4th Construction Specialty Conference
4e Conférence spécialisée sur la construction

Montréal, Québec
May 29 to June 1, 2013 / *29 mai au 1 juin 2013*

Image Dataset Development for Construction Equipment Recognition

Humaira Tajeen[1] and Zhenhua Zhu[1]
[1]Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada

Abstract: The automatic recognition of various construction operational resources such as materials and equipment is always necessary to achieve the full automation in construction. Although many object recognition methods have been developed so far, the datasets used to validate these methods are limited to the few categories of objects in natural scenes. As a result, it is unclear whether the methods can be used to recognize the operational resources at construction sites. In order to fill this gap, this paper proposes to create a standardized dataset of construction equipment images to evaluate existing recognition methods. Specifically, thousands of images are taken from multiple construction sites. The images contain a wide range of construction equipment from different manufacturers, such as Caterpillar, Volvo, Deere, Komatsu, and Hitachi. In each image, the Equipment of Interest (EOI) has been annotated. The annotations include the type of the equipment and the labeling of various equipment components, such as bucket, stick, boom, cab, tracks, wheels, etc. The annotations can be used as a ground truth to test existing object recognition methods. The experiments show that the existing recognition methods can be evaluated in a standard, unbiased, and extensive way, when they are adopted for the recognition of construction equipment with the dataset developed in this paper.

1. Introduction

The construction industry has been transformed as one of the largest industrial sectors in Canada (Historica-Dominion, 2012). As a continuous growing industry, it has been searching for efficient methods which can improve the productivity and quality of construction work. Automation in construction has been proposed for that purpose (Demsetz, 1990). Automation, which means in general the replacement of human labor by control systems and machineries for the manufacturing of goods and services, has now been increasingly adopted in the construction industry (Demsetz, 1990). The potential of automation have been explored to overcome the prevailing problems of poor quality and low productivity work in the construction industry. Automation can improve the consistency of construction operations by reducing operation cycle time and performing the tasks that are beyond human capabilities in size, weight, speed, etc. Also, it can enhance the safety of construction workers by replacing them for difficult and tedious physical work and in hazardous construction environments. Thus, automation, if properly adopted, can increase the speed, consistency, safety and quality of construction work in the construction industry.

So far, much automation work has been developed in the construction industry, and a large portion of the work was created on the basis of using construction site images. The construction site images record the as-built status of the project under construction, and capture daily job site activities, which help construction engineers/managers monitor and control the sites remotely and dynamically. Take project productivity analysis as an example. The traditional analysis was executed manually, which was slow, inefficient and error-prone (Davidson and Skibniewski, 1995). The construction site images can be used to indicate the state of construction operational resources at the site, which makes the automated construction productivity analysis possible (Azar and McCabe, 2012; Gong et al. 2011).

In order to fully utilize construction site images for the automation of construction work, one critical step is to automatically recognize various construction objects from the site images, such as material, worker, and equipment. This is not an easy task, considering the fact that construction sites are characterized as being dirty, disorderly, and cluttered with tools, materials, and debris. Also, construction objects in the site images are typically occluded partially, which makes the recognition even more difficult and challenging.

Currently, there are many object recognition methods available. Most of them were developed by researchers in the field of computer vision. For example, Dalal and Triggs (2005) relied on the histograms of oriented gradients (HOG) for the recognition of human. Felzenszwalb et al. (2010) developed a discriminatively trained part-based model for the recognition of generic objects, such as bicycles, cars, etc. In order to measure the performance of existing recognition methods, several image datasets have been created. Typically, a dataset compiles thousands of object images. The objects in the images are first identified manually as the ground truth. The recognition results with the methods are compared with the manual identification results. This way, the recognition accuracy of the methods can be determined. All existing datasets that are publicly available to validate the effectiveness of existing recognition methods only contain limited categories of objects in natural scenes. As a result, it is unclear about the recognition performance of existing methods to recognize construction objects at construction sites.

This paper proposes to create a standardized dataset of construction equipment images to evaluate the construction equipment recognition performance of existing recognition methods. Construction equipment is one of main operational resources in executing construction tasks and is frequently involved in most of the construction operations (Azar and McCabe, 2011). The successful equipment recognition will provide a solid foundation to automate multiple equipment monitoring and control tasks. The process for creating the construction equipment image dataset follows two main steps. First, thousands of images taken from multiple construction sites are collected. The images contain different types of construction equipment from the manufacturers, such as Caterpillar, Volvo, Deere, Komatsu, and Hitachi. Then, the Equipment of Interest (EOI) in each image has been annotated. The annotations include the type of the equipment contained and the labels of corresponding equipment components, such as bucket, stick, boom, cab, tracks, wheels, etc. The annotations can be used as a ground truth to test existing object recognition methods with slight modifications. So far, an existing object recognition method (Felzenszwalb et al. 2010) has been implemented and tested with the dataset. The experiments show that the method can be evaluated in a standard, unbiased, and extensive way, when it is adopted for the recognition of construction equipment with the dataset developed in this paper.

2. Background

The recognition of objects from images has been considered as a challenging task. This is especially true for the recognition of three-dimensional (3D) objects in the world, since an object may have multiple poses, be partially occluded, and/or experience different environmental light conditions (Yang 2009; Ulrich and Steger, 2008). So far, several methods have been proposed for the recognition of 3D objects. Based on the recognition cues adopted, these methods can be broadly classified into three categories: (1) the geometry-based category, (2) the appearance-based category, and (3) the feature-based category (Yang 2009). The geometry-based recognition methods rely on the object shape, and other object properties, such as color and texture, are not used. The appearance-based methods typically consider the object surface reflectance properties as recognition cues, and the feature-based methods use the object visual features that are locally invariant (Matas and Obdrzalek, 2004). Currently, there are several datasets available to evaluate the performance of these methods for the recognition of the objects such as people, car, bicycle, etc.

2.1 Geometry-based Category

In the geometry-based methods, an object is represented by a model of 3D geometric primitives (e.g. boxes, spheres, cylinders, etc.) or 2D shapes/contours. The primitives, shapes, or contours are typically organized hierarchically. When such model is created, the recognition of an object can be performed by determining the geometric similarity between that object model and all the geometric information that can be retrieved from an image (Pope, 1994). So far, there are several methods that can be used to check the geometric similarity, including the hierarchical chamfer matching (Borgefors, 1988), geometric hashing

(Lamdan and Wolfson, 1988), and shape-based matching (Steger, 2001). Also, the similarity can be determined using the Hausdorff distance transform (Rucklidge, 1995) or generalized Hough transform (Ballard, 1981).

The geometry-based methods are robust for the recognition of the objects partially occluded or under cluttered background. They are invariant to lighting and pose variations (Matas and Obdrzalek, 2004). However, the effectiveness of the methods is heavily dependent on the reliable extraction of geometric primitives, but not all the primitives could be detected. For this reason, the geometry-based methods are typically restricted to use to recognize the objects that have easily identifiable components (Matas and Obdrzalek, 2004). Also, the geometry-based methods, in general, require high computational load and lead to long computation time (Ulrich and Steger, 2008).

2.2 Appearance-based Category

The appearance-based methods are developed following the idea of "remembering all possible appearances” of an object. Typically, the methods consist of two phases. In the first phase, an appearance model is constructed on the basis of a set of reference images that includes the object's different views under different orientation and illumination conditions. The second phase is the recall phase. In this phase, the parts of a test image are first extracted through image segmentation. Then, the recognition is performed by matching the extracted parts of the test image with the model (Matas and Obdrzalek, 2004). There are several appearance-based methods available. For example, Murase and Nayar (1995) relied on the image Eigen values to recognize objects with different viewpoints and illumination variation. Swain and Ballard used image histograms (Swain and Ballard, 1991). This way, object recognition is converted to the problem of matching two histograms. The effectiveness of the appearance-based methods has been successfully demonstrated when recognizing objects in the scenes without occlusions or under the black background (Nayar et al. 1996).

There are two main advantages about the appearance-based methods. First, the methods do not require any user-provided models (Matas and Obdrzalek, 2004). The models could be automatically generated from the training images. In addition, the methods are effective for the recognition of the objects under variable illumination and viewpoint conditions (Yang, 2009). The main limitation of the appearance-based methods lies in their sensitivity to object occlusions and cluttered background. Therefore, they are not always robust. Also, the methods suffer from a lack of invariance to similarity transformations such as scale or rotation. In order to be invariant to changes in illumination and viewpoint conditions, the recognition using the appearance-based methods requires dealing with all variations of the object appearance, which is computationally expensive (Dorkó and Schmid, 2004).

2.3 Feature-based Category

In the feature-based methods, an object is represented by its local features, such as the surface patches, corners, or other interest points with intensity discontinuity. The features are typically invariant to scale, illumination and affine transformation (Yang, 2009), which can be extracted with certain feature descriptors, including the Scale-Invariant Feature Transform (SIFT) (Lowe, 1999), the Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005), and the Speeded Up Robust Features (SURF) (Bay et al. 2006). When the features of an object are extracted, the object recognition is to locate the features of the object in a test image. Specifically, all the local features are extracted from the test image with the same feature descriptor that has been used to extract the object features. Then, the matches between the features from the object and the image are determined. If the number of matched features reaches a satisfactory level, the presence of the object in the test image is confirmed.

Under the feature-based methods, the presence of an object in an image can be estimated as long as a few key features of the object are matched to the image features. Therefore, the feature-based object recognition methods are robust, especially when objects experience occlusions or are under clutter background (Lowe, 1999). In addition, the object features can be automatically extracted and learned from a set of training images. There is no requirement for the user-provided models. Moreover, the objects can be recognized under an unknown background, and image segmentation is not necessary

(Matas and Obdrzalek, 2004). The reliance on the features that are invariant to scale, illumination and affine transformation, makes the methods can recognize objects, even if they are under varying viewpoint and illumination conditions (Matas and Obdrzalek, 2004).

2.4 Datasets Available for Evaluating Object Recognition

Although the tremendous progress has been made towards object recognition, most existing recognition methods are still sensitive to large illumination variations and heavy occlusions (Yang 2009). In order to evaluate the performance of existing object recognition methods, several datasets have been created, such as PASCAL VOC dataset, UIUC dataset, CALTECH dataset, MIT-CSAIL dataset, INRIA Person dataset, CMU PIE dataset, YALE dataset, and UMIST dataset. These datasets include a large number of image collections with the ground truth annotations of certain objects. For example, the PASCAL VOC dataset includes twenty visual object classes (i.e. person, bird, cat, cow, dog, horse, sheep, aeroplane, bicycle, boat, bus, car, motorbike, train, bottle, chair, dining table, potted plant, sofa, and TV/monitor) (Everingham et al. 2010). The UIUC dataset was developed by the researchers at the University of Illinois at Urbana-Champaign. So far, it only includes the cars with side views (Agarwal et al. 2004). The CALTECH dataset was developed at the California Institute of Technology. It consists of aeroplanes, cars, human faces, motorbikes, etc. (Fei-Fei et al. 2006). The INRIA dataset was created as a part of the research work in human detection, and therefore the images in the dataset are those of people at upright positions (Dalal and Triggs, 2005). Similar to the PASCAL VOC dataset, the MIT-CSAIL dataset contain multiple object classes (e.g. bicycle, bottle, apple, bookshelf, car, chair, desk, sofa, building, door, and window) and the scenes viewed from offices or at streets (Russell et al. 2008).

These publicly available datasets provide a common ground truth for evaluating the performance of existing object recognition methods, which drives the recent development of object recognition research. However, there are several issues restricting the use of the current datasets to evaluate existing methods for the recognition of construction equipment. First, the datasets only contain limited object classes, and none of them include construction equipment. Second, the images in the datasets provide a small range of variability regarding the position of the object in the image. The relative position and orientation of the object-of-interest with respect to the camera tend to be typical. An object tends to be centered in an image and presented in stereotype pose (Ponce et al. 2006). Finally, most images in the datasets have little or no occlusion and background clutter (Ponce et al. 2006). As a result, it is unknown whether or not existing recognition methods can be used for on-site construction equipment recognition. In order to answer this question, it is necessary to create a new dataset, which covers typical construction equipment under realistic site conditions: multiple pieces of equipment working together with illumination variation and partial occlusion by debris and materials.

3. Objective and Scope

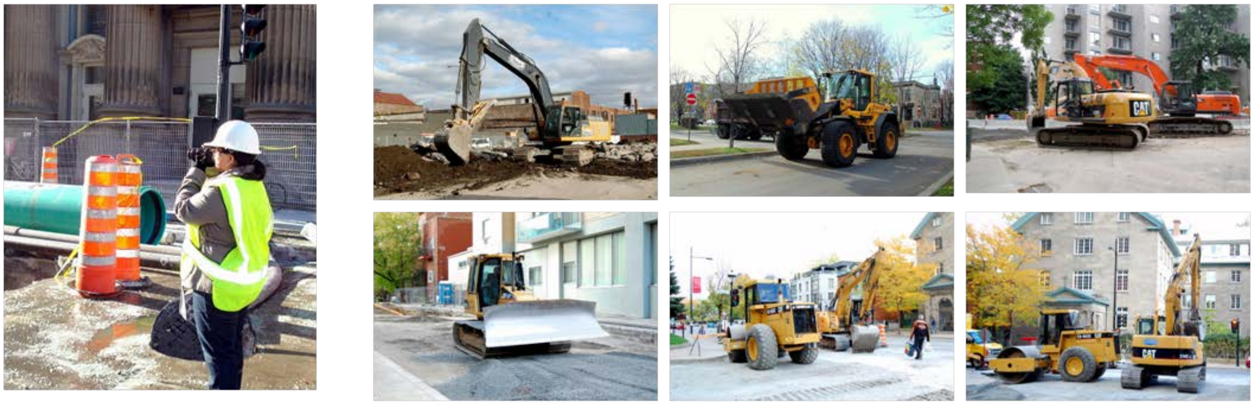
The main objective of this paper is to develop a dataset which comprises the images with different types of construction equipment. The equipment in the images are manually labeled as "excavator", "loader", etc. The labels establish the ground truth for construction equipment recognition. This way, the dataset can be used to evaluate the performance of existing object recognition methods, when they are implemented and applied to recognize construction equipment at construction sites.

In order to test the dataset effectiveness, the dataset is used to evaluate the performance of construction equipment recognition with an existing object recognition method (Felzenszwalb et al. 2010). The method was built upon the discriminatively trained deformable part models, which was awarded the PASCAL VOC "Lifetime Achievement" Prize in 2010 (Everingham et al. 2010). The results show the feasibility of using the dataset developed in this paper to measure the construction equipment recognition performance with existing object recognition methods.

4. Dataset Development for Construction Equipment Images

4.1 Image Collection

In order to create a construction equipment image dataset, multiple construction sites around Montreal have been selected as the image collection sources. A high-resolution camera, Nikon D40, which can produce an image with the maximum resolution of 3008 x 2000 pixels, was used in order to ensure good quality of images. The images of construction equipment were captured at real construction sites (Figure 1a). For example, the images were taken under different illumination conditions. The equipment has different poses. Multiple pieces of the equipment work together. One is partially occluded by another, or by debris and materials. Examples of the collected images are illustrated in Figure 1(b).



(a) Image collection

(b) Examples of collected images

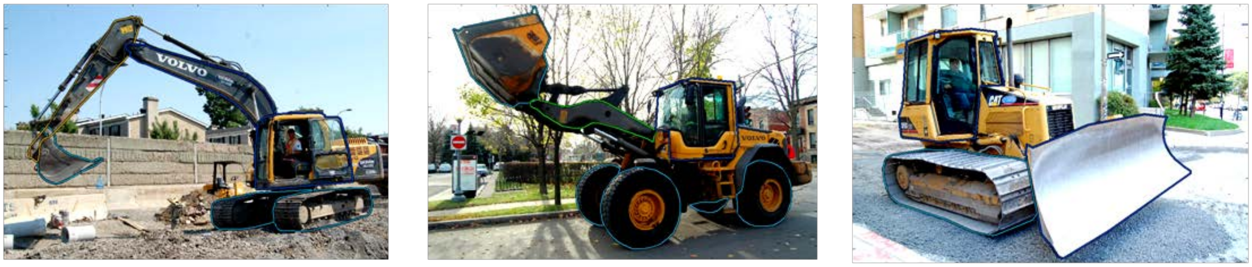
Figure 1: Image collection from construction sites

The EOI compiled in the image dataset is divided into 3 main categories: 1) excavating and lifting (excavator, backhoe), 2) loading and hauling (loader, dozer), and 3) compacting and finishing (roller, grader). For each class of equipment, hundreds or thousands of images were collected in order to ensure a wide range of image variations in terms of pose, viewing angle and illumination change. During the image collection process, the focus was placed on capturing the images with the following factors: 1) equipment from different manufactures (Caterpillar, Volvo, Deere, Komatsu, Hitachi, Case etc.); 2) different models and sizes within the same class of equipment (large, medium and small); 3) variation in equipment poses; 4) changes in viewing angles (front, rear, left, right, etc.); and 5) various illumination conditions (different period of day time). Construction equipment is commonly composed of various articulated parts. These parts undergo drastic pose variations under operation. Therefore, special attention was placed to capture the images of the EOI, when it is working.

4.2 Image Annotations

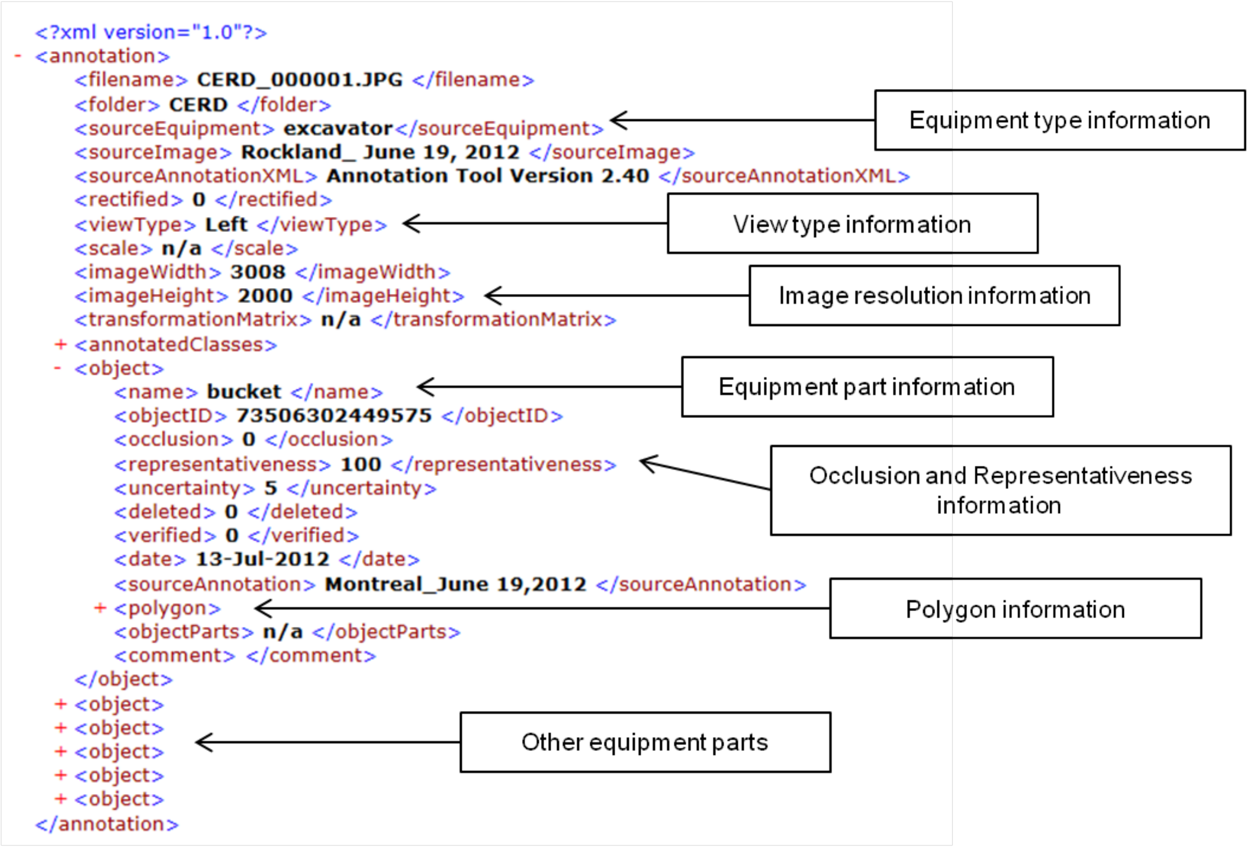
The images of construction equipment, collected from different construction sites, are compiled in the dataset. Each image is manually annotated. The annotation includes the identification of main components for each EOI. For instance, a typical excavator is composed of a bucket, stick, boom, cab on a rotating platform and undercarriage with tracks. A loader usually comprises a bucket, arm, cab and wheels. A roller generally contains a front roller, cab and wheels, and a dozer is composed of a blade, cab and tracks. Considering these components, the annotations for excavator, loader and dozer are illustrated respectively in Figure 2(a), (b) and (c).

In addition to the components, other information about the equipment is also included in the annotation. For example, the annotation indicates the type of the equipment, and the view of the equipment in the image. The occlusion and representativeness tags in the annotation means the percentage of the equipment part that has been occluded and the percentage of the equipment part that has not been truncated. Moreover, the annotation could be used to answer the questions like which construction equipment image is being annotated (image name), and what is the resolution (width and height) of the image. One example of the annotation files is illustrated in Figure 3.



(a) excavator (b) loader (c) dozer

Figure 2: Equipment parts identification with polygons



```
  <?xml version="1.0"?>
- <annotation>
      <filename> CERD_000001.JPG </filename>
      <folder> CERD </folder>
      <sourceEquipment> excavator</sourceEquipment>
      <sourceImage> Rockland_ June 19, 2012 </sourceImage>
      <sourceAnnotationXML> Annotation Tool Version 2.40 </sourceAnnotationXML>
      <rectified> 0 </rectified>
      <viewType> Left </viewType>
      <scale> n/a </scale>
      <imageWidth> 3008 </imageWidth>
      <imageHeight> 2000 </imageHeight>
      <transformationMatrix> n/a </transformationMatrix>
    + <annotatedClasses>
    - <object>
          <name> bucket </name>
          <objectID> 73506302449575 </objectID>
          <occlusion> 0 </occlusion>
          <representativeness> 100 </representativeness>
          <uncertainty> 5 </uncertainty>
          <deleted> 0 </deleted>
          <verified> 0 </verified>
          <date> 13-Jul-2012 </date>
          <sourceAnnotation> Montreal_June 19,2012 </sourceAnnotation>
        + <polygon>
          <objectParts> n/a </objectParts>
          <comment> </comment>
      </object>
    + <object>
    + <object>
    + <object>
    + <object>
    + <object>
  </annotation>
```

Equipment type information
View type information
Image resolution information
Equipment part information
Occlusion and Representativeness information
Polygon information
Other equipment parts

Figure 3: Example of an annotation file

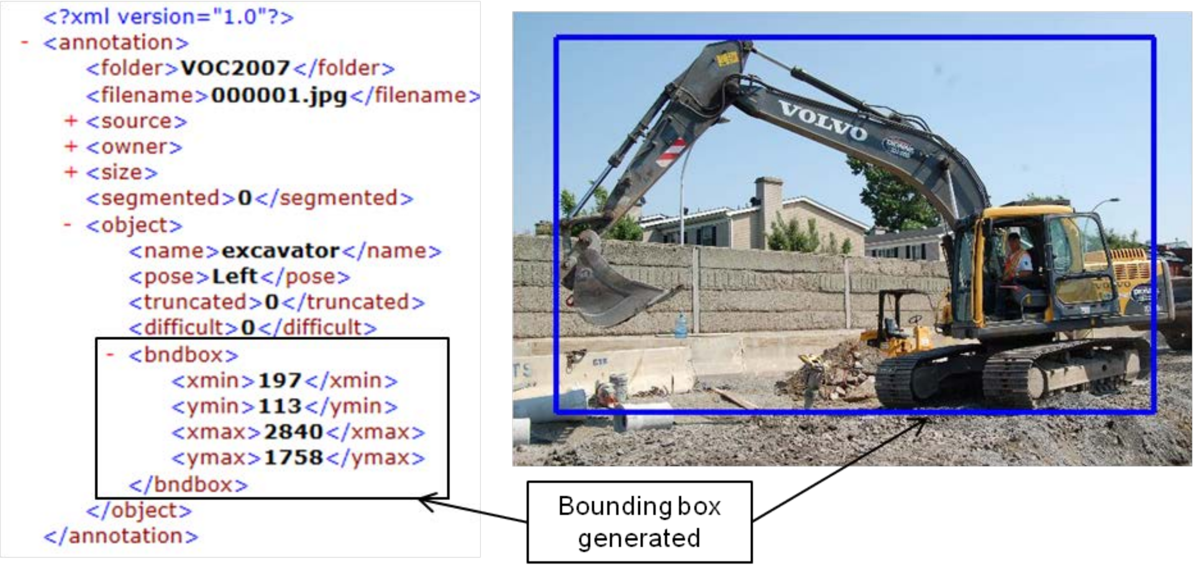
When the annotation files for all the collected construction equipment images are created, the dataset is organized into two folders. One folder is for the images and the other is for the annotations. Each image file in the image folder has its corresponding annotation file in the annotation folder, and vice versa. The relationships between the image file and annotation file are indicated by their file names. An image file and an annotation file will have the same name, if they are related. For example, when there is an image file ‘CERD_000001’, correspondingly, there is an annotation file ‘CERD_000001’.

5. Evaluation of Construction Equipment Recognition with the Dataset

So far, an object recognition method developed by Felzenszwalb et al. (2010) has been selected, and its performance for the recognition of construction equipment has been evaluated with the dataset created in this paper. The method requires the input of the images and the bounding boxes to indicate the equipment in the images as a ground truth. Under the method, the images and bounding boxes are first used to create the construction equipment recognition models through the supervised training. When the models are generated, they can be tested to recognize construction equipment for any given image.

5.1 Dataset Conversion

In order to meet the input requirements of the method developed by Felzenszwalb et al. (2010), slight conversions have to be made. The main idea of the conversion is to produce the new image annotation files that can be read by the method based on the annotation information contained in the dataset. Specifically, for each new image annotation file, the information about the equipment type and image (e.g. file name, image width, image height, etc.) is directly retrieved from the dataset and then transferred to the new file. As for the bounding box of the equipment, the polygons that represent equipment parts are extracted, and the coordinates of the polygon points are compared with each other. The comparison results indicate the maximum and minimum polygon point coordinates in x- and y- directions. This way, the top-left and bottom-right corners of the bounding box can be determined, and the bounding box can be created. Figure 4 shows one example of the annotation information conversion results. The left part is the annotation file produced based on the annotation information contained in Figure 3, and the right one shows the bounding box for the equipment in the image.



```
  <?xml version="1.0"?>
- <annotation>
      <folder>VOC2007</folder>
      <filename>000001.jpg</filename>
    + <source>
    + <owner>
    + <size>
      <segmented>0</segmented>
    - <object>
          <name>excavator</name>
          <pose>Left</pose>
          <truncated>0</truncated>
          <difficult>0</difficult>
        - <bndbox>
              <xmin>197</xmin>
              <ymin>113</ymin>
              <xmax>2840</xmax>
              <ymax>1758</ymax>
          </bndbox>
      </object>
  </annotation>
```

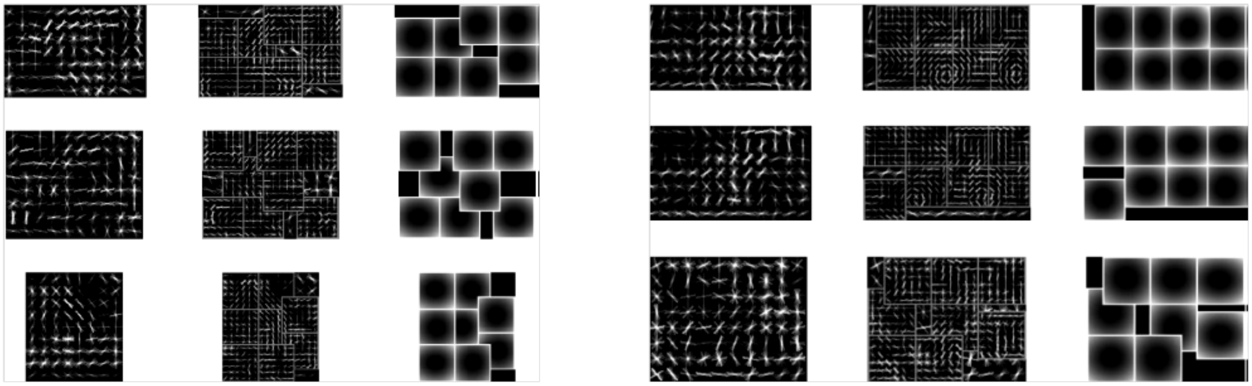
Bounding box generated

Figure 4: Annotation information conversion

5.2 Recognition Models Training and Recognition Results

When the annotation files in the dataset are converted, the dataset can be used by the method (Felzenszwalb et al. 2010) to train the models for construction equipment recognition. Each type of the construction equipment has its own recognition model. Figure 5 shows the examples of the recognition models for the excavator and loader, which are trained by the method with the dataset.

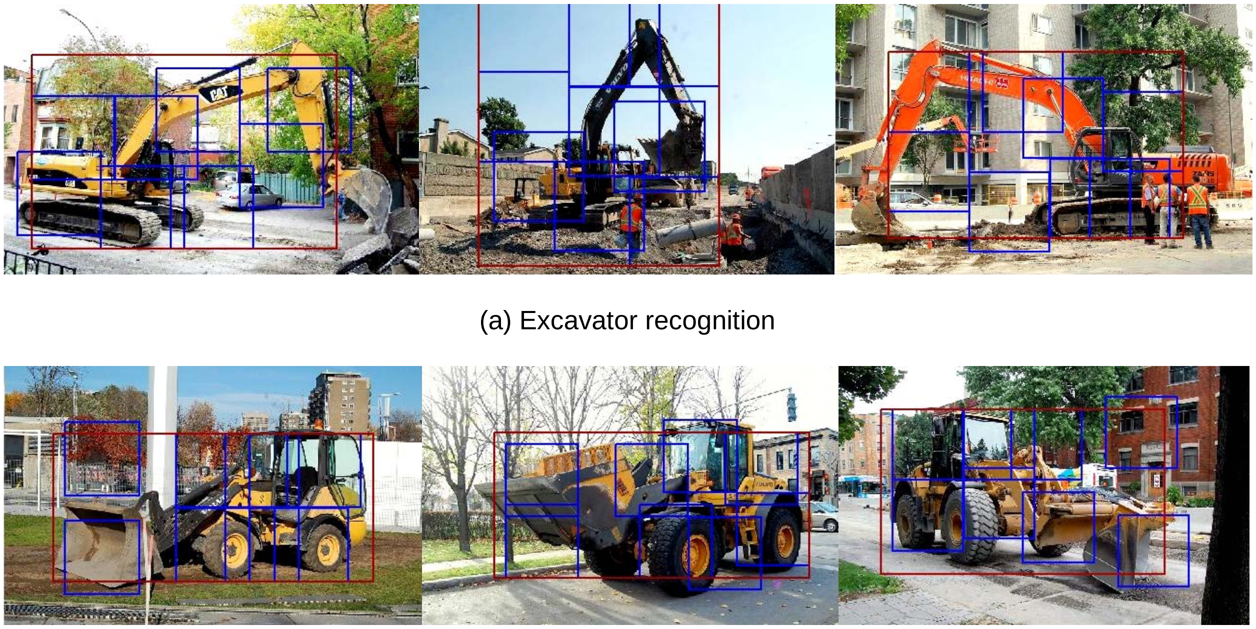
When the recognition models are generated and trained, the models can be used to recognize the construction equipment for a given test image. Figure 6 shows the examples of the recognition results using the trained recognition models to recognize the excavator and loader. In the recognition results, the equipment parts are first recognized by the corresponding equipment recognition model (blue boxes in Figure 6). Then, a bounding box is generated to cover the equipment (red boxes in Figure 6).



(a) Excavator recognition model

(b) Loader recognition model

Figure 5: Recognition models trained with the dataset



(a) Excavator recognition

(b) Loader recognition

Figure 6: Recognition of construction equipment

6. Conclusions and Future Work

The ultimate goal of construction automation is to enable an engineer or manager to control and manage construction tasks with full autonomous operations. In order to achieve this goal, it is necessary to automatically recognize various operational objects at construction sites (e.g. construction equipment). So far, there are many object recognition methods that have been developed in computer vision. The effectiveness of the methods has been evaluated by common datasets. However, existing datasets have several issues restricting their use to evaluate existing recognition methods to recognize construction equipment. As a result, it is unknown whether or not existing recognition methods can be used for on-site construction equipment recognition. In order to answer this question, it is necessary to create a new dataset, which covers typical construction equipment under realistic site conditions: multiple pieces of equipment working together with illumination variation and partial occlusion by debris and materials.

This paper creates a standardized dataset which contains thousands of construction equipment images with manual annotations. The dataset is expected to evaluate the construction equipment recognition performance for existing object recognition methods. The images in the dataset include a wide range of construction equipment from different manufacturers and with different poses, views, and degrees of occlusions. So far, one object recognition method has been tested with the developed dataset. The results show that the recognition method can be evaluated in a standard, unbiased, and extensive way. The future work will focus on testing more recognition methods with the dataset. A comparison between these recognition methods will be made to select an appropriate one for construction equipment recognition.

Acknowledgements

This work is in part supported by the SEED funding from the Vice President's Research and Graduate Studies (VPRS) at Concordia University, Montreal, Canada. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the VPRS.

References

Agarwal, S., Awan, A., and Roth, D. 2004. UIUC Image Database for Car Detection. *http://cogcomp.cs.illinois.edu/Data/Car/*

Azar, E. R., and McCabe, B. 2011. Automated Visual Recognition of Dump Trucks in Construction Videos. *Journal of Computing in Civil Engineering*, *26*(6): 769-781.

Azar, E. R., and McCabe, B. 2012. Vision-based Recognition of Dirt Loading Cycles in Construction Sites. *Construction Research Congress*, Purdue University, West Lafayette, IN, USA, 1042-1051.

Ballard, D. H. 1981. Generalizing the Hough Transform to Detect Arbitrary Shapes. *Pattern Recognition, 13*(2): 111-122.

Bay, H., Tuytelaars, T., and Gool, L.V. 2006. Surf: Speeded Up Robust Features. *Computer Vision-ECCV 2006*, *9th European Conference on Computer Vision,* Graz, Austria, 3951:404-417.

Borgefors, G. 1988. Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm. *Transactions on Pattern Analysis and Machine Intelligence, IEEE,* 10(6): 849-865.

Dalal, N., and Triggs, B. 2005. Histograms of Oriented Gradients for Human Detection. *IEEE Conference on Computer Vision and Pattern Recognition,* San Diego, CA, USA, 1: 886-893.

Davidson, I.N., and Skibniewski, M.J. 1995. Simulation of Automated Data Collection in Buildings. *Journal of computing in civil engineering,* 9(1): 9-20.

Demsetz, L. 1990. Automated Construction? *Construction Dimensions.*

Dorkó, G., and Schmid, C. 2004. Object Class Recognition using Discriminative Local Features. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

Everingham, M., Gool, L. V., Williams, C., Winn, J., and Zisserman, A. 2010. The PASCAL Visual Object Classes Challenge Workshop 2010. *http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2010/*

Fei-Fei, L., Fergus, R., and Perona, P. 2006. Caltech 101. *http://www.vision.caltech.edu/Image_Datasets/*

Felzenszwalb, P.F., Girshick, R.B., McAllester, D. and Ramanan, D. 2010. Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(9): 1627-1645.

Gong, J., Caldas, C.H. and Gordon, C. 2011. Learning and Classifying Actions of Construction Workers and Equipment using Bag-of-Video-Feature-Words and Bayesian Network Models. *Advance Engineering Informatics, 25*(4): 771-782.

Historica-Dominion. 2012. Construction Industry. *http://www.thecanadianencyclopedia.com*

Lamdan, Y. & Wolfson, H.J. 1988. Geometric Hashing: A General and Efficient Model-based Recognition Scheme. *2nd International Conference on Computer Vision*, Tampa, Florida, USA, 238-249.

Lowe, D.G. 1999. Object Recognition from Local Scale-Invariant Features. *7th International Conference on Computer Vision,* Kerkyra, Corfu, Greece, *2*: 1150-1157.

Matas, J. and Obdrzalek, S. 2004. Object Recognition Methods Based on Transformation Covariant Features. *12th European Signal Processing Conference (EUSIPCO 2004),* Vienna, Austria.

Nayar, S.K., Nene, S.A. and Murase, H. 1996. Real-time 100 object recognition system. *IEEE International Conference on Robotics and Automation, 3*: 2321-2325.

Ponce, J., Berg, T., Everingham, M., Forsyth, D., Hebert, M., Lazebnik, S., Schmid C., Russell B.C., Torralba, A., Williams C.K.I., Zhang J. and Zisserman A. 2006. Dataset issues in object recognition. *Toward category-level object recognition*, 29-48.

Pope, A.R. 1994. Model - based Object Recognition - A Survey of Recent Research. Technical Report TR-94-04. University of British Columbia.

Rucklidge, W.J. 1995. Locating Objects using the Hausdorff Distance. *5th International Conference on Computer Vision*, Boston, Massachusetts, USA, 457-464.

Russell, B.C., Torralba, A., Murphy, K.P., and Freeman, W.T. 2008. LabelMe: A Database and Web-based Tool for Image Annotation. *International Journal of Computer Vision, 77*(1): 157-173.

Steger, C. 2001. Similarity Measures for Occlusion, Clutter, and Illumination Invariant Object Recognition *Pattern Recognition*, 2191: 148-154.

Swain, M.J. and Ballard, D.H. 1991. Color Indexing. *International Journal of Computer Vision, 7*(1): 11-32.

Ulrich, M. and Steger, C. 2008. Performance Comparison of 2D Object Recognition Techniques. *Photogrammetric Computer Vision*.

Yang, M.H. 2009. Object Recognition. Liu L. and Ozsu M. T. (Ed.), *Encyclopedia of Database Systems,* 1936-1939.