Technical Disclosure Commons

Defensive Publications Series

November 2019

Using Computer Vision To Label And Search A Physical Space

Hauke Heibel

Damon Kohler

Stefan Hinterstoisser

Martin Bokeloh

Jürgen Sturm

Using Computer Vision To Label And Search A Physical Space

ABSTRACT

Effective operation of a warehouse requires keeping track of the location of various assets within the physical environment. As various sensors are carried through the warehouse environment by operators, range data collected by the sensors over time can be used to reconstruct 2D and 3D representations of the space. This disclosure describes techniques to estimate the locations of Point-Of Interest (POIs) and Regions-Of-Interest (ROIs) within a physical environment such as a warehouse. The location estimates are generated using a combination of 2D visual search of images containing text labels and barcodes, 2D/3D environment reconstruction using sensor data, and estimated trajectory of sensors. Computer vision techniques are applied to visual data which is obtained from operational processes that generate images, such as feeds from stationary cameras, images from moving cameras, photos of the environment, etc.

KEYWORDS

- Computer vision
- Warehouse
- Indoor map
- Warehouse Management System (WMS)
- Autonomous Moving Robots (AMR)
- Points-Of-Interest (POIs)
- Regions-Of-Interest (ROIs)

BACKGROUND

Effective operation of a warehouse requires keeping track of the location of various assets, such as boxes, shelves, doors, forklifts, aisles, etc., within the physical environment. For instance, an operator searching for a specific box may be guided by signs above the aisles then along shelves and finally on the box itself. As various sensors, such as LIDAR, RGB, RGBD, time-of-flight (ToF), etc. are carried through the warehouse environment by operators, range data collected by the sensors over time can be used to reconstruct 2D and 3D representations of the space.

One example of such an approach is Simultaneous Localization and Mapping (SLAM). Geometric representation of the environment can take many forms, such as point cloud, mesh, etc. Additional semantic information important to understanding the environment and its contents can be attached to the geometric representation regardless of the specific form of the representation. The locations and regions within the physical space from which such semantic information is extracted are known as Points-Of-Interest (POIs) and Regions-Of-Interest (ROIs), respectively.

DESCRIPTION

This disclosure describes techniques to estimate the locations of POIs and ROIs within a physical environment such as a warehouse. Location estimates are generated using a combination of 2D visual search of images containing text labels and barcodes, 2D/3D environment reconstruction using sensor data, and estimated trajectory of sensors.

Visual data is obtained from operational processes that generate images, such as feeds from stationary cameras, images from moving cameras, photos of the environment, etc. Computer vision techniques are applied to each of the images to detect barcodes as well as text via Optical Character Recognition (OCR) techniques. Further, computer vision techniques are used to detect bounding boxes surrounding the text. These bounding boxes represent ROIs. Similar text detected in successive images from a given camera is used to mark corresponding bounding boxes along the sensor trajectory. The use of multiple images to detect POIs and ROIs can help improve accuracy via triangulation.

The SLAM approach can then be applied to estimate the position of the bounding boxes in 3D. Information associated with an image, such as the detected text or barcode, the location of the camera, the angle and field of view of the camera, etc. is fused to detect the location of the text or barcode label within the physical space. For instance, observing the same text in different image views can help localize it within the 3D space with reasonable accuracy resulting in a 3D representation of the environment that shows the locations of the labels located within it.

The POIs and ROIs detected via the techniques described herein can be cross-referenced with other systems, such as Warehouse Management Systems (WMS) used to manage the space and the assets it contains. The cross-referencing can enable Autonomous Mobile Robots (AMRs) operating within the environment to be guided by semantic POI and ROI target locations in addition to their typical operation that uses only the cartesian coordinates of the physical space. Moreover, the cross-referencing enables the WMS to be kept up to date with the physical reality as the semantic POI and ROI information is generated, detected, and updated continually as new input is acquired dynamically from the various data sources.

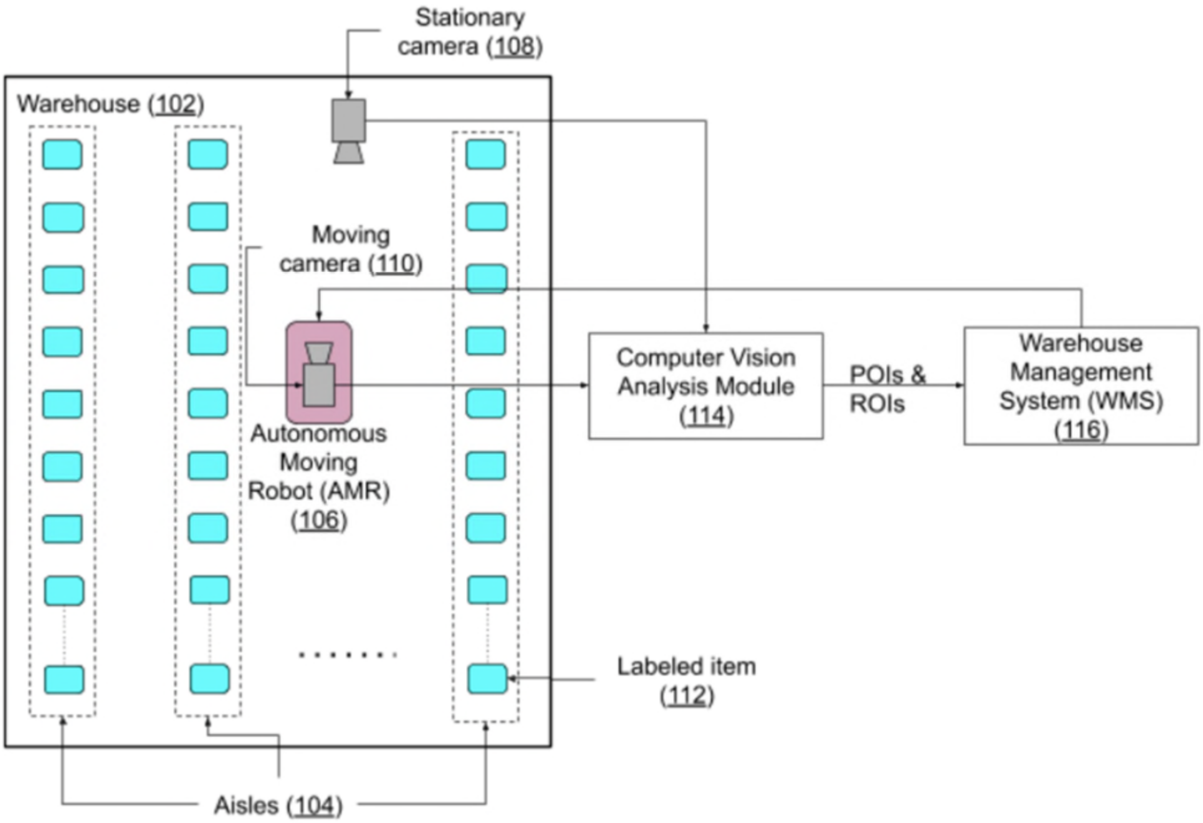

Fig. 1

Fig. 1 shows a two-dimensional representation of a physical space used as a warehouse (102). The warehouse includes several aisles (104) with labeled items (112) arranged on shelves within each aisle. Images of the warehouse space and the items within it are generated continually by stationary cameras (108) mounted within the space and/or moving cameras (110) mounted on Autonomous Moving Robots (AMR) (106) that operate within the warehouse to perform tasks such as placing or retrieving items. Images generated by the cameras are processed by a computer vision analysis module (114) to detect text and barcodes and generate POIs and ROIs of interest. The generated POIs and ROIs are cross-referenced with the corresponding information in the WMS (116). The cross-referenced semantic information of the POIs and ROIs is in turn used along with cartesian coordinates to help guide the AMRs operating within the warehouse.

The techniques described in this disclosure can improve the quality, accuracy, and user experience of locating items within a physical space, such as warehouses used in retail. Augmenting WMS information with POIs and ROIs within the space enhances the localization of the items in the environment and can make the space more readily searchable. In addition, the POIs and ROIs can enable AMRs operating within the space to navigate using semantic information in addition to their typical operation that relies on cartesian coordinates. Moreover, in the course of routine operation, the AMRs provide ambient data that can help automate keeping the WMS information up to date, thus improving on approaches that involve explicitly verifying the state of the warehouse at specific locations.

The described techniques can be applied in any physical space that requires locating and tracking physical objects, especially via AMRs. The techniques can also be extended to maps of the physical world that involve real-world views and augmented reality (AR) features.

CONCLUSION

This disclosure describes techniques to estimate the locations of Point-Of Interest (POIs) and Regions-Of-Interest (ROIs) within a physical environment such as a warehouse. The location estimates are generated using a combination of 2D visual search of images containing text labels and barcodes, 2D/3D environment reconstruction using sensor data, and estimated trajectory of sensors. Computer vision techniques are applied to visual data which is obtained from operational processes that generate images, such as feeds from stationary cameras, images from moving cameras, photos of the environment, etc. The use of multiple images to detect POIs and ROIs can help improve accuracy via triangulation. The different types of information associated with an image, such as the detected text or barcode, the location of the camera, the angle and field of view of the camera, etc., is fused to detect the location of the text or barcode label within the physical space. The POIs and ROIs detected via the techniques described herein can be cross-referenced with other systems, such as WMS. The techniques described in this disclosure can improve the quality, accuracy, and user experience of locating items within a physical space.