



Calibrando una cámara digital

R. A. Montalvo y P. H. Rivera *

Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Físicas, Lima, Perú

Recibido 01 setiembre 2020 – Aceptado 04 noviembre 2020

Resumen

La visualización de áreas, volúmenes, topologías y perspectivas de diferentes objetos que se encuentran en nuestro entorno son realizados por nuestros ojos y cerebro que se han perfeccionado a lo largo de millones de años. Este ensamble entre los ojos y el cerebro debe ser emulado por la tecnología para que una inteligencia artificial tome una decisión autónoma basada en la información proporcionada por los sensores y la visión artificial. La fidelidad de la información proporcionada por el entorno es fundamental para una buena decisión. En este contexto, la miniaturización de las cámaras y las lentes ópticas involucra distorsiones que modifican las imágenes adquiridas de las áreas, volúmenes, topologías y perspectivas de los objetos. En el presente trabajo, hacemos una calibración de las cámaras ópticas considerando las distorsiones radiales y tangenciales y usando las librerías del OpenCV calculamos los parámetros intrínsecos y extrínsecos que permiten corregir las imágenes adquiridas.

Palabras clave: Cámara digital, visión computacional, OpenCV, geometría proyectiva, calibración óptica.

Calibrating a digital camera

Abstract

The visualization of areas, volumes, topologies and perspectives objects that stay in our environment are making by ours eyes and brain, which have a long evolution through millions years ago. This ensemble between eyes and brain must be simulated by technology in which an artificial intelligence entity take an autonomous decision based on data information acquired by sensors and computer vision. The fidelity of acquired information from the environment is the fundamental key for a good decision. In this context, the cameras and optical lens miniaturization involves distortions which modify the acquired images of areas, volumes, topologies and perspectives objects. In the present work, we make a calibration process of optical cameras which involves radial and tangential distortions, and using OpenCV libraries we calculate the intrinsics and extrinsics parameters which permit us to obtain correct caquired images.

Keywords: Digital camera, computational vision, OpenCV, projective geometry, optical calibration.

1. Introducción

En los últimos veinte años, con el surgimiento de los teléfonos y las cámaras de video para seguridad y vigilancia, la industria de estos dispositivos se ha masificado y las dimensiones de los mismos se han reducido considerablemente. Desde el punto de vista de la óptica física, la reducción del tamaño de las lentes incorporan distorsiones radiales y tangenciales en las mismas, disminuyendo su calidad óptica. Por ello, tales distorsiones deben ser corregidas cuando se utilizan estos dispositivos para si-

tuaciones de análisis reales en el procesamiento de imágenes, en el reconocimiento de patrones, en visión artificial y por ende en inteligencia artificial, robótica y sistemas autónomos.

Hace veinte años también, enero de 1999, aparece en internet el OpenCV, acrónimo de *Open Computer Vision*, que reúne los esfuerzos iniciales de la Intel Corp. en el área de la visión artificial, luego con el aporte de la comunidad internacional ha devenido en una gran librería sobre visión artificial. Hasta la fecha, el OpenCV reúne algunos miles de códigos de algoritmos sobre vi-

*priverar@unmsm.edu.pe

sión computacional escritos en lenguaje de programación C++ y con interfaces aplicativos para Python y Java [1].

Por otro lado, en los primeros meses del 2020, para la detección de personas contagiadas de virus SARS-Cov2, en los medios de transporte masivo, se han instalado cámaras en el espectro visible e infrarrojo para detectar a las personas con temperaturas asociadas a la fiebre provocada por el virus. Para la identificación de las personas en el espectro visible se necesita ajustar los parámetros del sistema óptico para una detección certera de las personas, cuyas deficiencias han sido mencionadas anecdóticamente en las redes sociales.

En el presente trabajo, hemos realizado la calibración de varias cámaras de video y este es un testimonio del procedimiento para calibrar una cámara bajo el modelo de cámara *pinhole* o estenopeica usando las librerías del OpenCV como proceso inverso de la teoría que se desarrolla en la próxima sección para encontrar el detalle del algebra utilizado para extraer todas las distorsiones que las imágenes poseen al ser adquiridas con cámaras de bajo costo. El presente reporte comienza con la teoría del modelo de cámara estenopeica, la determinación de la matriz intrínseca y extrínseca. Describimos el procedimiento experimental y la calibración propiamente dicha usando el Python y Jupyter.

2. Teoría

La teoría sobre la que se basan los análisis de la cámara digital ha desenterrado del polvo la vieja teoría de la cámara estenopeica o *pinhole camera* cuyas primeras observaciones se remontan a 500 años antes de nuestra era cuando los chinos -los Escritos Mozi (500 AC)-, y los griegos -*Problemas Aristotelianos* (300 AC-600 DC)-, describen sus observaciones cualitativas. Un grabado explicando la inversión de la imagen a través de la abertura (*pinhole*) se muestra en el templo Virupaksha, localizado en Hampi, India y construida en el año 700 de nuestra era [2].

2.1. Parámetros internos

La cámara estenopeica está conformada por un cubo completamente cerrado con un agujero en uno de los lados por donde ingresa la luz y en el lado contrario sobre una lámina fotosensible se forma la imagen del(os) objeto(s) externo(s) a la cámara.

La cámara estenopeica se denomina en español porque proceden del griego $\sigma\tau\acute{\epsilon}\nu\omega = steno = estrecho$ y $\acute{o}\pi\acute{\eta} = ope = abertura$ [3], en el inglés se conoce como *pinhole camera*.

Esta cámara es conocida desde hace mucho tiempo, pero la versión moderna que se utiliza en la enseñanza de la óptica en las escuelas primarias y secundarias

del mundo se debe a Ibn al-Haytham, hombre de ciencias musulmán del siglo X; la formación de imágenes fue analizada geoméricamente por Shen Kuo (1031-1095) y la optimización del tamaño del agujero fue iniciada por József Petzval y fue mejorada por Lord Rayleigh cuya expresión final está dada por $d = 1.9\sqrt{f\lambda}$, donde d , f y λ son el diámetro, el foco y la longitud de onda de la luz, respectivamente [3]. La primera fotografía obtenida por una cámara estenopeica lo describe David Brewster en 1856 [2].

Para el análisis de las cámaras ópticas modernas, en primera aproximación, se utiliza el modelo de la cámara estenopeica o *pinhole camera*. La posición de la abertura estrecha o agujero de la cámara coincide con el centro óptico de la lente de la cámara que se encuentra al frente de la misma, aún si el lente está conformado por un sistema de lentes como un zoom óptico, éste tiene una distancia focal efectiva que es considerada como el foco del sistema [5]. La distancia desde el agujero hasta el lado opuesto donde se forma la imagen es el foco o distancia focal del sistema que en el caso de la cámara digital coincide con la posición del sensor CCD (CMOS), -acrónimos de *Charge Coupled Device (Complementary Metal Oxide Semiconductor)*. La línea recta que atraviesa el agujero hasta el lado opuesto y perpendicular al mismo en la cámara estenopeica coincide con el eje óptico de la cámara digital.

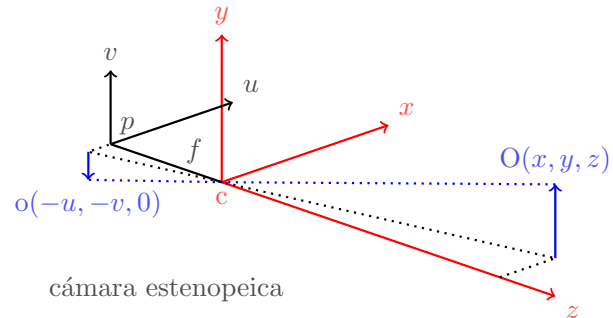


Figura 1: El modelo de la cámara estenopeica.

La geometría sobre la formación de la imagen en una cámara se muestra en la figura 1. El objeto O cuyas coordenadas respecto al sistema de coordenadas cuyo origen están en el centro óptico de la lente son (x, y, z) . La luz incidente sobre el objeto O se desplaza a través del origen de coordenadas y forma la imagen de menor tamaño o , invertida y focalizada en el sensor CCD (CMOS), cuyas coordenadas están descritas en el sistema coordenado 2D $(-u, -v)$. Debe observarse que el plano $u - v$ es paralelo al plano $x - y$ y ambas se encuentran a una distancia f una de otra. El eje z del sistema coordenado del objeto coincide con el eje óptico del sistema. Inicialmente consideramos que el eje óptico pasa por el centro del sensor

CCD(CMOS), es decir, el origen del sistema coordenado $u - v$ de la imagen es $(0,0)$.

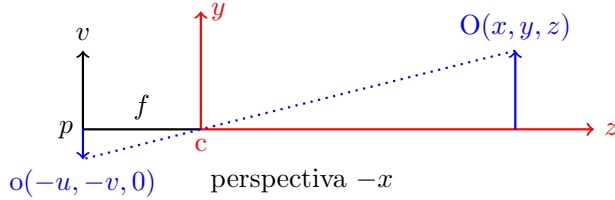


Figura 2: Proyección del modelo de la cámara estenopeica en el plano $y - z$ y observada desde la dirección $-x$.

Considerando la proyección de la Figura 1 sobre el plano $y - z$ y observándola desde la perspectiva $-x$, como se muestra en la Figura 2, se observan dos triángulos semejantes a lo largo de la hipotenusa que conecta los puntos del objeto y su imagen invertida, luego la relación entre los catetos de ambos triángulos se expresa como

$$\frac{-v}{f} = \frac{y}{z} \Rightarrow v = -f \frac{y}{z}. \quad (1)$$

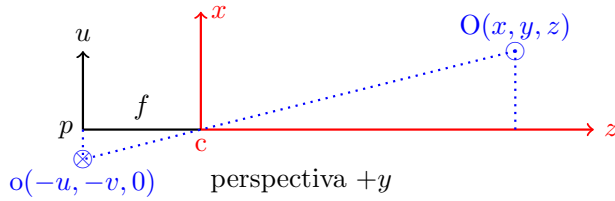


Figura 3: Proyección del modelo de la cámara estenopeica en el plano $x - z$ y observada desde la dirección $+y$.

De la misma forma, observando la Figura 3, los dos triángulos semejantes producen

$$\frac{-u}{f} = \frac{x}{z} \Rightarrow u = -f \frac{x}{z}. \quad (2)$$

Expresando las dos últimas ecuaciones de manera matricial, relacionando las coordenadas 3D con las 2D, obtenemos

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad (3)$$

por tanto, la Ec.(3) representa la primera aproximación para el análisis de una cámara digital basado en el modelo de la cámara estenopeica, cabe notar que la matriz en la que se encuentra el foco y el término $1/z$ son las que contienen la información sobre la transformación de un plano del objeto 3D al plano de la imagen [4].

También se puede considerar la imagen virtual del objeto delante del punto de referencia, posición del agujero

de la cámara estenopeica, la imagen es menor, derecha y se encuentra a una distancia f del agujero de la cámara. En esa consideración la transformación de las coordenadas 3D a 2D están expresadas como

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (4)$$

Para el análisis de una cámara digital en segunda aproximación, debemos considerar los defectos que se incorporan en la cámara durante el proceso de fabricación. En primer lugar, se observa que el término de la matriz $A_{22} = af$ es ligeramente diferente al elemento $A_{11} = f$, donde $a = 1$ para el caso de las matrices Ec. (3) y (4), durante la primera aproximación, pero para cámaras digitales que se producen para el mercado, $a \neq 1$. Más aún, los inconvenientes industriales en alinear el centro óptico de la lente con el centro del sensor CCD(CMOS), hace que el centro del sensor CCD(CMOS) se encuentre con un corrimiento respecto al origen del sistema coordenado (c_x, c_y) de modo que los elementos matriciales $A_{13} = c_x$ y $A_{23} = c_y$, esto es inevitable debido a que los fabricantes de las cámaras y las lentes utilizadas por las cámaras son diferentes, aún cuando se ha establecido los estándares como el C mount (montaje C) para las dimensiones de las ranuras (32 por 25.4 mm) de las roscas hembras(machos) de ajuste, en aluminio endurecido tanto para las cámaras como para las lentes, respectivamente, con un diámetro de 25.4 mm [6]. En tercer lugar, si el ángulo entre u y v del sistema coordenado de la imagen es $\neq \pi/2$ se define el *skewness* $\gamma \neq 0$, la matriz resultante en la segunda aproximación dado por

$$A = \begin{pmatrix} f & \gamma & c_x \\ 0 & af & c_y \\ 0 & 0 & 1 \end{pmatrix}, \quad (5)$$

se denomina la matriz intrínseca de la cámara, donde los elementos de matrices se denominan parámetros intrínsecos de la cámara [7].

2.2. Parámetros externos

La proyección de un punto 3D (x, y, z) sobre un plano 2D está definido por la recta que parte del punto 3D pasa a través de un centro de proyección e incide e intersecta un plano 2D, este proceso es natural porque ocurre en nuestros ojos y el plano de intersección está dado por $(x/z, y/z, 1)$. Esta triada se reduce a $(x/z, y/z)$ en coordenadas homegéneas o coordenadas proyectivas, que fue presentada por primera vez en 1827 por August Ferdinand Möbius en su trabajo titulado *Barycentrische Calcul* [8, 9].

Estas coordenadas proyectivas se usan en geometría proyectiva así como las coordenadas cartesianas se usan en geometría euclideana. La gran ventaja que poseen la

geometría proyectiva es que las coordenadas de los puntos de un objeto 3D, inclusive los puntos en el infinito, pueden ser expresados a través de coordenadas finitas y las fórmulas que involucran las coordenadas homogéneas son más simples y más simétricas que las fórmulas que involucran las coordenadas cartesianas [8, 9].

Para el análisis de la subsección anterior, hemos considerado que el objeto es observado desde el sistema de coordenadas de la cámara, -centro principal de la lente o posición del agujero en el modelo de la cámara estenopeica-, y proyectadas hasta el sistema de coordenadas de la imagen. Pero, el objeto puede tener su propio sistema 3D de observación que incorpore las rotaciones y traslaciones del propio cuerpo respecto a este sistema coordenado. Asimismo, desde la geometría proyectiva dicho sistema cartesiano 3D puede ser proyectado desde un sistema de coordenadas proyectivas que involucra las rotaciones y traslaciones de la siguiente forma [9],

$$\mathbf{y} \sim \mathbf{C}\mathbf{x} \quad (6)$$

donde \mathbf{x} es la representación de un punto 3D que en coordenadas homogéneas [8] es un vector 4D, \mathbf{y} es la representación de la imagen de este punto en la cámara estenopeica mediante un vector 3D, \mathbf{C} es la matrix de la cámara de dimensiones 3×4 y el símbolo \sim significa la igualdad entre diferentes elementos dentro de la geometría proyectiva [10].

Se demuestra que la matrix \mathbf{C} está conformada por la matrix de rotación \mathbf{R} y el vector de traslación \mathbf{t} , de la forma

$$\mathbf{C} \sim (\mathbf{R}|\mathbf{t}) \sim \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} & t_x \\ R_{yx} & R_{yy} & R_{yz} & t_y \\ R_{zx} & R_{zy} & R_{zz} & t_z \end{pmatrix}, \quad (7)$$

esta matrix transforma los puntos dados por \mathbf{x} en los puntos \mathbf{y} que explícitamente están dados por

$$\mathbf{x} \equiv \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} \quad \wedge \quad \mathbf{y} \equiv \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad (8)$$

sincronizando nuestras ecuaciones, las componentes x , y y z del vector \mathbf{y} son las componentes observadas en las Ecs. (3) y (4).

En las últimas dos décadas, la proliferación de cámaras de pequeño porte en laptops, teléfonos celulares y cámaras de vigilancia con lentes de tamaño extremadamente reducido incorporan en las imágenes distorsiones. Las más importantes suelen ser las distorsiones radiales y tangenciales, cuyas correcciones han sido establecidas

en librerías como el OpenCV [11]. Las cuales se obtienen como un proceso de análisis de problema inverso a partir del proceso de la calibración de las cámaras que permite obtener los elementos de las matrices intrínsecas y extrínsecas de las cámaras y con estos parámetros y con una nueva imagen de interés se obtiene los parámetros de distorsión radiales y tangenciales que permiten corregir los puntos de las imágenes de la forma siguiente [11],

$$x_c = x(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (9)$$

$$y_c = y(1 + k_1r^2 + k_2r^4 + k_3r^6), \quad (10)$$

donde x_c e y_c son las coordenadas corregidas de la distorsión radial, $r^2 = x^2 + y^2$, k_1 , k_2 y k_3 son los parámetros de corrección radiales,

$$x_c = x + [2p_1xy + p_2(r^2 + 2x^2)] \quad (11)$$

$$y_c = y + [p_1(r^2 + 2y^2) + 2p_2xy], \quad (12)$$

donde x_c e y_c son las coordenadas corregidas de la distorsión tangencial, p_1 , p_2 son los parámetros de corrección tangenciales.

El OpenCV proporciona los parámetros de corrección radiales y tangenciales a través de un vector que posee el siguiente orden

$$\text{Distorsion Coef} = [k_1, k_2, p_1, p_2, k_3]. \quad (13)$$

3. Resultados y discusión

Hemos utilizado una cámara Thorlabs DCC1545M-GL con un sensor CMOS de 1280×1024 pixeles, tamaño del pixel cuadrado es de $5.2 \mu\text{m}$. El lente utilizado tenía una distancia focal nominal de 25 mm. Hemos utilizado el modelo del tablero de ajedrez de 8×8 cuadrados para el proceso de calibración, cada cuadrado del tablero impreso tiene 25 mm de lado, se selecciona el cuadrado de 7×7 cuadrados, dejando los cuadrados fronterizos fuera del proceso de reconocimiento. El OpenCV recomienda hacer esto, y el OpenCV tiene un algoritmo que permite calcular los puntos de intersección de los cuadrados de forma automática usando un *notebook* que se usa en el entorno Jupyter y Python [13], desarrollado por el OpenCV [11]. El programa de Matlab denominado *Camera Calibration Toolbox for Matlab* desarrollado por Jean-Yves Bouguet [12] permite seleccionar cada uno de los puntos de intersección manualmente, lo que hace de este método muy tedioso cuando se utilizan para la calibración 50 imágenes o más.

El resultado de la calibración se encuentra en la determinación de la matrix de la cámara y los coeficientes de distorsión dados como

```
camera matrix:
[[ 5.07433786e+03  0.00000000e+00 -2.54399606e+02]
```

```
[ 0.00000000e+00  5.18310921e+03  6.11468707e+02]
[ 0.00000000e+00  0.00000000e+00  1.00000000e+00]]
distortion coefficients: [-1.65722894e+00  1.30069808e+01 -8.12361166e-03
9.50135201e-02 -7.35736992e+01]
```

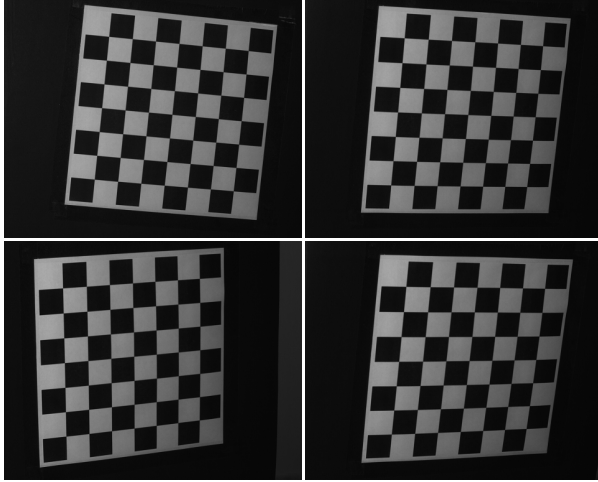


Figura 4: Se muestran cuatro imágenes de las 50 de un tablero de ajedrez que ha sido rotado, trasladado y torsionado para cada toma.

Estos resultados están en pixeles, multiplicando por $5.2 \mu\text{m}/\text{pixel}$ y dividiendo entre 1000 se obtiene los parámetros de cámara en mm.

$$\begin{pmatrix} 26.385 & 0 & -1.321 \\ 0 & 26.952 & 3.177 \\ 0 & 0 & 1 \end{pmatrix} \quad (14)$$

se observa que el centro del sensor y el centro óptico de la lente están separados por $(-1.321, 3.117)$ mm, el foco nominal de la lente es 25 mm y el foco efectivo calibrado es 26.385 mm. Los coeficientes de distorsión tangencial son $k_1 = -8.62 \mu\text{m}$, $k_2 = 6.765 \mu\text{m}$, $k_3 = -38.26 \mu\text{m}$ y los coeficientes de distorsión tangenciales son $p_1 = 42.24 \mu\text{m}$ y $p_2 = 49.40 \mu\text{m}$. Con estos parámetros se realiza la corrección a la imagen que se desee, aquí hemos denominado una imagen de control. En la Figura 5 se muestran la imagen con distorsión y la imagen con las distorsiones corregidas.

El proceso de calibración puede hacerse de manera autoconsistente en el sentido de ir eliminando las figuras de calibración que dan mayor error en la determinación de los parámetros y usando los parámetros determinados del corrimiento entre el centro óptico del sistema de lentes y el centro del sensor. Este problema es difícil de corregir puesto que dependen del ensamblaje del fabricante de la lente y las ranuras del C mount (montaje C) [6] que usan los fabricantes de las cámaras y de los lentes, a pesar de tener una medida estándar, siempre los corrimientos del orden de las décimas de μm existen.

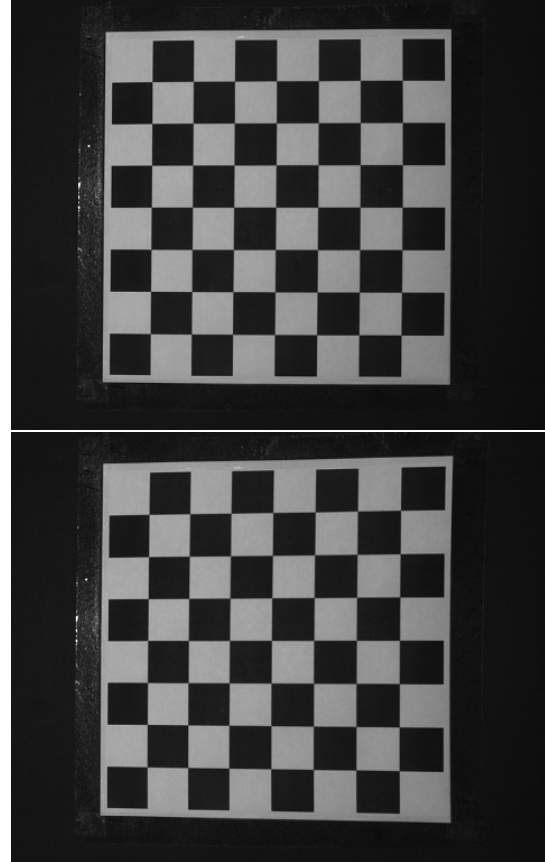


Figura 5: Se muestran, en la parte superior, la imagen con distorsiones tangenciales y radiales. En la parte inferior, la imagen con las distorsiones corregidas.

Este proceso de calibración se tiene que realizar siempre que se coloque una lente nueva a la cámara. Y si se usa lentes con focos variables se tiene que se seleccionar un foco dado, ajustar bien los tornillos de focalización, calibrarlo, usarlo en la adquisición de imágenes y corregir las mismas de las distorsiones mencionadas en este trabajo. Si se mueve uno de los tornillos de focalización, varía el foco, en este caso, hay que realizar todo el proceso de ajuste, calibración, adquisición y corrección de las distorsiones, nuevamente.

4. Conclusiones

El análisis de la cámara estenopeica usando las coordenadas homogéneas permite entender la reducción de la geometría 3D a 2D en el plano del sensor de las cámaras

digitales y permite calcular los parámetros intrínsecos de la cámara. Los parámetros externos están asociados a la matriz de rotación y el vector de traslación de los objetos con respecto al sistema de referencia del cual se observan los objetos. Con estos parámetros calculados se procedió a disminuir las distorsiones radiales y tangencial de las imágenes adquiridas.

5. Comentarios finales

Hemos aprendido una técnica de calibrar las cámaras y los lentes de un sistema de adquisición de imágenes usando las librerías del OpenCV. Esto se aplica a todo sistema óptico tal como un microscopio, telescopio,

espectrómetro óptico y el sistema de adquisición de imágenes que poseen estos equipos. El grupo Lab2D4, ha dado prioridad a este trabajo en la medida de que existe el interés de construir nuestro propios equipamientos de adquisición de datos en física mesoscópica y la fidelidad de los datos o de la información es uno de las prioridades del Lab2D4.

6. Agradecimientos

Los autores agradecen el soporte financiero proporcionado por el VRIP-UNMSM a través de los contratos B17131161, B18131561 y B19131021. P.H.R. agradece el apoyo financiero del RENACYT-CONCYTEC.

Referencias

- [1] <https://docs.opencv.org/4.1.1/d1/dfb/intro.html>.
- [2] https://en.wikipedia.org/wiki/Pinhole_camera.
- [3] https://es.wikipedia.org/wiki/Cámara_estenopeica.
- [4] https://en.wikipedia.org/wiki/Pinhole_camera_model.
- [5] Fowles, G. (1989). *Introduction to Modern Optics*, Second edition, Dover Publications, Inc., New York.
- [6] https://en.wikipedia.org/wiki/C_mount.
- [7] https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html.
- [8] https://en.wikipedia.org/wiki/Homogeneous_coordinates.
- [9] Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*, draft September 3, 2010. Publicado por Springer.
- [10] https://en.wikipedia.org/wiki/Pinhole_camera_model.
- [11] https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_calib3d/py_calibration/py_calibration.html.
- [12] http://www.vision.caltech.edu/bouguetj/calib_doc/index.html#system.
- [13] <https://jupyter.org>.