2018 B.Tech CSE Batch
Literature Review Document - Team 5B

SpotLight *- Making conversations stay focused.*

Approved problem statement

Intelligently adjust the Field Of View (FOV) of an Integrated/ External Camera by taking advantage of the wide angle prospect and cropping into areas of interest based on Faces, Objects, and other factors, Enhancing the Area of Focus without moving the device physically.

Literature Review

Foundation:

The idea behind is to use the wide-angle front camera with applied machine learning to keep the person/people in frame as they move, allowing a person to go hands-free or move about during a video call or when recording oneself in ease.
The operation beneath is to use a custom model built for adjusting crop factor for the sensor along with Face-recognition and Object tracking to allow focus on both people as well as objects for consistent and clear communication.
The hypothesis being that users can stay more confident and needn't worry about whether focus or positioning in the capture frame of the video to a higher degree without much setup and time needed.

Discussion:

1. Video object detection with a convolutional regression tracker.
Pros:

- We have created an object tracker for the video object detection task, which can be easily inserted into a well trained image object detector. Without harming the performance of an object detector, the tracking functionality can be implanted into the model.

- Our tracker is light-weighted, memory efficient, and computationally efficient, as it re-uses the features from an object detector. Our tracker is compatible with the deep features extracted for the object detection purpose.
- Our new tracker performs in adaptive scales according to the sizes of different objects being tracked, which could cope with large object size variation in a video.
- We have designed a new video object detection pipeline to combine the advantages from both object detection and object tracking. With better bounding box proposals and linkages through time, we improve the performance with better effectiveness and efficiency.

Cons:

- The detector finds new objects, while the tracker follows the objects and provides better boxes for linking and re-scoring. Without re-scoring, the detection results could be incorrect or weak.

2. Deep neural networks for low cost eye-tracking.

Pros:

- Improves the targeting capability of the system in detecting living subjects by scanning for eyes over the view of the camera.
- The effective region of interest(ROI) is minimised to just the eye, thus reducing sustained higher order computing required for longer intervals.
- The gaze detection improves the prediction of the movements in the frame by accounting for the inertia of the subject for better tracking, greatly reducing the miss ratio in certain times.

Cons:

- Labelling is an additional step to be performed for mapping related spatial coordinates for better accuracy and minimal overhead.
- The eye-tracking can be disturbed at times when direct Infrared light source hits the frame making detection of eye harder.
- The field of view has to produce a rectangular output in defined ratios such as 16:9 or 3:2 with possible amount of reduction in anomalies in video such as distortion, focus miss etc.

3. Three-step action search networks with deep Q-learning for real-time object tracking.

Pros:

- Provides an efficient and acceptably accurate method that is real-time to meet our demands of cropping into video in real-time.
- Works considerably faster than existing algorithms at a loss of acceptable accuracy.

Cons:

- The algorithm even though the best available at the moment does require significant computational power to run in real-time.
- The usage of CNN must be researched further and its implementations must be looked into.

4. Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies.

Pros:

- The advantages of eye tracking and the change it made in the modern world
- Giving a brief overview about the different kind of eye tracking techniques
- How can ML be an important part in eye tracking

Cons:

- The techniques which were good but were complex in structure.
- The usage of ML is highlighted but no examples were shown

5. Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision.

Pros:

- The methods to calculate the displacement of the objects were accurate.
- The implementation of AIOS algorithm with OTC and PU scheme.
- The validation examples of the experiments performed.

Cons:

- No works are mentioned about the extension of this method to space multi-DOF movement measurement.

Guide approval of selected papers:

S #	Name	Roll Number	Paper Title	URL	IEEE/ACM Publication	Year	Scopus Journal with impact factor/ Transaction/ conference	How is the selected paper relevant for your project?
1	Ch. Mounish	AM.EN.U4CSE18113	Video object detection with a convolutional regression tracker.	https://doi.org/10.1016/j.isprsjprs.2021.04.004	Elsevier	2021	**ISPRS Journal of Photogrammetry and Remote Sensing** Volume 176, June 2021, Pages 139-150	Video object detection is a fundamental problem that needs to be addressed as a part of our project, using a convolutional regression tracker bypasses the need to use a large number of video labels for correlated frames.

2	T. Vijay Tanmay	AM.EN.U4CSE18153	Deep neural networks for low cost eye-tracking.	https://doi.org/10.1016/j.procs.2020.09.041	Elsevier	2020	**Procedia Computer Science** Volume 176, 2020, Pages 685-694	Provides insights towards understanding of tracking moving objects based on eye movements. Techniques involved helped us in what to focus on and avoid for the project along with deeper understanding on object tracking.
3	K. Subhash Reddy	AM.EN.U4CSE18126	Three-step action search networks with deep Q-learning for real-time object tracking	https://doi.org/10.1016/j.patcog.2019.107188	Elsevier	2020	**Pattern Recognition** Volume 101, May 2020, 107188	A method that enables us to track an object in a bounded box effectively in real time.
4	M. Likith	AM.EN.U4CSE18136	Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies	https://doi.org/10.1016/j.eswa.2020.114037	Elsevier	2021	**Expert Systems with Applications** Volume 166, 15 March 2021, 114037	A deeper dive into object tracking algorithms with niche over Eye based tracking techniques and tools with emphasis over integration of Algorithms into hardware based solutions. Implementation being a core part, understanding the ways of doing it helped us better grasp over requirements.
5	M. Likith	AM.EN.U4CSE18136	Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision	https://doi.org/10.1016/j.ymssp.2019.01.046	Elsevier	2019	**Mechanical Systems and Signal Processing** Volume 124, 1 June 2019, Pages 111-123	Understanding the tracking prospects for collecting data. Prediction algorithms for better tracking and precision combined with missed tracking. Retro fitting the computer coordinates into real-time for better understanding and portability over the existing Frame of view.

1. Video object detection with a convolutional regression tracker.

Key Points:

- Designed a new video object detection pipeline to combine the advantages from both object detection and object tracking. With better bounding box proposals and linkages through time, we improve the performance with better effectiveness and efficiency.
- Tracker is light-weighted, memory efficient, and computationally efficient, as it re-uses the features from an object detector.
- Created an object tracker for the video object detection task, which can be easily inserted into a well trained image object detector.

Limitations:

- Major disadvantage is that without re-scoring images cannot be identified correctly as due to misinterpretation.

2. Deep neural networks for low cost eye-tracking.

Key Points:

- The region of interest being the eye of the subject helps maintain better data about the subject for enhanced tracking and observation.
- The ROI is much smaller compared to regular object based tracking solutions, leading to much lower pixel processing over each frame once the initial processing is done.
- The employance of gaze detection helps improve the tracking points by predicting next moves helping predict subject movements in real time and avoid false lock and drops.
- The translation from eye to computer coordinates help maintain a virtual view for better understanding of subjects position thus, drawing actuators when reached a certain threshold or condition.

Limitations:

- The major disadvantage comes with the labelling step required, which adds an additional computation step into the pipeline which never comes into the output.
- The frame shape for capture is limited to rectangular shapes for mapping coordinates and thus cant be applied to wider angle footage from drones and 360-degree cameras.
- Input sources with blemishes such as not in focus, or being hit by direct infrared source effects the performance of eye detection as eye-tracking is based on infrared input from pre-processing step.

3. Three-step action search networks with deep Q-learning for real-time object tracking.

Key Points:

- Sliding windows and candidate sampling are widely used to object tracking in videos, they are not real-time.
- Three step decision process achieves a faster (real-time) localization (framing) of the object in real-time.
- The three step action search network (TSAS) takes into consideration only three subsets from the region of interest (ROI), using fewer candidate locations achieves faster localization.

- A convoluted neural network agent is formulated to interact with the video.
- Three step action search network, starts from the centre of the target in the (t-1)th frame and moves a stride D, we have nine new locations to evaluate and the location with the most confidence is the origin for the tth frame. Now the process is repeated for two more iterations while halving the stride value.

Limitations:

- The applied method does seem to be quite computationally heavy, further optimisation is required to make it work on common hardware.
- The paper does not dive into CNN and Deep Q-Learning, further efforts in this front are necessary.

4. Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning Technologies.

Key Points:

- Eye tracking is the process of tracking the movement of the eyes to know exactly where a person is looking and for how long.
- The usage of different methods of eye tracking in software field.
- Importance of scleral search coil technique, infrared oculography (IOG), electrooculography (EOG), and video oculography (VOG).
- Conglomeration of ML and AI into Eye tracking which gives a precise result.

Limitations:

- The paper does not give a standard support to the four important techniques mentioned.
- The paper did not mention about any deployed models that use ML for eye tracking.

5. Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision.

Key Points:

- A vision based method for realtime in-plane displacements tracking of the precision positioning stage has been developed in this paper.
- An AIOS algorithm was developed in which an OTC strategy and a PU scheme were used for real time processing of the microscopic image sequences. The proposed method does not need artificial markers or special patterns.
- The method can be used as an alternative approach for the motion performance evaluation and close-loop control of the PPSs as well as a supplementary visual sensor for longterm micro-motion monitoring.

Limitations:

- No works are mentioned about the extension of this method to space multi-DOF movement measurement.

<u>Summarized survey of all the papers:</u>

S. No	Name of Student	Title of paper	Links
1	Ch. Mounish	Video object detection with a convolutional regression tracker.	Link
2	T. Vijay Tanmay	Deep neural networks for low cost eye-tracking.	Link
3	K. Subhash Reddy	Three-step action search networks with deep Q-learning for real-time object tracking	Link
4	M. Likith	Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies	Link
5	M. Likith	Realtime in-plane displacements tracking of the precision positioning stage based on computer micro-vision	Link