3D Transformation Matrix Calculation and Pixel Intensity Normalization for the Dual Focus Tracking System

Kenneth Sutherland, Toshiyuki Hamada, Masayori Ishikawa, Naoki Miyamoto, Masahiro Mizuta, Hiroyuki Date & Hiroki Shirato

VOLUME 39 · NUMBER 6 · DECEMBER 2019
ISSN 1609-0985 · e-ISSN 2199-4757

Journal of Medical and Biological Engineering

AN OFFICIAL JOURNAL OF TAIWANESE
SOCIETY OF BIOMEDICAL ENGINEERING

Springer

Springer

ORIGINAL ARTICLE

3D Transformation Matrix Calculation and Pixel Intensity Normalization for the Dual Focus Tracking System

Kenneth Sutherland[1] · Toshiyuki Hamada[2] · Masayori Ishikawa[1,4] · Naoki Miyamoto[1,5] · Masahiro Mizuta[3] · Hiroyuki Date[4] · Hiroki Shirato[1,5]

Abstract

Purpose We provide details of the dual-focal 3D tracking and pixel intensity calibration (DuFT) system used to record gene expression in targeted areas on the skin of freely moving mice.

Methods The accuracy of the 3D position calculation was determined by placing a scintillator on the calibration disk at various known locations within the recording cage. The height of the disk was varied from the bottom ($Z = 0$ cm) to the top ($Z = 30$ cm). The distance from the central axis range from near the center ($R = 1$ cm) to near to edge ($R = 5$ cm).

Results The mean deviation between the known and calculated position was $.31 \pm 0.16$ mm. The maximum deviation was less than $.86$ mm.

Conclusion The results indicate that the location of a scintillator within the recording cage imaged with two cameras can be calculated with submillimeter accuracy. We hope that our methods can be applied to improve automatic (even real-time) tracking of various animals in vivo.

Keywords Circadian rhythm · In vivo imaging · Fluorescence imaging · Multi-camera tracking · 3D tracking · 3D interpolation

✉ Toshiyuki Hamada
toshi-ha@iuhw.ac.jp

✉ Masayori Ishikawa
masayori@med.hokudai.ac.jp

1 Global Station for Medical Science and Engineering, Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, Sapporo, Japan

2 International University of Health and Welfare, Ibaraki, Japan

3 Laboratory of Advanced Data Science, Information Initiative Center, Hokkaido University, Sapporo, Japan

4 Graduate School of Health Sciences, Hokkaido University, Sapporo, Japan

5 Hokkaido University Hospital, Sapporo, Japan

1 Introduction

Gene expressions in many tissues of the body show circadian rhythm [1]. To examine the functional role of these genes, including clock genes *Period* (*Per*) *1*, *Per2* and *Bmal1* for the circadian system throughout the body, previous reports have indicated that the mice have to be sacrificed in experiments in order to measure basal mRNA levels from slice samples [2, 3]. Other methods of obtaining in vivo clock gene expression involve placing the mice on their backs for monitoring [4] or require anesthesia [5–7]. Such invasive restrictions of movement may affect the natural circadian rhythm that is being studied.

A method to measure gene expression in vivo in freely moving mice was subsequently developed by our group, which we call dual focus tracking (DuFT). In DuFT, bioluminescent signals representing *Per1* expression are produced following intraperitoneal administration of luciferin through an application system into *Per1-luciferase (Per1-luc)* genetically modified mice. Scintillators attached to the skin on the abdomen or the head are used to detect and track regions

of interest (ROIs); namely the skin, OB, ears and cerebral cortices.

Although tracking is usually performed later during post-processing on captured video files, tracking can be performed in real-time during the experiment. The advantage of real-time tracking is that video files do not have to be stored, greatly reducing the amount of output data. Real-time tracking can be performed continuously 24/7, recording subtle changes that can be missed when recording each half hour. The disadvantage of real-time tracking is that if there is a problem, such as poor choice of pattern match model, the data cannot be re-processed and must be completely discarded. This is why we chose to perform analysis on captured video files. The high speed of modern computer hardware makes processing time essentially a non-issue for real-time data capture and analysis.

2 Materials and Methods

The video capture software controls two cameras with a timer running on a personal computer. Video is stored in uncompressed audio video interleave (AVI) files. Video data can be acquired continuously 24 h per day, but is normally set to acquire only 10 min each half hour. This is sufficient to track changes in gene expression over several days and reduces the amount of video data. Our experiments lasted from several days to two weeks. After the video files (96 10-min pairs of video files per day) have been captured, tracking and analysis is performed.

Our tracking software is based on the real-time tumor tracking (RTRT) system used in clinical X-ray cancer treatment [9], developed by Mitsubishi Corp., Japan. RTRT utilizes stereo fluoroscopy to track a tiny (usually 2 mm diameter) spherical gold marker surgically implanted near the tumor. The marker serves as a surrogate for tumor motion, usually in the lungs or abdomen. Two X-ray cameras (flat panel or image intensifier) capture a pair of images. Two straight lines passing from the two X-ray point sources through the images of the marker on the left and right cameras are calculated. The 3D coordinates of the marker are determined by calculating the intersection of these two lines. If the lines do not intersect exactly, the point nearest the two lines is chosen.

In the DuFT system, instead of spherical gold markers, scintillators are used to locate the position of a ROI on the mouse's body. A representative image of a scintillator is used as a model for pattern matching. Because the intensity of pixels within the ROI depends on the distance from the camera, the 3D position of the ROI is necessary to normalize the pixel intensity. Pre-computed normalization tables are used for this purpose. The creation of these tables is described later in this paper.

In addition to the video camera timer, the recording stage light is controlled with another timer, normally 12 h of light and 12 h of darkness. Because mice are nocturnal animals, they are more active when the lights are off. In order to track the mice in darkness, an infrared light source and two highly sensitive EM-CCD cameras (Hamamatsu Photonics, Japan) are used. The cameras must be water-cooled to reduce heat noise from the electronics.

Our method is unique from other automatic mouse tracking applications [10, 11] in that it can determine the 3D coordinates within the mouse cage of a ROI on the body of freely moving mice. Dankert et al. have developed a method based on machine vision for automatically measuring aggression and courtship in fruit flies (*Drosophila melanogaster*) [12]. Khan et al. developed an outdoor 3D visual tracking system for the study of spatial navigation and memory in rhesus monkeys [13]. To the best of our knowledge, ours is the first system that can detect the 3D position of a scintillator placed on freely moving mice and record normalized pixel values within a ROI based on the distance between the ROI and the camera.

We used part of the Mitsubishi RTRT source code for the DuFT system, namely the file *doutaimath.cpp*, included with this paper as Supplementary Software. The code should be general enough to be applicable for a wide variety of triangulation problems in which the 3D coordinates of a point needs to be determined from a pair of 2D coordinates. In this paper, which is a follow-up to our earlier paper [8], we concentrate on the 3D perspective transformation matrix calculation and pixel intensity normalization tables. We describe in detail the mathematical theory behind the 3D perspective transformation matrix calculation and pixel value normalization algorithm. Our hope is that our methods can be applied to other 3D tracking applications, even to applications not related to mice or tumor tracking.

3 Creating the Perspective Transformation Matrix for the RTRT System

In this section, we describe in detail the process of perspective transformation in the RTRT system, on which DuFT is based. Details of the RTRT calibration process have not previously been published in the scientific literature. In order to calculate three-dimensional coordinates, it is necessary to accurately know the spatial coordinates around the isocenter of the rotating gantry of the therapeutic X-ray linac within the patient treatment room. In general, since the perspective transformation is represented by a matrix having 12 elements, if a correspondence can be obtained with respect to the spatial coordinates and fluoroscopic image coordinates, a point in space can be obtained from two arbitrary 2D points on a pair of fluoroscopic images.
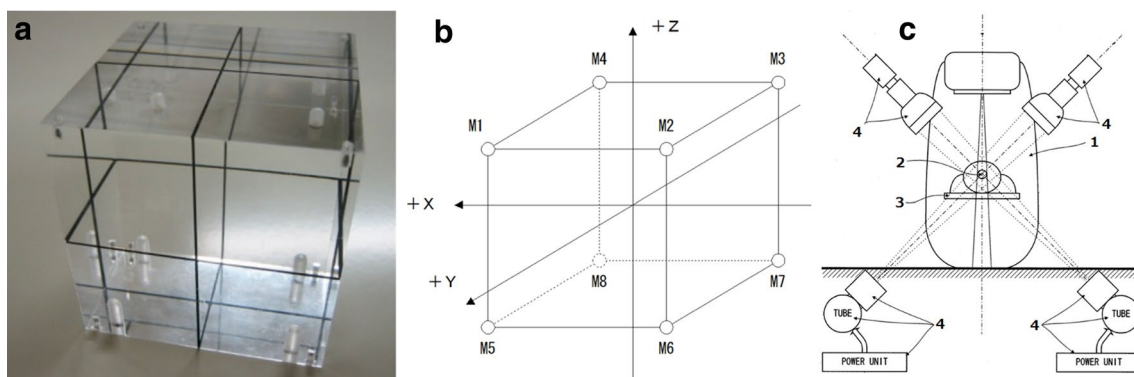
Fig. 1 Calibration of the RTRT system. **a** The RTRT perspective transformation matrix calibration cube. **b** The treatment room coordinate system with marker positions [14]. **c** Schematic of an RTRT equipped treatment room with (1) linac, (2) isocenter, (3) patient couch and (4) Two X-ray power units, X-ray generators and image intensifiers [9]

For calibration, the RTRT system uses an acrylic cube with eight gold markers embedded in each corner, 58 mm on each side (Fig. 1a). The cube is placed at the isocenter and two fluoroscopy images are obtained. The user specifies the 2D coordinates of the markers on two fluoroscopic images by clicking on the images. Because the actual 3D coordinates of the markers embedded in the calibration cube are known, the transformation matrix can be computed.

The acrylic calibration cube is placed at the isocenter of the linac gantry. Referring to Fig. 1b, the extension of the line connecting M4 and M6 is seen through fluoroscope camera 1; M3 and M5 are seen through fluoroscope camera 2; the extension of the line connecting M1 and M7 passes through fluoroscope camera 3; and fluoroscopic camera 4 exists on the extension of the line connecting M2 and M8. A schematic of an RTRT equipped treatment room (showing two of the four fluoroscopy units) is shown in Fig. 1c.

The coordinate transformation calibration dialog box (Fig. 2) in the RTRT software is used to specify six calibration points. A perspective transformation matrix is required for each fluoroscope camera. The perspective transformation matrix creation procedure for the RTRT system is as follows:

(1) Select the X-ray camera for which you want to create a matrix.
(2) Acquire a calibration image from the fluoroscope camera.
(3) Click on the 6 marker points on the calibrator image following the guidance text.
(4) Calculate and save the transformation matrix in a text file.
(5) Repeat for each remaining camera.

The user clicks on six marker points visible around the edge of the calibration cube (Fig. 2a) using the coordinate system guide (Fig. 2c) and calibration step guide (Fig. 2d).
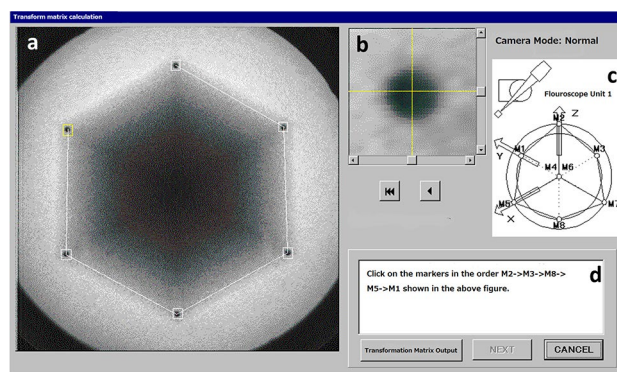

Fig. 2 The RTRT coordinate transformation calibration dialog box [14]. **a** Real-time fluoroscopy image of calibration cube with gold markers visible around the edge. **b** Detailed mouse click position of selected (yellow) marker. **c** Isocenter coordinate system schematic. **d** Calibration step guidance text

Points are specified in clock-wise order. The exact mouse click position on the marker can be adjusted by moving the cross-hairs shown in Fig. 2b. The two markers obfuscated at the center of the calibration cube should roughly align with the currently selected fluoroscope unit. The user does not click on these two points. When specification of the 6 periphery points is completed the transformation matrix is computed and can be stored in text form (Windows INI format). The procedure is repeated for fluoroscope units 2 to 4.

Note that some RTRT systems utilize two fluoroscope units instead of four. In this case, the X-ray source and cameras must be moved to avoid the linac gantry. The transformation matrices must thus be recomputed for each fluoroscope camera position, so four transformation matrices are still required even though there are only two fluoroscope units.

A mathematical description of how 3D coordinates are obtained from a pair of 2D coordinates can be found

in Supplementary Note 1 [14] and a C source code implementation developed by Mitsubishi Corp. is available upon request. To use the source code, a matrix containing the six 2D coordinates that the user specified by clicking on the marker positions on the fluoroscope image is passed to the function *m_calc()*. The known 3D coordinates of all eight markers are specified in the source file *doutaimath. cpp*. Note that the order in which the 2D coordinates are specified should correspond to the order in which the 3D positions are specified in the source code, so that each pair of 2D positions corresponds to the correct 3D marker position. A three by four matrix is returned and should be stored by the user program. In principle, the transformation matrices only need to be computed once, when the fluoroscope units are installed, but in practice, the transformation matrices are re-calibrated as part of regularly scheduled maintenance of the RTRT system.

4 Determination of 3D Coordinates from Two 2D Coordinates

Once the transformation matrices have been computed and stored, the system is ready to convert a pair of 2D coordinates of a gold marker, determined by pattern matching, into a single 3D coordinate. The two perspective matrices corresponding to the two fluoroscope units, along with the pair of 2D coordinates from the left and right cameras are passed to *sp_pos_calc()*, which returns the resulting 3D coordinates.

The function *sp_pos_calc2()* can also be called to compute the common perpendicular. This is the nearest distance between the two lines projected from the X-ray source to the marker point on the left and right camera images, if they do not intersect exactly. This value is useful for determining the accuracy of the transformation. In general, a small value is better, indicating a near-perfect interception of the two projection lines.

Note that the RTRT system processes video data in real time at 30 frames per second, so the calculations must complete in less than 33 ms. For applications (like DuFT) that post-process captured video, the computation speed is less critical.

5 Coordinate System Transformation Matrix Calibration for DuFT

An over-view of the DuFT system is described in [8]. Briefly, two high-sensitive EM-CCD cameras (Hamamatsu Photonics, Japan) are mounted and their angles are adjusted so that their focal point is approximately 1 cm above the center of the mouse cage. Infrared LED lights are mounted above the recording stage inside the imaging box.
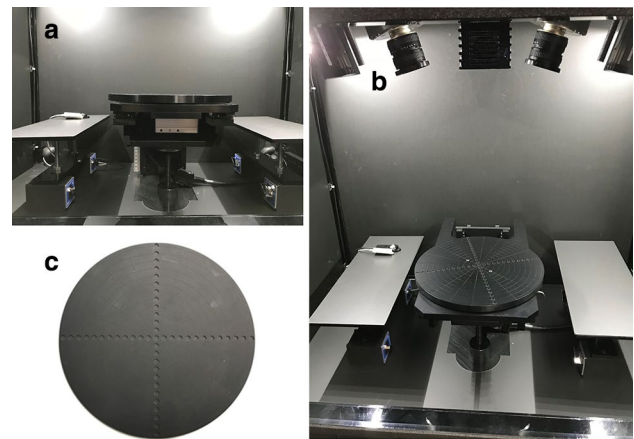

Fig. 3 The DuFT calibration disk within the recording box. **a** Side view, **b** oblique view and **c** top view

The DuFT System was developed jointly by Nakamura Service Inc. Sapporo, Japan (http://www.nakamura-service.com/) and Mateo Inc. Sapporo, Japan, in collaboration with Hokkaido University. The isocenter is regarded as the focal point of the two cameras. Instead of a calibration cube, a calibration disk is used to measure the location of calibration points. The disk is placed on a stand at the center of the cage and can be rotated and raised up and down (Fig. 3a, b). The disk contains several wells which hold scintillator material on a radiant spaced each centimeter from the center (Fig. 3c). Normally, only a single scintillator is placed in a well on the calibration disk, and multiple images are obtained by rotating and changing the height of the disk.

To compute the coordinate transformation matrix, two images (from the left and right cameras) are obtained with the scintillator positioned at eight locations, at the maximum radius (R = 5 cm) rotating 90° at the bottom (Z = 0 cm) and top (Z = 4 cm), analogous to the eight gold markers embedded in the corners of the RTRT calibration acrylic cube. A total of 16 calibration images are acquired. The 2D coordinates of these images are located using pattern matching (Fig. 4), and used along with the known 3D coordinates of the calibration positions (specified in the source code) to compute the coordinate transformation matrix.

The scintillator positions are located by pattern matching. A representative image of a scintillator is used as a pattern match model. The model is searched in eight pairs of images and the 2D coordinates are recorded. These coordinates, along with the known 3D positions of the eight calibration points are used to compute a transformation matrix for each camera. For clarity, only four of the eight pairs of calibration points, taken at Z = 45 mm above the cage floor, are shown.
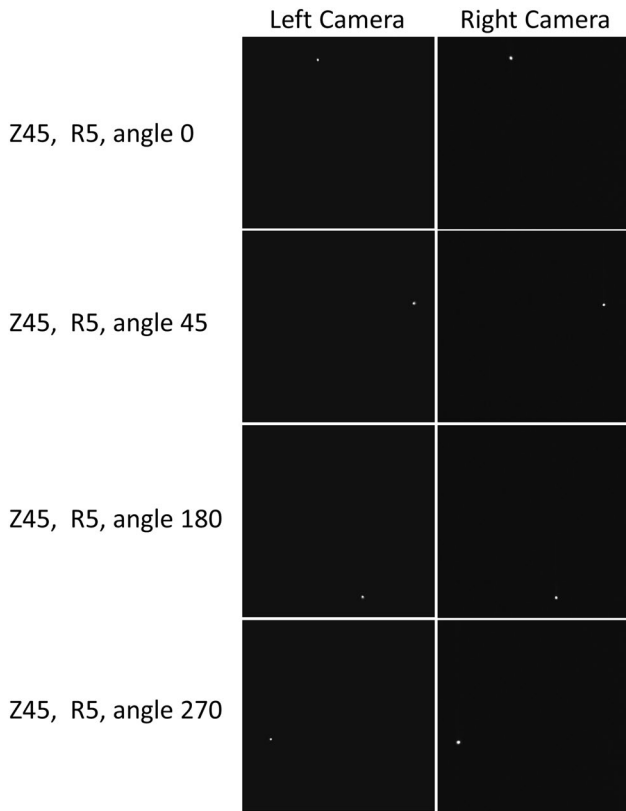
Left Camera   Right Camera

Z45, R5, angle 0

Z45, R5, angle 45

Z45, R5, angle 180

Z45, R5, angle 270

Fig. 4 Locating calibration points in left and right camera images

6 Calibration of Pixel Intensity Normalization Tables

The apparent pixel intensity of a ROI depends on its distance from the camera; the intensity decreases with the square of the distance. In order to obtain the normalized pixel intensity, which does not depend on the position of the ROI within the cage, images of a scintillator are captured at numerous calibration points distributed throughout the observing cage.

A single scintillator is placed in a well on the calibration disk and pairs of images are captured at each calibration point. The calibration disk can be rotated an angle (A), the distance (radius) of the well holding the scintillator from the center (R) and raised up and down within the cage (Z) determine the coordinate system (A, R, Z). Calibration points are taken every 15° (24 positions per radius), each cm from the center (5 positions) and each cm from the bottom to the top of the cage (5 positions), plus five positions on the vertical central axis for a total of 605 pairs of calibration images. The calibration points form a cylinder illustrated in [8] Supplementary Fig. 5a–c.

These images are automatically read by a program that generates the normalization tables. The position at the focal point of the two cameras ($R = 0$, $Z = 1$ cm) is assigned a relative intensity of exactly 1. The relative pixel intensity of scintillators at other points is stored in a table. In general,
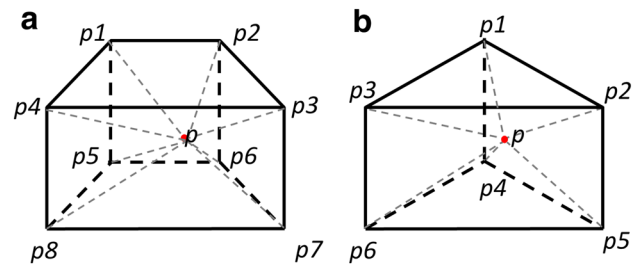
Fig. 5 Recording space interpolation method. **a** Interpolation trapezoid assuming $R = 1$ to 5 and $Z = 0$ to 4 cm. The distance from the arbitrary point (p) to each of the eight calibration points ($p1$ to $p8$) forming an isosceles trapezoid volume is computed and used to linearly interpolate the pixel intensity. **b** Interpolation isosceles triangle volume if the point is near the central vertical axis (R < 1)

positions that are nearer to the camera will have a value greater than 1 and positions further from the camera will have a value less than 1. For convenience, a separate table is used to store the relative pixel intensities of the five calibration points along the central vertical axis of the cage ($R = 0$, $Z = 0$ to 4 cm). The pixel value intensity calibration tables are used to convert raw pixel values into normalized pixel values. Test results of the pixel intensity normalization can be found in [8] Supplementary Fig. 5d–f.

7 Interpolation of Raw Pixel Intensity to Obtain Normalized Pixel Intensity

Assuming the point is within the calibration cylinder, but not at the central vertical axis ($R = 1$ to 5, $Z = 0$ to 4), an arbitrary ROI point (p) will lie within an isosceles trapezoid volume formed by eight calibration points ($p1$ to $p8$) (Fig. 5a). The distance from p to the eight calibration points is computed. The raw intensity value is interpolated linearly based on these eight distance values to yield a normalized pixel intensity value according to the following algorithm:

Let p be an arbitrary point within the trapezoid volume formed by eight ($n = 8$) calibration points $p1$ to $p8$. Let v be the raw pixel intensity of point p and $r1$ to $r8$ are the calibration values of point $p1$ to $p8$, respectively. Furthermore, let $d1$ to $d8$ be the geometric distances from point p to calibration point $p1$ to $p8$, respectively. Let d_{min} be the distance from point p to the nearest calibration point, r_{min} be the calibration value and i_{min} be the index of that point. Finally, let $v'_{sum} = 0$. Then;

For each i in n ($i \neq i_{min}$),

$$d_{tot} = d_i + d_{min},$$

$$\alpha = \frac{d_i}{d_{tot}},$$

$$v'_i = \alpha r_{\min} + (1 - \alpha)r_i,$$

$$v'_{sum} + = v'_i.$$

Then the interpolated intensity value *v'* can be calculated with

$$v' = \frac{v}{\frac{v'_{sum}}{n-1}}.$$

If the point is near the central vertical axis (R < 1), then the point will lie within an isosceles triangle instead of a trapezoid volume (Fig. 5b). In this case, six calibration points (*n* = 6) instead of eight are considered and the interpolation algorithm is the same.

If the point lies outside the calibration cylinder, for example, when the mouse is in a corner of the cage, the frame should be discarded. Frames for which Z < 0 or greater than the top of the cage are always rejected because the mouse obviously cannot maneuver below the cage floor or above the ceiling. This is usually caused by an error in pattern matching. A number of other rejection criteria can be specified so that only video frames for which a reliable normalized pixel value are considered for further analyses, as described in the following section.

8 Pattern Matching

Matrox Imaging Library (MIL) [15] version 9 is used for pattern matching. A small (usually 16 by 16 pixels) representative image of the scintillator is used as the pattern model. The model is loaded and pre-processed at startup to improve pattern matching efficiency at run-time. A single model image can be used for both left and right camera images, or separate models can be created to improve pattern matching accuracy.

Two scintillator sizes, large (3 mm diameter) and small (2 mm diameter) were used in our experiments. Large scintillators were used for the skin ROI and small scintillators were used for the head ROI (around the OB). For each frame of video, three instances each of the large and small scintillators were searched for within the left and right camera images. Large scintillators were searched for first, so they were numbered 1 to 3. Next, three instances of the small scintillator were search for, numbered 4 to 6.

Because the approximate actual sizes of the triangles formed by the three scintillators on the mouse skin and head are known, pattern matching misses can be detected by rejecting video frames in which the triangle fails to meet a number of user specified conditions. For example, the base of the skin triangle should be within a certain range of values (measured in pixels). If the measured value is out of this range, the pattern recognition has probably misidentified one or more of the three scintillators and the frame should be rejected. The value of these thresholds must be determined by the user by observing the pattern matching progress and re-running the video tracking. Once reasonable thresholds are determined for an experiment setup, assuming that the size of the mouse and placement of the scintillators are fairly constant, these threshold values can be used in subsequent experiments.

For our experiments, ten minutes of video were captured each half hour over a 4 day period, so 96 pairs (from left and right cameras) of video files were generated. Video was captured at 2 frames per second, so each video file contained 1200 frames. Although many frames had to be discarded due to pattern matching errors, etc., usually enough good frames were found in order to compute an average normalized pixel intensity value and standard deviation for each hour of the experiment.

9 Results

The accuracy of 3D position calculation was determined by placing a scintillator on the calibration disk at various known locations within the recording cage. The height of the disk was varied from the bottom (Z = 0 cm) to the top (Z = 30 cm). The distance from the central axis range from near the center (R = 1 cm) to near to edge (R = 5 cm). The results are summarized in Table 1.

The mean deviation between the known and calculated position was .31 ± 0.16 mm. The maximum deviation was less than .86 mm. The results indicate that the location of a scintillator within the recording cage imaged with two cameras can be calculated with sub-millimeter accuracy.

10 Discussion and Conclusions

There are several limitations concerning this research. First, our system can only track a single animal. Tracking two or more mice may be possible, but the problem of over-lapping positions leading to misidentification is non-trivial [16]. One possible solution would be to place small scintillators on one test subject and large scintillators on another.

Second, the user interface for the 3D transformation matrix calculation is still under development. Presently, the source code (*doutaimath.cpp*) must be modified, and the program recompiled in order to specify the 3D positions of the calibration scintillators. In the future, we intend to add a dialog box, in which the eight 3D calibration positions can be specified by the user. Furthermore, the output of the software is primarily a single comma separated value (.csv) file

Table 1 Distance (in mm) from actual and calculated position

Z (cm)	R (cm)	Angle			
		0°	90°	180°	270°
0	1	0.575938	0.437850	0.480479	0.376609
	2	0.554148	0.401749	0.469195	0.355798
	3	0.512997	0.280294	0.461750	0.236870
	4	0.485206	0.158468	0.251871	0.120195
	5	0.398948	0.208542	0.075437	0.457449
1	1	0.371877	0.231496	0.381951	0.142660
	2	0.356882	0.235951	0.385983	0.102531
	3	0.311305	0.260189	0.250399	0.265019
	4	0.354713	0.206182	0.183660	0.275149
	5	0.206512	0.169250	0.165956	0.577699
2	1	0.118089	0.143325	0.315951	0.21384
	2	0.172633	0.259534	0.352222	0.206674
	3	0.115780	0.233663	0.372008	0.260975
	4	0.148344	0.162739	0.380180	0.413052
	5	0.134070	0.092160	0.858486	0.431881
3	1	0.098036	0.283643	0.316664	0.431408
	2	0.072870	0.280306	0.370202	0.466659
	3	0.141897	0.338631	0.419679	0.600403
	4	0.251526	0.258550	0.585170	0.511896
	5	0.452027	0.291637	0.655845	0.590674

containing the 3D positions and normalized average pixel values. Further data analysis is performed with a suite of console programs that we have developed for our research. We need to combine the capture, tracking and analysis programs into a single, coherent application.

When tracking without scintillators, the OB, left and right ears, left and right cortex, etc., must be specified as offsets within the head model image. At present, these values have to be typed into a text input file (Windows INI format), which is burdensome for the user. In the future, our intention is to develop a graphical interface for specifying the location, size and other properties of each ROI.

Although only eight pairs of images are required to compute the two 3D perspective transformation matrices, we had to acquired 605 pairs of images in order to generate the pixel intensity normalization tables (see [8] Supplementary Fig. 5). Such highly detailed tables may not be necessary. Especially near the central vertical axis of the cage (R = 1 cm), because the calibration points were generated by rotating the disk each 15 degrees, the 24 calibration points are packed closely together (2.6 mm between points) on the plane of the disk. Around the outer edge of the disk (R = 5 cm), on the other hand, the 24 points are spaced 13.1 mm apart. We should distribute the calibration points more evenly within the cage, roughly every 10 mm, reducing the number of pixel value calibration points.

Details of the 3D perspective transformation matrix calculation and pixel value calibration tables have been described in this paper. We hope that our methods can be applied to improve automatic (even real-time) tracking of various animals in vivo confined to a cage or enclosure. Future applications may include tracking of larger animals such as monkeys [13], or even smaller test subjects such as fruit flies [12].

Finally, our method of tracking and observing freely moving animals provides a means of more ethical research in that test subjects do not have to be sacrificed to obtain quantitative data, reducing the overall number of animals used. More sophisticated research tools such as DuFT should be employed by all research institutions conducting animal experiments [17].

Compliance with Ethical Standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical Approval Institutional Review Board approval was obtained.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Young, M. W., & Kay, S. A. (2001). Time zones: A comparative genetics of circadian clocks. *Nature Reviews Genetics, 2,* 702–715.
2. Abe, M., et al. (2002). Circadian rhythms in isolated brain regions. *Journal of Neuroscience, 22,* 350–356.
3. Yamazaki, S., et al. (2000). Resetting central and peripheral circadian oscillators in transgenic rats. *Science, 288,* 682–685.
4. Tahara, Y., et al. (2012). In vivo monitoring of peripheral circadian clocks in the mouse. *Current Biology, 22,* 1029–1034.
5. Sakamoto, A., et al. (2005). Influence of inhalation anesthesia assessed by comprehensive gene expression profiling. *Gene, 356,* 39–48.
6. Ohe, Y., Iijima, N., Kadota, K., Sakamoto, A., & Ozawa, H. (2011). The general anesthetic sevoflurane affects the expression of clock gene mPer2 accompanying the change of NADt level in the suprachiasmatic nucleus of mice. *Neuroscience Letters, 490,* 231–236.

7. Zhang, W., et al. (2001). Rapid in vivo functional analysis of transgenes in mice using whole body imaging of luciferase expression. *Transgenic Research, 10,* 423–434.

8. Hamada, T., Sutherland, K., Ishikawa, M., Miyamoto, N., Honma, S., Honma, K., et al. (2016). In vivo imaging of clock gene expression in multiple tissues of freely moving mice. *Nature Communications, 7,* 11705.

9. Shirato, H., et al. (2000). Physical aspects of a real-time tumor-tracking system for gated radiotherapy. *International Journal of Radiation Oncology Biology Physics, 48,* 1187–1195.

10. de Chaumont, F., Dos-Santos Coura, R., Serreau, P., Cressant, A., Chabout, J., Granon, S., et al. (2012). Computerized video analysis of social interactions in mice. *Nature Methods, 9*(4), 410.

11. de Chaumont, F., Dallongeville, S., Chenouard, N. & Olivo-Marin, J.-C. (2010) Tracking multiple articulated objects using physics engines: Improvement using multiscale decomposition and quadtrees. *Proceedings IEEE International Conference on Image Processing (Hong Kong).*

12. Dankert, H., Wang, L., Hoopfer, E. D., Anderson, D. J., & Perona, P. (2009). Automated monitoring and analysis of social behavior in Drosophila. *Nature Methods, 6,* 297–303.

13. Khan, Z., Herman, R., Wallen, K., & Balch, T. (2005). An outdoor 3D visual tracking system for the study of spatial navigation and memory in rhesus monkeys. *Behavior Research Methods, 37,* 453–463.

14. Mitsubishi Digital Radiograph RTRT Instruction Manual revision C. (2003).

15. Matrox Electronic Systems, Matrox Imaging Library (MIL) 9 User Guide, Manual, no. Y10513-301-0900. (2008).

16. Khan, Z., Balch, T., & Dellaert, F. (2005). MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27,* 1805–1819.

17. Editorial. (2011). Animal rights and wrongs. Nature *470,* 435

Supplementary Note 1

Mathematical description of the determination of 3D coordinates in the RTRT system [1]

Here, we explain in detail the calculation of the translation and rotation angle calculation using three landmark points. First, the following two basic coordinate systems are defined:

    (1) Treatment room coordinate system {A},
    (2) CT coordinate system {B}.

The CT coordinate system is represented by a simple translation with respect to the treatment room coordinate system, assuming that the translation amount is known. That is,

$$
{}^{A}_{B}T = \begin{bmatrix} 1 & 0 & 0 & \mathit{offset}_X \\ 0 & 1 & 0 & \mathit{offset}_Y \\ 0 & 0 & 1 & \mathit{offset}_Z \\ 0 & 0 & 0 & 1 \end{bmatrix}.
$$

Here, (offset_x offset_y offset_z) is the origin position of the CT coordinate system in the treatment room coordinate system. Given this, the equation can be expressed as

$$
{}^{A}P = {}^{A}_{B}T\,{}^{B}P,
$$

where ^{B}P is an arbitrary point in the CT coordinate system and ^{A}P is that point transformed by $^{A}_{B}T$ into the treatment room coordinate system by a known vector *offset*. Points ^{B}P and ^{A}P are represented in homogeneous coordinates; that is, the translation can be expressed with matrix multiplication. Next, the landmark frame $\{C\}$ is defined from three points *P*, *Q* and *R* as follows using two vectors $\overrightarrow{QP}$ and $\overrightarrow{QR}$;

$$
\overrightarrow{{}^{B}Z_C} = \frac{\overrightarrow{QP} \times \overrightarrow{QR}}{\left| \overrightarrow{QP} \times \overrightarrow{QR} \right|}
$$

$$
\overrightarrow{{}^{B}Y_C} = \frac{\overrightarrow{QP} + \overrightarrow{QR}}{\left| \overrightarrow{QP} + \overrightarrow{QR} \right|}
$$

$$
\overrightarrow{{}^{B}X_C} = \frac{\left(\overrightarrow{QP} + \overrightarrow{QR}\right) \times \left(\overrightarrow{QP} \times \overrightarrow{QR}\right)}{\left| \left(\overrightarrow{QP} + \overrightarrow{QR}\right) \times \left(\overrightarrow{QP} \times \overrightarrow{QR}\right) \right|}
$$

where $\left|\overrightarrow{QP}\right|$ is the length of vector $\overrightarrow{QP}$ and $\overrightarrow{QP} \times \overrightarrow{QR}$ is the exterior product of vectors $\overrightarrow{QP}$ and vector $\overrightarrow{QR}$.
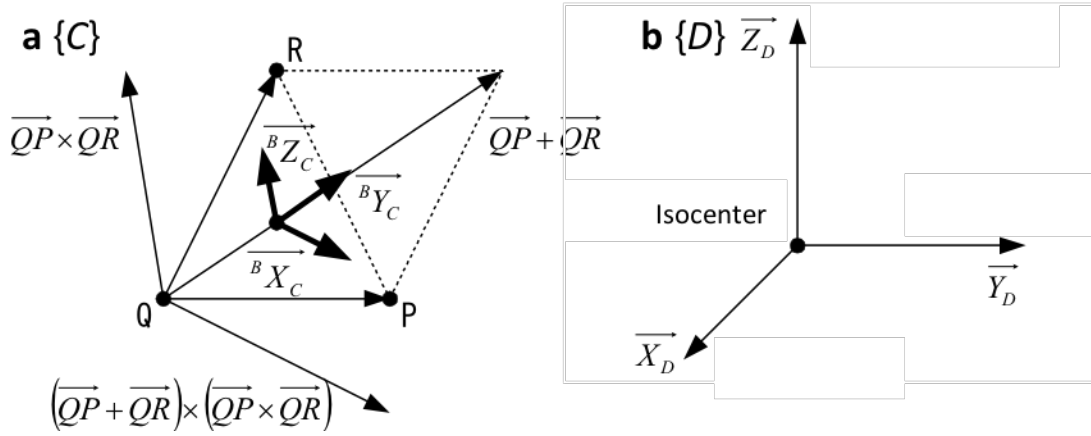
Supplementary Figure 1. The landmark frame $\{C\}$ (a) and the tumor frame of reference {D} (b).

The landmark frame of reference {C}, with the origin the center of gravity of Q, P and R, can be visualized as shown in Supplementary Fig. 1a. Furthermore, the tumor frame of reference {D} is defined as shown in Supplementary Fig. 1b. These frames of reference are defined for an ordinary linac. That is, the relationship between the treatment room coordinate system and the tumor frame of reference is related by a known set of conditions.

One of the relations, which we denote as $_{D}^{A}R$, can be obtained by setting the gantry angle φ, the rotation angle of the patient couch about the gantry isocenter ψ and the collimator angle θ, and performing the rotations in the following order:

(1) Collimator rotation

$$_{}^{A}R_{\mathrm{Z}}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
ϑ rotation about the ^{A}Z axis

(2) Gantry rotation

$$_{}^{A}R_{\mathrm{Y}}(\phi) = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix}$$
φ rotation about the ^{A}Y axis

(3) Couch rotation

$$^{A}R_{Z}(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ Equivalent to ψ rotation about the ^{A}Z axis

Therefore,

$$^{A}_{D}R = {}^{A}R_{Z}(y)\,{}^{A}R_{Y}(j)\,{}^{A}R_{Z}(q)$$

$$= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \cos\phi\cos\phi\cos\theta - \sin\phi\sin\theta & -\cos\phi\cos\phi\sin\theta - \sin\phi\cos\theta & \sin\phi\cos\phi \\ \cos\phi\sin\phi\cos\theta + \cos\phi\sin\theta & -\cos\phi\sin\phi\sin\theta + \cos\phi\cos\theta & \sin\phi\sin\phi \\ -\sin\phi\cos\theta & \sin\phi\sin\theta & \cos\phi \end{bmatrix}.$$

That is, when $^{A}_{D}R$ is known, the corresponding gantry angle φ, couch rotation angle ψ, and collimator angle ϑ can be obtained. According to the above equation,

$$\phi = \pm\cos^{-1}(r_{33}).$$

There are therefore two solutions to be considered, and it is necessary to divide the problem into subsequent cases. Here, let

$$C_{2} = \cos\phi = r_{33},\ \ C_{1} = \sin\phi.$$

Furthermore,

$$C_{4} = \cos\phi = \frac{r_{13}}{C_{1}},\ \ C_{3} = \sin\phi = \frac{r_{23}}{C_{1}},\ \ C_{6} = \cos\theta = -\frac{r_{31}}{C_{1}},\ \ C_{5} = \sin\theta = \frac{r_{32}}{C_{1}}.$$

Given that,

in case $C_{3} > 0$, $\qquad\qquad\qquad \phi = +\cos^{-1}(C_{4})$,

in case $C_{3} < 0$, $\qquad\qquad\qquad \phi = -\cos^{-1}(C_{4})$,

in case $C_{3} = 0$, when $C_{4} \geq 0$, $\qquad \phi = 0$,

and when $C_4 < 0$, $\qquad \phi = \pi$,

and

in case $C_5 > 0$, $\qquad\qquad\qquad \theta = + \cos^{-1}(C_6)$,

in case $C_5 < 0$, $\qquad\qquad\qquad \theta = - \cos^{-1}(C_6)$, $\qquad\qquad$ (1)

in case $C_5 = 0$, when $C_6 \geq 0$ $\qquad \theta = 0$,

and when $C_6 < 0$ $\qquad \theta = \pi$,

can be obtained. However, when the gantry angle is $n\pi$, that is when $\sin\varphi = \sin n\pi = 0$, ψ and ϑ in $_D^A R$ have the relationship of only a phase difference, which is shown below. Here, we consider the collimator rotation angle ϑ by adopting the planned value of the couch rotation angle ψ. If $\psi = \psi_{PLAN}$, then

$$C_4 = \cos\phi_{PLAN} \text{、} \quad C_3 = \sin\phi_{PLAN}$$

and we get

$$_D^A R = \begin{bmatrix} C_2 C_4 \cos\theta - C_3 \sin\theta & -C_2 C_4 \sin\theta - C_3 \cos\theta & 0 \\ C_2 C_3 \cos\theta + C_4 \sin\theta & -C_2 C_3 \sin\theta + C_4 \cos\theta & 0 \\ 0 & 0 & C_2 \end{bmatrix}.$$

Therefore, if we set C6 and C5 as follows,

$$C_6 = \cos\theta = \frac{r_{11}C_4 + r_{21}C_3}{C_2} \text{、} \quad C_5 = \sin\theta = -\frac{r_{12}C_4 + r_{22}C_3}{C_2},$$

we can compute θ with the same set of conditions described in equation (1).

In addition, when the gantry angle is around $n\pi$, ψ and ϑ tend to be drastically influenced by small 2D marker position errors. This is normally caused by the user not clicking exactly on the marker in the fluoroscope image. If the difference from $n\pi$ is small, an approximation that rounds down and regards it as exactly $n\pi$ is therefore practical. The orientation of $_D^A T$ after the rotational displacement of the tumor frame can be calculated as follows if the landmark frame $_C^A T$, after the corresponding rotational displacement, is known. $_C^A T$ is obtained by actual measurement of three landmark points, and the rotational displacement can be obtained by making it correspond to the

orientation of $_C^A T$ within the plan landmark frame. If $_{D'}^A T$ is found, the gantry angle, couch rotation angle, and collimator angle corresponding to this can be obtained. Furthermore, $_D^A T$ can be obtained by adding an amount of parallel translation to $_D^A R$.

First, the plan landmark frame $\{C\}$ is defined according to the CT coordinate system $\{B\}$. That is,

$$_C^B T \quad = \begin{bmatrix} \overrightarrow{^B X_C} & \overrightarrow{^B Y_C} & \overrightarrow{^B Z_C} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Accordingly,

$$_C^A T \quad = {}_B^A T \, {}_C^B T .$$

Next, the measured landmark frame $\{C'\}$ is obtained according to the treatment room coordinate system $\{A\}$,

$$_{C'}^A T \quad = \begin{bmatrix} \overrightarrow{^A X_{C'}} & \overrightarrow{^A Y_{C'}} & \overrightarrow{^A Z_{C'}} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

By these means,

$$_{D'}^A T \quad = {}_{A'}^A T \, {}_{D'}^{A'} T$$

$$= {}_{C'}^A T \, {}_C^A T^{-1} \, {}_D^A T$$

can be calculated.

(References)

[1] "Mitsubishi Digital Radiograph RTRT Instruction Manual revision C," 2003.