2017 3rd International Conference on Control, Automation and Robotics

Multi Sensorial Data Fusion for Efficient Detection and Tracking of Road Obstacles for Inter-Distance and Anti-Colision Safety Management

Ahmad K. Aijazi, Paul Checchin, Laurent Trassoudaine
Institut Pascal, CNRS UMR 6602
Université Clermont Auvergne
Clermont-Ferrand, France
e-mail: kamalaijazi@gmail.com, paul.checchin@univ-bpclermont.fr, laurent.trassoudaine@univ-bpclermont.fr

***Abstract*—In this paper we present an automatic obstacle detection and tracking system for efficient inter-distance and anti-collision management that fuses both 3D LiDAR and 2D image data. The obstacles are first detected both in LiDAR scans and camera images and the data are then fused together. Even though LiDAR based detections are very accurate they are slower than image based detections. Hence, the proposed method helps in obtaining the state estimates more quickly with good accuracy. The unique fusion technique presented uses the detected object's geometrical information to extract the depth information at each image scan which is then corrected at each LiDAR scan. The results evaluated on real data demonstrate the prowess as well as the applicability of the proposed method which can be used for different vehicle safety applications**

Keywords-obstacle detection; tracking; data fusion; LiDAR; image processing; vehicle safety

I. Introduction

Efficient detection and tracking of road obstacles is fundamental for autonomous driving and vehicle safety applications such as obstacle avoidance, collision mitigation and inter-distance management. The different obstacle detection approaches need to run in (near) real-time so that necessary evasive actions can be performed. Recently, dense LiDAR based systems have been employed [1]-[3] for this purpose (e.g. Google's self driving car). Even though they are fairly accurate providing a rich 3D representation of the environment, they usually constitute high precision moving parts and acquire large amount of data, which make them computationally costly and difficult to process in near real time. On the other hand image (camera) based systems have also been employed [4], [5]. They are not only more cost effective, they also operate at a faster rate but they have a difficulty in dealing with changes in illumination and bad weather conditions etc. To improve reliability and accuracy of the obstacle detection systems, many researches focus on multi-sensor fusion approaches [6]-[8]. In order to provide better performances, we also propose in this paper an innovative fusion scheme.

In this work we present a method for an automatic anti collision distance management system for road highways which fuses both 3D LiDAR and 2D image data. The obstacles are first detected both in LiDAR scans and camera images. The high frequency 2D image data help accurately estimate the inter-distance between potential obstacles and the driving vehicle in between the relatively slower 3D LiDAR scans. This unique fusion technique uses the detected object's geometrical information to extract the depth information at each instance of the image based detection which is then corrected at each LiDAR scan. The resulting solution helps to better manage the safety inter-distance on the highways, especially at high speeds. The results are evaluated on real data to demonstrate its effectiveness as well as its technical strength.

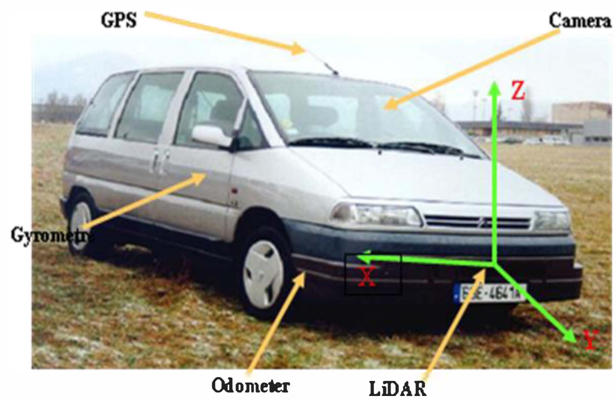


Figure 1. The experimental vehicle equipped with different sensors.

II. Detection of Road Lanes and Boundaries

In order to determine whether the detected obstacles offer a potential threat to the driving vehicle or not, one needs to localize them on the road. For this purpose it is imperative to determine the road boundaries as well as the road lanes. Where road boundaries help limit the detected objects to be analyzed as obstacles, the road lanes help determine which of these obstacles offer more of a threat. In order to detect road lanes and boundaries we use the method presented in [9]. The method combines three different algorithms including a lateral consistent detection based on extracted features and longitudinal consistent detection that uses M-estimators to estimate road shape.

These image based methods are quite effective however they can be hampered in case of occlusions due to road traffic, deteriorating road conditions and sometimes poor visibility due to bad weather conditions. So in our work we also use the proprioceptive information to get a rough prediction of the road ahead. This information is obtained with the help of on-board sensors like gyrometer and steering angle sensor as well as the odometer as shown in Fig. 1. In

978-1-5090-6088-7/17/$31.00 ©2017 IEEE

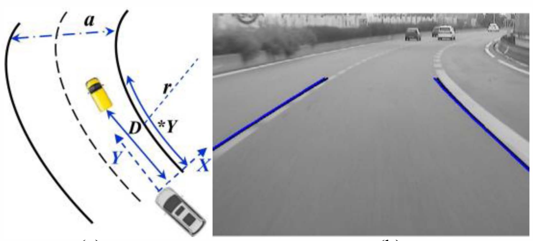
this method, each road segment is considered as an arc of a circle with center (r, 0) having a very large radius, and equal width (which is usually the case in highways and motorways). Considering the vehicle position in the center of one of the lanes, the road points (belonging to the boundaries as well as the center lanes) are extrapolated as shown in Fig. 2 using (1), (2) & (3):

$$\begin{cases} Yr_n &= \frac{1}{r}\left((r-\frac{a}{2})Y_n\right) \\ Xr_n &= r-\frac{1}{r}\left((r+\frac{a}{2})(r-X_n\right) \end{cases} \quad (1)$$

$$\begin{cases} Y_n &= Y_{n-1}+\overset{*}{y} \\ X_n &= r-\sqrt{r^2-Y_n^2}+X_{n-1} \end{cases} \quad (2)$$

$$\begin{cases} Yl_n &= \frac{1}{r}\left((r+\frac{a}{2})Y_n\right) \\ Xl_n &= r-\frac{1}{r}\left((r-\frac{a}{2})(r-X_n\right) \end{cases} \quad (3)$$

Here X_n, Y_n are the 2D road coordinates of the center road lane marking while Xr_n, Yr_n and Xl_n, Yl_n are that of the right and left respectively, in the vehicle frame of reference (shown in Fig. 1). 'a' is the road width and $^*y$ is the predicted distance. The Radius 'r' is given by $\frac{v}{\omega}$. Where ω is the angular velocity measured by the gyrometer and ν is the linear velocity measured by the odometer. If the vehicles encounter a turn in the opposite direction the radius becomes '$-r$' and the center ($-r$, 0).



(a) (b)

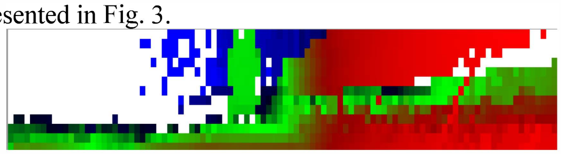
Figure 2. Detection and prediction of road boundaries and lanes ahead of the driving vehicle.

Once the road boundaries and lanes are determined, all obstacles detected within these are considered as potential threat and are further analyzed.

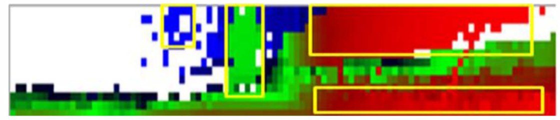
III. Detection & Tracking of Obstacles (Vehicles) on the Road

A. Obstacle Detection Using 3D LiDAR

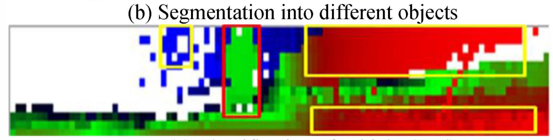
The objects on the road are detected by the 3D LiDAR scanner mounted on the vehicle as shown in Fig. 1. The objects are detected using the method proposed in [10]. In this method a depth based 3D image is first created and then segmented into objects using a region growing algorithm. These objects are then classified as vehicles and non vehicles using geometrical properties. The different steps are presented in Fig. 3.



(a) Construction of 3D image



(b) Segmentation into different objects



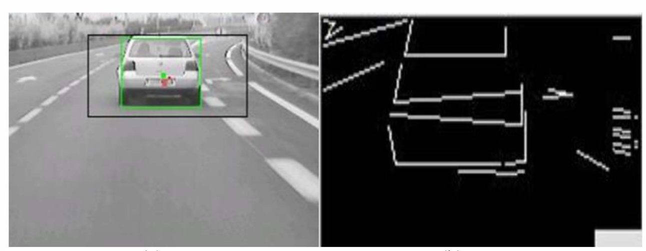
(c) Classification of vehicles(Red)

Figure 3. Detecting vehicles in 3D LiDAR scans.

Even though the detection of objects are quite reliable, for most high end 3D laser scanner, the scanning frequency of the scanner is relatively low followed by the processing time of the large amount of 3D data. As a result the detection frequency based on the LiDAR alone may not always be sufficient for obstacle detection and inter-distance management on highways, especially at fast speeds. For example in our case the vehicle detection rate (acquisition & processing) is about 400ms. This implies that at a speed of 130km/hr the detection is every 14m which may not be enough to avoid accidents in certain situations. This is why we fuse 2D image data.

B. Obstacle Detection Using 2D Image Data

In order to detect the same vehicles in the 2D images obtained from the onboard camera as shown in Fig. 4, we first select a region of interest (ROI) around each vehicle. The camera image and the 3D LiDAR image are matched together using extrinsic calibration and homography [11]. The center of the object detected in the 3D scan is projected onto the 2D image. This serves as the initial center point of the object in the 2D image. Based on this center an ROI (2 x object dimension) is initialized. Inside this ROI we use a Canny edge detector as presented in [12] to segment out line segments. These characteristic segments, corresponding to the back of the detected vehicle, include a number of horizontal segments with few vertical segments joining them from the ends, as shown in Fig. 4.



(a) (b)

Figure 4. Detecting vehicles in 2D camera image. In (a) black box represent the ROI where as the Green box represent the Bounding box. Red and green spots are the object centers detected by LiDAR and camera respectively. (b) Shows the line segments within the ROI.

These segments are then grouped/linked together to form characteristic polygonal shaped objects and then classified as vehicles using geometrical information. Every time the vehicle is detected in the image, the center and the ROI are updated for the next image.

Usually the image based detection is much quicker (acquisition & processing) for example in our case about 40ms. However, the lack of depth information makes it difficult to estimate the safety distance necessary for our application. In order to obtain the best of both methods we fuse the information from both sensors.

C. Obstacle Tracking: Fusion of 3D-2D Data

When driving on the road all potential obstacles needs to be continuously monitored till the threat is over. For this purpose, we need to continuously track each of these detected obstacles to ascertain if the status of a certain obstacle changes from potential to imminent threat or vice versa and so as to take effective safety measures. Efficient tracking is also necessary in road scenarios as due to traffic sometimes there are intermittent occlusions and so the track of obstacles is lost. So in order to track these detected obstacles we use an extended Kalman filter as also presented in [13]. A constant velocity model with acceleration as noise is considered.

The state vector is given as $x = [\,x \;\; Vx \;\; y \;\; Vy \;\; z \;\; Vz\,]^T$ where x, y, z and Vx, Vy, Vz are positions and velocities of the detected obstacle respectively. The measurement vector is given as $z = [\,u \;\; v \;\; y_I\,]^T$. Where u and v are the obstacle position (in pixels) in 2D image while y_I is the inferred distance.

The estimation $\mathbf{X}_\oplus$ and covariance $\mathbf{P}_\oplus$ at every stage is given as:

$$\mathbf{x}_\oplus = \mathbf{x}_\ominus + \mathbf{W}\nu \tag{4}$$

$$\mathbf{P}_\oplus = \mathbf{P}_\ominus - \mathbf{W}\mathbf{S}\mathbf{W}^T \tag{5}$$

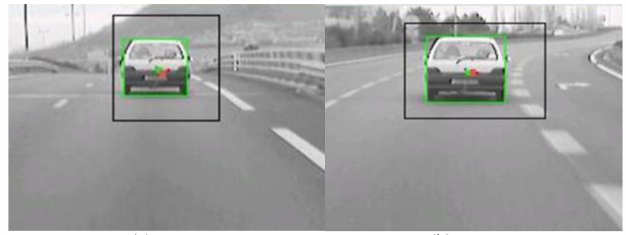
where innovation $\nu = (\mathbf{z} - \mathbf{H}\mathbf{x}_\ominus)$ and the covariance $\mathbf{S} = \mathbf{H}\mathbf{P}_\ominus\mathbf{H}^T + \mathbf{R}$. Here $\mathbf{W}$ is the gain while $\mathbf{H}$ is the Jacobean of the measurement matrix $\mathbf{H}_\mathbf{X}$.

In the measurement vector $\mathbf{Z}$, the distance 'y_I' needs to be inferred as in images there is no depth perception. We can either calculate this distance using a constant velocity model along with its variance. However, a simple prediction may not be the best choice. This is why we calculate the distance 'y_I' in terms of the size of the obstacle as viewed in 2D.

So in order to calculate the distance 'y_I' of the obstacle we use the evolution of their geometrical size as perceived by the driving vehicle in each camera image. Every time the obstacle is detected in the 3D LiDAR scan, not only the position and velocities are corrected but also the real distance is associated with the size of the bounding box of the obstacle detected in the 2D image. Then until the next LiDAR scan is available, the distance 'y_I' is estimated based on the changing size of the bounding box as expressed in (6).

$$y_I = f(\text{3D object distance, 2D size of bounding box}) \tag{6}$$

As the obstacle (vehicles) is a non-deformable body i.e. the size remains inversely proportional to the distance as shown in Fig. 5. The algorithm is summarized in Fig. 6.



(a) (b)

Figure 5. The change in size of the bounding box of the detected vehicle (obstacle) is shown. Red & green spot shows the obstacle center detected by LiDAR and camera image respectively. In (a) the obstacle is far and so the size of the bounding box is small where as in (b) as we approach the obstacle the size of the box increases.

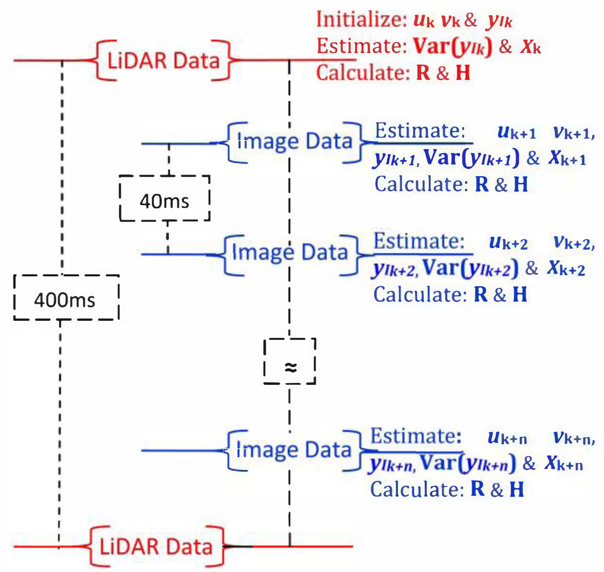


Figure 6. An overview of the Algorithm

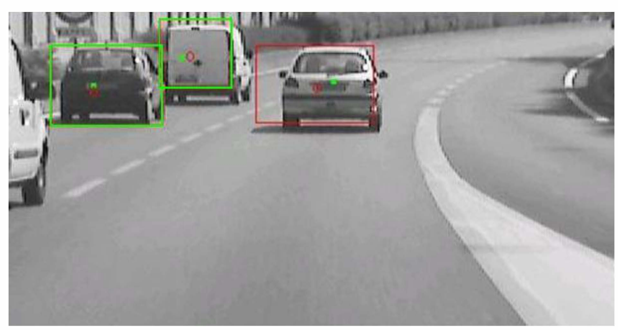


Figure 7. Detection and tracking of multiple obstacles. In red we have immediate threat while in Green we have potential threats.

This unique fusion of both 2D-3D data helps in the accurate estimation of inter-distances of potential threats at a fast rate. An obstacle moving in the same lane with an inter-distance corresponding to collision time < 2s, is considered as imminent threat and prompts an appropriate safety measure where as all other obstacles (even moving in the neighboring lanes), considered as potential threats, are also closely monitored/analyzed as shown in Fig. 7. The collision time is given as: $C_t = \frac{D}{V_r}$ where D is the estimated distance while V_r is the relative velocity.

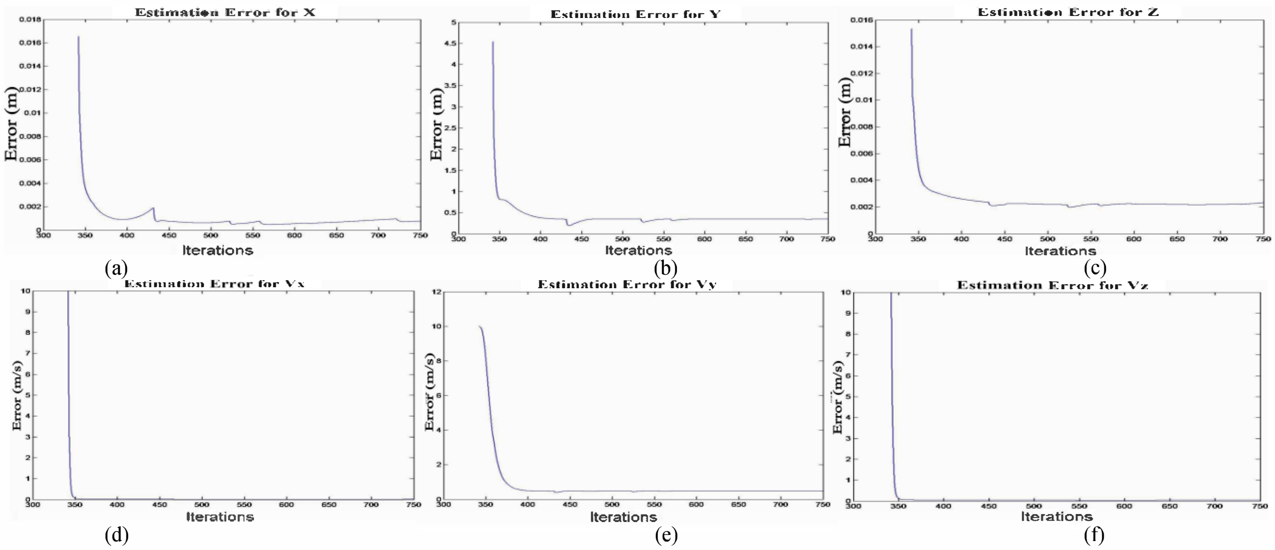


Figure 8. (a), (b) & (c) show the estimation error for position while (d), (e) & (f) present the estimation error for velocities.

IV. Experiments and Results

The proposed method was validated on real data obtained from our experimental vehicle (see Fig. 1) driving (about 40km) on the A-75 Auto-route in France. The combined detection results are presented in Table I using different standard evaluation metrics as described in [14]. Ground truth was obtained by manually annotating via visual inspection.

TABLE I: Detection Results Expressed in Different Evaluation Metrics

ACC	Accuracy	0.919
PPV	Positive Predictive Value	0.969
NPV	Negative Predictive Value	0.899
FDR	False discovery Rate	0.031
F_1	F_1 measure	0.915

The value of NPV< PPV shows that there are sometimes more false negatives (FN). A large number of these detection failures were due to complete or partial occlusions in case of heavy traffic. This highlights the importance of the tracking phase.

The tracking results for a particular obstacle are presented in Fig. 8. The figure clearly shows that the distance as well as the velocities of the obstacle (*x, y, z, Vx, Vy, Vz*) in our experimental vehicle's frame of reference are very well estimated. The results converge very quickly (<10 iterations) with an error of 0.0008 m to 0.4 m for the distances *(x, y, z)* and between 0 and 0,3 m/s for the velocities (*Vx, Vy, Vz*). We also find negligible difference between the estimated and the real value of 'y_I'. The improved estimate/inference using our method (using (6)) over the standard constant velocity model based prediction can be clearly seen in Fig. 9. The results clearly demonstrate the efficacy as well as the utility of the proposed method.

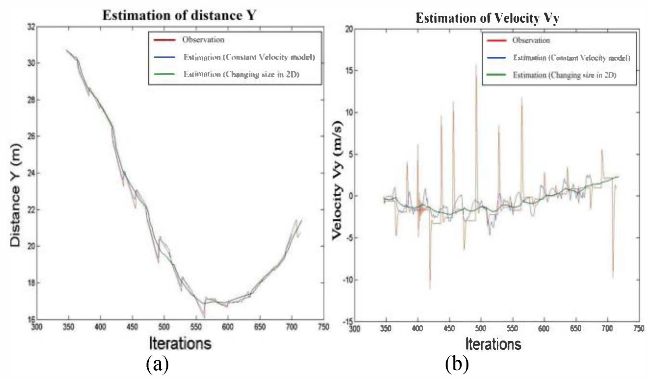


Figure 9. (a) & (b) show the estimation of obstacle distance Y and velocity Vy for both prediction and calculation based method respectively.

V. Conclusion

In this work we present an automatic obstacle detection and tracking method that fuses information from both LiDAR and camera to obtain state estimates of multiple of multiple obstacles in the road environment not only more quickly but also more accurately. The obstacles are detected both in LiDAR scans and camera images and the data is then fused together. The road lanes and boundaries are first determined/extracted to not only limit the number of obstacles to be detected and tracked but also to help better classify them into imminent and potential threats, based on their position on the road and estimated collision time.

The results evaluated on real data not only demonstrate the technical strength but also the utility of the proposed method for different vehicle safety applications such as inter-distance management and anti-collision mitigation.

References

[1] D. Streller, K. Furstenberg, and K. Dietmayer, "Vehicle and objectmodels for robust tracking in traffic scenes using laser range

images,"in Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on, 2002, pp. 118–123.

[2] Carlino, A.; Altomare, L.; Darin, M.; Visintainer, F. & Marchetto, A. Schulze, T.; Müller, B. & Meyer, G. (Eds.) Automotive LIDAR-Based Strategies for Obstacle Detection Application in Rural and Secondary Roads. Advanced Microsystems for Automotive Applications 2015: Smart Systems for Green and Automated Driving, Springer International Publishing, 2016, 89-100

[3] Asvadi, Alireza, et al. "3D Lidar-based static and moving obstacle detection in driving environments: An approach based on voxels and multi-region ground planes. " Robotics and Autonomous Systems 83 (2016): 299-311.

[4] M. Qing, V.-D. Hoang, and K.-H. Jo, "Localization and tracking of same color vehicle under occlusion problem," in Mecatronics-REM, 2012, pp. 245–249.

[5] P. Lenz, J. Ziegler, A. Geiger, and M. Roser, "Sparse scene flowsegmentation for moving object detection in urban environments," inIntelligent Vehicles Symposium (IV), 2011.

[6] C. Wang, H. Zhao, F. Davoine, and H. Zha, "A system of automated training sample generation for visual-based car detection," inIntelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, Oct 2012, pp. 4169–4176.

[7] G. Alessandretti, A. Broggi, and P. Cerri, "Vehicle and guard rail detection using radar and vision data fusion," Intelligent Transportation Systems, IEEE Transactions on, vol. 8, no. 1, pp. 95–105, March 2007.

[8] L. Kurnianggoro, Wahyono, D. C. Hern´ndez, and K.-H. Jo, "Camera and laser range finder fusion for real-time car detection," in International Conference on Industrial Electronics (IECON), 2014, pp. 3419–3424.

[9] Douret, J., Labayrade, R., Laneurit, J., & Chapuis, R. (2005, May). A Reliable and Robust Lane Detection System based on the Parallel Use of Three Algorithms for Driving Safety Assistance. In MVA (pp. 398-401).

[10] C. Blanc, L. Trassoudaine, Y. LeGuilloux, and R. Moreira " Track to track fusion applied to road obstacle detection," in 7th International Conference on Information Fusion, Stockholm, June 28 1st july 2004.

[11] Ciclo, X. X. V. "Laser and Camera Intercalibration Techniques for Multi-Sensorized Vehicles." (2013).

[12] M. Xie, L. Trassoudaine, J. Alizon, M. Thonnat, and J. Gallice "Active and intelligent sensing of road obstacles : Application to the european Eureka Prometheus project". In 4th International Conference on Computer Vision, Berlin, May 1993.

[13] Todoran, Horaţiu George, and Markus Bader. "Extended Kalman Filter (EKF)-Based Local SLAM in Dynamic Environments: A Framework." Advances in Robot Design and Intelligent Control. Springer International Publishing, 2016. 459-469.

[14] Vihinen, M.: How to evaluate performance of prediction methods? Measures and their interpretation in variation effect analysis. BMC Genomics 13 (2012) S2