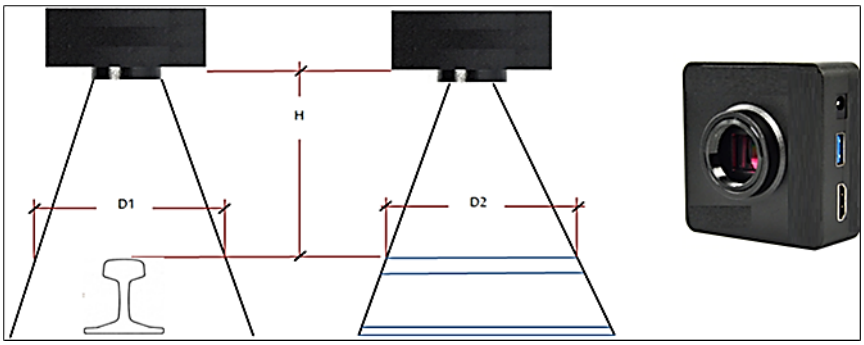

Figure 4 VRU. Right: Camera used. Left: cameras fixing in rail running and cross directions

Cameras can also be fixed obliquely in running and cross directions, providing bigger ground footprint and showing more details of the rail sides. This also means cameras with less frame per second rate can be used, where the ground foot print becomes wider with oblique images. Figure (5) illustrates the advantages of using oblique images.

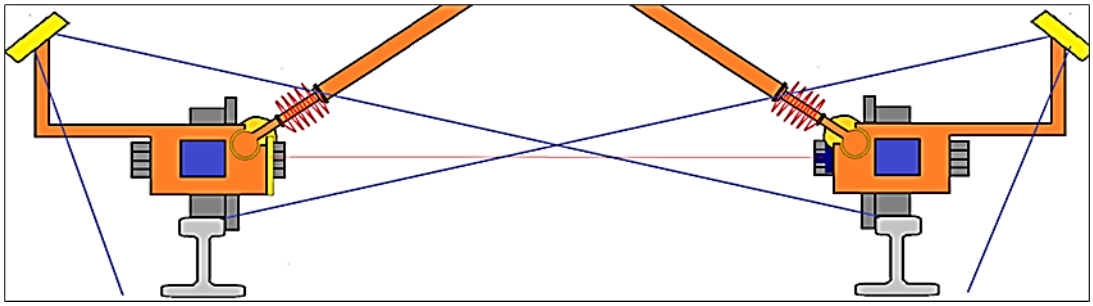

Figure 5 Oblique images for rail sides coverage

However, providing clear high-definition undistorted photos in high speed It needs to know the factors that influence the image blurring and use the right adjustment. The most important factors that affect the image quality in high speeds and vibrations are ISO sensitivity, shutter speed and aperture size. In high speeds and vibration levels, high shutter speed should be used for good image quality. However, increasing shutter speed means less light can reach the camera CCD, giving darker images. To provide the required amount of light with high shutter speed, the camera aperture size must be increased. But again, this will affect the depth of field, which is the distance between the nearest and farthest items in a picture that show acceptably sharp in the image. ISO sensitivity, which is an indicator of the CCD light sensitivity, can also be decreased to overcome the need of lighter when increasing shutter speed. Decreasing ISO sensitivity makes the CCD more sensitive to light which means small amount of light will be enough for processing the image. However, this can affect the image quality giving bigger grain images. In general, it is important for all these parameters to be balanced for VRU to be useful. Image pyramid and integral images can be utilizing for reducing the size of stored images and speeding up extracting and searching techniques [11, 12].

2.2.2 VRU: Sensors Description

The two cameras used in the system design concept are locally modified cameras that are enhanced for industrial and researching purposes to provide 48 megapixels, 4K ultra-HD, 120 frames per second videos and to have high shock and vibration resistant. The cameras are provided with 28 mm focal length, 0.0061 mm pixel size, HDMI & USB output for real-time downloading, external power supply, and have the ability to be connected with GPS receiver. The cameras have high ISO level, wide aperture size and extremely high shutter speed, which helps to provide higher image quality. This camera with 1m height from the rail surface can provide 1m* o.85 m ground coverage in running and cross directions, respectively. With this height, pixel ground footprint of 0.2 mm can be obtained, providing full vision coverage for rails with high recognition ability even in high speeds.

2.3 Laser-Based Measurements Unit (LBMU)

2.3.1 LBMU: Description & Aim

LBMU is an important part in the monitoring system, where it has been designed to measure rail gauge, cross-level, and warp and detecting rail cracks and fractures. The unit is divided into two subunits, namely Single Point Laser subunit (SPL) and Multi-Spot Laser (MSL) subunit. Figure (6) shows the different work methods of SPL and MSL measurements.

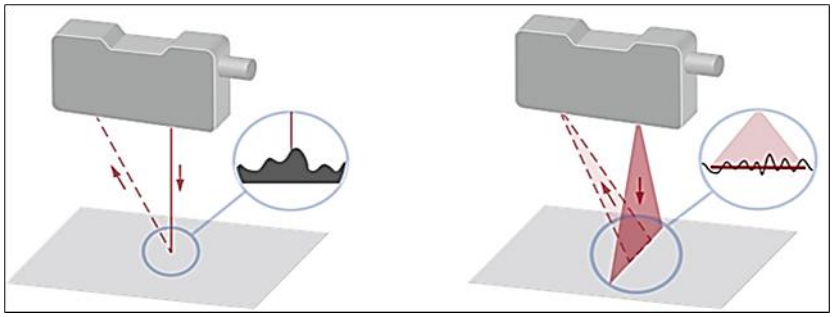

Figure 6 Methodology of SPL & MSL measurements

SPL consists of two high-frequency high-vibration-resistant precise single point laser measurement sensors. The laser sensors are connected to the data logging and synchronizing unit to record the laser measurements and color intensity values in GPS time. SPL works on measuring rail gauge and cross-level at the same time. The design concept of SPL is

based on fixing the laser sensors on a simple Automatic Adjusted Wheeled Carrier Frame (AAWCF), which changes according to rail gauge changes. AAWCF will be described in details in section (3-5). Laser sensors are fixed facing each other with displacement in running direction of 10 cm, and each sensor is provided with colored coded target that are exactly facing the laser beam. Using that, the rail gauge can be measured using the precise SPL technique, and the cross-levels can be determined in high accuracy using the intensity value of the codded target. The coded targets can be provided with wipers to remove water, snow, ice, etc. Using two SPL helps to provide more reliable measurements and easily detecting the problem in the unit functionality. Design concept of SPL can be shown in figure (7) & (8).

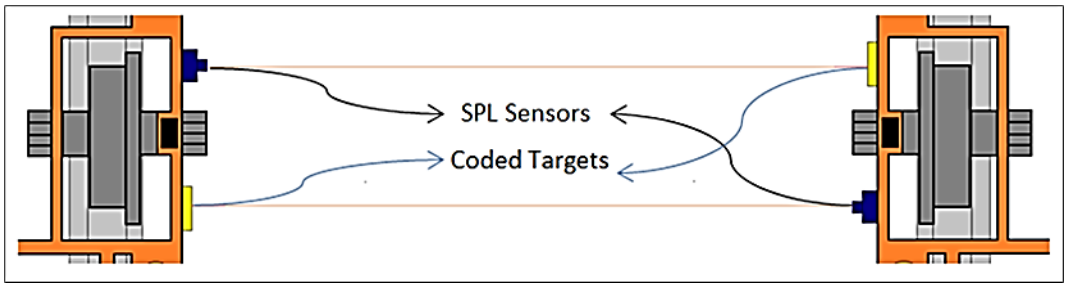

Figure 7 SPL subunit design concept top view

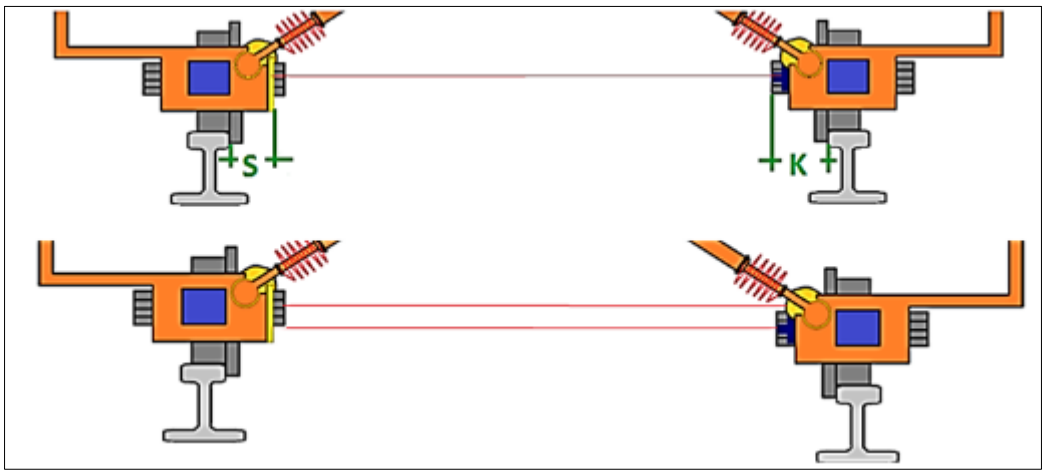

Figure 8 SPL subunit design concept cross section. Up: zero cross-level. Down: non-zero cross-level

As mentioned above, AAWCF works on making the two wheels moving exactly on the rails even with inconstant rail gauge. As clear from the shown design concept, rail gauge can be measured by adding the SPL measurements to the two distances (S) and (D), which are the distance between the internal rail face and the center of laser sensor, and the distance between the coded target and the other internal rail face, respectively. The coded target is a rectangular aluminum plate (25 cm * 10 cm) divided horizontally to 250 strips (1 mm/strip) with different colors and known laser intensity values. These strips work as a digital leveling staff with accuracy of 1 mm, where the vertical distance between each strip and the zero-level strip is known. Then, the intensity values that are recorded with each distance measurement is used to determine the corresponding strip and consequently, the cross-level between the two rails becomes known. Using RNS to determine the precise relative distances between each two cross-levels, rail-track wrap parameter can also be determined. With 2 KHz data rate, one gauge measurement and one cross-level measurement can be observed every 2.5 cm with train speed of 160 km/h, which are adequately enough for detecting any related rail irregularities.

MSL consists of two high-frequency high-vibration-resistant precise multi-spot laser measurement sensors, one for each rail. The laser sensors are connected to the data logging and synchronizing unit to record the MSL measurements in GPS time. The aim of MSL is to detect rail fractures and surface cracks. The design concept of MSL is based on fixing the MSL sensors on the AAWCF, facing the rail surface on a certain height. AAWCF and the right sensor-to-rail height guarantee MSL to project exactly on the rain top surface. MSL design concept is illustrates in figures (9) & (10).

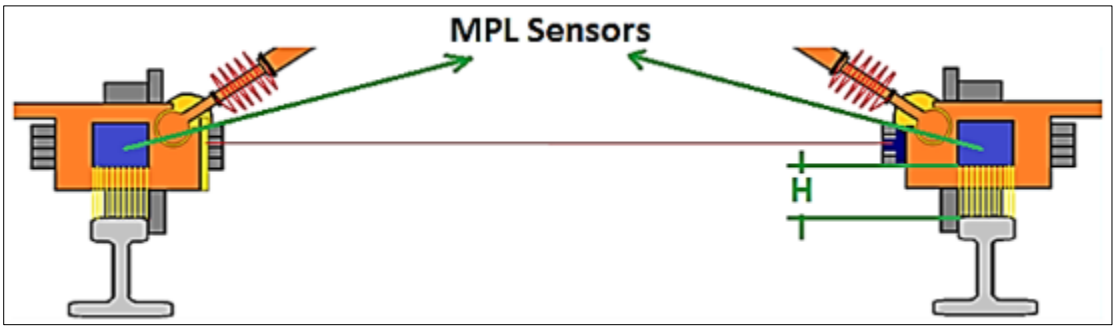

Figure 9 MSL subunit design concept: cross section

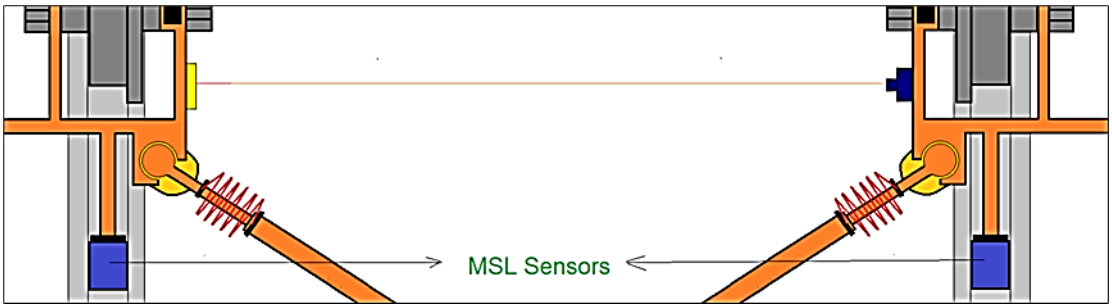

Figure 10 MSL subunit design concept: top view

MSL sensors are able to provide detailed and precise cross section for the rail top surface with date rate of 1.5 KHz. Each cross section includes 600 measurements distributed regularly on the width of rail surface (usually 72 mm) at a rate of one point every 0.12 mm, which is sufficient for dealing with invisible small surface cracks. These precise and measurements are used in the processing unit to determine rail fractures and surface cracks by comparing the obtained rail surface to the optimal rail surface dimensions. Statistical testing is used to match the two surfaces, finding out the unmatched parts in each cross section, which might point to rail cross crack or surface fracture. Also, the average of all MSL measurements is used to detect easily the deep rail fractures, where the observed height is compared to the certain height of the MSL sensor. Figure (11) shows the cross-rail surface matching and the average MSL techniques. Using this technique, rail expansion gaps might also be detected, but the vision recording unit helps to deal with such misdetections. Detecting cross irregularities in sequential cross sections points to longitudinal cracks and fractures. See figure (12). With 1.5 KHz data rate, one cross section can be observed in less than 3 cm with train speed of 160 km/h in each train pass. Using the same rail by different rains or by the same train, which are all provided with this monitoring system can guarantee covering the whole rail completely in short time.

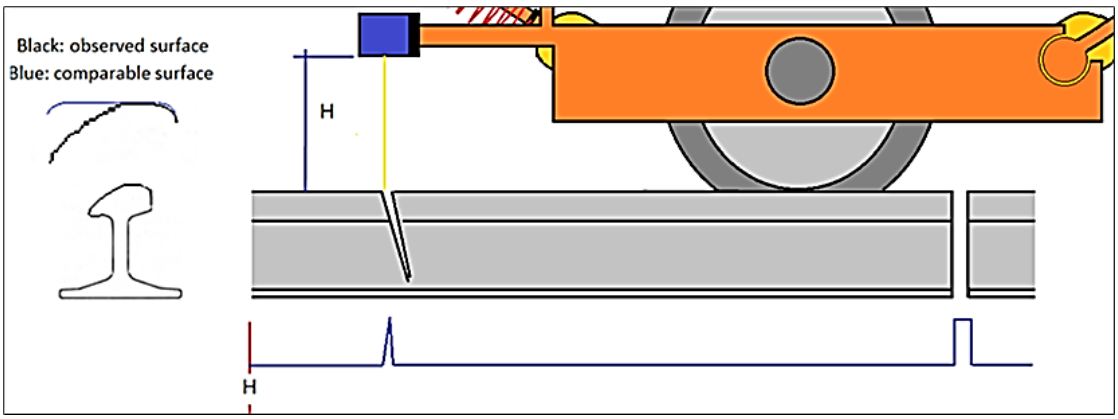

Figure 11 Left: Cross rail surface matching. Right: Average multi-spot laser

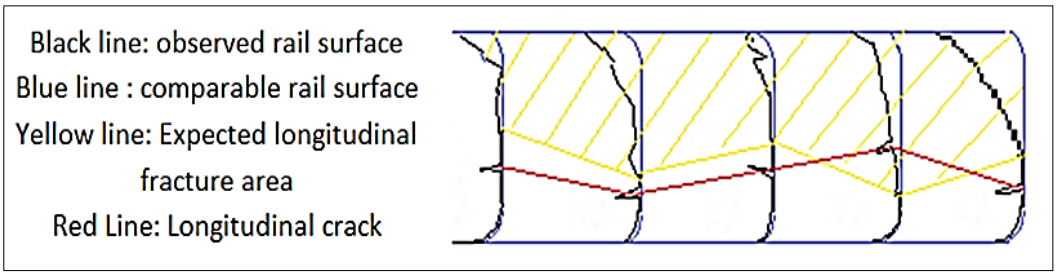

Figure 12 Detecting longitudinal cracks and fractures using sequential cross sections

2.3.2 LBMU: Sensors Description

- High-Performance Laser Distance Sensors (Baumer-OM70)

Baumer OM70 is a single point laser sensor, which provide high performance even under difficult environmental conditions as ambient light, robust and reliable observations even on demanding surfaces using extra fine line beam, 0.7µm resolution, 1.5m measuring range, ±0.06% linearity deviation, 0.3µm repeat accuracy, ±24µm linearity error, digital interface, easy integration with other sensors, stable measurement signals, extremely easy adjustment to specific requirements, real time analyses, easy connection to different levels of the automation pyramid, and 2 KHz measuring frequency. The sensor can be adjusted to record the laser intensity value for each measurement, which is used to determine the differences in cross-level. See figure (13).

- Multi-Spot Laser Distance Sensors (Baumer MESAX 70)

Baumer MESAX 70 is a multi-spot laser sensor, that provides high vibration resistant, Ethernet interface, up to 600 single measurements on a 72 mm wide line for highest reliability, high accuracy even under varying ambient light conditions, easy installation and integration with other sensors, stable and reproducible measurements even on unevenly shiny and extremely rough surfaces, three integrated measuring modes (mean, minimum, maximum), 2µm resolution and 1540 Hz data frequency. See figure (13).

Figure 13 Left: SPL sensor (Baumer OM70); Right: MSL sensor (Baumer MESAX 70)

2.4 Data Logging & Synchronizing Unit (DLSU)

2.4.1 DLSU: Description & Aim

DLSU works on collecting and synchronizing all the data recorded by the different sensors used in the monitoring system in the same time frame. As GPS has one of the most precise time frames, u-blox GPS receiver is used as a base for providing precise synchronization when GPS is on; otherwise, precise timing interpolation can be used. The main aim of GPS-based time synchronizing is to link the recorded data with their geodetic positions and video frame. Figure (14) shows how the system units are linked to DLSU. Table (2) shows type of data collected from each sensor and the data loaded to Data Processing Unit (DPU).

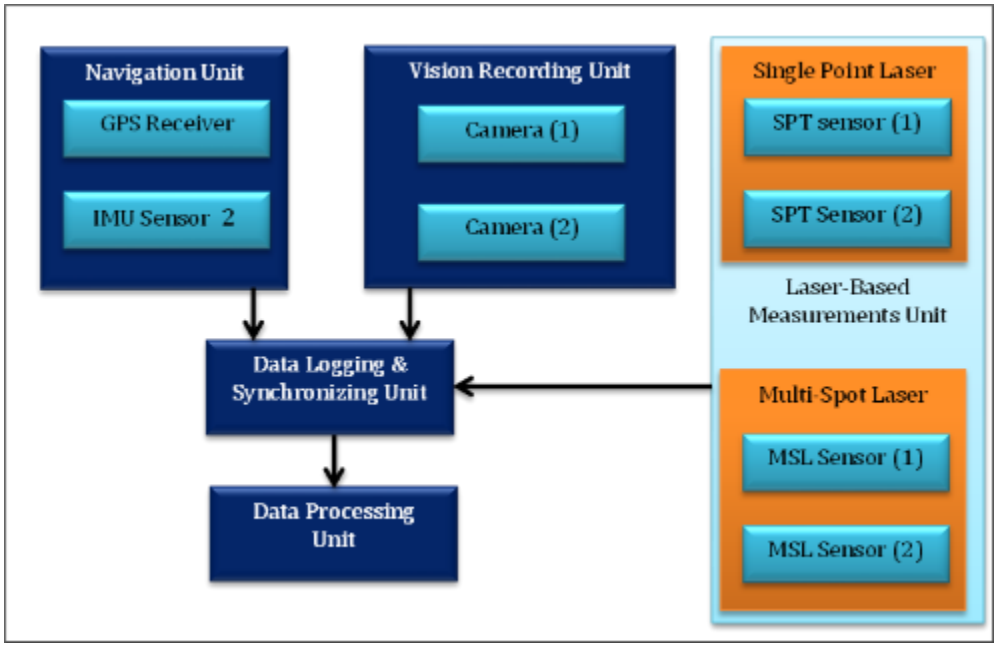

Figure 14 The system units connection with DLSU

Table 2 DLSU: data collected from other units; data frequency & data exported to DPU

System Unit	Sensor	Collected Data	Data Frequency	Data Exported to DPU
NU	GPS receiver	Raw GPS observation (obs. file) & navigation messages (nav. file) Ready SA-GPS (GPS Time & 3D coordinates)	1 Hz	Raw GPS observation (obs. file) & navigation messages (nav. File) Ready SA-GPS (GPS Time & 3D coordinates)
	IMU sensor	Raw IMU file (self-generated time frame, 3D gyro measurements, 3D accelerometer measurements)	1 KHz	Raw IMU file (GPS time frame, 3D gyro measurements, 3D accelerometer measurements)
VBU	2 Cameras	Videos	60 - 120 fps	Videos + start and end recording time in GPS time frame
LBMU	2 SPL sensors	Ready SPL files (self-generated time frame, distance measurements, laser intensity values)	2 KHz	Ready SPL files (GPS time frame, distance measurements, laser intensity values)
	2 MSL sensors	Ready MSL files (self-generated time frame, distance measurements)	1.54 KHz	Ready MSL files (GPS time frame, distance measurements)

2.4.2 DLSU: Sensors Description

One of the suggested data logger and synchronizer that can be used in the system is Precise Time Data Logger, which can be adjusted for the introduced monitoring system for connecting all sensors, collecting all data, and synchronizing operation. PTDL has been used by the author in [13, 14, 15, 16]. PTDL is based on u-blox GPS receiver used in NU for precise GPS timing. All data recorded in this sensor is timed in GPS time frame when satellite signals are available and in precise relative interpolated timing when GPS signals are hidden. All sensors used in the monitoring system can be

connected to data logger, namely external u-blox GPS receiver, IMU sensor, two SPL sensors, two MSL sensors, two cameras, and the internal GPS receiver. Figure (15) shows PTDL sensor.

Figure 15 DLSU: data logger and synchronizer sensor

2.5 Data Processing Unit (DPU)

Table (3) summarizes the main functions of DPU, describing the processing techniques, and showing the unit outputs. Processing techniques can be divided into two steps. The first one deals with raw observations and files, giving direct positioning, navigation, distance measurements, and photos in GPS time frame. The second step is for statistical testing and comparing with specifications to decide whether there are rail-track irregularities.

Table 3 DPU: the first main steps of data processing

Unit	Data Imported from DLSU	Processing Technique	Outputs
NU	- Raw GPS observation (obs. file) & navigation messages (nav. file) - Ready SA-GPS (GPS Time & 3D coordinates) - 2 Raw IMU file (GPS time frame, 3D gyro measurements, 3D accelerometer measurements)	Matlab self-developed algorithms: - SA-GPS /CP-DP integration - IMU raw data to INS measurements (relative 3D displacements) - Studying the relative displacements of each MEME-INS solution in cross & vertical directions, two cross displacements from the 2 IMUs are compared and the same for vertical displacements, finding out the sudden changes using RMSE values, comparing RMSE values with the allowable vibration levels to find out expected rail irregularities - CP-DP/MEMS-INS integration - Note: averaged MEM-INS is in-use - EA-GPS/ RNS integration Note: averaged RNS is in-use	- SA-GPS Solution - 1 Hz (GPS Time, X, Y, Z) - EA-SAP Solution - 1 Hz (GPS Time, X, Y, Z) - Average RNS Solution - 1 KHz (GPS Time, ΔX, ΔY,ΔZ) - 2 MEMS-INS Solution - 1 KHz (GPS Time, ΔX, ΔY,ΔZ) - (X, Y, Z, "H alignment", RMSR value of cross direction, photo No. (GPS time) that shows the problem area) - (X, Y, Z, "V alignment", RMSR value of vertical direction, photo No. (GPS time) that shows the problem area) - Averaged EA-NS Solution - 1 Hz (GPS Time, X, Y, Z) - Report of (GPS OFF/ON) Times for locating low accuracy positioning areas
VBU	- Videos + start & end recording time in GPS time frame	Matlab self-developed algorithm: - Video to frames extraction, automatically enhancing & synchronizing with GPS time frame	- List of Enhanced extracted photos, titled by GPS time frame

LBMU	- Ready 2 SPL files (GPS time frame, distance measurements, laser intensity values)	Matlab self-developed algorithm: - Determining the average rail gauge value from the 2 files at the same GPS time frame, comparing the results with the gauge ranges that stated in the specifications, & finding out the irregularities - Converting the intensity values to differences in height, taking the average value at the same GPS time, comparing the results with cross-level allowable ranges that stated in the specifications, & finding out the irregularities - Determining the changing rate of the obtained consecutive cross-level values in a certain distance, and comparing the results with warp allowable rates that stated in the specifications, & finding out the irregularities	- (X, Y, Z, "Gauge", gauge irregularity value mm, photo No. (GPS time) that shows the problem area) - (X, Y, Z, "Cross-level', cross-level irregularity value mm, photo No. (GPS time) that shows the problem area) - (X, Y, Z, "Wrap", wrap rate, photo No. (GPS time) that shows the problem area), where GPS data is for the mid of the wrap distance.
LBMU	- Ready 2 MSL files (GPS time frame, distance measurements)	Matlab self-developed algorithm: - Comparing the averaged value of the 600 measurements of each cross section with the certain sensor-to-rail height & finding out the irregularities - Matching the 600 measurements of each cross section with the optimal rail surface measurements by determining the RMSE value and comparing this value with the allowable limits to find out the irregularities - Finding out the strange and odd values in the 600 measurements of each cross section even if the RMSE value is regular, studying the odd values in the nearby consecutive cross sections to find out any possible longitudinal cracks	- (X, Y, Z, " Rail Fracture", height value mm, photo No. (GPS time) that shows the problem area) - (X, Y, Z, " Cross Rail Fracture", RMSE mm, photo No. (GPS time) that shows the problem area) - (X, Y, Z, " longitudinal cracks" , photo No. (GPS time) that shows the problem area)

2.6 Automatic Adjusted Wheeled Carrier Frame (AAWCF)

AAWCF is the main part of the monitoring system, where all sensors are fixed on this frame. AAWCF has been designed to include two steel wheels dimensionally matched with the real train wheels to simulate the actual train movements and vibration. These two wheels are fixed on steel frame which has the ability to automatically increase and decrease the cross distance between the two frame wheels to comply with the actual rail gauge. AAWCF uses 4 hydraulic spring dampers, which work to ensure the wheels remain on the right and left rail-tracks, whether in the case of gauge expansion or convergence. Dampers work also to absorb the longitudinal frame movements due to the cross movements. Dampers are used for pushing and pulling AAWCF in trains round-trips and for the vertical wheels' stability on the rails, especially in the case of vertical alignment irregularities in high speeds. AAWCF is connected to the main train body using two connection points, which are shown in the following design figures. As clear from the design, the sensors are fixed directly on the steel frame without any vibration absorbers, except those coming with the laser sensors and cameras to extending the sensors life and do not affect the works accuracy. The 2 IMU sensors are fixed directly on the steel frame exactly on the top of the axle shaft of wheels to experience the actual vibration levels of train movements on the rails, where the methodology of using such sensors is based on investigating the vibration behaviors on cross and vertical directions for rail alignment irregularities. Dampers should be chosen carefully to not be tight or loose but keeping the wheels exactly on the rails. Tight pistons can create an extra fraction between the rail and the prominent rim of the system wheels, increasing the vibration level, and sliding the wheels off rails. Loose pistons also are not recommended as the system will not be able to automatically adjustable with increasing and decreasing the actual rail gauge. The system is designed to stay on the rails in all cases of rail-track irregularities, namely gauge, cross-level, wrap. Fractures, cracks, and vertical & horizontal alignments, as it has a wider range of gauge flexibility than that allowed train to stay on the rail track. Figures (16), (17) & (18) show the whole system design concept, including top, side, and cross views, respectively. Figures (19) & (20) show side and front views of the recommended places of the system on the train.

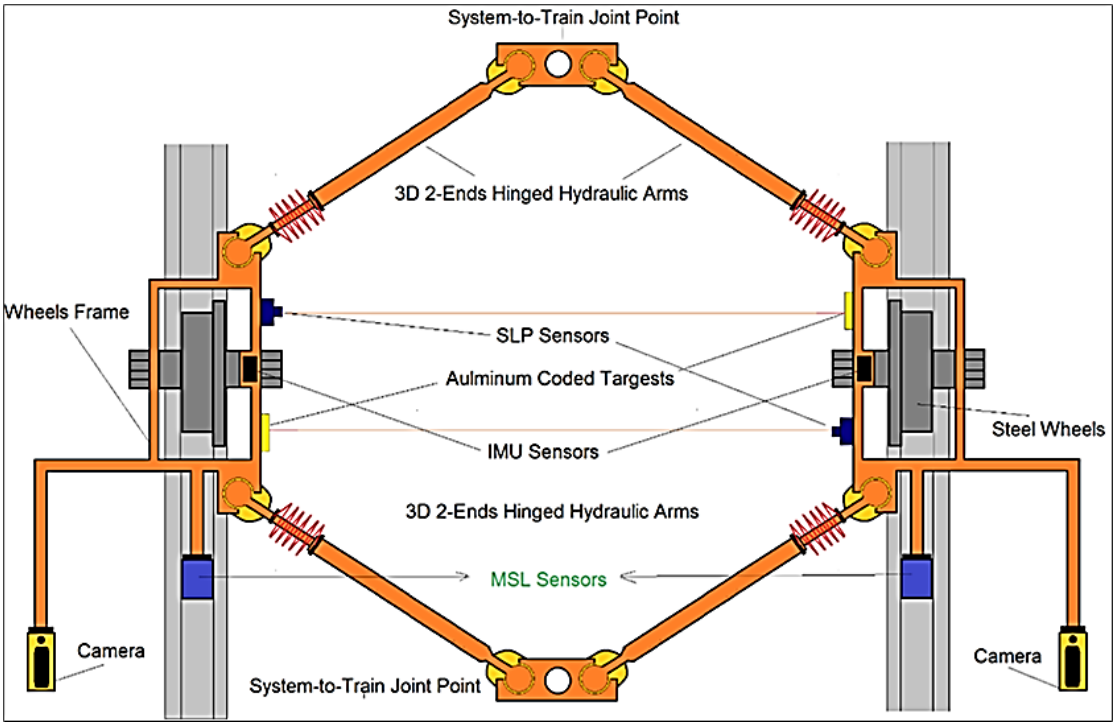

Figure 16 Design concept of the whole system: top view

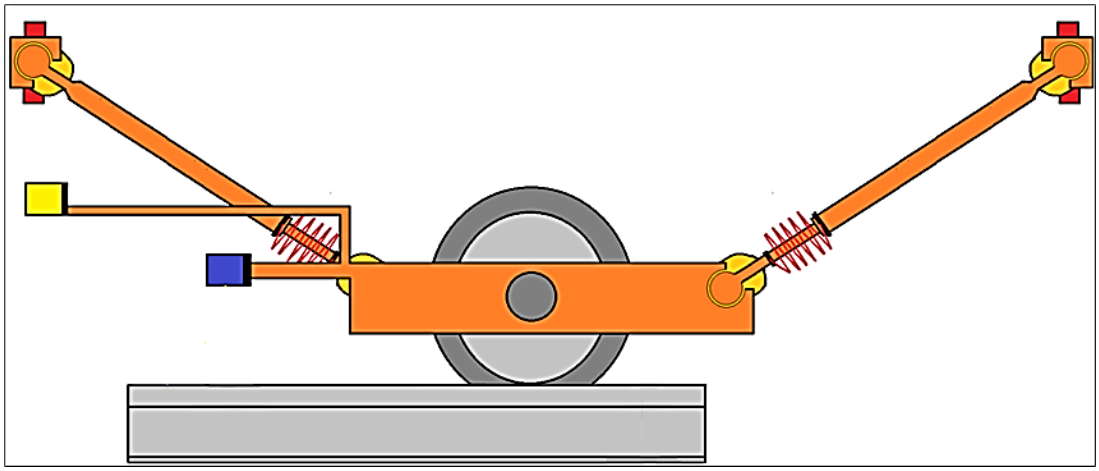

Figure 17 Design concept of the whole system: side view

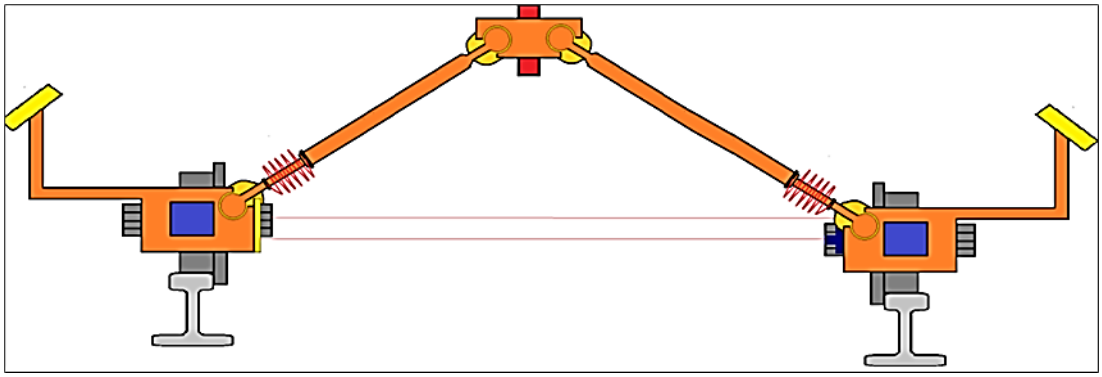

Figure 18 Design concept of the whole system: cross view

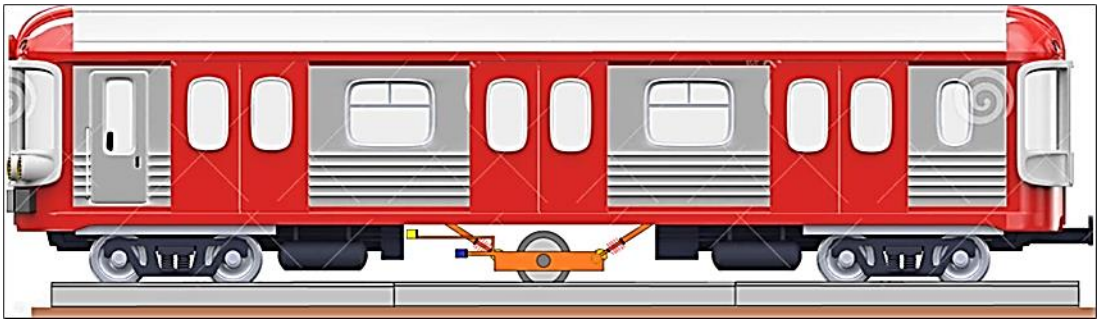

Figure 19 The recommended place of the system on the train: side view

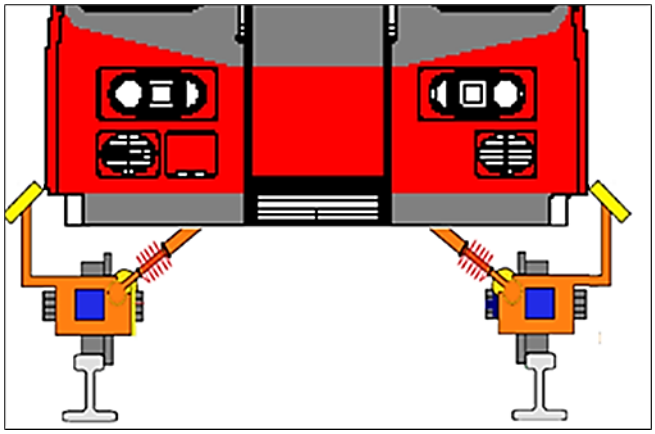

Figure 20 The recommended place of the system on the train: front view

3 Initial Trails & Further Works

As the design concept of this system is still waiting for supporting adoption to manufacture AAWCF, evaluating the system performance and the design concept of each unit in actual environment is difficult. Initial trails have been carried out for evaluating the design ideas using GNSS/INS testing train of Nottingham Geospatial Institute (NGI), where all the sensors used in the system have been connected together, giving initial results. As the maximum train speed is 30 km/h, it is unfair to judge the success of the ideas, where high speed and high vibrations play a major role in the final decision.

Figure 21 GNSS/INS testing train of NGI with MSL, SPL, IMU, camera, and GPS

Also, AAWCF is the main part in the design concept and the performance of whole system is based mainly on the functionality of this part. Thus, individual testing for the sensors as well as techniques have been carried out using the available capabilities as a first step of applied the whole design concept, providing cost-effective permanent monitoring

system for rail-track irregularities. Figures (21), (22) & (23) illustrate parts of the initial trails carried out by the author using GNSS/INS testing train at NGI, Nottingham University.

Figure 22 DLSU, u-blox GPS receiver, & DPU

Figure 23 Testing the cameras for different speeds

4 Conclusion

This paper aimed to present a novel design concept of low-cost every-day monitoring system for rail-track problems. The introduced design is theoretically overcome the currant in-use designs in terms of providing permanent accurate observations for detecting rail-track irregularities for all railway network during service time using low-cost integrated sensors. The paper has investigated the rail-track common irregularities by identifying the issues, presenting challenges, and highlighting the limitations of currently in-use rail-track monitoring techniques. Then, the potential of using cost-effective sensors and intelligent techniques for detecting the rail-track geometric parameters and rail irregularities has been theoretically studied. This has included using low-cost GPS, MEMS-INS & high frequency single-point and multi-spot laser distance sensors for detecting rail surface cracks, fractures, cross-level**,** wrap, gauge and vertical & horizontal rail alignments. Different levels of GPS/MEMS-INS integration have been covered, illustrating the advantages of each solution and the optimal utilization in the system. Vision-based monitoring using low-cost HD non-metric digital cameras has also be investigated, showing the optimal specifications required for fulfill the system requirements. Data logging and GPS-time-based synchronization method has been presented as a part of the design concept of the suggested system. The processing techniques and unit outputs have been illustrated in details, showing how each unit works for detecting the rail-track irregularities, providing the geodetic positioning and the vision-based assessing. The main part of the system, which is AAWCF has been discussed in more details, showing simple and clear different views for all details and integrated sensors. It is recommended to carry out this idea, by building up and evaluating the performance of AAWCF in real railway environments, evaluating the design concept of each unit individually to know the weaknesses of each technique, and then evaluating the whole system in all expected railway cases to find out the advantages and limitations of this design concept.

References

[1] Agarwal M, Chandra S. Railway Engineering. 2nd ed. Oxford University Press India; 2013.

[2] British Standards Institution. BS EN 13848-1:2019: Railway applications. Track. Track geometry quality - Part 1: Characterization of Track Geometry. London BSI. Available from (https://www.en-standard.er/bs-en-13848-1-2019-railway-track-track-geometry-quality-characterization-of-track-geometry/).

[3] British Standards Institution. BS EN 13848-2:2019: Railway Applications. Track. Track geometry quality - Part 2: Measuring Systems - Track recording vehicles. London BSI. Available from (https://standards.globalspec.com/std/14344279/BS%EN%2013848-2)

[4] Amami M. Enhancing Stand-Alone GPS Code Positioning Using Stand-Alone Double Differencing Carrier Phase Relative Positioning. Journal of Duhok University (Pure and Eng. Sciences). 2017; 20(1): 347-355.

[5] Amami M. The Integration of Stand-Alone GPS Code Positioning, Carrier Phase Delta Positioning & MEMS-Based INS. International Research Journal of Modernization in Engineering Technology and Science. March 2022; 4(3): 700-715.

[6] Amami M. The Advantages and Limitations of Low-Cost Single Frequency GPS/MEMS-Based INS Integration. Global Journal of Engineering and Technology Advances. Feb. 2022; 10(2): 018-031.

[7] Amami M. The Integration of Time-Based Single Frequency Double Differencing Carrier Phase GPS/ Micro-Elctromechanical System-Based INS. International Journal of Recent Advances in Science and Technology. Dec. 2018; 5(4): 43-56.

[8] Amami M, Smith M, Kokkas N. Low-Cost Vision Based Personal Mobile Mapping System. ISPRS- International Archives of The Photogrammetry, Remote Sensing and Spatial Information Sciences. 2014; XL-3/W1: 1-6.

[9] Amami M. Low-Cost Vision Based Personal Mobile Mapping System [Ph.D dissertation]. University of Nottingham; 2015.

[10] Amami M. Testing Patch, Helix and Vertical Dipole GPS Antennas with/without Choke Ring Frame. International Journal for Research in Applied Sciences and Engineering Technology. Feb. 2022; 10(2): 933-938.

[11] Amami M. Speeding up SIFT, PCA-SIFT & SURF Using Image Pyramid. Journal of Duhok University, [S.I]. July 2017; 20(1): 356-362.

[12] Amami M. Fast and Reliable Vision-Based Navigation for Real Time Kinematic Applications. International Journal for Research in Applied Sciences and Engineering Technology. Feb. 2022; 10(2): 922-932.

[13] Amami M. Comparison Between Multi & Single Epipolar Geometry-Based Filters for Optical Robot Navigation. International Research Journal of Modernization in Engineering Technology and Science. March 2022; 4 (3):476-485.

[14] Amami M. Comparison Between Multi Epipolar Geometry & Conformal 2D Transformation-Based Filters for Optical Robot Navigation. International Journal for Research in Applied Sciences and Engineering Technology. March 2022; 10(3): 388-398.

[15] Amami M. Comparison Between Multi Epipolar Geometry & Affine 2D Transformation-Based Filters for Optical Robot Navigation. International Journal for Research in Applied Sciences and Engineering Technology. March 2022; 10(3): 399-409.

[16] Amami M. Multi and Single Epipolar Geometry-Based Filters Vs. Affine and conformal 2D Transformation-Based Filters. Global Journal of Engineering and Technology Advances. March 2022; 10(3): 032-051.