A review of computer vision for railways

1. Camera hardware and data requirements

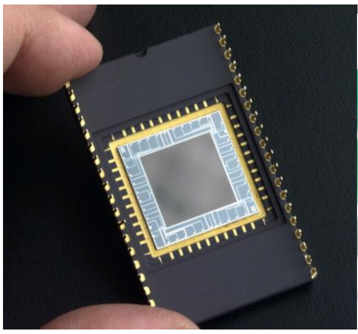
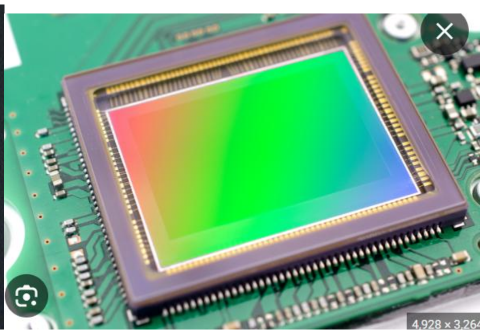
   a. General

: camera lens need two considerations=Focal length and Aperture size

- Focal length decides the angle of view. Small focal length means wider-angle view.

- Aperture size(조리개 크기) decides the size of the opening, which means it decides how much light can enter the lens.

1) Technology type : standard image sensor, lidar, sensors + lens
2) Image sensor type:
   - Charge-Coupled Device(CCD)
   - Complementary Metal-Oxide-Semiconductor(CMOS)

(Image sensor changes light into electric signal.)

: CCD

1. consists of an array of capacitors connected in rows.
   (빛이 닿으면 픽셀에서 픽셀로 순차적으로 이동시켜 읽어낸다)
2. CCD has high resolution.
3. CCD has high cost, high energy , low speed.
4. They are good for line scanning. -> long railway images monitoring, catenary wire monitoring

(catenary wire)

: CMOS

1. consists of matrix of photo sensors(in rows and columns)
2. Difference with CCD is that the photo sensors do not

communicate their charge with their neighbours, instead communicate directly with the register.
(픽셀에서 받은 빛을 바로 픽셀 단위에서 증폭하여 읽는다)
3. high speed, low energy, widespread in recent years.
4. Used in overhaead line equipment, fasteners checking, with vehicle mounted cameras

3) Sensor Spectral Response (빛의 파장)
   1. It means the sensitivity to different wave lengths. Classical sensors are usually more sensitive in visible light(가시광선) , other sensors(lidar or thermal camera) are more sensitive in other range of lights(자외선, 적외선)→ good for invisible phenomenon detection.(열기 같은 기차의 상태에 이용)
   2. (ex) mount a thermal camera on a personnel carrier to monitor rail temperature or use thermal camera to monitor brake discs and steel axles in railway vehicles.

4) Sampling Frequency and Pixel Count
   1. frequency means how many pictures exists in a 1 second(fps).(초당프레임 수)
   2. Pixel count means how many pixels in a picture.
   3. (ex) monitoring level crossings(철도 건널목) may use lower resolution cameras, (human,cars are big) while monitoring tunnel cracks use higher resolution.(crack is small)
   4. (ex2) level crossings need immediate data processing because it is real-time monitoring. It means the data volume should be small. In contrast, tunnel lining analysis is less time sensitive. Data can be stored on local hard drives.

5) Calibration(보정)
   1. When images are taken by camera, INTRINSIC calibration and EXTRINSIC calibration is needed. It means image needs to be re-arranged(correction) in order to get 'real image'.
   2. Intrinsic calibration = calibration for intrinsic characteristics like focal length, sensor resolution

3. Extrinsic calibration = calibration for extrinsic characteristics like position and orientation(direction) of the camera. It is done with patterns like chessboard pattern or color like white balance.
5. (ex) Quantitive information(like 균열 폭) needs these calibration.
6. (ex) However, Qualitive information(like 사람이 있다/없다) doesn't need calibration. Because it is not considering real objects.

6) Recording format(녹화 형태)
   1. It means the form of the data. The format of the data can be different depending on where the data will be used.
   2. (ex) If the data is used for ballast endoscope(자갈 도상) monitoring, the data is analyzed by human. Therefore the format should be more suitable to human analysis, not computer analysis.

7) Data storage(데이터 저장 방식)
   1. This means how the data will be stored.
   2. Railway monitoring generate large datasets → This cause storage challenges → edge computing (데이터를 센터로 보내지 않고 끝단(edge)에서 처리한다)
   3. (ex) edge devices can locally process the data on board (not sending all data to processing center). → this reduce the processing and storage pressure on data center.

8) Image recording location
   1. Collected images needs to be stored with location data in order to align the datasets along the route. Usually GPS is used for this.
   2. However if the rail way is on remote area, it is hard to use GPS to get location information. → We need other alternatives for this.

9) Region of Interest

   1. ROI is a subset of an image in which the monitored phenomenon is expected to be located. This area can be fixed region or shiftable region.

   2. (ex) In a image there is an area that cracking is expected to

located.

10) Data Failsafes
    1. This means we have to detect errors from hardware.
    2. (ex) In rail buckling monitoring, hardware failures must be detectable.
    3. If there is power fauilure, this can be detected using communication gateway.
    4. If there is a recording error, this can be detected and assisted by deploying multiple cameras.

11) Data security
    1. This is about cyber security.

b. Hardware performance requirements

: Railway can be a harsh environment. Devices should go through extreme environment.

1) Contaminates(오염)

   : contamination happens frequently during train operation.

   (ex) A mechanical shutter can be automatically opened and shut when required.
2) Lighting
3) Weather conditions
4) Mechanical/Electrical interface

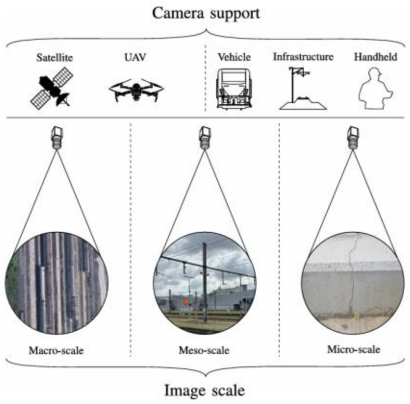
   : Vibrations occur in trains. System needs to be capable of this.
5) Power supply

   : Imaging hardware typically has higher power requirements compared to other mechanical sensing methods.(data quantity is big) Current solutions(like solar panels) are difficult to provide enough power to record images at high frequency.
6) Deployment/Installation
   1. The deployment of imaging systems should have minimal negative impact on scheduling of trains.(열차의 스케줄)
   2. (ex) The installation and maintenance times should be short in order to not disrupt the train schedule.

3. (ex) The hardware should be moduler.(사람의 접근이 어렵기때문)

2. Camera setup types

: This means the type of camera settings like fixed, handheld, mounted on vehicle. In order to clarify which way is suitable, we should see the size of the ROI(Region of Interest)



It means that the macro-scale→large zones are monitored, micro-scalde→small zones are monitored.

a) Vehicle borne cameras

: This is attaching camera on the train.

1. Pros(장점)

: don't have to access to rail to setup cameras.

: entire route can be monitored using a small number of cameras.

: It can detect both meso(중간크기) and micro scale(작은 크기).

2. Concicitons(조건)

: It have to handle harsh conditions.(contamination of lens, various light conditions)

3. Type of vehicle
   1. in-service trains = it provides a frequent and low-cost way to collect data.
   2. dedicated measurement trains. =Trains only for inspecting rails.
   3. self-propelled and road-rail vehicles.

4. Type of camera

- For rail or catenary(전선) inspection (대부분의 경우)→Use laser-scanning camera-based system.

📌 **의미**

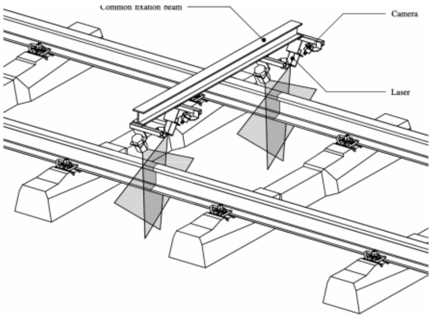
- **Laser-scanning camera** = 레이저 스캐너 + 카메라를 결합한 장치
- **기능**: 물체 표면에 레이저를 쏘고, 반사된 빛을 카메라가 감지해서 → **거리·형상·3D 구조**를 측정하는 시스템

📐 **원리**

1. **레이저 빔 발사** → 대상(레일, 터널 벽, 시설물)에 닿음
2. **반사광 수집** → 카메라(광센서)가 반사된 빛을 감지
3. **시간차/위치 계산** → 빛이 돌아오는 시간(Time-of-Flight)이나 삼각측량 원리로 **정밀한 거리 정보** 획득
4. **3D 형상 재구성** → 레일 표면, 균열, 구조물 변형 등을 고해상도로 복원

- 장점: can get 3D profile of the rail.(카메라는 2D data)
- 조건: need to consider train speed.

5. Location and number of cameras
   - commonly on the roof of the vehicle.(line monitoring)(카메라 위치가), or under-carriage(track structure monitoring), or the front of the vehicle(route monitoring)
   - (ex)



multiple cameras and lasers attached on the under-carriage. Both sides of the cameras and lasers(왼쪽 오른쪽) are coupled in an all in one device.
: depending on the application, other sensors may be used with the cameras to get additional data.

   - (ex) cameras in the centre of all in one devices are also used to detect defects on fastenings, sleepers and even trackbed.

6. Datum and Device registration.
   : It means the technology that can send the location of images.
   In order to do that, GPS coordinates or curvilinear abscissa(곡선좌표계) are needed.

b) Unmanned Aerial vehicle Cameras(UAV)

: this is the way that can use camears on drones.

1) Pros(장점)

   : It can hover over rails, It can detect both macro and meso scales.

2) Usages(활용)

   : It can be used for track inspection.

   : Detecting vegetation grouth and artefacts on the line.

   : Surveying sites after incidents

   : Inspecting overhead wires.

3) Challenges

   : weather. Short battery life, aviation regulation

   : need to do post processing(후처리)

   (ex) camera lens aberration, GPS location error.

c) Helicopter

1) Pros

   : It doesn't get affected by weather, Longer operating ranges, Longer time flying.

2) Challenges

   : Can not use in tunnels

d) Satellite imaging cameras

: Two technologies are used for satellite imaging, Optical imagery and Radar imaging.

The difference : the type of wavelengths(파장) used and their active or passive characteristic.

1) Optical remote sensing techniques : measure the light naturally emitted by earth.(visible light 가시광선 이용)(지구에서 반사된 빛을 인식한다)= this means passive

   It can achieve VHR(Very High Resolution).

   This can be used for assessing the risk of vegetation(like trees) falling on railway.

   -challenge: Images can not be seen by cloud. Climate counts.

2) SAR(Synethetic aperture rada) : use specific wavelegths of energy(가시광선이 아닌 다른 파장을 이용한다), It emits wave energy and measure the reflection. =active이다.

   - How it operates : It operates in this way : Interferograms are the diffence between two SAR images taken at different times.(one is pictures that are taken in same time but in different angle/location The other is pictures that are taken in the same location(angle), but in different time(after satellite rotations earth one more time) → determine the variation(change) of the distance between the satellite and the measured object

- Usage: civil engineering structure's shape change, sink (time difference method), recognization of 3D-land shape(location difference method)
- (ex) bridge collapse detection, landslide detection

: There are two subsets in SAR.

a. PS(persistent scatter)
   : PS means persistent reflection of the wave(균일한 전파 반사)
   In Urban area, the PS is high.
b. SBAS(small baseline subsets)
   : SBAS is for non urban area, where PS is low.
c. Combining the two.
   : use for detecting landslide in railway. Or sudying bridge collapse.

3) Qualitative or Quantitative Imaging
   : This means choosing the right resolution for different purposes.
   (ex) when the purpose is to approximate the construction location, it is qualitative inspection. = no high resolution needed
   (ex) in quantative technique, high resolution is needed.

4) Weather
   : Bad weather like clouds and rain can hinder the accuacy of INSAR
   It refract the beam(파장의 굴절)→phase delay(위상 변화로 감지)→ surface deformation

5) Surface water, snow ,vegetation
   : when a tree or vegetation is covering the surface, or after rain water is covering rails, it ican affect the accuracy of SAR.

e) Handheld Cameras

1) Type of devices

: Many handheld devices are used

(ex) DSLR camera : It has been used to take high resolution images of track compotents.

(ex) smart phones : smartphones has GPS inside, lidar sensors, it can directly communicate with cloud server.

(ex) portable webcam

2) Datum and Device Registration

: This is about how to get location of the cameras.

If cameras are attached on trains, it is difficult to get location of the camea.

f) Way-side cameras

1) Rolling stock monitoring(=train monitoring)

: high resolution camera systems can be placed in the way side → monitor the overheating and ball bearings.

This is cost effective because it doesn't need to be fitted to every different trains for attachment.

2) Track monitoring:

Wayside cameras are good for applications that need semi-continuous Monitoring(주기적 감시, 계속 감시가 아니다) because it needs a lot of power supply for continuous monitoring.

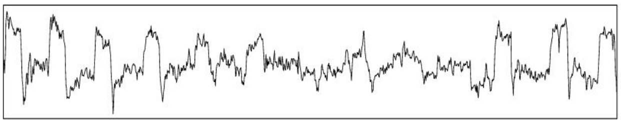
Pysical placement of the hardware is a problem. Most of cameras are located in high place as well as hardware. And it must resist vibration.

3. Image processing algorithms

a) Classical methods

1) Grey histogram detection

(a)

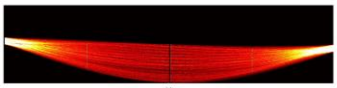


(b)

: grey histogram(밝기) variation on the columns of an image to monitor rail defects along the track with a vehicle borne camera. Each bar represents average value (밝기값) for a rail section. This process enables easy recognization of defects on rails.

2) Image Thresholding
   : After the rail changes into grey scale, Each pixel will be set either white or black, depending on specific threshold(기준값 보다 위/아래로 구분)
   (Ex) Researchers used this method in detecting the contact between pantograph and catenary. (추론) When the spark or flame errupts between pantograph and catenary because of contact miss, The algorithm can notify this using graey scale and image thresholding.
3) Edge detection
   1. This techiqe is used to detect image imperfections or discountinuities.
   2. After doing grey scale filter, → edges are spotted by the computation of the gradient between adjacent pixels.(밝기 차이)
   3. The edge is where dark and bright part meets(어두운 부분과 밝은 부분의 경계를 찾는다 .) That edge is an area that defect may exist.

4) Hough transform



   1. This is a grphical representation of the Hough matrix.
   2. Result of Hough transform of an image is not a image, but a graphical representation of the matrix

5) Colour detection
    1. This method is making certain color on a image look bright while other colors look dark. → This enaables to identify specific colour.
    2. (ex) Researchers used RGB-HSV colour transformation method in order to distinguish color of traffic lights.
6) Stereovision
    1. when using multiple independent cameras to take pictures of the same objext, it is an image combination.
    2. an image combination that leads to a computation of the depth of the object in both images. → a 3D point cloud can then be reconstructed for each pixel in a specific zone covered by both cameras.
    3. (ex) Use this stereoscopic technique to reconstruct the surface of the rail to dedece(estimate) three-dimensional infromation about rail surface defects.
7) Digital image correlation
    : This is the comparison between images taken using the same camera position. It compares pixel by pixel.

" In conclusion, Edge detection, Hough transforms, Gabor Filters and colour detection are useful for the detection and identification of an object in an image. Thresholding and digital image correlation are useful for object change detection. Grey histogram detection or stereovision can be suitale for both. "

Also combination of these techniques can lead to better result. In addition, These classical processing techniques can be used to post-process the images to prepare tham for the use of AI techniques (like YOLO).

b) AI-Based algorithms
    In recent years deep learning has shown to offer advantages compared to the traditional image processing.
    1) Object detection based methods

They are commonly divided into either one-stage or twostage detectors.

There are two types.

a. One-stage detector
   : one stage detector executes the prediction directly without generating potential regions or areas within an image that are likely to contain objects of interest .
b. Two-stage detector
   : assigns categorical labels to inputs as well as the refinement in the proposal stage.

➔ Biggest differece(가장 큰 차이는) : whether the model generats potential regions that are likely to contain objects.

c. The most used approaches in railway applications
   : R-CNN(Region-base convolutional neural network) and YOLO.

(ex) A researcher made AED-YOLO network to locate small components in the catenary system.(전차선 시스템에서 작은 부품을 탐지 한다.) → check component loss, break.

2) Semantic Segmentaion based methods
   : It is assigning a specific class label to every pixel within a given image, effectively representing each pixel according to its corresponding category. This process typically uses three steps.
   (checks every pixels and clearify which object the pixel is included)
   1. First, classification is used to recognise various categories present within the image.(이미지 속에 어떤 물체가 있는지 확인)
   2. Second, localisation is used to identify target objects, often accompanied by the creation of bounding boxes.
      (이미지 속에서 물체의 위치를 파악)
   3. Lastly, segmentaion is used to group pixels with matching categories.
      (모든 픽셀을 어떤 물체에 속하는지 파악)

c) Decision making

1. It is crucial to decide how we will use the result of computer vision.
2. There is SIL (Safety Integrity Levels).
   SIL 1 is lowest safety level (not urgent, like crack detection)
   SIL 4 is highest safety level (urgent, very important, like component loss detection)

3. Real-time or Delayed decision making?

a) Delayed decision making: It is performed off-site and it needs assistance of engineers. → This is good for where decision making process is not time sensitive.

b) Real time decision making eliminates human error. It is fast, direct action possible.

4. RAILWAY APPLICATIONS

a) Track applications

:This is about detection of objects within an image.

(ex) A methodology that can detect track components.

-The argorithm uses periodicity(주기) of the occurrence of these track components such as the tie, tie plate, rail, anchor or fastening components

-The detection is done on trolley -borne images.(카메라를 철도 검사용 소형 차량에 부착하여 검사)

b) Rail surface defects

: Railway wheels roll, slip and slide along the rail causing a range of rail defects.

-A technique able to detect the rail corrugation using a line scanning camera mounted on a maintenace trolley.

-If surface defects are located on the top of the rail head, a two-dimensional image can be used to detect them since the camera can focus on the top without monitoring the sides of the rail head. However due to the rail profile geomery and conical wheel profiles, wheel/rail contact is usually not a straight line on the rail head surface. Therefore to detect three-dimensional wear → laser scanning techniquer was used, CCD camera takes images of the track → beam is perpendicular to the track and illuminates the rail section, the camera has a different angle→This gives visibility of the intersection of the rail and the laser beam.

-A challenge with classical image processing is the high computatinal cost → developed the MOLOv3, THis is tiny model that combined YOLO v3-tiny detection head.

-To improve detection accuacy of rail surface defects→ multilevel rail surface defect detection network.Using Differential box-counting and Grabcut algorithm , Gaussian model parameter identification and defect segmentation. YOLO v2 was used for locating and detection of rail surface defects.

c) crossing nose damage(선로 교차로 파손)

: Railway crossing=trains change lines.

High impact forces and wear along the crossing.

3) Cossings have short lifespan compared to plain rail.
4) To detect this change, Under-carriage cameras can be use.

(ex) When performing image comparison, the image should be fixed position.

(ex)Use heatmap

d) Wheel profiles

: Due to wheel rail contact the wheel can also develop defects that negatively affect structural integrity or reduce ride quality.

(ex) used CCD camera and 2 laser beams fixed to the track to monitor wheel state when the vehicle passed through the measurement device.

(ex) Used a camera with laser beams → reconstruct 3D profiles of wheels.

(ex) Developed an algorithm combining images taken by a camera sensor with the 3D reconstruction of the wheel profile, with lasers. Using AI based YOLOv5 algorithm

e) Wheel rail contact position

: Trains are self-steering. → Wheel constantly shuffles along the rail surface.(the area that wheel contact with railroad moves)

5) Monitoring the motion of the wheel on the rail and the contact point is for ride quality.

    (ex) Using stereo cameras below train on the wheel-rail contact location(바퀴로 철로가 닿는 부분에 카메라 설치)

f) Clip and fastner absence

: Railway clips connect the rail and sleeper. → This clip can be loosen. → use under carriage cameras

:It takes too many time to check all clips by human force. → use computer vision

(ex) trained VGG16(이미지 분류기) for the detection and recognition of broken fasteners.

(ex) Use YOLO v3 to generate initial detection results → domain logic based hybrid model (DLHM) was proposed to improve detection performance

(ex) developed a deep anomaly detection ststem

: U-Net model → segmentation and denoising routines was trained → a deep learning based method was adopted using the feature map obtained from the U-Net encoder.

g) Track slab defects

: detecting concrete slab.

(ex) Used crack images from hand held cameras and UAVs to develop a deep learning.

(ex) multiscale information to extract crack feature shape.

h) Ballast detection(자갈)

: It is necessary to check whether ballast shoulder heights are sufficient. → this can be done using computer vision from vehicle-borne cameras.

: stress that is forced on ballast can cause ballast degradation.(break, stones become smaller) → photogrammatery

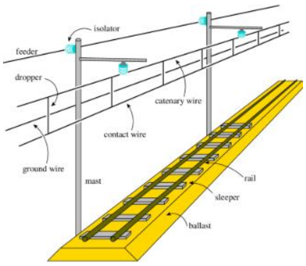
5. Overhead line applications

a) Pantograph

: There is loss of contact between the pantograph and catenary(열차 지붕에 있는 전원장치와 전선)

6) To detect this, monitored the height of the contact point by images taken by camera at the top of the vehicle. → use Hough transform, Canny edge detection and other filters In order to detct line in images. → Get the height of the contact point.

7) Overhead line



6. Other applications

1) Earthworks

   : This is for inspecting

2) Vegetation Management
3) Tunnel Subsidence
4) Tunnel Inspection
5) On-Track Safety Hazards
6) Level-Crossings
7) Trespass