

DETECCIÓN Y DESCRIPCIÓN DE PUNTOS CARACTERÍSTICOS EN IMÁGENES MULTIESPECTRALES UTILIZANDO ESQUEMAS CLÁSICOS

Wilfrido Rolando Guerrero Albán
Facultad de Ingeniería en Electricidad y Computación
Escuela Superior Politécnica del Litoral
Campus Gustavo Galindo Km 30.5, vía Perimetral
Apartado 09-01-5863, Guayaquil, Ecuador
rguerrer@espol.edu.ec

Resumen

En este trabajo se estudió el comportamiento de algoritmos usados en visión por computador para la detección y descripción de puntos de interés cuando son usados en el caso de imágenes multiespectrales. Para esto se procedió a buscar con criterio las variables de los algoritmos SIFT y SURF que se puedan adaptar para poder obtener una mayor cantidad de puntos de interés y a su vez un mayor número correspondencia entre imágenes de espectro visible (VS) e imágenes de espectro infrarrojo (IR). A lo largo del documento se podrá observar que variables fueron adaptadas para cada algoritmo, teniendo así, un resultado favorable en la etapa de detección de puntos de interés al momento de ser ejecutado, es decir un incremento en la detección de puntos para las imágenes de espectro infrarrojo, pero un resultado poco favorable en la etapa de descripción, ya que las correspondencias de puntos al momento de ser comparadas entre las imágenes VS-IR fueron muy pocas.

Palabras Claves: *Detección y Descripción de puntos de interés en imágenes Multiespectrales.*

Abstract

In this work we study the behavior of algorithms used in computer vision for key point detection and description when they are used in multispectral imaging. In this study we search for those variables of SIFT and SURF algorithms that can be adapted in order to have more key points as well as to increase the number of correspondences between images of the visible spectrum (VS) and images of infrared spectrum. Through this document we will present the variables that were adapted, improving the key point detection stage by increasing the number of points detected in the infrared spectrum; unfortunately, this improvement was not in the description stage, since key point correspondence between VS-IR remain lower.

Keywords: *Detection and description of key points in multispectral images.*

1. Introducción

Sin duda alguna, uno de los aspectos más estudiados en aplicaciones de Visión por Computador es la correspondencia entre dos imágenes; este tipo de tareas es utilizado en un gran número de aplicaciones, como por ejemplo el reconocimiento de objetos, patrones, seguimiento de objetos, entre otras.

El establecimiento de correspondencia se lo realiza en tres etapas: detección de puntos de interés, descripción de los mismos y puesta en correspondencia.

Cuando se busca establecer correspondencias entre dos imágenes de distinta banda espectral, los métodos convencionales dejan de ser confiables a medida que dichas bandas se alejan.

En este documento se expone la posibilidad de utilizar esquemas clásicos de Detección y Descripción, para obtener la correspondencia entre dos imágenes multiespectrales, adaptando aquellas variables que favorecen dicha puesta en correspondencia.

2. Análisis del caso

Como ya se ha mencionado, existen algoritmos que permiten la detección y descripción de puntos entre imágenes en el espectro visible, el propósito del presente trabajo es evaluar la robustez cuando aplicamos estos algoritmos entre una imagen en el espectro visible y una imagen en el espectro infrarrojo como se muestra en la Figura 1.



Figura 1. Imágenes en diferentes espectros VS-IR

Al aplicar los algoritmos de detección con sus parámetros iniciales, obtenemos mayor cantidad de puntos característicos o de interés en la imagen visible, a diferencia que, en la imagen infrarroja se aprecia muy pocos puntos detectados como se muestra en la Figura 2.



Figura 2. Detección de puntos en imágenes VS-IR

Al momento de aplicar un algoritmo para describir los puntos de interés detectados y mediante una matriz de comparación de puntos entre las dos imágenes, se puede observar poca correspondencia de puntos, como se muestra en la Figura 3.

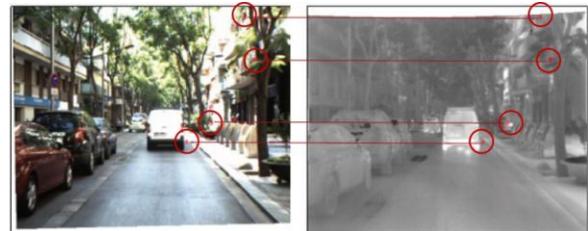


Figura 3. Correspondencia de puntos de interés entre dos imágenes.

2.1 Análisis de la solución

Como posible solución ante este problema se propone variar con criterio ciertos parámetros iniciales de los algoritmos en forma independiente en cada espectro, esperando así obtener menor diferencia entre el número de puntos característicos encontrados y a su vez mayor correspondencia de puntos.

2.2 Objetivos generales

Evaluar el comportamiento de ciertos esquemas clásicos de detección y descripción de puntos de interés entre imágenes en el espectro visible e imágenes en el espectro infrarrojo.

2.3 Objetivos específicos

Determinar si es posible usar esquemas clásicos en el caso de imágenes de espectros diferentes, y en tal caso, determinar la metodología a seguir para obtener soluciones robustas.

2.4 Justificación

En la actualidad hay muchos campos donde se puede aplicar la visión por computador, y estos algoritmos nos ayudan al análisis para el reconocimiento de un objeto o de una escena. Uno de los principales motivos para el desarrollo de este documento es tratar de usar estos algoritmos para detectar y describir puntos de interés en imágenes de diferentes espectros, para así contribuir al campo de la investigación y seguir desarrollando aplicaciones que permitan un análisis multispectral.

3. Detectores de puntos de interés

La detección de puntos de interés corresponde a la primera etapa durante el procesamiento de una imagen, que implica la individualización de puntos, que por sus características ayudan en su conjunto a definir el objeto contenido en una imagen. Se pretende que los detectores sean eficientes respecto de los recursos computacionales que utilizan, esto influye directamente en la cantidad de cuadros por segundo que se pueden procesar y en las plataformas de hardware sobre las cuales se puede implementar el procesamiento en tiempo real en caso de ser necesario [1].

4. Descripción de puntos de interés

Un detector ideal de puntos de interés localiza dichos puntos repetidamente a pesar del cambio de punto de vista siendo, fiable ante transformaciones de la imagen. Un descriptor ideal obtiene de cada punto de interés información relevante y distintiva de la región que lo rodea, de manera que la misma estructura pueda ser reconocida si es encontrada en otra imagen. Implica un verdadero desafío describir de manera eficiente puntos de interés en forma estable y compacta, con representaciones robustas invariantes a escala, rotación, diversas transformaciones y el ruido [2].

5. Esquemas clásicos de detección y descripción

Existen los denominados esquemas clásicos, que permiten llevar a cabo un eficiente procesamiento en la detección y descripción de puntos característicos, tales

como los algoritmos SIFT [3], SURF [4], FAST [5] entre otros. Cabe destacar que estos han sido propuestos para el caso de imágenes en el espectro visible (VS).

6. Análisis en imágenes multispectrales

A continuación se analiza como caso particular el comportamiento de dos algoritmos SIFT y SURF. Primero se evaluarán los algoritmos como detectores de puntos de interés y se realizarán las comparaciones correspondientes entre los resultados de ambas imágenes de diferentes espectros. Como segundo paso se evaluarán los mismos algoritmos como descriptores y de igual manera se realizarán las comparaciones entre los resultados obtenidos.

7. Algoritmo SIFT

Este algoritmo está representado por la siguiente función, con sus respectivos parámetros [6]: `sift(int nfeatures, int nOctaveLayers, double contrastThreshold, double edgeThreshold, double sigma)`.

- **nfeatures:** Número de las mejores características para almacenar.
- **nOctaveLayers:** Número de capas en cada octava.
- **contrastThreshold:** Umbral del contraste.
- **edgeThreshold:** Umbral para filtrar características de borde similar.
- **Sigma:** Constante gaussiana.

Los parámetros iniciales con que ejecutaremos el algoritmo son los siguientes:

`nfeatures = 0, nOctaveLayers = 3, contrastThreshold = 0.04, edgeThreshold = 10, sigma = 1.6`

7.1 Detección de puntos de interés (SIFT)

Al ejecutar el algoritmo con los parámetros anteriormente mencionados en distintas imágenes VS e IR obtenemos los resultados mostrados en la Tabla 1 y Tabla2, respectivamente.

Tabla 1. Detección de puntos de interés con valores iniciales para una imagen VS (SIFT)

| Imagen VS | Puntos de interés |
|----------------|-------------------|
| 0000026_VSRRTC | 3337 |
| 0000032_VSRRTC | 3425 |
| 0000059_VSRRTC | 3273 |
| 0000098_VSRRTC | 3310 |

Tabla 2. Detección de puntos de interés con valores iniciales para una imagen IR (SIFT)

| Imagen IR | Puntos de interés |
|----------------|-------------------|
| 0000026_IR1CTC | 370 |
| 0000032_IR1CTC | 382 |
| 0000059_IR1CTC | 381 |
| 0000098_IR1CTC | 398 |

Como se puede observar el algoritmo SIFT se comporta de manera diferente en una imagen VS que en una imagen IR, obteniendo mayor cantidad de puntos de interés en la imagen VS como se muestra en la Tabla 3 y como se puede ilustrar en la Figura 4.

Tabla 3. Diferencia en la detección de puntos de interés entre imágenes VS-IR (SIFT)

| | Imagen VS | Imagen IR | Diferencia |
|--------------------------------------|-----------|-----------|------------|
| Promedio de puntos de interés | 3336 | 382 | 2954 |



Figura 4. Ilustración de puntos encontrados.

Cabe destacar que las imágenes IR corresponden al mismo escenario que las imágenes evaluadas en el espectro visible (VS). Ambas imágenes fueron tomadas en el mismo instante usando un trigger externo que sincroniza la adquisición de las mismas. Ambas imágenes tienen el mismo tamaño 640x480 y las lentes de las cámaras fueron dispuestas de manera tal que los sensores adquieran la misma información. Las imágenes a procesar ya se encuentran rectificadas lo cual garantiza que un punto de interés detectado en la imagen VS se encuentra en la misma fila en la imagen IR.

7.2 Descripción de puntos (SIFT)

Al ejecutar el algoritmo SIFT como descriptor analizaremos si los puntos detectados sobre ambas imágenes concuerdan en sus coordenadas y, como se comento anteriormente, como las imágenes están rectificadas los puntos deberán pertenecer a la misma fila. Para esto el descriptor utiliza un vector de características de longitud 128 para cada punto de interés detectado en ambas imágenes, como se ilustra en las Figuras 5 y 6.

| | 1 | 2 | 3 | 4 | 5 | 6 | . | . | . | . | . | . | . | . | . | . | . | . | 128 |
|-----|----|-----|-----|-----|-----|----|-----|---|---|---|---|---|---|---|---|---|---|---|-----|
| Vs1 | 0 | 150 | 95 | 135 | 15 | 32 | 180 | - | - | - | - | - | - | - | - | - | - | - | 99 |
| Vs2 | 55 | 10 | 145 | 5 | 44 | 31 | 175 | - | - | - | - | - | - | - | - | - | - | - | 123 |
| Vs3 | 76 | 17 | 97 | 13 | 87 | 71 | 34 | - | - | - | - | - | - | - | - | - | - | - | 9 |
| . | | | | | | | | | | | | | | | | | | | |
| . | | | | | | | | | | | | | | | | | | | |
| . | | | | | | | | | | | | | | | | | | | |
| Vsn | 45 | 150 | 89 | 135 | 165 | 89 | 179 | - | - | - | - | - | - | - | - | - | - | - | 38 |

Figura 5. Vectores característicos para una imagen VS

| | | | | | | | | | | | | | | |
|-----|----|-----|-----|-----|----|-----|-----|---|---|---|---|---|---|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | . | . | . | . | . | . | . | 128 |
| Ir1 | 12 | 60 | 87 | 135 | 3 | 198 | 18 | - | - | - | - | - | - | 49 |
| Ir2 | 98 | 32 | 12 | 56 | 98 | 43 | 69 | - | - | - | - | - | - | 12 |
| Ir3 | 66 | 34 | 154 | 78 | 32 | 89 | 2 | - | - | - | - | - | - | 89 |
| . | | | | | | | | | | | | | | |
| . | | | | | | | | | | | | | | |
| Im | 56 | 183 | 67 | 198 | 15 | 99 | 189 | - | - | - | - | - | - | 68 |

Figura 6. Vectores característicos para una imagen IR

Para poder obtener la correspondencia de puntos entre ambas imágenes crearemos una matriz que contenga la distancia euclidiana entre cada descriptor, mediante la ecuación:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

La matriz distancia de los descriptores de ambas imágenes quedaría como se muestra en la Figura 7.

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|---|---|---|---|---|---|-----|
| | Vs1 | Vs2 | Vs3 | Vs4 | Vs5 | Vs6 | . | . | . | . | . | . | . | VSn |
| Ir1 | 35 | 12 | 5 | 140 | 54 | 23 | 67 | - | - | - | - | - | - | 56 |
| Ir2 | 56 | 21 | 67 | 69 | 78 | 34 | 76 | - | - | - | - | - | - | 87 |
| Ir3 | 44 | 13 | 68 | 123 | 76 | 98 | 34 | - | - | - | - | - | - | 98 |
| . | | | | | | | | | | | | | | |
| . | | | | | | | | | | | | | | |
| Im | 87 | 34 | 156 | 143 | 78 | 92 | 143 | - | - | - | - | - | - | 12 |

Figura 7. Matriz de distancia de los descriptores

De esta matriz se escoge la menor distancia y buscamos sus correspondientes coordenadas y las comparamos, dando como resultado la información mostrada en la Tabla 4.

Tabla 4. Correspondencias detectadas y correspondencias correctas usando los parámetros iniciales (SIFT)

| Imágenes VS - IR | Correspondencias Detectadas | Correspondencias Correctas |
|------------------|-----------------------------|----------------------------|
| 0000026 | 5 | 4 |
| 0000032 | 5 | 3 |
| 0000059 | 10 | 2 |
| 0000098 | 10 | 3 |

Como resultado de las imágenes analizadas se obtiene que en promedio se detectan unas 7 correspondencias. Sin embargo de ese promedio solo 3 son correctas como se muestra en la Figura 8.



Figura 8. Comparación de imágenes con sus puntos de correspondencia.

7.3 Variación de parámetro en algoritmo SIFT

En base al estudio presentado y a los resultados mostrados en la sección anterior se propone un método para estudiar cómo se puede mejorar el número de correspondencia de puntos; este método consiste en variar uno de los parámetros del algoritmo SIFT en la imagen IR y analizar el número de correspondencias correctas. Aplicaremos nuevamente el proceso anteriormente descrito, pero variando el parámetro contrastThreshold a un valor de 0.02, y obtenemos los resultados mostrados en la Tabla 5.

Tabla 5. Detección de puntos de interés para una imagen IR con valor contrastThreshold = 0.02 (SIFT)

| Imagen IR | Puntos de interés |
|----------------|-------------------|
| 0000026_IR1CTC | 1079 |
| 0000032_IR1CTC | 1127 |
| 0000059_IR1CTC | 1055 |
| 0000098_IR1CTC | 1120 |

Como podemos observar al variar el parámetro del algoritmo para la imagen IR, existe un incremento en la cantidad de puntos de interés detectados y una diferencia menor a lo anteriormente calculado, esto lo podemos visualizar en la Tabla 6 y de manera ilustrativa en la Figura 9.

Tabla 6. Diferencia entre la cantidad de puntos detectados entre las imágenes VS-IR (SIFT)

| | Imagen VS | Imagen IR | Diferencia |
|--------------------------------------|-----------|-----------|------------|
| Promedio de puntos de interés | 3336 | 382 | 1095 |



Figura 9. Incremento de puntos de interés en imagen IR.

Como conclusión a esto podemos decir que mientras menor sea el valor de `contrastThreshold` mayor es el número de puntos de interés detectados. Pero esto no garantiza una buena correspondencia de puntos como se presenta en la Tabla 7 y de manera ilustrativa en la Figura 10.

Tabla 7. Correspondencias detectadas y correspondencias correctas entre imagen VS-IR con valor `contrastThreshold=0.02` (SIFT)

| Imágenes VS - IR | Correspondencias Detectadas | Correspondencias Correctas |
|------------------|-----------------------------|----------------------------|
| 0000026 | 12 | 3 |
| 0000032 | 3 | 2 |
| 0000059 | 8 | 5 |
| 0000098 | 11 | 4 |



Figura 10. Correspondencia de puntos para un valor de `contrastThreshold = 0.02`.

Como resultado de las imágenes analizadas se obtiene que en promedio se detectan 8 correspondencias, de las cuales solo unas 3 son correctas, en la Figura 10 se muestra un ejemplo ilustrativo.

8. Algoritmo SURF

Este algoritmo está representado por la siguiente función, con sus respectivos parámetros [6]: `surf` (`double hessianThreshold`, `int nOctaves`, `int octaveLayers`, `bool extendet`, `bool upright`).

- **hessianThreshold:** Umbral de la constante Hessiana para la detección de puntos.
- **nOctaves:** Número de pirámides.
- **OctavesLayers:** Número de capas dentro de cada octava.
- **extendet:** Bandera extendida del descriptor.
- **upright:** Bandera de rotación.

Los parámetros iniciales con que ejecutaremos el algoritmo son los siguientes:

`hessianThreshold = 500`, `nOctaves = 4`,
`OctavesLayers = 2`, `extendet = true`, `upright = false`.

8.1 Detección de puntos de interés (SURF)

Al ejecutar el algoritmo con los parámetros anteriormente mencionados en distintas imágenes VS e IR obtenemos los resultados presentados en la Tabla 8 y Tabla 9, respectivamente.

Tabla 8. Detección de puntos de interés con valores iniciales para una imagen VS (SURF)

| Imagen VS | Puntos de interés |
|----------------|-------------------|
| 0000026_VSRRTC | 1649 |
| 0000032_VSRRTC | 1637 |
| 0000059_VSRRTC | 1531 |
| 0000098_VSRRTC | 1508 |

Tabla 9. Detección de puntos de interés con valores iniciales para una imagen IR (SURF)

| Imagen IR | Puntos de interés |
|----------------|-------------------|
| 0000026_IR1CTC | 231 |
| 0000032_IR1CTC | 247 |
| 0000059_IR1CTC | 264 |
| 0000098_IR1CTC | 282 |

Como se puede observar el algoritmo SURF se comporta de manera diferente en una imagen VS que en una imagen IR, obteniendo mayor cantidad de puntos de interés en la imagen VS como se muestra en la Tabla 10 y como se puede ilustrar en la Figura 11.

Tabla 10. Diferencia en la detección de puntos de interés entre imágenes VS-IR (SURF)

| | Imagen VS | Imagen IR | Diferencia |
|-------------------------------|-----------|-----------|------------|
| Promedio de puntos de interés | 1581 | 256 | 1325 |



Figura 11. Comparación de imágenes con sus puntos de interés.

8.2 Descripción de puntos (SURF)

Al ejecutar el algoritmo SURF como descriptor analizaremos si los puntos detectados sobre ambas imágenes concuerdan en sus coordenadas y. Para esto, al igual que en la sección 7.2, el descriptor utiliza un vector de características de longitud 128 para cada

punto de interés detectado en ambas imágenes, como se muestra en la Figura 5 y 6 de la sección 7.2.

Para poder obtener la correspondencia de puntos entre ambas imágenes crearemos una matriz que contenga la distancia euclidiana entre cada descriptor, mediante la ecuación:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

La matriz distancia de los descriptores de ambas imágenes es similar a la ilustrada en la Figura 7 de la sección 7.2. Una vez realizado este procedimiento obtenemos los resultados mostrados en la Tabla 11.

Tabla 11. Correspondencias detectadas y correspondencias correctas entre imágenes VS-IR (SURF)

| Imágenes VS - IR | Correspondencias Detectadas | Correspondencias Correctas |
|------------------|-----------------------------|----------------------------|
| 0000026 | 3 | 1 |
| 0000032 | 6 | 1 |
| 0000059 | 2 | 1 |
| 0000098 | 8 | 1 |

Como resultado de las imágenes analizadas se obtiene que en promedio se detectan unas 4 correspondencias, de las cuales, en promedio, solo 1 es correcta; la Figura 12 muestra una ilustración de un ejemplo en el cual solo 1 correspondencia es correcta.



Figura 12. Comparación de imágenes con sus puntos de correspondencia.

8.3 Variación de parámetro en algoritmo SURF

En base al estudio presentado y a los resultados mostrados en la sección anterior se propone un método para estudiar cómo se puede mejorar el número de correspondencia de puntos; este método consiste en variar uno de los parámetros del algoritmo SURF en la imagen IR y analizar el número de correspondencias correctas. Haremos el mismo proceso anteriormente descrito, pero variando el parámetro `hessianThreshold` a un valor de 100, y obtenemos los resultados mostrados en la Tabla 12.

Tabla 12. Detección de puntos de interés para una imagen IR para una valor de `hessianThreshold = 100` (SURF)

| Imagen IR | Puntos de interés |
|----------------|-------------------|
| 0000026_IR1CTC | 747 |
| 0000032_IR1CTC | 797 |
| 0000059_IR1CTC | 811 |
| 0000098_IR1CTC | 839 |

Como podemos observar al cambiar los valores de `hessianThreshold` del algoritmo en el caso de la imagen IR, existe un incremento en la cantidad de puntos de interés detectados y una diferencia menor a lo anteriormente calculado en promedio, esto lo podemos visualizar en la Tabla 13 y en la ilustración de la Figura 13.

Tabla 13. Diferencia en la detección de puntos entre las imágenes VS-IR (SURF)

| | Imagen VS | Imagen IR | Diferencia |
|--------------------------------------|-----------|-----------|------------|
| Promedio de puntos de interés | 1581 | 798 | 1095 |



Figura 13. Detección de puntos de interés para un valor de `hessianThreshold = 200`.

En función de esto podemos decir que mientras menor sea `hessianThreshold`, mayor cantidad de puntos de interés se van a obtener. Pero así mismo esto no nos garantiza que exista una buena correspondencia de puntos, como se muestra en la Tabla 14 y de manera ilustrativa en la Figura 14.

Tabla 14. Correspondencias detectadas y correspondencias entre imágenes VS-IR para un valor de `hessianThreshold = 100`

| Imágenes VS - IR | Correspondencias Detectadas | Correspondencias Correctas |
|------------------|-----------------------------|----------------------------|
| 0000026 | 9 | 1 |
| 0000032 | 5 | 2 |
| 0000059 | 3 | 1 |
| 0000098 | 8 | 1 |



Figura 14. Correspondencia de puntos para valores de `hessianThreshold = 200`.

9. Conclusiones

- Los algoritmos anteriormente analizados ejecutados con sus parámetros iniciales en las imágenes IR, detectan pocos puntos de interés y a su vez una muy baja correspondencia.
- Al variar ciertos parámetros iniciales en cada algoritmo, `contrastThreshold` para SIFT con valor de 0.02 y `hessianThreshold` para SURF con valor de 100 obtenemos un incremento en la detección de puntos de interés.
- Al hacer la correspondencia entre las imágenes VS-IR, pocos son los puntos que coinciden.

10. Recomendaciones

- Añadir algún filtro a la imagen original, como un ecualizador de histogramas que genere una imagen más llana y separada posible.
- Realizar la variación de otros parámetros.
- Realizar una variación acoplada de varios parámetros.

11. Bibliografía

- [1] Marcelo Daniel Pistarelli. Detección y correspondencia de características en imágenes multiespectrales. Proyecto final de carrera Ingeniería en Informática, Universidad Autónoma de Barcelona, 4 de Febrero de 2013, Capítulo 2, páginas 8 - 9.
- [2] Marcelo Daniel Pistarelli. Detección y correspondencia de características en imágenes multiespectrales. Proyecto final de carrera Ingeniería en Informática, Universidad Autónoma de Barcelona, 4 de Febrero de 2013, Capítulo 3, páginas 19.
- [3] D.G. Lowe, "Object Recognition from Local Scale Invariant Features". In Proceedings of the IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999, pp. 1150–1157.
- [4] H. Bay, T. Tuytelaars, L.J. Gool, "SURF: Speeded up Robust Features", In Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, 7–13 May 2006, pp. 404–417.
- [5] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection", European Conference on Computer Vision, May, 2006.
- [6] OpenCV. "Feature detection and description"
[Online]
http://docs.opencv.org/modules/nonfree/doc/feature_detection.html [2013]