

Visión por computadora en un robot móvil tipo oruga

Uriel González Escalona, Elsa Rubio Espino, Juan Humberto Sossa Azuela

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Ciudad de México,
México

a150158@sagitario.cic.ipn.mx, {erubio,hsossa}@cic.ipn.mx

Resumen. En este artículo se realiza un estudio comparativo entre tres tipos diferentes de detectores y descriptores de rasgos sobre un robot móvil tipo oruga. Los detectores y descriptores son SIFT, SURF y ORB. Las métricas utilizadas como comparación son: tiempo de procesado, resistencia al ruido e invarianza a la rotación. Los experimentos muestran que la técnica de ORB es la más rápida obteniendo resultados similares en la resistencia al ruido y a la invarianza a la rotación. Por lo cual, para robots móviles utilizados en ambientes no estructurados, donde la velocidad de procesado es un factor primordial, esta técnica es la más recomendable.

Palabras clave: robot oruga, visión por computadora, ORB.

Computer Vision in Mobile Tracked Robot

Abstract. In this article a comparative study is carried out between three different types of detectors and feature descriptors on a mobile tracked robot. The detectors and descriptors are SIFT, SURF and ORB. The metrics used as a comparison are: processing time, resistance to noise and invariance to rotation. The experiments show that the ORB technique is the fastest obtaining similar results in the resistance to noise and the invariance to the rotation. Whereby, for mobile robots used in unstructured environments, where the speed of processing is a prime factor, this technique is the most recommendable.

Keywords: tracked robot, computer vision, ORB.

1. Introducción

La visión por computadora fue diseñada para poder simular el sistema de visión humano, donde una de las tareas principales es el reconocimiento del entorno. Lo anterior, significa que es posible saber que objetos se pueden localizar

en una imagen, para lo cual existe una gran diversidad de técnicas que tratan de resolver este problema.

Con el avance de las investigaciones se desarrollaron conceptos como extracción de esquinas, detectores de bordes, reconocimiento de texturas, etc. Estos conceptos ayudaron a mejorar el entendimiento que puede tener una computadora acerca de lo que es un objeto, y por lo tanto poder reconocer el entorno que se le muestra. Para poder extraer estas características se desarrollaron técnicas capaces de distinguirlas y a su vez, de ser lo menos invariantes a aspectos como son el ruido y los cambios de iluminación.

Los robots móviles que utilizan la visión por computadora como herramienta para poder moverse e interactuar con su ambiente, deben poseer algoritmos que sean rápidos y eficientes, ya que se procura que el procesamiento de la imagen se realice en tiempo real [11]. Por lo cual, se debe especificar qué características del entorno son necesarias y cuáles pueden ser desechadas ahorrando tiempo de procesamiento.

2. Trabajo previo

2.1. Visión por computadora

Se han realizado diferentes estudios sobre la eficiencia de los extractores de rasgos algunos de ellos son SIFT, SURF, FAST, ORB, CenSurE, START donde se busca observar la eficiencia, tiempo de procesado y resistencia al ruido [9,6].

Los robots móviles para poder realizar tareas de forma tele-operada o de forma autónoma, necesitan de herramientas que le permitan reconocer el entorno en el que se encuentran, si el operador no puede observar el entorno y reaccionar a este, se le adapta al robot técnicas de visión por computadora que le permiten reconocer el entorno, estas técnicas necesitan ser procesadas en el menor tiempo posible para poder ejecutarse y dar al robot los elementos necesarios para su funcionamiento [8,17].

2.2. Robot móvil tipo oruga

Los avances de la tecnología han hecho posible el diseño y construcción de robots móviles para un amplio rango de aplicaciones. Estos sistemas son diseñados para facilitar algunas tareas como son la inspección y reconocimiento de lugares remotos, donde el uso de los robots liberan al operador a exponerse a situaciones de riesgo.

Los robots tipo oruga de todo terreno, integran bandas laterales para su desplazamiento, ofreciendo una mejor tracción que los robots de ruedas, especialmente en terrenos no estructurados como son la arena y la grava. Por ejemplo, el robot utilizado para la inspección de túneles cableados [19], el cual es un robot capaz de subir peldaños de hasta su propia altura, gracias a la tracción que ofrecen las orugas.

De igual forma, algunas investigaciones han analizado el uso de estos tipos de robot [12], abriendo el debate sobre los tipos de locomoción y las superficies en

donde se desplazan, lo cual ha demostrado las ventajas de los robot tipo oruga sobre otros robots con ruedas.

El uso de sensores incrementa el desempeño y fomenta el diseño de algoritmos avanzados, permitiendo al robot navegar de forma autónoma u optimizar complejas trayectorias de desplazamiento reduciendo así, el tiempo y la energía utilizada [8,17].

3. Detectores y descriptores de rasgos

3.1. Transformación de características invariable de escala

ransformación de características invariable de escala (Scale Invariant Feature Transform, SIFT) es un método para extraer características distintivas de una imagen. Este método, fue publicado por David G. Lowe en 2004, y es invariante a escalamiento y rotación; puede utilizarse para realizar una relación entre diferentes vistas de un objeto o escena [10], para lo cual utiliza cuatro principales aspectos:

- Scale-space extrema detection: el cual como primer paso, busca sobre todos las escalas y rotaciones de las imágenes para identificar puntos de interés que sean invariantes a escala y rotación.
- Keypoint localization: donde a cada punto candidato se le asigna un modelo que designa la localización y escala.
- Orientation assignment: una o mas orientaciones son asignadas a cada punto clave basado en su dirección de gradiente.
- Keypoint descriptor: los gradientes locales de la imagen son medidos a una escala seleccionada de la región que rodea al punto clave.

3.2. Rasgos robustos y acelerados

Rasgos robustos y acelerados (Speeded up robust features, SURF) es un detector y descriptor de puntos invariantes a escala y rotación, el cual utiliza convoluciones integrales de imagen por imagen usando como base la matriz Hessiana para la identificación de los detectores, además del uso de un descriptor basado en la distribución de los datos [16].

Para poder ser invariante a la rotación se identifican orientaciones posibles del punto de interés, al cual se le calculan las respuestas Haar-wavelet en las direcciones X y Y en un radio de 6 pixeles alrededor del punto, las cuales son representadas como vectores.

En la extracción del descriptor, se construye una región cuadrada centrada en el punto de interés con la orientación previamente asignada con un tamaño de ventana de 20 pixeles; posteriormente se subdivide en regiones de 4×4 pixeles a las cuales se le aplica un vector de descripción.

3.3. Oriented FAST and rotated BRIEF

El detector y descriptor de puntos característicos ORB [15], se desarrolla con la idea de poder tener un descriptor que sea mas rápido y utilice menos procesamiento que SIFT teniendo un rendimiento parecido, para lo cual se utiliza el detector de puntos característicos FAST [13,14], el cual utiliza un círculo de 16 pixeles al rededor del pixel a clasificar. En la Fig. 1, se observa la intensidad de los 16 pixeles comparada con el pixel a clasificar.

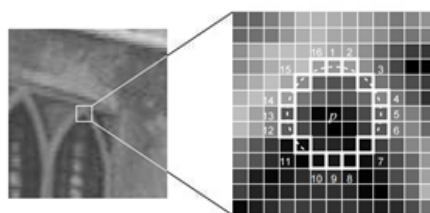


Fig. 1: Círculo de Bresenham.

Una vez que se han detectado las esquinas en la imagen, se utiliza BRIEF [3] para asignar un descriptor. Este descriptor es un código binario. La ventaja de utilizar BRIEF es que al generarse un código binario este puede compararse para observar similitud utilizando la distancia de Hamming, la cual es computacionalmente óptima [2].

4. Uso de la visión por computadora en robots móviles

Los robots móviles tele-operados utilizados para la inspección, planeamiento de rutas o recolección de objetos, necesitan poder buscar de forma eficiente los obstáculos y el terreno viable para moverse, para lo cual el uso de técnicas de visión por computadora son muy necesarias [7,1]. El desempeño de este tipo de robots, es primordial ya que algunas tareas específicas deben realizarse con el menor número de fallos posibles [18]; todo esto para poder realizar el procesamiento a la velocidad requerida [4,5]. En un robot la capacidad de procesamiento está limitada a su sistema embebido, por lo cual los algoritmos que se utilicen deben ser lo menos complejos posible.

5. Experimentos

El robot usado para este trabajo es un robot todo terreno con un chasis totalmente metálico (Fig. 2). Tiene seis ruedas con suspensión independiente que mejora la tracción y la absorción de impactos. El robot contiene un espacio interior donde se alojan los componentes como son los motores eléctricos, el

sistema embebido Raspberry PI, una cámara OV5647 de 5 MPx y un sensor ultrasónico HC-SR04, los cuales son utilizados para poder controlar al robot de forma remota. Además se adaptó al robot un brazo robótico con 4 grados de libertad, el cual permite al robot poder tomar los objetivos en el ambiente de búsqueda.

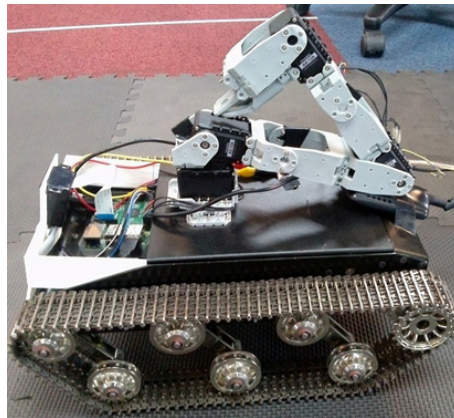


Fig. 2: Robot móvil tipo oruga T-Rex.

Las características principales del robot y el entorno de búsqueda en el cual se desempeña se muestran en la Tabla 1.

Los objetivos como el mostrado en la Fig. 3, son objetos rectangulares de tamaño 10×5 centímetros que se precargan en el sistema para su posterior búsqueda en el entorno, el objeto se encuentra sobre el terreno de prueba del robot el cual deberá detectar al objeto por medio de sus puntos característicos.



Fig. 3: Figura de muestra.

El ejercicio consiste en la activación remota del robot, el cual podrá buscar de forma autónoma el objetivo seleccionado en un rango entre 12 y 60 centímetros, poder tomarlo con el brazo robótico y poder posteriormente localizar el depósito y dejar el objeto.

Tabla 1: Características del robot tipo oruga T-Rex.

Parámetro	Valor
Velocidad	0.5 m/s
Ángulo de visión	60 grados
Peso	4.5 Kg
Comunicación	Wi-Fi
Objetivo	
Ángulo de búsqueda	180 grados
Distancia de búsqueda	60 centímetros
Peso máximo	100 gramos
Depósito	
Ángulo de búsqueda	360 grados
Distancia de búsqueda	60 centímetros

Se realizaron pruebas para determinar el tiempo promedio de búsqueda del objeto en un ambiente de trabajo real (Fig. 4), el cual consiste en terrenos no preparados, donde la superficie de movimiento puede ser arena, tierra o gravilla.



Fig. 4: Objeto en un entorno de búsqueda.

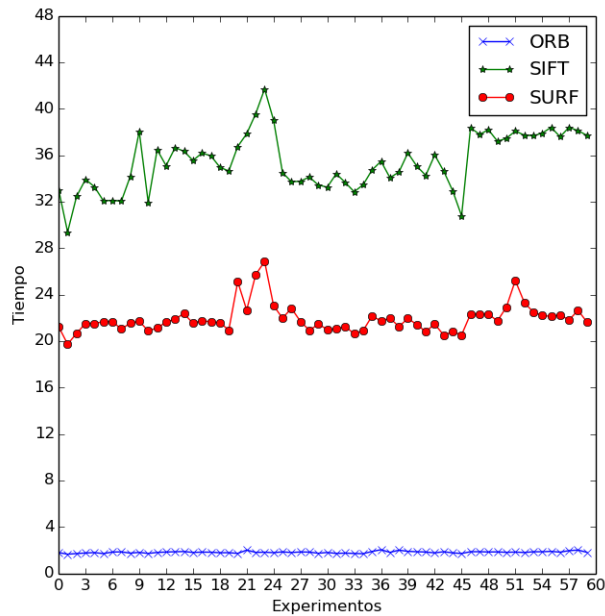


Fig. 5: Tiempo de ejecución en el sistema Raspberry Pi.

Por estudios anteriores, se conoce que otro tipo de locomociones como lo pueden ser las ruedas sobre este tipo de terrenos, presentan problemas serios de desplazamiento. Posteriormente a las imágenes encontradas por el robot, se les añade ruido del tipo Gaussiano en un porcentaje del 5, 10, 15 y 20%; esto con el fin de poder observar la resistencia al ruido que presentan los algoritmos. Finalmente, se gira el objeto en un rango entre 0 grados y 45 grados con pasos de 15 grados para observar el número de puntos característicos que encuentran los diferentes métodos y detectar si todavía es posible reconocer el objeto en el ambiente de búsqueda.

En ambientes no preparados el número de puntos característicos encontrados por los algoritmos son mayores a 10,000 puntos, de los cuales los algoritmos deben poder identificar los puntos del objeto previamente seleccionado, y que se encuentren en un rango de entre 500 a 700 puntos.

Referente al tiempo de búsqueda (Fig. 5), se observa que el algoritmo de ORB se ejecuta en un tiempo constante mucho menor que SIFT o SURF, los cuales tienen resultados variantes respecto a la distancia a la que se encuentra el objeto del robot, lo cual ayuda a que los puntos característicos necesarios para la detección de los objetos sean procesados de forma más rápida, con lo cual el tiempo de ejecución del robot en su tarea programada disminuya.

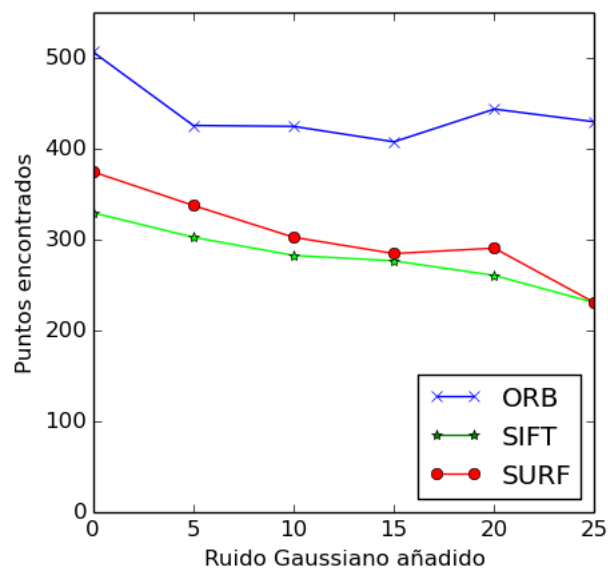


Fig. 6: Resistencia al ruido.

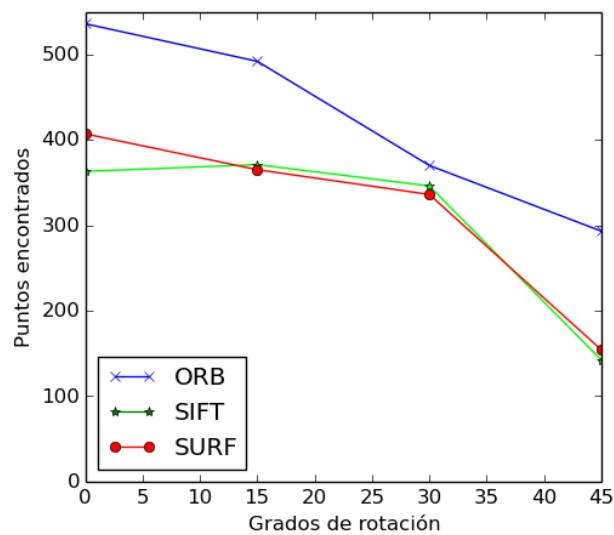


Fig. 7: Rotación del objeto.



Fig. 8: Prueba de ejecución del robot móvil tipo oruga T-Rex.

En la segunda prueba, a las imágenes encontradas por el robot se le agrega ruido Gaussiano en porcentajes de 5%, 10%, 15%, 20% y 25% (Fig. 6), esto con el fin de poder observar la resistencia al ruido de los algoritmos ya que en

ambientes con partículas en el aire como son la arena y tierra, la cámara puede verse obstruida. Se puede observar que el algoritmo de ORB para todos los casos encuentra más de 400 puntos característicos, mientras que SIFT y SURF no encuentran más de 400 puntos característicos aún sin ruido y esta cantidad disminuye hasta menos de 250 puntos característicos cuando el ruido Gaussiano es mayor a 25 %, por lo cual se observa que ORB es más resistente al ruido.

Finalmente se realizaron pruebas para observar la invarianza a la rotación del objetivo, esto en un rango de 0 grados a 45 grados con pasos de 15 grados (Fig. 7). Cuando el objetivo no tiene rotación (0 grados), el número de puntos encontrados por los algoritmos son 530, 405 y 361 respectivamente, estos números decaen si la rotación aumenta y si la cantidad de puntos característicos decae más de la mitad de cuando se está a 0 grados, se dice que el sistema no puede encontrar al objetivo, por lo cual para los casos en donde los métodos de SIFT y SURF tienen una rotación mayor a 45 grados, éstos no pueden localizar al objetivo.

Tabla 2: Tiempo promedio de ejecución.

Ángulo del objetivo	Tiempo promedio
-90 grados	298.3 s
-45 grados	225.1 s
0 grados	95.6 s
45 grados	153.8 s
90 grados	211 s

Dado que el algoritmo de ORB muestra un mejor desempeño en el tiempo de procesado, se obtiene una cantidad mayor de puntos característicos aún en presencia de ruido y para cuando el objeto presenta una rotación de hasta 45 grados este algoritmo puede detectarlo. Por lo que ORB es el método que se decidió implementar en el robot tipo oruga. Para observar el desempeño del robot al realizar la tarea de buscar el objetivo, recogerlo y dejarlo en el depósito como se muestra en la Fig. 8, se realizaron una serie de experimentos donde se mide el tiempo de ejecución y el porcentaje de acierto. El terreno de prueba consiste en una superficie de tierra y arenisca con una inclinación no mayor a 20 grados, los objetos se colocaron de forma aleatoria en un rango de 180 grados frente al robot con una inclinación no mayor a 40 grados y el depósito se posicionó en un rango de 360 grados al rededor del robot a una distancia entre 30 y 60 centímetros.

En la Fig. 9 se muestra el tiempo obtenido de los resultados de una serie de pruebas, de acuerdo a la gráfica generada, el tiempo de ejecución varía dependiendo de la posición del objeto respecto del robot.

Por lo que es posible observar que el tiempo promedio para esta tarea es proporcional a la posición del objetivo. Lo anterior se puede ver en la Tabla 2.

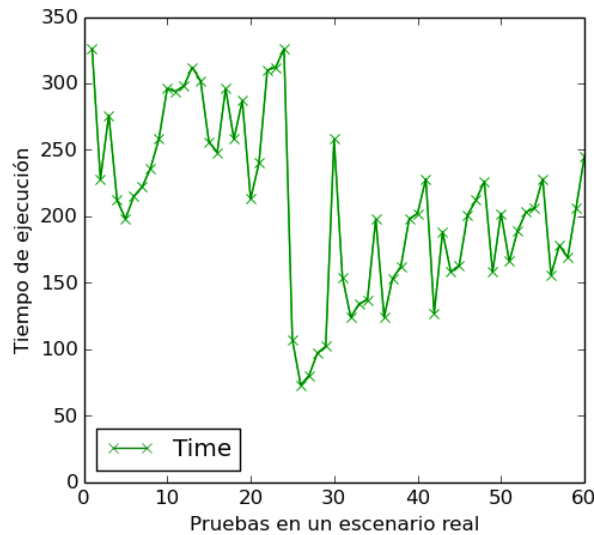


Fig. 9: Tiempo de ejecución del robot móvil tipo oruga T-Rex.

Si el objeto se encuentra frente al robot, se tiene un tiempo mínimo; cuando el objeto se localiza a la derecha del robot el tiempo aumenta en un rango aproximado del 60 %, pero si el objeto se encuentra a la izquierda del robot, el tiempo aumenta hasta un 200 %.

6. Conclusiones y trabajo futuro

De acuerdo a los resultados de los experimentos, el detector y descriptor de rasgos que tiene un procesamiento más rápido es ORB, el cual tiene un promedio de búsqueda de puntos clave de 1.7621 segundos; mientras que SIFT es de 34.2433 segundos y SURF de 21.6925 segundos, además de que de acuerdo a las pruebas realizadas con ruido se observó que ORB puede encontrar más relaciones que SIFT o SURF aún cuando el ruido es mayor. En cuanto a la rotación del objeto, los tres algoritmos decaen de manera similar, por lo que se elige el detector y descriptor de rasgos ORB para su uso en el robot oruga ya que aún con interferencias se puede encontrar el objeto en su entorno de búsqueda (Fig. 10).

El robot móvil tipo oruga es capaz de poder detectar, tomar y depositar un objeto en terrenos no preparados en un tiempo óptimo el cual para los casos de mayor complejidad no es mayor a 300 segundos. Además de presentar un margen de error no mayor al 8 %, por lo cual este robot constituye una herramienta de búsqueda con un alto índice de acierto.

En el futuro se propone el uso de técnicas de redes neuronales para poder procesar los rasgos de los objetos a buscar y así tener un rango de búsqueda

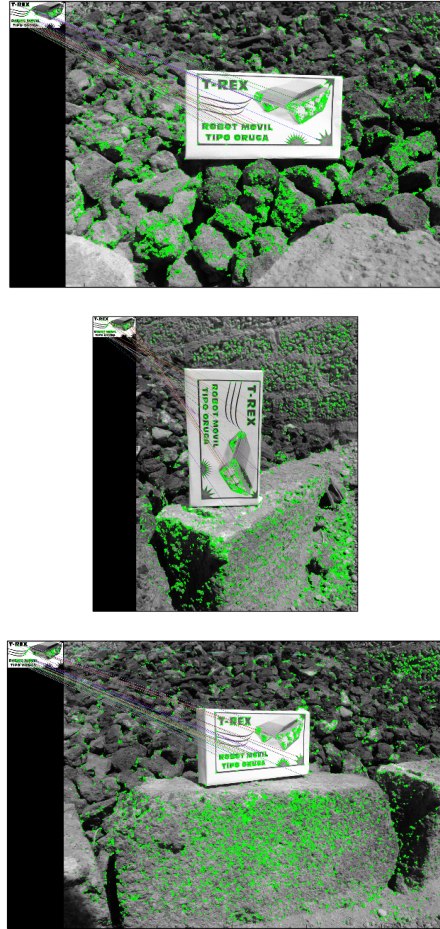


Fig. 10: Relaciones entre el objeto de prueba y el ambiente real.

mayor, con lo cual el robot será capaz de poder seguir rutas en las cuales se encuentren los objetivos, y que el robot pueda recolectarlos de manera secuencial de acuerdo a la distancia entre el objetivo y el robot. Además de la integración de sensores que permitan al robot maniobrar diferentes tipos de objetos con tamaños y formas diferentes.

Agradecimientos. E. Rubio y H. Sossa agradecen el soporte provisto por el Instituto Politécnico Nacional y la Secretaria de Investigación y Posgrado (SIP) en la realización de esta investigación. Este trabajo fue apoyado económicamente por SIP-IPN (20171548 y 20170901) y CONACYT (155014 y 65). U. Escalona agradece al CONACYT por la beca para realizar sus estudios de Maestría.

Referencias

1. Al-Mutib, K., Mattar, E., Alsulaiman, M.: Implementation of fuzzy decision based mobile robot navigation using stereo vision. *Procedia Computer Science* Volume 62 (2015)
2. Atallah, M.J., Duket, T.W.: Pattern matching in the hamming distance with thresholds. *Information Processing Letters* Volume 111 (2011)
3. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: Brief: Binary robust independent elementary features. *European Conference on Computer Vision* (2011)
4. Conde, B.C., Fuentes, S., Caron, M., Xiao, D., Collman, R., Howell, K.S.: Development of a robotic and computer vision method to assess foam quality in sparkling wines. *Food Control* Volume 71 (2017)
5. Cunha, J., Ferreira, R., Lau, N.: Computer vision and robotic manipulation for automated feeding of cork drillers. *Materials & Design* Volume 82 (2015)
6. Ferreira, F., Veruggio, G., Caccia, M., Bruzzone, G.: A comparison between different feature based methods for rovs vision based speed estimation. *IFAC Proceedings Volumes* (2010)
7. Gryaznov, N., Lopota, A.: Computer vision for mobile on-ground robotics. *25th DAAAM International Symposium on Intelligent Manufacturing and Automation* (2014)
8. Huang, C., Zhou, W.: A real-time image matching algorithm for integrated navigation system. *Optik - International Journal for Light and Electron Optics* (2014)
9. Kashif, M., Deserno, T., Haak, D., Jonas, S.: Feature description with sift, surf, brief, briesk or freak? a general question answered for bone age assessment. *Computers in Biology and Medicine* (2015)
10. Lowe, D.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60 (2004)
11. Miura, J., Ikeda, S.: A simple modeling of complex environments for mobile robots. *Int. J. Intelligent Systems Technologies and Applications* (2009)
12. Nastaran, T.A., Majid, M.: Driving performance analysis of a tracked mobile robot on different terrains. *International Conference on Robotics and Mechatronics* (2015)
13. Rosten, E., Drummond, T.: Machine learning for high speed corner detection. *European Conference on Computer Vision* (2006)
14. Rosten, E., Porter, R., Drummond, T.: Faster and better: A machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2010)
15. Rublee, E., Rabaud, V., Koolige, K., Bradski, G.: Orb: an efficient alternative to sift or surf. *International Conference on Computer Vision* (2011)
16. Tinne, H.B., Gool, T.L.V.: Surf: Speeded up robust features. *European Conference on Computer Vision* (2006)
17. Vincent, I., Sun, Q.: A combined reactive and reinforcement learning controller for an autonomous tracked vehicle. *Robotics and Autonomous Systems* (2012)
18. Zhao, Y., Gong, L., Huang, Y., Liu, C.: A review of key techniques of vision-based control for harvesting robot. *Computers and Electronics in Agriculture* (2016)
19. Zou, M., Bai, H., YangWang, Yu, S.: Mechanical design of a self-adaptive transformable tracked robot for cable tunnel inspection. *International Conference on Mechatronics and Automation* (2016)